From *Handbuch der Physik*, Bd. V, *Grundlagen der Mechanik. Mechanik der Punkte und starren Körper*, Chap. 2, Springer, Berlin, 1927.

THE PRINCIPLES OF DYNAMICS

BY

L. Nordheim, Göttingen.

Translated by D. H. Delphenich

TABLE OF CONTENTS

I. Introduction.

1. Historical overview and literature. – After **Newton** had proposed the fundamental laws of mechanics, the ambition of mathematicians and physicists was then to find a fruitful expression for the laws of mechanics and a summary of them in terms of so-called *principles*. The goal of that initiative was two-fold: First of all, one wished to raise those principles to the status of axioms, so to replace and extend the Newtonian axioms. In a certain sense, they should summarize all of mechanics in the smallest-possible number of laws. Secondly, however, the original Newtonian laws proved to be unwieldy and partly unsatisfactory in some of the more complicated cases, such as auxiliary conditions, etc., and one would then have to find rules that would allow one to derive the equations of motion in each case simply and uniquely.

The first, more-logical, problem was already treated in the previous chapter, and for that reason, it will be touched upon only when it becomes necessary to show its connection. By contrast, here it will be the question of exhibiting the equations of motion that takes the foreground. In that way, one will simultaneously gain an important viewpoint for their integration.

The principles of mechanics can be divided into two main groups, namely, differential principles and integral principles. The former is the oldest one. The principle of virtual displacements has already been used by **Stevin**, **Galilei**, and in particular, **Johann Bernoulli**, *et al.*, in some simple cases, and was employed by **Lagrange** more generally as a basis for statics. The transition to dynamics was the contribution of **d'Alembert**, who also gave his name to the well-known principle, while once more, its systematic implementation was mainly carried out by **Lagrange**. One must further thank **Gauss** for the principle of least constraint, which represented the most far-reaching differential principle, while an intermediate formulation was found by **Jourdain** in recent years.

The development of the integral principles ran parallel to that. The first of them to appear historically was the one that is ordinarily named after **Maupertuis**, although it was first formulated mathematically by **Euler**, namely, the principle of least action. The greatest advance on it was achieved by **Hamilton** and **Jacobi**, who we also have to thank for the systematic theory of the integration of mechanical equations. One of the very interesting attempts took the form of **Hertz**'s mechanics, which sought to get by without the concept of force. **Hertz** was likewise the first to systematically address the non-holonomic systems. In general, he still did not succeed in completely clarifying the questions that were connected with them, which was first achieved by **Hölder**.

We shall refer to *Enzyklopädie* article of **Voss** ([1]) for the literature concerned with the older work. Here, we shall generally cite only the more recent literature (i.e., since 1900). **Boltzmann** ([2]) has written a thorough textbook, and there is also a brief modern presentation by **Schaefer** ([3]). Naturally, along with them, the principles are treated, more or less, in all textbooks on mechanics

([1]) **A. Voss**, *Enzykl. d. Math. Wiss.*, Bd. IV, Art. 1, pp. 1.
([2]) **L. Boltzmann**, *Vorlesungen über die Prinzipe der Mechanik*, Bd. I and II, Leipzig, 1897 and 1904.
([3]) **Cl. Schaefer**, *Die Prinzipe der Dynamik*, Leipzig, 1919.

([1]). For the history of mechanics, one must cite, above all, the book by **Mach** ([2]), as well as the one by **Haas** ([3]).

2. General overview. – In the following presentation, much weight will be given to, on the one hand, addressing all of the cases that occur in practice, i.e., giving instructions on how to exhibit the equations of motion. On the other hand, the connections between the known principles shall be clarified, i.e., we will show the extent to which they relate to each other or include new assumptions, respectively. Naturally, one cannot "prove" a principle, but one can only either show directly that it produces the equations of motion that are known to be correct from experiments or that it leads back to another one. It would also become unnecessary then for one to discuss all of the special cases in detail. Rather, it would suffice to do that in the most convenient forms and otherwise only refer to their possibility.

Those principles should probably be distinguished from the general laws of mechanics that are often misleadingly referred to as principles, such as the laws of energy, center of mass, and areas. However, they can be derived from the fundamental equations and are therefore consequences of them, while the converse is not true. It is only upon extending them to the full scope of physics that they will take on the status of principles.

The two main branches, viz., differential and integral principles, also correspond to essentially two separate domains of application. The mechanics of the differential principles, whose main points were developed by **Lagrange** ([4]), with its various types of constraint conditions and the introduction and calculation of reaction forces, is the instrument that was given to engineering mechanics. By contrast, the mechanics of the integral principles, which is mainly connected with the names of **Hamilton** ([5]) and **Jacobi** ([6]), is the method that is best-suited to the treatment of free mass-points, and therefore astronomy and the mechanics of atoms, so to modern physics. In fact, the recent development of theoretical physics is very closely linked with Hamiltonian mechanics, in the macroscopic context of the theory of relativity, as well as in the microscopic world of quantum theory. That is why we shall discuss it thoroughly in the next chapter (of this *Handbuch*).

Even though currently it is the goal of theoretical physics to reduce all of mechanics to atomic phenomena, naturally, that path is unsuited to the applications in engineering, so it will become necessary to further develop each branch of mechanics independently. However, in and of itself, that reduction can be regarded as having been resolved in an apparently satisfactory way as of now,

---

([1]) Here, we shall cite only the ones that are appropriate to our special topics: For general mechanics: **P. Appell**, *Traité de mécanique rationelle*, 3rd ed., t. 1-V, Paris, 1911-1926. **E. T. Whittaker**, *Analytical Dynamics*, 2nd ed., Cambridge, 1917, also in German translation by **F.** and **K. Mittelsten-Scheidt**, Berlin, 1924. Moreover, for the differential principles, there is also **C. H. Müller** and **G. Prange**, *Allgemeine Mechanik*, Hannover, 1923.

([2]) **E. Mach**, *Die Mechanik in ihrer Entwicklungen*, 7th ed., Leipzig, 1912.

([3]) **A. E. Haas**, *Die Grundgleichungen der Mechanik auf Grund ihrere geschichtlichen Entwicklung*, Leipzig, 1914.

([4]) **J. L. Lagrange**, *Mécanique analytique*, Paris, 1811-15.

([5]) **W. R. Hamilton**, Phil. Trans. (1834), pp. 307, and *ibid.* (1835), pp. 95. For his optical investigations: Trans. Irish Acad. **15** (1828), pp. 69; *ibid*. **16** (1830), pp. 93.

([6]) **C. G. J. Jacobi**, cf., esp., *Vorlesungen über Dynamik*, held from 1842-1843, edited by **A. Clebsch**, 2nd ed., Berlin, 1884.

and indeed for the theory of elasticity, by the theory of crystal lattices ([1]), and for hydrodynamics, by the kinetic theory of gases ([2]).

In what follows, we shall restrict ourselves to the mechanics of discrete mass-points. The mechanics of rigid bodies can always be derived from that by the introduction of auxiliary conditions, i.e., constraints that are characterized by the properties of the rigid bodies. By contrast, as was shown in the previous chapter, the transition to the mechanics of deformable continua is much more complicated and requires one to add some new axioms. What one usually refers to as the general theory of mechanics is therefore the mechanics of mass-points and rigid bodies, to which we shall restrict ourselves here.

For the sake of symmetry and simplicity of notation, unless stated to the contrary, in what follows, the coordinates of the individual mass-points shall be numbered consecutively, so for instance, the x, y, z coordinates of the first mass-point shall be denoted by x_1, x_2, x_3, those of the second by x_4, x_5, x_6, etc. The generalized force components X_i and the masses m_i shall be treated likewise. Naturally, for the latter, one always has $m_{3n} = m_{3n+1} = m_{3n+2}$. Unless stated to the contrary, we shall imagine that the X_i are independent of the velocities $\dot{x}_i$. For the sake of intuition, we will also use a configuration space of coordinates whose dimension is equal to the number of degrees of freedom f. When summing over all coordinates, we shall use i, k, or l as summation indices.

II. – Differential principles.

a) Statics.

3. The principle of virtual displacements for free systems. – If we have a system of completely-free mass-points then the conditions for equilibrium, so the ones under which all mass-points will remain at rest under the influence of forces, will read:

$$X_i = 0 \qquad (i = 1, \ldots, l). \tag{1}$$

Those f equations can be formally combined into a single one when one multiplies each of them by entirely arbitrary quantities δx_i, sums over all coordinates, and demands that:

$$\sum_i X_i \, \delta x_i = 0 \, . \tag{2}$$

The demand that the expression on the left-hand side should vanish for all arbitrary values of δx_i represents the principle of virtual displacements ([3]), and it is obviously completely equivalent to

([1]) See **M. Born**, *Atomtheorie des festen Zustandes (Dynamik der Krystallgitter)*, Leipzig, 1923; also in *Enzykl. d. Math. Wiss.*, Bd. V, art. 25, pp. 3.

([2]) **L. Boltzmann**, *Vorlesungen über Gastheorie*; **D. Hilbert**, Math. Ann. **72** (1912), pp. 562. **D. Enskog**, *Kinetische Theorie der Vorgänge in mäßig verdünnten Gasen*, Diss. Uppsala, 1917; **J. H. Jeans**, *The Dynamical Theory of Gases*, 4th ed., Cambridge, 1925; German translation by **R. Fürth**, Braunschweig, 1926.

([3]) For the meaning of the term "virtual," see no. **5**.

the system of equations (1) for free mass-points since one can, e.g., take all δx_i, except for δx_1, to be zero and then conclude from $X_1\ \delta x_1 = 0$ that $X_1 = 0$ due to the arbitrariness of X_1 .

Initially, that is a purely-formal result. However, one can also regard the δx_i as infinitesimal displacements of the i^{th} coordinate. $\sum_i X_i\,\delta x_i$ will also represent the work done by the force when it imparts that displacement on the mass-point then. Equation (2) will then be referred to as the *principle of virtual work*, and it can be expressed as follows:

A system of free mass-points is in equilibrium if and only if the work done under any arbitrary infinitesimal displacement vanishes.

The formulation is still completely meaningless at this point. However, it will prove to be convenient when the degrees of freedom in the system are subject to any sort of constraints.

4. Holonomic, non-holonomic, scleronomic, and rheonomic constraints. – The considerations of the previous number were valid only for free mass-points. (1) and (2) were completely equivalent at the time. However, constraints are often prescribed between the coordinates. Such constraints can have various forms:

1. The constraints include only the coordinates of the points, so they have the form:

$$\varphi_r\,(x_1, \ldots, x_f) = 0 \qquad\qquad (r = 1, \ldots, g). \tag{1}$$

Let the number of constraints be g. This simplest, so-called *holonomic*, form for the equations of constraint is realized by all types of rigid constraints and all motions on completely-smooth strings and surfaces. Furthermore, the constraints (1) do not contain time t explicitly, so one calls then *scleronomic*.

2. The constraints also include derivatives of the coordinates with respect to time explicitly, so they have the form:

$$\varphi_r\left(x_1,\ldots,x_f\,;\frac{dx_1}{dt},\ldots,\frac{dx_r}{dt}\right) = 0\ . \tag{2}$$

Hertz called them *non-holonomic*. That type includes, e.g., motions that arise from rolling without slipping, or restricting the motion of a blade to a surface (e.g., ice skating). In almost all cases that occur in practice, the constraints are linear in the derivatives, so they will have the form:

$$\sum_i a_{ir}(x_1,\ldots,x_f,t)\,dx_i + a_r\,dt\ = 0 \qquad (r = 1, \ldots, g). \tag{3}$$

Naturally, they differ from the constraints (1) only when those differential expressions are not integrable, i.e., they cannot be reduced to finite equations of constraint by any process of

integration. In order for that to be true, it is necessary (but not always sufficient) that one must have:

$$\frac{\partial a_{ir}}{\partial x_k} \neq \frac{\partial a_{kr}}{\partial x_i}, \qquad \frac{\partial a_{ir}}{\partial t} \neq \frac{\partial a_r}{\partial x_i}, \quad \text{resp.,}$$

for at least one pair of coefficients.

3. The constraints can also include time explicitly. An example of that type is motion on a moving surface. **Boltzmann** called them *rheonomic*. They can be holonomic, as well as non-holonomic. The most general case is then that of the non-holonomic-rheonomic constraint equations. An example of that would be rolling on a moving surface.

4. The last possibility that remains is the one in which inequalities exist, e.g., when the degrees of freedom of a mass-point are restricted by an impermeable surface in space. Inequalities can also be non-holonomic-rheonomic, so they can be put into the form:

$$\sum_i a_{ir}\, dx_i + a_r\, dt \leq 0$$

in all of the cases that come under consideration.

5. The principle of virtual displacements for constrained systems. – In order exhibit the conditions for equilibrium when such constraint equations exist, we now require an additional axiom. For free mass-points, the equilibrium conditions can be derived from the requirement that the work done under an arbitrary infinitesimal displacement should vanish. However, arbitrary infinitesimal displacements are no longer possible now since the degrees of freedom in the system are restricted by constraints. The obvious generalization of the principle is to now demand that the work done should vanish only under all infinitesimal displacements that are compatible with the equations of constraint since all other displacements are excluded from the outset. Naturally, that implies a hypothesis that can only be confirmed by experiment.

The displacements that are restricted in that way are distinguished from the most general displacements by referring to them as *virtual* (i.e., possible) displacements. For example, when a point is constrained by a surface, the virtual displacements will be the ones under which it will remain on that surface.

In order to obtain the analytical expression for the principle, we initially restrict ourselves to the finite, holonomic constraint equations (1) in no. **4**. If we always write the infinitesimal displacements with a δ, as in no. **3**, in order to distinguish them from the differential symbols d and ∂, which is an abbreviation that shall be established throughout this article, then the virtual displacements δx_i for the constraint equations (1) in no. **4** will be subject to the conditions:

$$\varphi_r (x_1 + \delta x_1, \ldots, x_r + \delta x_r) = 0 ,$$

which says just that the varied configuration will also satisfy the equations of constraint. Since the displacements should also be infinitesimal, when one develops that into a **Taylor** series and truncates it after the first term, while recalling (1) in no. **4**, one will have:

$$\sum_i \frac{\partial \varphi_r}{\partial x_i} \delta x_i = 0 \qquad (r = 1, \ldots, g). \tag{1}$$

The demand that the displacements should be infinitesimal is essentially that those equations should be linear in the δx_i.

That must be combined with the demand of the principle of virtual displacements:

$$\sum_i X_i \, \delta x_i = 0 \,, \tag{2}$$

in which the left-hand side again represents the work done under such a displacement.

Equations (1) and (2) can be combined into one equation with the use of undetermined multipliers λ_r that are usually called ***Lagrange*** *factors* since one multiplies each of equations (1) by λ_r and adds them to (2). One will then get:

$$\sum_i \left(X_i + \sum_r \lambda_r \frac{\partial \varphi_r}{\partial x_i} \right) \delta x_i = 0 \,. \tag{3}$$

Of the components of the virtual displacements δx_i in that, precisely $(f - g)$ of them can be chosen arbitrarily, while the rest of them can be determined by the auxiliary condition (1). One now imagines that the multipliers λ_r have been chosen in such a manner that the components of just the last g factors vanish. Naturally, it is irrelevant how one chooses them. $(f - g)$ terms will then remain in (3) that are now multiplied by completely arbitrary displacements. Therefore, all of their coefficients must vanish by themselves. The artifice of the Lagrange factors then allows one to infer consequences from equations (3) in the same way as when the δx_i are completely arbitrary, and one will obtain the $f = 3n$ equations:

$$X_i + \sum_r \lambda_r \frac{\partial \varphi_r}{\partial x_i} = 0 \tag{4}$$

as the equilibrium conditions. Together with the g constraint equations (1), they will suffice to determine the $3n + g$ unknowns x_i, λ_r, so they will establish equilibrium completely.

Now, it is easy to see from the argument that led up to equations (4) how one must proceed when one has non-holonomic and rheonomic constraint equations. For non-holonomic-rheonomic constraints, a relationship between the components of the virtual displacements that is analogous to (1) is already included in equations (2.a) of no. **4** (in which one temporarily sets $a_r \equiv 0$), namely:

$$\sum_i a_{ir}\,\delta x_i = 0\,, \tag{5}$$

and equations (4) will obviously be true when takes the corresponding coefficients a_{ir} in place of $\partial\varphi_r / \partial x_i$. h holonomic constraints can also enter into that, such that we will get:

$$X_i + \sum_r \lambda_s\, a_{ir} + \sum_s \mu_r \frac{\partial \varphi_r}{\partial x_i} = 0 \tag{4.a}$$

as the most-general form of the equilibrium conditions. They represent f equations, and together with the $g + h$ constraint equations, they will suffice to determine the $f + g + h$ unknowns x_i, λ_r, μ_s. Naturally, the system of equations (4.a) will have many solutions, i.e., several equilibrium configurations will exist.

If time appears explicitly in the equations of constraint (viz., the rheonomic case) then one will get the equilibrium conditions from the argument that the system must be in equilibrium under the influence of all forces and constraints. That is, if we replace t with any numerical value then equations (4) [(4.a), resp.] must be satisfied. That means that we can employ our principle formally unchanged when we treat time as a constant parameter. However, in so doing, we should observe the following point: Time will also be varied for a true displacement that might result in the time interval dt. The true displacements will then satisfy constraints of the forms:

$$\sum_i \frac{\partial \varphi_r}{\partial x_i}\, dx_i + \frac{\partial \varphi_r}{\partial t}\, dt = 0$$

or

$$\sum_i a_{ir}\, dx_i + a_r\, dt = 0\,,$$

resp. Therefore, they will not belong to the virtual displacements that arise by dropping the term in dt, so for holonomic constraints, they will satisfy equations (1), while for non-holonomic constraints, they will satisfy the corresponding ones (1.a). The supporting surfaces that move in a given way in the rheonomic case can then be regarded as being at rest at any particular moment for the virtual displacements. We can base that convention, which initially seems very hard to understand, upon the fact that one derives the rheonomic case from the scleronomic one by passing to the limit when one lets the mass of the supporting surface go to infinity. The reaction of the motion of the system in question to the supporting body will then, in fact, vanish, and a constraint between them that was originally scleronomic will go to a rheonomic one on the system by itself ([1]).

With that, the principle of virtual displacements is exhibited in its greatest-possible generality. It will break down only for non-holonomic auxiliary conditions, which cannot be put into the Pfaffian form (2.a) of no. **4**.

([1]) See **E. Delassus**, Darboux Bull. (3) **45** (1921), pp. 231.

6. The meaning of the Lagrange factors. – One can obviously imagine that the equations of constraint that the mechanical system is subject to are replaced with suitable forces, namely, the so-called "guiding forces" or "constraint forces," which are arranged to be precisely such that when they are in effect, the free motion of the system will coincide with constrained motion of the system. Conversely, from the axiom of action and reaction, the forces of constraint must be equal and opposite to the action of the forces that the mass-points exert upon their constraints. Naturally, the determination of the reaction forces has great importance for engineering, even though it is not necessary for one to calculate the actual motion.

One will see that this introduction of guiding forces is always possible immediately when one compares equations (1) in no. **1** for free motions and equations (4) [(4.a), resp.] in no. **5** for constrained ones. The latter will go to the former when one regards the expressions:

$$Z_i = \sum_r \lambda_r \, a_{ir} + \sum_s \mu_s \frac{\partial \varphi_r}{\partial x_i} \tag{1}$$

as forces that correspond to the desired guiding forces that were introduced just now. The equilibrium conditions will then say that the impressed forces (as one refers to the external forces) should be equal and opposite to the guiding forces.

In the case of a mass-point that is constrained to a surface, so there is a constraint of the form $\varphi(x, y, z) = 0$, it is very simple to make the constraint force more intuitive. The components here are proportional to the direction cosines to the normal to the surface, as one will see when it is represented by:

$$\frac{\frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y}, \frac{\partial \varphi}{\partial z}}{\sqrt{\left(\frac{\partial \varphi}{\partial x}\right)^2 + \left(\frac{\partial \varphi}{\partial y}\right)^2 + \left(\frac{\partial \varphi}{\partial z}\right)^2}} .$$

The constraint force is then perpendicular to the guiding surface. In the general case, the constraints (1) [(1.a), resp.] of no. **5** say that the constraint forces can do no work under virtual displacements. Naturally, one can also elevate that theorem to an axiom and then derive equations (3), (4) in no. **5** from it.

7. Inequalities as auxiliary conditions. – We must now consider the case in which one or more inequalities:

$$\varphi(x_i, t) \leq 0 \tag{1}$$

are prescribed, in addition to the usual auxiliary conditions ([1]). We shall preserve the definition of a virtual displacement for them, as well, namely, that they represent small displacements δx_i of

([1]) A thorough discussion of inequalities in kinetics, as well, is given by **E. Delassus**, Ann. sci. Ec. Norm. Sup. (3) **34** (1917), pp. 95.

the coordinates that are compatible with the constraints. When inequalities of the form (1) exist, they will also be subject to inequalities of the form:

$$\sum_i a_{ir}\,\delta x_i \leq 0 \tag{1.a}$$

then. In the holonomic case, we will again have:

$$a_{ir} = \frac{\partial \varphi_r}{\partial x_i}, \tag{1.b}$$

but we can also admit non-holonomic constraints (1.a) that can therefore not be represented in the integrated form (1). A special characteristic of inequalities is that in general the opposite of a virtual displacement, which arises by inverting the signs of the δx_i, is not admissible for them.

The extension of the principle of virtual displacements to these cases was achieved by **Fourier**. Obviously, one can no longer postulate the vanishing of virtual work now.

We can divide the virtual displacements into two groups: In the first group, we have the ones for which the equality sign in (1.a) is valid everywhere, and we refer to them as *equality displacements*. In the second group, we have *inequality displacements*, to which the < sign applies. Now, it is initially clear that a configuration of the system can be an equilibrium configuration when it is already one under equality displacements since the admissibility of inequality displacements indeed means the lifting of a restriction on the degrees of freedom. We now assume that we have found an equilibrium configuration for all equality displacements, so we will have:

$$\sum X_i\,\delta x_i = 0$$

for them. However, if we add the inequality displacements, and thus extend the scope of the virtual displacements, then from (1.a), we can certainly find ones for which:

$$\sum X_i\,\delta x_i \neq 0\ .$$

Now one can easily see that the sum of all the virtual works done under all of the motions that actually occur must be > 0. Namely, each mass-point will follow the resultant of the forces that act upon it in those motions. If we then take the displacements to have those directions then the X_i and δx_i will always have the same sign, and $\sum X_i\,\delta x_i$ will always be positive for those displacements. None of the virtual displacements for which the sum of the virtual works is < 0 can correspond to a motion that actually occurs, but probably the displacements for which the work proves to be positive. Obviously, one must include that state of affairs in the calculations in such a way that one no longer takes the inequality sign in the principle of virtual displacements, but the ≤ sign, so one demands that:

$$\sum X_i\,\delta x_i \leq 0\ . \tag{2}$$

Naturally, that Ansatz is yet another hypothesis that cannot be proved, but only verified.

Based upon (2), as before, one can write the conditions for equilibrium with the help of the Lagrange factors in the form:

$$X_i + \sum_r \lambda_r \, a_{ir} = 0 \, .$$

In that way, the factors λ_r that belong to the inequalities will be subject to the single restriction that they must be < 0, since:

$$\sum_i X_i \, \delta x_i = - \sum_r \lambda_r \sum_i a_{ir} \, \delta x_i$$

would certainly be negative then as a result of equations (1.a).

b) Kinetics.

8. D'Alembert's principle. – We now move on to kinetics, so the simple equations of motion (1) in no. **3** will no longer be valid for a system of mass-points, but in their place, we will have the Newtonian equations of motion:

$$X_i - m_i \, \ddot{x}_i = 0 \qquad (i = 1, \ldots, f). \tag{1}$$

As **Jacob Bernoulli** once pointed out, and **d'Alembert** then exploited systematically in his mechanics, they can then be treated in an entirely analogous manner to the equilibrium equations when one also regards the quantities $m_i \, \ddot{x}_i$ as forces, namely, the so-called *inertial forces*. The inertial forces will then be in equilibrium with the impressed forces.

We initially combine equations (1) into a single relation with the help of arbitrary displacements δx_i, as we did in statics:

$$\sum_i (X_i - m_i \, \ddot{x}_i) \, \delta x_i = 0 \, . \tag{2}$$

That formulation, which is again identical to (1), goes back to **Lagrange**. If we now introduce auxiliary conditions, which can have all of the forms in no. **4**, then ***d'Alembert**'s principle* postulates that equation (2) will preserve its validity when the δx_i fulfill the same constraints that they would in statics. One can again replace that axiom with the demand that the guiding forces must do no work, or also with the following one, which reduces to statics by an analogous argument that is due to **d'Alembert**.

For a free system under the combined action of the impressed forces X_i and certain guiding forces Z_i, the accelerations that actually occur would enter in. Those guiding forces are determined by:

$$X_i - m_i \, \ddot{x}_i = Z_i \, , \tag{3}$$

in which $\ddot{x}_i$ are the accelerations that actually occur. The forces of reaction Z_i are now neutralized by the constraints. The forces Z_i must then yield equilibrium by themselves at each moment when one considers the constraints. For that reason, one also refers to them as *lost forces*. From the principle of virtual displacements, one will then have:

$$\sum_i Z_i \, \delta x_i = 0 \, , \tag{4}$$

which is identical to (2), from (1). The dynamical problem has been reduced to a problem in statics then. D'Alembert's principle can then be expressed as follows:

A system of mass-points moves under the influence of arbitrary forces and constraints in such a way that it will always keep the lost forces in equilibrium. They will then do no work.

Here, we would like to follow the general terminology and always refer to formula (2) as d'Alembert's principle.

Naturally, the argument that was laid out above includes an essentially new assumption, and one cannot refer to that reduction as a proof of d'Alembert's principle from the principle of virtual displacements.

One will then get the equations of motion from (2) precisely as one did in no. **5**. They read:

$$X_i - m_i \, \ddot{x}_i + \sum_r \lambda_r \, a_{ir} + \sum_s \mu_s \frac{\partial \varphi_s}{\partial x_i} = 0 \, . \tag{5}$$

They are the so-called ***Lagrange*** *equations of the first kind*, which are also valid for non-holonomic-rheonomic constraints. It should be pointed out once more that time is not varied here, even in the rheonomic case, since one treats an equilibrium problem at each moment.

The guiding forces themselves will again be:

$$Z_i = \sum_r \lambda_r \, a_{ir} + \sum_s \mu_s \frac{\partial \varphi_s}{\partial x_i} \, , \tag{6}$$

and they will fulfill the relation (4), as they must, which one will see immediately with the help of the auxiliary conditions (1) and (1.a) in no. **5**.

If inequalities are again present then the equality sign in (2) can no longer be satisfied. In the spirit of no. **7**, in its place, one must demand that:

$$\sum_i (X_i - m_i \, \ddot{x}_i) \, \delta x_i \leq 0 \, . \tag{7}$$

That condition will not always determine the motion uniquely. We shall return to it later in the context of **Gauss**'s principle ([1]).

9. Generalized coordinates. The Lagrange equations of the second kind – Naturally, the constant attention that has to be given to auxiliary conditions is very inconvenient in many problems. Using the **Lagrange** process, one then seeks to eliminate the auxiliary conditions by introducing suitable new parameters and then exhibit the equations in terms of the independent parameters. Likewise, it is often only natural to introduce other coordinates, e.g., polar coordinates, and even for free systems.

For holonomic constraints, regardless of whether they are scleronomic or rheonomic, that will encounter no difficulties. Namely, if g such constraints exist then a corresponding number of coordinates can be eliminated with their help, and the system can now be described by $f = 3n - g$ independent parameters q_k that are now completely free. They are called the *generalized Lagrange coordinates*, and the number of them f is called the *degrees of freedom* in the system. One calls the derivatives of q_k with respect to time:

$$\frac{dq_k}{dt} = \dot{q}_k$$

the *generalized velocity components.*

By contrast, if the system is non-holonomic then there will exist h relations between the differentials, in addition, i.e., at any moment, the motion cannot have all of the directions in the configuration space of the q_k, but it will be restricted to $(f - h)$-dimensional manifold, while the system itself can assume an f-parameter manifold of configurations.

Our problem is now to obtain the equations of motion in terms of generalized coordinates. That will be initially accomplished with *holonomic* systems. From what was said, the coordinates of the system points can be expressed with the help of generalized configuration parameters and time here:

$$x_i = x_i\,(q_1, \ldots, q_f, t) \qquad (i = 1, \ldots, 3n). \tag{1}$$

We shall assume that those equations determine the x_i uniquely in terms of the q_k. From d'Alembert's principle, the equations of motion:

$$m_i\,\ddot{x}_i = X_i + Z_i \qquad (i = 1, \ldots, 3n), \tag{2}$$

in which X_i are the impressed forces and Z_i are the still-unknown constraint forces. If we multiply those equations in succession by $\partial x_i / \partial q_k$ and sum over all coordinates then that will give:

$$\sum_i m_i\,\ddot{x}_i \frac{\partial x_i}{\partial q_k} = \sum_i (X_i + Z_i) \frac{\partial x_i}{\partial q_k}\,. \tag{3}$$

([1]) For d'Alembert's principle, cf., **G. Hamel**, Zeit. techn. Phys. **3** (1922), pp. 181; **E. Delassus**, C. R. Acad. Sci. **156** (1913), pp. 205. See also Chap. 1, no. **24**, of this volume of the *Handbuch*.

However, one has the following identities between the x_i and $\dot{x}_i$:

$$\left.\begin{aligned} \dot{x}_i &= \sum_i \frac{\partial x_i}{\partial q_k}\dot{q}_k + \frac{\partial x_i}{\partial t}, \\ \frac{\partial \dot{x}_i}{\partial \dot{q}_k} &= \frac{\partial x_i}{\partial q_k}, \end{aligned}\right\} \tag{4.a}$$

so

$$\left.\begin{aligned} \dot{x}_i \frac{\partial x_i}{\partial q_k} = \ddot{x}_i \frac{\partial \dot{x}_i}{\partial \dot{q}_k} &= \frac{d}{dt}\left(\frac{\partial x_i}{\partial q_k}\right) \\ &= \frac{d}{dt}\left(\dot{x}_i \frac{\partial \dot{x}_i}{\partial \dot{q}_k}\right) - \dot{x}_i\left(\sum_s \frac{\partial^2 x_i}{\partial q_k\, \partial q_s}\dot{q}_s + \frac{\partial^2 x_i}{\partial q_k\, \partial t}\right) \\ &= \frac{d}{dt}\left\{\frac{\partial}{\partial \dot{q}_k}\left(\frac{\dot{x}_i^2}{2}\right)\right\} - \frac{\partial}{\partial q_k}\left(\frac{\dot{x}_i^2}{2}\right). \end{aligned}\right\} \tag{4.b}$$

Now:

$$\tfrac{1}{2}\sum m_i \dot{x}_i^2 = T \tag{5}$$

is the kinetic energy of the system. Naturally, it can also be represented as a function of q_k, $\dot{q}_k$, and t, and in fact, from (4.a), it will once more be a quadratic function of the $\dot{q}_k$ that is homogeneous in them when the constraints do not depend upon time explicitly. In the other case, T will decompose into $T_0 + T_1 + T_2$, where T_s is a function of degree s in the $\dot{q}_k$.

Equations (3) will then go to:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} = \sum_i (X_i + Z_i)\frac{\partial x_i}{\partial q_k} \qquad (k = 1, \ldots, f). \tag{3.a}$$

The right-hand side of this has been brought into a form in which the x_i no longer appear. If we carry out a motion under which the coordinate q_k goes to $q_k + \delta q_k$, so it belongs to the virtual displacements, with our previous terminology, then the x_i will go to:

$$x_i + \frac{\partial x_i}{\partial q_k}\delta q_k ,$$

and the work done by all of the forces that are applied to the system under such a displacement will be equal to:

$$\sum_i (X_i + Z_i)\frac{\partial x_i}{\partial q_k}\delta q_k .$$

Now, d'Alembert's principle says precisely that the work done by constraint forces should vanish under all virtual displacements, i.e., one must also have:

$$\sum_i Z_i \, \delta x_i = \sum_i Z_i \frac{\partial x_i}{\partial q_k} \delta q_k$$

for them, and therefore due to the arbitrariness in δq_k, one must also have:

$$\sum_i Z_i \frac{\partial x_i}{\partial q_k} = 0 \, .$$

The right-hand side of (3.a) will then reduce to:

$$\sum_i X_i \frac{\partial x_i}{\partial q_k} = Q_k \, . \tag{6}$$

However, since we know the applied external forces as functions of x_i and possibly t, the Q_k will be known functions of q_k and t by means of (1), and one refers to them as *generalized force components* or *force components in the direction of* q_k . However, it should be emphasized that the Q_k do not necessarily have the dimensions of forces, namely, that will be the case only when the q_k have the dimension of a length. However, the product $Q_k \, \delta q_k$ will always have the dimension of work.

We finally obtain the ultimate form of the equations of motion in general coordinates in the form of the so-called ***Lagrange*** *equations of the second kind:*

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} = Q_k \qquad (k = 1, \ldots, f). \tag{7}$$

The constraint equations and the reaction forces have been eliminated from them. The latter, which are very important in engineering (support stresses and material stresses), can be determined when one uses either the equations of the first kind or intermediate form by initially introducing one or more parameters in such a way that when they are set equal to constants, they will produce the corresponding constraint equation precisely. One might introduce, e.g., polar coordinates, for a rotation around an axis and then fix the radius. One can once more determine the reaction forces after later recalling those constraints, as one does with the Lagrange equations of the first form ([1]) (On this subject, cf., also no. **12**.)

In the event that the external forces have a potential $U(x_i)$, so one has:

([1]) See, e.g., **F. Paulus**, Wiener Ber. **119** (1910), pp. 1669.

$$X_i = -\frac{\partial U(x_1,\ldots,x_{3n})}{\partial x_i},$$

one will get:

$$Q_k = \sum_i X_i \frac{\partial x_i}{\partial q_k} = -\sum_i \frac{\partial U}{\partial x_i}\frac{\partial x_i}{\partial q_k} = -\frac{\partial U(q_1,\ldots,q_f)}{\partial q_k}$$

by means of (1). In the event that U is independent of the velocities, one can put (7) into the simpler form (which was already exhibited by **Euler**):

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0, \tag{8}$$

in which:

$$L = T - U \tag{9}$$

is the so-called *kinetic potential* or the ***Lagrangian** function*, and is a known function q_k, $\dot{q}_k$, and t.

10. Systems with forces that depend upon velocity. – It can happen that forces that depend upon the velocities act upon a system. Examples of that type of force are forces of friction, which we will treat in no. **14**. However, there can also be forces that depend upon velocities and for which the simple form of the Lagrange equations will still be preserved.

Namely, if Q_k has the form:

$$Q_k = -\frac{\partial U}{\partial q_k} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_k}\right), \tag{1}$$

in which U is a given function of the q_k and $\dot{q}_k$, then that will be the case because one will have:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} = -\frac{\partial U}{\partial q_k} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_k}\right),$$

and with the definition of the kinetic potential:

$$L = T - U,$$

we come back to the form (8) in no. **9**.

A very important example of that is the motion of an electrically-charged particle, e.g., an electron under the influence of arbitrary electric and magnetic forces ([1]). If $\mathfrak{H}$ is the magnetic field strength vector, $\mathfrak{E}$ is the electric field strength vector, $-e$ is the charge of the electron, and $\mathfrak{v}$ is the velocity vector then the force $\mathfrak{K}$ will be:

$$\mathfrak{K} = -\frac{e}{c}[\mathfrak{v}\,\mathfrak{H}] - e\,\mathfrak{E}\,, \tag{2}$$

in vector notation. That so-called Lorentz force can be derived from the generalized potential:

$$U = -\frac{e}{c}\mathfrak{A}\,\mathfrak{v} - e\,\varphi = \mathfrak{A}_x\,\dot{x} + \mathfrak{A}_y\,\dot{y} + \mathfrak{A}_z\,\dot{z} - e\varphi\,, \tag{3}$$

in which $\mathfrak{A}$ and φ are the vector and scalar potential of the electromagnetic field, from which one can always (so, e.g., even for light waves) determine the field from:

$$\left.\begin{aligned} \mathfrak{E} &= -\operatorname{grad}\varphi - \frac{1}{c}\frac{\partial \mathfrak{A}}{\partial t}, \\ \mathfrak{H} &= \operatorname{rot}\mathfrak{A}. \end{aligned}\right\} \tag{4}$$

One easily confirms the validity of the Ansatz (3), as follows: With the prescription (1), one gets the force components as:

$$\begin{aligned} X &= \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_k}\right) - \frac{\partial U}{\partial q_k} = \frac{e}{c}\frac{d\mathfrak{A}_x}{dt} - \frac{e}{c}\frac{\partial \mathfrak{A}\,\mathfrak{v}}{\partial x} + e\frac{\partial \varphi}{\partial x} \\ &= \frac{e}{c}\left(\frac{\partial \mathfrak{A}_x}{\partial x}\dot{x} + \frac{\partial \mathfrak{A}_x}{\partial y}\dot{y} + \frac{\partial \mathfrak{A}_x}{\partial z}\dot{z} + \frac{\partial \mathfrak{A}_x}{\partial t}\right) - \frac{e}{c}\left(\frac{\partial \mathfrak{A}_x}{\partial x}\dot{x} + \frac{\partial \mathfrak{A}_y}{\partial x}\dot{y} + \frac{\partial \mathfrak{A}_z}{\partial z}\dot{z}\right) + e\frac{\partial \varphi}{\partial x} \\ &= -\frac{e}{c}\left[\dot{y}\left(\frac{\partial \mathfrak{A}_y}{\partial x} - \frac{\partial \mathfrak{A}_x}{\partial y}\right) - \dot{z}\left(\frac{\partial \mathfrak{A}_x}{\partial z} - \frac{\partial \mathfrak{A}_z}{\partial x}\right)\right] - \frac{e}{c}\frac{\partial \mathfrak{A}_x}{\partial t} + e\frac{\partial \varphi}{\partial x}\,, \end{aligned}$$

and therefore, from (4), one will have:

$$X = -\frac{e}{c}[\mathfrak{v}\,\mathfrak{H}]_x - \mathfrak{E}_x\,,$$

which agrees with (2). If one would like to consider the variation of mass according to the theory of relativity then the:

([1]) See, e.g., **J. Frenkel**, *Lehrbuch der Elektrodynamik*, Berlin, 1926, Chap. 10, pp. 317, *et seq.* For the general conditions for the existence of a kinetic potential, see H. von Helmholtz, Crelles Journal für Math. **100** (1896), pp. pp. 137; **A. Mayer**, Leipziger, Ber. **74** (1898); **A. Hirsch**, Math. Ann. **50** (1898), pp. 429.

$$T = \frac{1}{2m}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2m}\mathfrak{v}^2$$

in the Lagrangian function will be replaced with the relativistic expression ([1]):

$$m_0 c^2 \left(1 - \sqrt{1 - \frac{v^2}{c^2}}\right).$$

The complete Lagrangian function for an electron will then read:

$$L = m_0 c^2 \left(1 - \sqrt{1 - \frac{v^2}{c^2}}\right) + \frac{e}{c}\mathfrak{A}\mathfrak{v} - e c \ . \tag{5}$$

Another example of that type of force is defined by the older elementary law of electrodynamics that was due to Weber. In it, the generalized potential between two particles at a distance of r from each other is:

$$U = \frac{1}{r}\left(1 + \frac{\dot{r}^2}{c^2}\right),$$

and the corresponding force is:

$$X_r = \frac{1}{r^2}\left(1 - \frac{\dot{r}^2 - 2\dot{r}\ddot{r}}{c^2}\right).$$

11. Generalized impulse. Cyclic coordinates. Law of energy. – One calls the quantities:

$$p_k = \frac{\partial L}{\partial \dot{q}_k} \tag{1}$$

the *generalized impulses* that belong to the coordinates q_k. (In fact, e.g., for rectangular coordinates, if one has:

$$T = \tfrac{1}{2}\sum_i m_i \dot{x}_i^2,$$

and therefore when the potential V does not depend upon velocity:

$$p_k = \frac{\partial L}{\partial \dot{x}_k} = m_i \dot{x}_i$$

([1]) See Chap. 10, no. **4**, of this volume of the *Handbuch*.

will be the i^{th} component.) They play a very important role in all of mechanics. The Lagrange equations, in the form:

$$\frac{dq_k}{dt} = \frac{\partial T}{\partial q_k} + Q_k \,,$$

then say that the temporal change in each impulse component is equal to the change in the kinetic energy with respect to the associated position coordinate plus the force that is exerted upon it.

That will immediately imply some integrals of the Lagrange equations in some very important cases.

We initially assume that the system has a kinetic potential L, so the Lagrange equation will have the form (8) in no. **9**. Now, if some of the coordinates q_k – e.g., q_r – do not appear in L, but only their derivatives, then one calls those coordinates *cyclic.* The corresponding Lagrange equations reduce to:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_r}\right) = 0 \tag{3}$$

for them, so:

$$\frac{\partial L}{\partial \dot{q}_r} = \text{const.} = \beta_r \,, \tag{2.a}$$

in which the β_r are integration constants. In (2.a), we already have intermediate integrals of the equations of motion.

The impulses that belong to cyclic coordinates are constant then. Special cases of the law are, e.g., the laws of impulse and areas (see Chap. **7** and **8**).

It can often be the case that time t does not appear explicitly in L. In that case, one can likewise *always* give an integral of the equations of motion because one will then have:

$$\frac{dL}{dt} = \sum_k \frac{\partial L}{\partial \dot{q}_k}\ddot{q}_k + \sum_k \frac{\partial L}{\partial q_k}\dot{q}_k = \sum_k \frac{\partial L}{\partial \dot{q}_k}\ddot{q}_k + \sum_k \dot{q}_k \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = \frac{d}{dt}\left(\sum_k \frac{\partial L}{\partial \dot{q}_k}\dot{q}_k\right).$$

That will give:

$$\frac{d}{dt}\left(\sum_k \frac{\partial L}{\partial \dot{q}_k}\dot{q}_k - L\right) = 0 \,,$$

so

$$\sum_k \frac{\partial L}{\partial \dot{q}_k}\dot{q}_k - L = \text{const.} = E$$

will be an integral of the Lagrange equations. That integral will be the *energy integral* when T is a homogeneous quadratic function of the velocities, and the potential energy does not depend upon the velocities because in that case, from Euler's theorem on homogeneous functions:

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

so

$$E = \sum_k \frac{\partial T}{\partial \dot{q}_k} \dot{q}_k - T + U = 2T - T + U = T + U .$$

One can always reduce the number of parameters with the help of such integrals. We shall refer to Part III for that topic, in which that question is discussed from a general viewpoint.

12. Adding auxiliary conditions to the Lagrange equations of the second kind ([1]). – The methods of no. **9** will break down for non-holonomic conditions, and one must then proceed with the auxiliary conditions separately. For holonomic equations of constant, it can also prove to be useful to eliminate only some of the auxiliary conditions by introducing suitable coordinates but preserve the other ones. Now let the q_k no longer be free coordinates then, but ones that are subject to auxiliary conditions that might take the general non-holonomic form:

$$\sum_k a_{rk} dq_k + a_r \, dt = 0$$

or

$$\sum_k a_{rk} \dot{q}_k + a_r = 0 \qquad (r = 1, \ldots, g), \tag{1}$$

in which one will have:

$$a_{rk} = \frac{\partial \varphi_r}{\partial q_k} , \qquad a_r = \frac{\partial \varphi_r}{\partial t}$$

in the holonomic case.

Naturally, the a_{rk} and a_r are assumed to be functions of the q_k and t. We once more imagine that the constraints have been replaced with forces of constraint Z_k that will enter into the right-hand sides of equations (7) in no. **9**. Now, we know from d'Alembert's principle that those forces of constraint will do no work under virtual displacements. The work that they do:

$$\delta A = \sum_k Z_k \delta q_k$$

must then vanish for all displacements that satisfy the constraints:

$$\sum_k a_{rk} \delta q_k = 0 .$$

However, it will again follow from this that:

([1]) **N. M. Ferrers**, Quart. J. Math. **12** (1871), pp. 1. Also **C. Neumann**, Leipziger Ber. **40** (1888), pp. 22; **A. Verkandt**, Monatsber. Math. Phys. **4** (1892), pp. 31.

$$Z_k = \sum_k \lambda_k \, a_{rk} \qquad (r = 1, \ldots, g),$$

in which λ_r once more represent Lagrange factors. The equations of motion will then read:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} = Q_k + \sum_k \lambda_k \, a_{rk} \; . \tag{2}$$

Along with the auxiliary conditions (1), there will always be exactly $f + g$ equations then for the $f + g$ quantities q_k, λ_k . However, it is not possible to reduce the number of independent parameters in the unvaried form of the Lagrange equations with the help of non-holonomic auxiliary conditions. That can be first achieved with method that will be given in the next subsection.

13. The extended Lagrange equations of the second kind for quasi-coordinates. – From time to time, it is preferable to replace the velocity components $\dot{q}_k$ in the Lagrange equations with linear combinations of them that we would like to denote by $\overset{*}{\pi}_\rho$, and we define them by the equations:

$$\overset{*}{\pi}_\rho = \sum_k \alpha_{rk} \, \dot{q}_k \qquad (\rho = 1, \ldots, f), \tag{1}$$

in which the α_{rk} can be functions of the q_k and t. Initially, the number of $\overset{*}{\pi}_\rho$ shall be equal to the number of $\dot{q}_k$ themselves in that, such that our considerations shall initially relate to only holonomic coordinates since the q_k are supposed to be ordinary coordinates. In analogy to (1), we denote the linear combinations of the differentials by:

$$\delta \pi_\rho = \sum_k \alpha_{rk} \, \delta q_k \; . \tag{1.a}$$

If the Pfaffian expressions (1.a) are integrable, so perhaps the following integrability conditions exist:

$$\frac{\partial \alpha_{\rho k}}{\partial q_i} = \frac{\partial \alpha_{\rho i}}{\partial q_k}, \tag{2}$$

then they will define true coordinates π_ρ as functions of the q_k . However, that is not possible in the other case, and one will then call the $d\pi_\rho$ the differentials of *quasi-coordinates*. For that reason, we shall employ the asterisk notation for $d\pi_\rho \,/\, dt$ in order to emphasize the fact that we are not dealing with a true derivative.

One employs such quasi-coordinates in, e.g., the Euler equations of the mechanics of rigid bodies, in which one describes the motion of the system in terms of the components of the angular

velocities around three axes that are fixed in the body, namely, the principal axes of inertia, which are connected with non-integrable linear combinations of the time derivatives of the Euler angles.

In order to arrive at the equations of motion in terms of such quasi-coordinates, we proceed as follows. By inverting equations (1) (we assume that this is possible, so we are excluding the singular cases in which the determinant of the α_{rk} vanishes), we will get:

$$\dot{q}_k = \sum_k \beta_{k\rho} \overset{*}{\pi}_\rho , \tag{3}$$

in which

$$\sum_k \beta_{k\rho}\alpha_{\sigma k} = \begin{cases} 0 & \text{for} \quad \rho \neq \sigma \\ 1 & \text{for} \quad \rho = \sigma. \end{cases} \tag{4}$$

We multiply the Lagrange equations (7) of no. **9** by $\beta_{k\rho}$ in succession and sum over k :

$$\sum_k \beta_{k\rho} \left\{ \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} \right\} = \sum_k \beta_{k\rho} Q_k = \Pi_\rho . \tag{5}$$

Obviously, the Π_ρ in that is the generalized force component "in the direction of π_ρ," i.e., the work done by the external forces under a displacement will be equal to $\Pi_\rho\, \delta\pi_\rho$ only when $\delta\pi_\rho \neq 0$, while the work done by a general displacement will be:

$$\sum_\rho \Pi_\rho\, \delta\pi_\rho$$

since $\sum_k Q_k\, \delta q_k$ is the work done by external forces under an arbitrary displacement, and therefore $\sum_k Q_k\, \beta_{k\rho}\, \delta\pi_\rho$ will be the work done by the external forces under a displacement for which all of the $\delta\pi_\sigma$ vanish, with the exception of $\delta\pi_\rho$.

We can express T as a function of the q_k and $\overset{*}{\pi}_\rho$, instead of $\dot{q}_k$, with the help of (3), and we would like to denote that function by $\mathfrak{T}$ in order to distinguish the two. We will then have:

$$\frac{\partial T}{\partial \dot{q}_k} = \sum_\rho \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \frac{\partial \overset{*}{\pi}_\rho}{\partial \dot{q}_k} = \sum_\rho \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \pi_{\rho k}$$

and

$$\frac{\partial T}{\partial q_k} = \frac{\partial \mathfrak{T}}{\partial q_k} + \sum_\rho \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \frac{\partial \overset{*}{\pi}_\rho}{\partial q_k} = \frac{\partial \mathfrak{T}}{\partial q_k} + \sum_\rho \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \frac{\partial \alpha_{\sigma l}}{\partial q_k} \dot{q}_l ,$$

so

$$\Pi_\rho = \sum_k \beta_{k\rho} \left\{ \frac{d}{dt} \left(\sum_k \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \alpha_{\sigma k} \right) - \frac{\partial T}{\partial q_k} \right\}$$

$$= \frac{d}{dt} \left(\frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \right) + \sum_{k,\rho} \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \beta_{k\rho} \frac{d\alpha_{\sigma k}}{dt} - \sum_k \beta_{k\rho} \frac{\partial T}{\partial q_k}$$

$$= \frac{d}{dt} \left(\frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \right) + \sum_{k,l,\sigma} \beta_{k\rho} \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \dot{q}_k \left(\frac{\partial \alpha_{\sigma k}}{\partial q_l} - \frac{\partial \alpha_{\sigma l}}{\partial q_k} \right) - \sum_k \beta_{k\rho} \frac{\partial \mathfrak{T}}{\partial q_k}.$$

Now, the quantities:

$$\gamma_{\rho\sigma\tau} = \sum_{k,l} \beta_{k\rho} \beta_{l\tau} \left(\frac{\partial \alpha_{\sigma l}}{\partial q_k} - \frac{\partial \alpha_{\sigma k}}{\partial q_l} \right) \tag{6}$$

are independent of the motion of the system and depend upon only the relations between the differentials of the quasi-coordinates and those of the true coordinates. If the π_ρ were true coordinates then one would have:

$$\sum_k \frac{\partial \mathfrak{T}}{\partial q_k} \beta_{k\rho} = \sum_k \frac{\partial \mathfrak{T}}{\partial q_k} \frac{\partial q_k}{\partial \pi_\rho} = \frac{\partial \mathfrak{T}}{\partial \pi_\rho}. \tag{7}$$

We shall preserve that relationship for quasi-coordinates as well in order to simplify the notation. With that, we finally get the general form of the Lagrange equations, which was first derived by **Boltzmann** ([1]) and **Hamel** ([2]):

$$\frac{d}{dt} \left(\frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \right) + \sum_{\sigma,\tau} \gamma_{\rho\sigma\tau} \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\sigma} \overset{*}{\pi}_\sigma - \frac{\partial \mathfrak{T}}{\partial \pi_\rho} = \Pi_\rho. \tag{8}$$

If the π_ρ are themselves true coordinates then the quantities $\gamma_{\rho\sigma\tau}$ will all be zero from the integrability conditions (2). Equations (7) will then reduce to the usual Lagrange equations, as they must.

With that, we can also consider non-holonomic auxiliary conditions with no difficulty now. Namely, let the original coordinates q_k ($k = 1, \ldots, f$) be subject to g non-holonomic constraints:

$$\sum_k a_{sk} \dot{q}_k = 0 \qquad (s = 1, \ldots, g) \tag{9}$$

([1]) **L. Boltzmann**, Wiener Ber. **111** (1902), pp. 1603; also in *Werke*, Bd. III, 1904, pp. 682.

([2]) **G. Hamel**, Zeit. Phys. u. Phys. **15** (1904), pp. 1; cf., also Math. Ann. **59** (1905), pp. 416; there, one will also find a group-theoretic interpretation of the supplementary terms; cf., also *ibidem* **92** (1924), pp. 33.

so we take the $\alpha_{f-g,k}, \ldots, \alpha_{fk}$ in (1) (i.e., the coefficients of the last g equations) to be equal to a_{sk} precisely. Thus:

$$\overset{*}{\pi}_{f-g} = \sum_k a_{sk}\,\dot{q}_k\,, \qquad \delta\pi_{f-g} = \sum_k a_{sk}\,\delta q_k\,, \tag{10}$$

while the remaining ones can be chosen arbitrarily, with the single restriction that the determinant of all $\alpha_{\rho k}$ must be non-zero.

We now perform the same procedure as before, except that we take equations (2) of no **12** instead of equations (7) in no. **9**. Precisely the terms that are endowed with Lagrange factors will then drop out of the first $f - g$ of equations (5) by means of the relations (4). When one recalls the auxiliary conditions:

$$\overset{*}{\pi}_{\rho} = 0 \qquad (r = f - g, \ldots, f), \tag{11}$$

the last g of them will define precisely g relations for determining the λ_g . However, they will no longer be of any interest to us since obviously the remaining equations, together with (11), will already suffice to determine the motion of the system. With that, we have achieved our goal. Equations (8) are also true for non-holonomic systems then, and we can at least reduce the number of velocity components by means of the non-holonomic constraints.

Even more generally, we can also take expressions of the form:

$$\overset{*}{\pi}_{\rho} = \sum_k \alpha_{\rho k}\,\dot{q}_k + \alpha_{\rho}$$

in place of (1), which would be necessary, e.g., for non-holonomic-rheonomic systems. In that case, we formally introduce $q_{k+1} = t$ as a new coordinate, which means that we must naturally set $\dot{q}_{k+1} = 1$. We must then take:

$$\begin{aligned} \overset{*}{\pi}_{\rho} &= \sum_k \alpha_{\rho k}\,\dot{q}_k + \alpha_{\rho} \\ \overset{*}{\pi}_{\rho+1} &= \dot{q}_{k+1} = 1 \end{aligned}$$

as the equations of transformation. The argument above can then be adapted, word-for-word, except that everywhere that we sum over the coordinates, another term must be added to the summation that belongs to the new time coordinate. The form (8) of the equations of motion will then remain the same now and will thus subsume all of the cases that might come under consideration.

14. Forces of friction and collisions. – As an appendix, let us briefly discuss the modifications that must be made to the Lagrange equations when forces of a more general type act upon the system, such as forces of friction and collisions.

It will be assumed that the forces of friction are proportional to the velocities of their points of application. The equations of motion in Cartesian coordinates will then take the form:

$$m_i \ddot{x}_i = - k_i \dot{x}_i + X_i + Z_i ,$$

in which the k_i are functions of only the x_i. Naturally, the energy that is dissipated by the forces of resistance under an arbitrary displacement δx_i is equal to minus the work done by the forces of friction, so:

$$\delta E = \sum_i k_i \dot{x}_i \delta x_i . \tag{1}$$

We shall now go over to the generalized coordinates by multiplying each of the equations of motion by $\partial x_i / \partial q_k$ and summing over all coordinates. In that way, we will get:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} - Q_k = - \sum_i k_i \dot{x}_i \frac{\partial x_i}{\partial q_k} .$$

The right-hand side can be regarded as the derivatives of a single function since we have:

$$\frac{\partial x_i}{\partial q_k} = \frac{\partial \dot{x}_i}{\partial \dot{q}_k}$$

from equation (4.a) in no. **9**, so:

$$\sum_i k_i \dot{x}_i \frac{\partial x_i}{\partial q_k} = \sum_i k_i \dot{x}_i \frac{\partial \dot{x}_i}{\partial \dot{q}_k} = \frac{\partial F}{\partial \dot{q}_k} ,$$

in which:

$$F = \tfrac{1}{2} \sum k_i \dot{x}_i^2 \tag{2}$$

is the so-called ***Rayleigh*** *scattering function.* From (1), it measures the decrease in the system energy over time that results from the forces of friction. The equations of motion will then take the form:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} + \frac{\partial F}{\partial \dot{q}_k} = Q_k . \tag{3}$$

A suitable form can also be given to the Lagrange equations for collision forces, so for the effects of instantaneous impulses. If we denote all coordinates and velocities before the impact

with the index 0, and the ones after the impact with no index then the equations before the impact will be ([1]):

$$m_i (\dot{x}_i - \dot{x}_i^0) = S_i + Z_i$$

in Cartesian coordinates, in which the S_i are components of the external impulse relative to the i^{th} coordinate, and the Z_i are once more the constraint reactions. With the process in no. **9**, we will again get:

$$\sum_i m_i (\dot{x}_i - \dot{x}_i^0) \frac{\partial x_i}{\partial q_k} = \sum_i S_i \frac{\partial x_i}{\partial q_k} = P_k ,$$

since the terms with the reaction impulses must vanish, from d'Alembert's theorem. The impulse components P_k in the directions of the coordinates q_k are assumed to be known in that. One again converts the left-hand side as follows: One has:

$$\dot{x}_i \frac{\partial x_i}{\partial q_k} = \frac{1}{2} \frac{\partial x_i}{\partial q_k} (\dot{x}_i^2) , \qquad \dot{x}_i^0 \frac{\partial x_i}{\partial q_k} = \frac{1}{2} \frac{\partial x_i}{\partial q_k^0} (\dot{x}_i^0)^2 ,$$

in which the q_k (q_i^0, resp.) are the generalized velocity components before and after the impact. One can then put the equations of motion into the form:

$$\frac{\partial T}{\partial \dot{q}_k} - \left(\frac{\partial T}{\partial \dot{q}_k} \right)^0 = P_k . \tag{4}$$

As opposed to the usual Lagrange differential equations, those are finite equations for determining the $\dot{q}_k$ as functions of the $\dot{q}_i^0$. If one introduces the generalized impulse p_k (see no. **11**) then (4) will go to:

$$p_k - p_k^0 = P_k , \tag{5}$$

i.e., the difference between the impulses before and after the impact is equal to impulse that the impact exerts.

15. Gauss's principle of least constraint. – Now that we have discussed d'Alembert's principle and the Lagrange equations of the first and second kind that we get from it for all cases of practical importance, we shall turn to the remaining differential principles. Among them, **Gauss**'s principle of least constraint assumes an especially-distinguished position due to its simplicity and multifaceted utility. One arrives at it by the following argument:

([1]) Naturally, they are derived from the general equations by passing to the limit. On this, cf., e.g., **J. Tzènoff**, Math. Ann. **92** (1924), pp. 42. The case of the sudden introduction or removal of constraints is also treated there. See also **Beghin** and **Rousseau**, J. de math. (5) **9** (1903), pp. 21.

For a system of free mass-points, the expression:

$$Z = \sum_i m_i \left(\ddot{x}_i - \frac{X_i}{m_i} \right)^2 \tag{1}$$

will vanish as a result of Newton's equations of motion. If constraints of any sort exist then (1) can no longer be continually zero, but it must naturally be a function of the trajectory. Now **Gauss** demanded that this expression should be as small as possible for the actual motion, since it cannot vanish, i.e., Z shall be a minimum for the actual motion. That demand represents the *principle of least constraint* since one refers to the quantity Z as the *constraint*, following the lead of **Gauss**. In fact, $\ddot{x}_i - X_i / m_i$ is a sort of measure for the effect of the external constraints on the i^{th} coordinate.

One sees that as follows: Consider the system at two successive time-points t and $t + dt$, once under the influence of all auxiliary conditions, and then under the same initial conditions, but as a free system with no constraints that is still acted upon by the external forces.

In the second case, for a free system, the image point in the configuration space of x_i will go from the initial point A with the coordinates x_i to the point B with the coordinates:

$$x_i + \dot{x}_i \, dt + \frac{X_i}{2m_i} dt^2 ,$$

since the acceleration in the infinitesimal time interval dt can be regarded as constant, and indeed it will be equal to X_i / m_i, from Newton's equations (1) in no. **8**.

The segments AB and AC will then have the following components:

$$(AB)_i = \dot{x}_i \, dt + \frac{X_i}{2m_i} dt^2 ,$$

$$(AC)_i = \dot{x}_i \, dt + \frac{\ddot{x}_i}{2} dt^2 .$$

The segment BC, so the deviation of the path of the system that is produced by the constraints in comparison to the free motion, will then have the components:

$$(BC)_i = \frac{1}{2} \left(\ddot{x}_i - \frac{X_i}{2m_i} \right) dt^2 . \tag{2}$$

The larger that is, the greater the deviation from the free motion will be. Now, **Gauss** was led by the analogy to the principle of least squares in the theory of errors to suppose that the sum of the squares of those deviations would assume a value for the actual motion that is as small as possible.

However, a complication exists in that. When we have m mass-points with different masses, those deviations will obviously not be equivalent, since it is easier to impart a well-defined displacement upon a smaller mass that it is with a larger one. Now, in order to compare the segments, they must be assigned statistical weighting factors that will compensate for the differing masses, in the spirit of the theory of error.

One will obtain those factors by means of a continuity argument. If they are all initially equal to the masses then obviously the individual components (2) of the constraint will have the same character. Now let the ratios of the masses be rational, i.e., the m_i are whole-number multiples of a unit mass. One then imagines replacing each mass-point m_i with m_i mass-points of unit mass, but between which constraints exist, in such a way that they can always be found at the same location. In that way, the system will be extended to $\sum_i m_i$ points. However, at the same time, $\sum_i m_i - n$ constraints will be added to them. Similarly, let m_i external forces be uniformly applied to the elementary points. The new system will obviously be equivalent to the old one then and must also perform the same motions. If one now forms the sum of the squares of the constraints for the new system then they will now have the same character for the elementary masses, and the sum of their squares will have a well-defined meaning. The constraints in the elementary masses that belong to the same mass-point of the system will all be equal now as a result of the constraining conditions. Since the number of them is now m_i in each case, one will get the constraint on the total system by multiplying the square of each $(BC)_i$ by the corresponding mass and then summing over all coordinates. One will then arrive at precisely the expression (1) for the constraint. It is therefore established for rational ratios and must naturally have the same form in general, as well, on the grounds of continuity.

Naturally, that whole argument is purely heuristic and not a rigorous proof, which one could not carry out, anyway.

Just as for any extremal principle, one must also specify the manifold of comparison states in which the extremum is to be sought for Gauss's principle. Now, since Z is a function of the accelerations $\ddot{x}_i$, one must take the states that are being compared to be all motions that have the same position coordinates and velocity components at a given moment but can assume all values of the acceleration that are compatible with the constraints. Naturally, one must put that into a corresponding form that includes the second derivatives.

If the auxiliary conditions have the general non-holonomic form:

$$\varphi_r\,(x_i, \dot{x}_i, t) = 0 \tag{3}$$

then upon differentiating that with respect to time:

$$\dot{\varphi}_i = \sum_i \frac{\partial \varphi_r}{\partial x_i}\dot{x}_i + \sum_i \frac{\partial \varphi_r}{\partial \dot{x}_i}\ddot{x}_i + \frac{\partial \varphi_r}{\partial t} = 0\,. \tag{4}$$

From the rules of differential calculus, one will then get the condition for the relative minimum of Z when one introduces undetermined multipliers $2\,\lambda_r$:

$$\frac{\partial}{\partial \ddot{x}_i}\left(Z_i + \sum_r 2\,\lambda_r\,\dot{\varphi}_i\right) = \left(\ddot{x}_i - \frac{X_i}{m_i}\right)m_i + \sum_r \lambda_r \frac{\partial \dot{\varphi}_r}{\partial \dot{x}_i} = 0\,. \tag{5}$$

However, when the auxiliary conditions have the specialized form that is linear in the $\dot{x}_i$:

$$\dot{\varphi}_r \equiv \sum a_{ir}\,\dot{x}_i + a_r = 0\,, \tag{6}$$

they will be nothing but the well-known Lagrange equations of the first kind. Gauss's principle will then be completely equivalent to d'Alembert's for this case then. However, we see that, above and beyond that, it also allows one to treat the general form of the auxiliary conditions (2) of no. **4**. We correspondingly express Gauss's principle as follows:

A system of mass-points moves in such a way that the constraint (1) *will be a minimum at any moment in comparison to all states of motion with the same position and velocity, but with all possible accelerations that fulfill the constraint* (3) [(4), *resp.*].

16. Uniqueness of the principle. Singular cases. – With the help of Gauss's principle, one can also easily resolve the question that we still have yet to touch upon, namely, the question of whether the principle actually determines the motion uniquely ([1]).

Namely, except for some singular exceptional cases, the constraint will assume a minimum at only one location. The fact that it even possesses a minimum at all follows from its positive-definite character.

Now let $\ddot{x}_i = \ddot{\xi}_i$ be the location of a minimum, i.e., let:

$$Z(\ddot{\xi}_i + \delta\ddot{x}_i) > Z(\ddot{\xi}_i)$$

for all sufficiently-small $\delta\ddot{x}_i$ that are chosen in such a way that the quantities $\ddot{\xi}_i + \delta\ddot{x}_i$ themselves will satisfy the constraint equations (4) in no. **15**. However, that means that the increase $\delta\ddot{x}_i$ in the accelerations that is given by equations (4) and (6) of no. **15** must satisfy the equations:

$$\sum_i a_{ir}\,\delta\ddot{x}_i = 0\,, \tag{1}$$

which play an analogous role to the conditions (1) [(1.a), resp.] in no. **5** on the virtual displacements δx_i. Now, if there is a second location for the minimum with the value $\ddot{x}_i = \ddot{\eta}_i$ such that the $\ddot{\xi}_i$, as well as the $\ddot{\eta}_i$, fulfill the constraints (4) of no. **15** then the differences $\ddot{u}_i = \ddot{\xi}_i - \ddot{\eta}_i$ must satisfy the constraint that:

([1]) For this subsection, cf., **P. Staeckel**, Sitzungsber. Heidelberger Akad., Abt. A (1919), Abh. 11.

$$\sum_i a_{ir}\,\ddot{u}_i = 0\,. \tag{1.a}$$

Now, from (1) of no. **15**, one has:

$$Z(\ddot{\xi}_i + \delta\ddot{x}_i) = Z(\ddot{\xi}_i) + \sum_i m_i(\delta\ddot{x}_i) + 2\sum_i (m_i\,\ddot{\xi}_i - X_i)\,\delta\ddot{x}_i\,. \tag{2}$$

Thus, for a minimum, it is necessary that one must have:

$$\sum_i (m_i\,\ddot{\xi}_i - X_i)\,\delta\ddot{x}_i = 0\,.$$

However, from (1.a), we must also have:

$$\sum_i (m_i\,\ddot{\xi}_i - X_i)\,\ddot{u}_i = 0\,. \tag{3}$$

If we now set $\delta\ddot{x}_i$ equal to $\ddot{u}_i$ in (2) then it will follow that $Z(\ddot{\eta}_i)$ is greater than $Z(\ddot{\xi}_i)$, with no equality. However, we can likewise prove that conversely $Z(\ddot{\xi}_i) > Z(\ddot{\eta}_i)$. The assumption that there is more than one minimum would then lead to a contradiction. Uniqueness in the case of auxiliary conditions that take the form of inequalities was proved in a similar way by **Zermelo** ([1]). Since Gauss's principle is equivalent to d'Alembert's, uniqueness can also be proved for the latter. Even the proof that **Jacobi** gave in his *Dynamik* can be produced on the basis of d'Alembert's principle with the help of determinants, although the proof is somewhat more involved.

Up to now, we have excluded singular exceptional cases. Let all of the auxiliary condition be brought into the form:

$$\sum_i a_{ir}\,\dot{x}_i + a_r = 0$$

uniformly. We shall then call the configuration of the system *regular* when at least one of the g-rowed determinants in the coefficients matrix of the a_{ir}, a_r is non-zero. For a single auxiliary condition, not all of the quantities a_{ir}, a_r should vanish simultaneously then. D'Alembert's principle would break down in the other cases since the simple relation:

$$\sum_i a_{ir}\,\delta x_i = 0$$

would no longer exist for the virtual displacements then. In order to explain that, we shall consider the following simple example that was given by **Staeckel**:

A point of mass 1 is constrained to move on the surface of a cone:

([1]) **E. Zermelo**, Göttinger Nachr. (1899), pp. 306.

$$x_1^2 + x_2^2 = x_3^2 \ .$$

That condition will imply the equations:

$$x_1 \dot{x}_1 + x_2 \dot{x}_2 - x_3 \dot{x}_3 = 0 \ ,$$
$$x_1 \ddot{x}_1 + x_2 \ddot{x}_2 - x_3 \ddot{x}_3 + \dot{x}_1^2 + \dot{x}_2^2 - \dot{x}_3^2 = 0 \ ,$$

and the virtual displacements will be subject to the equations:

$$x_1 \delta x_1 + x_2 \delta x_2 - x_3 \delta x_3 = 0 \ . \tag{4}$$

Now let the mass-point be found to be at rest at the vertex of the cone at time *t*. Equation (4) for the virtual displacements will break down at that singular position. Rather, by their definition, namely, that they represent the motions that the mass-point can perform when it is in a position that is compatible with the constraints, they will be determined by:

$$\delta x_1^2 + \delta x_2^2 - \delta x_3^2 = 0 \tag{4.a}$$

now. The analogous relation for the acceleration components at that point will be true:

$$\ddot{x}_1^2 + \ddot{x}_2^2 - \ddot{x}_3^2 = 0 \tag{5}$$

since the mass-point can obviously remain on the surface of the cone only when the acceleration vector itself lies on that surface to begin with.

D'Alembert's principle will clearly break down here since the condition for virtual displacements (4.a) is no longer linear. By contrast, the principle of least constraint will make it possible to determine the motion. It then requires one to find the minimum of *Z* under the auxiliary condition (5), which is a soluble problem. In general, the motion is no longer determined uniquely now since it is double-valued, as one can easily confirm. Namely, if one lays a plane through the axis of the cone and places the force vector at the vertex of the cone then the motion will result along one of the two lines of intersection of the plane with the surface of the cone, and the calculation will also break down when the direction of the force is exactly parallel to the axis of the cone. In any event, one will see that only Gauss's principle is applicable at the singular locations. Naturally, one cannot prove its validity in that case either, but one must take it to be an axiom that it is even possible at all to subject such assumptions to any analytical treatment. However, a deeper discussion of the singular cases would lie beyond the scope of this presentation ([1]).

([1]) For that topic, cf., Chap. 1, nos. **33** and **34**, of this volume of the *Handbuch*.

17. Hertz's principle of the straightest path. – We can also interpret the expression for the constraint (1) in no. **15** geometrically when we restrict ourselves to force-free motion. Namely, we initially take a single mass-point with the coordinates x, y, z. Under the influence of any sort of holonomic or non-holonomic constraint equations, but no impressed forces, the path that it describes will be a space curve that we can describe in terms of the arc-length s as its parameter. In this case, from known laws of kinematics, the total acceleration can be decomposed into the tangential and normal accelerations:

$$\ddot{\mathfrak{r}}^2 = \ddot{x}^2 + \ddot{y}^2 + \ddot{z}^2 = \left(\frac{d^2 s}{dt^2}\right)^2 + \left(\frac{ds}{dt}\right)^4 \frac{1}{\rho^2},$$

in which $ds / dt = v$ is the speed of the mass-point, and ρ is radius of curvature of the path:

$$\frac{1}{\rho^2} = \left(\frac{d^2 x}{dt^2}\right)^2 + \left(\frac{d^2 y}{dt^2}\right)^2 + \left(\frac{d^2 z}{dt^2}\right)^2 .$$

The constraint will then take the form:

$$Z = m\left[\left(\frac{d^2 s}{dt^2}\right)^2 + \left(\frac{ds}{dt}\right)^4 \frac{1}{\rho^2}\right] . \tag{1}$$

If we now drop the external forces from the equations of motion (5) in no. **8**, with our current abbreviation, multiply them by $\dot{x}$, $\dot{y}$, $\dot{z}$, resp., and add them then from the constraints (1) and (1.a) in no. **5** for the virtual displacements in a scleronomic system, we will get:

$$\dot{x}\ddot{x} + \dot{y}\ddot{y} + \dot{z}\ddot{z} = \frac{1}{2}\frac{d}{dt}(\dot{s})^2 = 0 .$$

The tangential velocity $\dot{s}$ will then be constant for force-free, scleronomic systems, and therefore the velocity of the path, as well. Naturally, that also follows from the law of energy since the total energy consists of only kinetic energy. The constraint will then reduce to:

$$Z = m\frac{v^4}{\rho^2} + \text{const.},$$

and the principle of least constraint will now say that the constant factor is naturally irrelevant since the curvature $1/\rho$ of the path is a minimum for the actual motion. With that, we have ***Hertz's*** *principle of the straightest path* ([1]):

A force-free mass-point that is under the influence of constraints will move with constant velocity along those trajectories that have the smallest curvature and are allowed by the constraints.

The constraint requires that the point must move on a certain surface, so its trajectory must then be a geodetic line.

If we now have several mass-points then **Hertz** defined the curvature ρ of the trajectory of the system by:

$$\frac{1}{\rho^2} = \sum_i m_i \left(\frac{d^2 x_i}{ds^2} \right)^2 , \tag{2}$$

with the generalized definition of $\dot{s}$:

$$\dot{s}^2 = \sum_i m_i \dot{x}_i^2 . \tag{3}$$

Naturally, for one mass-point, the expression (2) will go to the usual curvature of the trajectory, except for the factor m. If we introduce t as a variable here then we will have:

$$\frac{d^2 x_i}{ds^2} = \frac{d}{ds}\left(\frac{\dot{x}_i}{\dot{s}} \right) = \frac{1}{\dot{s}} \frac{d}{dt}\left(\frac{\dot{x}_i}{\dot{s}} \right) = \frac{\dot{s}\,\dot{x}_i - \ddot{s}\,\dot{x}_i}{\dot{s}^3} ,$$

so

$$\frac{1}{\rho^2} = \frac{1}{\dot{s}^4} \sum_i m_i \ddot{x}_i^2 + \frac{\ddot{s}^2}{\dot{s}^6} \sum_i m_i \dot{x}_i^2 - \frac{2\dot{s}\,\ddot{s}}{\dot{s}^6} \sum_i m_i \dot{x}_i \ddot{x}_i ,$$

or when we observe (3):

$$\frac{1}{\rho^2} = \frac{1}{\dot{s}^4} \sum_i m_i \ddot{x}_i^2 - \frac{\ddot{s}^2}{\dot{s}^4} . \tag{4}$$

For force-free motion, we will again have that $\dot{s} = v$ is constant, so $\ddot{s} = 0$, and Gauss's principle for force-free motion will again go to the principle of least curvature, due to (4), so it has now been established for an arbitrary force-free system. Ordinarily, one employs Hertz's principle in an integrated form, which we will discuss in no. **26** in the context of the remaining integral principles.

([1]) **H. Hertz**, *Ges. Werke*, Bd. III: *Die Prinzipien der Mechanik*, Leipzig, 1892; See also **A. Brill**, *Vorlesungen zur Einführung in die Mechanik raumerfüllender Massen*, Leipzig, 1909; **F. Paulus**, Wiener Ber. (II.a) **125** (1916), pp. 835.

The starting point for **Hertz**'s mechanics was his ambition to eliminate the concept of force from mechanics completely, and in particular, forces at a distance, which he felt had been introduced into the laws of nature artificially. That was the starting point that he gave to his principle, whereas, in contrast to **Hertz**, **Boltzmann**'s ambition was to banish the constraints and replace them with suitable molecular forces.

In order to develop Hertzian mechanics further, one also addresses the problem of arriving at the general mechanical equations in the presence of, e.g., gravitational forces, by a corresponding combination of constraints (the introduction of ideal masses, resp.). That goal was probably still not achieved satisfactorily by **Hertz**. Nonetheless, his idea was really very deep, and in a certain sense, it was realized by the theory of gravitation in the form of the theory of relativity, in which the trajectory of a body that is small enough that it does not affect the gravitational field itself essentially will represent a geodetic line, but generally in a more general non-Euclidian space-time geometry.

18. Jourdain's principle. – In order to make the relationship and connection between Gauss's principle and that of d'Alembert emerge clearly, let the variation of the acceleration in the former one be actually performed. We then impart a small increase $\delta\ddot{x}_i$ to the components of the acceleration, while the coordinates and velocities remain unchanged, so we should have $\delta x_i = \delta\dot{x}_i = 0$. (One also calls that variation a ***Gaussian*** *variation*.) If we then perform it in (1) in no. **15** then according to the principle of least constraint, Z should not change. We must then have (see also no. 16):

$$\sum_i (m_i\,\ddot{x}_i - X_i)\,\delta\ddot{x}_i = 0 \tag{1}$$

for all arbitrary values of the $\delta\ddot{x}_i$ that are compatible with the equations of motion. Now, equation (1) has a completely similar form to d'Alembert's principle in no. **8**, except that the variations of the accelerations $\delta\ddot{x}_i$ appear in place of the displacements δx_i. However, since both of them are arbitrary quantities, except for the constraint equations, we can apply precisely the same reasoning to (1) that we did to d'Alembert's principle, which will naturally lead to the same result then since the auxiliary conditions will also take on a corresponding form, namely:

$$\sum_i \frac{\partial\varphi_r}{\partial x_i}\,\delta\ddot{x}_i = 0\,, \qquad \text{or} \qquad \sum_i a_{ri}\,\delta\ddot{x}_i = 0\,,\ \text{resp.},$$

since only the $\ddot{x}_i$ are in fact varied under the Gaussian variation. The two principles are then seen to be completely equivalent again. Obviously, differences between them can occur only for singular configurations when the virtual displacements no longer satisfy simple linear constraints, as was explained in no. **16**.

A comparison of d'Alembert's principle (2) in no. **8** and Gauss's principle in the form (1) will show immediately that there must be yet a third form for the general differential principle of

dynamics when we vary the velocities $\dot{x}_i$ to $\dot{x}_i + \delta\dot{x}_i$, but leave the coordinates unvaried, so we choose $\delta x_i = 0$. The corresponding differential principle must obviously read:

$$\sum_i (m_i \ddot{x}_i - X_i)\,\delta\dot{x}_i = 0 \tag{2}$$

then. **Jourdain** ([1]) had pointed out that form and derived the general Lagrange equations in the form that they took in no. **12** from d'Alembert's, Gauss's, and his own principle in the same way. That principle is obviously just as useful for auxiliary conditions that are nonlinear (in the velocities).

According to **Leitinger** ([2]), one will see the connection between the three principles very easily when one differentiates the expression for d'Alembert's principle:

$$\sum_i m_i \ddot{x}_i\,\delta x_i = \sum_i X_i\,\delta x_i$$

with respect to time. That will give:

$$\sum_i m_i \frac{d\ddot{x}_i}{dt}\delta x_i + \sum_i m_i \ddot{x}_i \frac{d\,\delta x_i}{dt} = \sum_i \frac{dX_i}{dt}\delta x_i + \sum_i X_i \frac{d\,\delta x_i}{dt}\ .$$

If one then takes $\delta x_i = 0$ after the differentiation, and one sets the freely-chosen quantities equal to:

$$\frac{d\,\delta x_i}{dt} = \delta\dot{x}_i$$

then one will get Jourdain's principle precisely, in which one chooses the comparison states to be just the ones with the same position, but varied velocities. One will get Gauss's formula (1) in precisely the same way by repeated differentiation, in which the comparison states are the ones that arise by varying the acceleration while fixing δx_i and $\delta\dot{x}_i$. Naturally, one can arrive at even more differential principles by further differentiation, but they have no practical value.

19. Appell's equations of motion. – Just as one will arrive at the Lagrange equations of the second kind upon starting from d'Alembert's principle with the use of generalized coordinates, one can also get there from Gauss's and Jourdain's principle. The most-general process of that type can now be performed quite uniformly on the three principles. In that way, one will come to a new and very remarkable form for the dynamical equations that was first exhibited by **Gibbs** ([3]) and

([1]) **Ph. E. B. Jourdain**, Quart. J. **39** (1908), pp. 251.
([2]) **R. Leitinger**, Wiener Ber. (II.a) **122** (1913), pp. 635; see also **A. Wassmuth**, *ibidem* **128** (1919), pp. 365.
([3]) **J. W. Gibbs**, Am. J. Math. **2** (1879), pp. 49.

Appell ([1]). It has the advantage over the Lagrange equations that is also useful with non-holonomic systems and coordinates, whereas the more complicated formulas in no. **13** must be introduced with the former.

We start from d'Alembert's principle in the form:

$$\sum_i m_i \ddot{x}_i \delta x_i = \sum_i X_i \delta x_i . \tag{1}$$

We now introduce the general Lagrange parameters q_k, which can be quasi-coordinates in their general form, as well as non-holonomic-rheonomic. Now let them be coupled with the x_i by the differential formulas:

$$dx_i = \sum_k \alpha_{ik} dq_k + \alpha_i dt , \tag{2}$$

which do not need to be integrable. Naturally, in the holonomic case, one would arrive at them by differentiating the relations:

$$x_i = x_i (q_k, t) .$$

In the non-holonomic case, we must write $\delta\pi_\rho$, instead of δq_k, with the notations of no. **13**, in which the $\alpha_{i\rho}$ are functions of any type of position parameters. The virtual displacements of the x_i are coupled with the q_k by:

$$\delta x_i = \sum_k \alpha_{ik} \delta q_k \tag{3}$$

since the last term in dt in equations (2) must drop out in each case from our definition of virtual displacement. One will get the corresponding relation between the accelerations upon differentiating (2) with respect to time:

$$\ddot{x}_i = \frac{d^2 x_i}{dt^2} = \sum_k \left(\alpha_{ik} \ddot{q}_k + \frac{d\alpha_{ik}}{dt} \dot{q}_k \right) + \frac{d\alpha_i}{dt} . \tag{4}$$

Naturally, in the holonomic case, one will have:

$$\frac{d\alpha_{ik}}{dt} = \sum_l \frac{\partial \alpha_{ik}}{\partial q_l} \dot{q}_l + \frac{\partial \alpha_{ik}}{\partial t} .$$

If one introduces the new variables into (1) with the help of (2), (3), and (4) then that equation will go to the differential relation:

$$\sum_k P_k \delta q_k = \sum_k Q_k \delta q_k . \tag{5}$$

([1]) **P. Appell**, C. R. Acad. Sci. Paris **317** (1899), pp. 317; See also his most-recent presentation in Mémorial des Sciences mathematiques, t. I, Paris, 1925. Numerous citations to the further literature are also given there.

P_k and Q_k are initially just abbreviations for the coefficients of δq_k on the left-hand (right-hand, resp.) side. From the discussion in no. **9** (no. **13**, resp.), the Q_k will then be the general force components in the directions of q_k.

Since the δq_k are completely free [we assume that the constraint equations can be eliminated entirely with the help of (2)], the general equations of motion that follow from d'Alembert's principle will read:

$$P_k = Q_k \qquad (k = 1, \ldots, f)\,. \tag{6}$$

Since each $\ddot{q}_k$ occurs in in only one term, it will now follow from (4) that:

$$\frac{\partial \ddot{x}_i}{\partial \ddot{q}_k} = \alpha_{ik}\,.$$

Therefore, one will have:

$$P_k = \sum_i m_i\, \ddot{x}_i\, \alpha_{ik} = \sum_i m_i\, \ddot{x}_i\, \frac{\partial \ddot{x}_i}{\partial \ddot{q}_k}\,.$$

If we finally introduce the expression:

$$S = \tfrac{1}{2} \sum_i m_i\, \ddot{x}_i^2 \tag{7}$$

then we will have:

$$P_k = \frac{\partial S}{\partial \ddot{q}_k}\,,$$

and we will then get the equations of motion in the **Appell** form from (6):

$$\frac{\partial S}{\partial \ddot{q}_k} = Q_k \qquad (k = 1, \ldots, f). \tag{8}$$

From their derivation, they will also be true for quasi-coordinates, and can then be written:

$$\frac{\partial S}{\partial \ddot{\pi}_\rho} = \Pi_\rho\,, \tag{8.a}$$

with the notations of no. **13**. One calls the quantity S the ***Appell*** *function*, or also the *energy of acceleration*. It has the same form as kinetic energy in rectangular coordinates, except that accelerations will appear in place of velocities.

If the constraint equation $\sum_k \alpha_{ik}\, \delta q_k = 0$ has still not been eliminated then it can be added once more with undetermined multipliers. The equations of motion will then read:

$$\frac{\partial S}{\partial \ddot{q}_k} = Q_k + \sum_k \lambda_k \, \alpha_{ik} \, . \tag{9}$$

The derivation that was just given can obviously be linked precisely with that of Jourdain's or Gauss's principle since when one looks at the respective manifolds of comparison states, one will see that:

$$\delta \dot{x}_i = \sum_k \alpha_{ik} \, \delta \dot{q}_k$$

for the Jourdain variations and:

$$\delta \ddot{x}_i = \sum_k \alpha_{ik} \, \delta \ddot{q}_k$$

for the Gaussian ones.

The various forms of the equations of motion will obviously come about when one gives various forms to the P_k ([1]). Formally, the Appell form is decidedly the simplest and most reasonable one. It generally has the disadvantage that the second derivatives appear in it explicitly, so the calculations with the Appell function in arbitrary coordinates can become correspondingly inconvenient.

By contrast, the Lagrange equations are distinguished by just the fact that when one exhibits them, the kinetic energy, which depends upon only the first derivatives, must be known as a function of the new parameters and velocities. Its form is less transparent in the non-holonomic case (nos. **12** and **13**). Which form of the equations is more convenient will then depend upon the special circumstances. For example, the derivation of the Euler equations for the motion is significantly simpler when one uses the Appell equations than it is when one uses the more-general Lagrange equations ([2]).

The connection between the Appell equations and Gauss's principle, in its original form [(1), no. **15**], is also very simple. If we introduce the expression:

$$\tfrac{1}{2} Z^* = S - \sum_k Q_k \, \ddot{q}_k \tag{10}$$

then from (8), the equations of motion will become simply:

$$\frac{\partial Z^*}{\partial \ddot{q}_k} = 0 \, .$$

That says: Z^* is an extremum with respect to the variations of the accelerations, just like the constraint in no. **15**.

In fact, the expression (10) differs from the constraint only by a term that is free of the second derivatives, so by a term that is irrelevant to the extremum, since from (4), one has:

([1]) For some of the mixed forms, cf., **G. Hamel**, Math. Ann. **92** (1924), pp. 33.

([2]) See, e.g., **Clemens Schaefer**, *Die Prinzipe der Dynamik*, pps. 37 and 75.

$$\sum_k Q_k \ddot{q}_k = \sum_i X_i \ddot{x}_i + \Phi(x_i, \dot{x}_i) ,$$

so

$$Z^* = 2\left(S - \sum_k Q_k \ddot{q}_k \right) = \sum_i m_i \ddot{x}_i^2 - 2\sum_i X_i \ddot{x}_i - 2\Phi = Z + \Phi ,$$

in which Φ and Ψ no longer depend upon $\ddot{x}_i$. One must then regard Z^* as the expression for the constraint in arbitrary coordinates.

20. True and varied motion. – The differential principles that were discussed here all require that a variation of the existing state of motion must be performed at a certain moment. That is not as intuitive in the case of d'Alembert's principle especially since the variation in that case has nothing to do with the actual course of motion. Rather, one reinterprets the state of the system artificially as being in a state of equilibrium, such that the question cannot even arise of what sort of meaning that the variation might have in regard to the further course of motion. Gauss's principle of least constraint seems most natural from that standpoint. It imposes certain accelerations upon the system at each moment that are given by just the constraint. The true motions are then selected from all possible accelerations, and the system will then proceed along the path that was affected in that way.

In order to gain a greater degree of intuition about d'Alembert's principle, we must summarize the virtual displacements for the total time evolution of the motion in some way. That will happen naturally when one imagines that a virtual displacement has been performed at each time-point, in the spirit of d'Alembert's principle. In and of itself, that displacement must satisfy only the constraint equations, but it does not have to be connected in any other way. However, one can choose it in a special way that will make it a continuous and sufficiently-differentiable function of time:

$$\delta x_i = \delta x_i(t) \qquad \text{or} \qquad \delta q_k = \delta q_k(t) \quad \text{in generalized coordinates, resp.}$$

In that way, every point of the true trajectory will be associated with a neighboring point in such a way that from our assumption, all of the neighboring points will lie along a smooth curve. The true path will then be associated with a varied one point-wise. Since time is not varied under virtual displacements, with our previous convention, we will also get a time scale for the neighboring path such that its points will always correspond at equal times, and one can then compare the varied and true paths. That is also the starting point for the different variational principles.

The meaning of this introduction of a continuously-varied path lies in the fact that under that assumption, the relation:

$$\frac{d}{dt}(\delta q_k) = \delta \frac{dq_k}{dt} = \delta \dot{q}_k \tag{1}$$

will exist for the *true* coordinates, which is proved directly. That says: The change in δq_k during the time interval *dt* is equal to the difference between the velocities along the true and neighboring

paths. Naturally, those two quantities would not be connected at all without the assumption of continuity, but now we can reason as follows: Let $q_k\,(t)$ be the motion along the true path, while $q_k^{(1)}(t)$ is the motion along the varied one. Hence, from the definition of δ, we will have $\delta q_k = q_k^{(1)} - q_k$ with no further discussion, so:

$$\frac{d}{dt}(\delta q_k) = \frac{dq_k^{(1)}}{dt} - \frac{dq_k}{dt} = \dot{q}_k^{(1)} - \dot{q}_k .$$

Moreover, by definition, we have:

$$\delta \dot{q}_k = \delta\left(\frac{dq_k}{dt}\right) = \dot{q}_k^{(1)} - \dot{q}_k ,$$

from which (1) will follow immediately.

That commutation relation is therefore trivial for true coordinates, but it will by no means be true for quasi-coordinates, which must always be observed in calculations with such things. One will get the corresponding relations as follows: Let the connection between the true and quasi-coordinates be given by:

$$\overset{*}{\pi}_\rho = \sum_k \alpha_{\rho k}\,\dot{q}_k , \qquad \delta\pi_\rho = \sum_k \alpha_{\rho k}\,\delta q_k ,$$

$$\dot{q}_k = \sum_\rho \beta_{k\rho}\,\overset{*}{\pi}_\rho , \qquad \delta q_k = \sum_\rho \beta_{k\rho}\,\delta\pi_\rho ,$$

as in no. **13**. We will then restrict ourselves to scleronomic systems, for the sake of simplicity. Hence, let the $\alpha_{\rho k}$ and $\beta_{k\rho}$ be independent of t. With that, we then define:

$$\frac{d}{dt}(\delta\pi_\rho) = \sum_k \alpha_{\rho k}\frac{d\,\delta q_k}{dt} + \sum_k \frac{d\alpha_{\rho k}}{dt}\delta q_k ,$$

$$\delta\overset{*}{\pi}_\rho = \sum_k \alpha_{\rho k}\,\delta\frac{dq_k}{dt} + \sum_k \delta\alpha_{\rho k}\frac{dq_k}{dt} .$$

Now, since the symbols $d\,/\,dt$ and δ commute for true coordinates, that will give:

$$\delta\overset{*}{\pi}_\rho - \frac{d}{dt}(\delta\pi_\rho) = \sum_k (\delta\alpha_{\rho k}\,\dot{q}_k - \frac{d\alpha_{\rho k}}{dt}\delta q_k) .$$

Now:

$$\dot{\alpha}_{\rho k} = \sum_k \frac{\partial\alpha_{\rho k}}{\partial q_l}\dot{q}_l , \quad \delta\alpha_{\rho k} = \sum_k \frac{\partial\alpha_{\rho k}}{\partial q_l}\delta q_l ,$$

so

$$\delta \overset{*}{\pi}_\rho - \frac{d}{dt}(\delta\pi_\rho) = \sum_{k,l} \dot{q}_k \frac{\partial \alpha_{\rho k}}{\partial q_l} \delta q_l - \sum_{k,l} \frac{\partial \alpha_{\rho k}}{\partial q_l} \delta q_k \dot{q}_l = \sum_{k,l} \left\{ \frac{\partial \alpha_{\rho k}}{\partial q_l} - \frac{\partial \alpha_{\rho l}}{\partial q_k} \right\} \dot{q}_k \delta q_l .$$

However, the term in brackets is just the integrability conditions on the coefficients $\alpha_{\rho k}$, so they will, in fact, vanish only for holonomic coordinates. If we also introduce the $\overset{*}{\pi}_\rho$, $\delta\pi_\rho$ themselves on the right-hand side now then with the relations (6) in no. **13**, that will give the commutation conditions for non-holonomic coordinates:

$$\left.\begin{aligned} \delta \overset{*}{\pi}_\rho &= \frac{d}{dt}(\delta\pi_\rho) + \sum_{k,l,\sigma,\tau} \left\{ \frac{\partial \alpha_{\rho k}}{\partial q_l} - \frac{\partial \alpha_{\rho l}}{\partial q_k} \right\} \beta_{k\sigma} \beta_{l\tau} \overset{*}{\pi}_\sigma \delta\pi_\tau \\ &= \frac{d}{dt}(\delta\pi_\rho) - \sum_{\sigma,\tau} \gamma_{\sigma\rho\tau} \beta_{k\sigma} \beta_{l\tau} \overset{*}{\pi}_\sigma \delta\pi_\tau . \end{aligned}\right\} \tag{2}$$

We now ask whether the varied motion satisfies the equations constraint at each moment, i.e., whether it represents a kinematically-possible motion.

That is by no means obvious, and a fundamental difference between holonomic and non-holonomic systems again exists here. In order to examine that question, we shall impose the condition that the varied path should satisfy the equations of constraint. The virtual displacements will satisfy equations (1.a) of no. **5** since time is indeed not varied (we once more confine ourselves to scleronomic systems), so:

$$\delta\varphi_r = \sum_k a_{kr} \delta q_k = 0 , \tag{3}$$

which will possess the form:

$$\delta\varphi_r = \sum_k \frac{\partial \varphi_r}{\partial q_k} \delta q_k = 0$$

for holonomic constraints. Naturally, the q_k in that must be true coordinates. One will initially get from (3) that:

$$\frac{d}{dt}(\delta\pi_\rho) = \frac{d}{dt}\left(\sum_k a_{kr} \delta q_k \right) = \sum_k \left(\sum_l \frac{\partial a_{kr}}{\partial q_l} \frac{dq_l}{dt} \right) \delta q_k + \sum_k a_{kr} \frac{d}{dt} \delta q_k = 0 . \tag{4}$$

That is just the condition for the virtual displacements to obey the constraint equations.

Now, if the varied motion is also supposed to satisfy the constraint equations in the course of motion then that would have to mean that for the varied coordinates $q_k + \delta q_k$, the increase $d\,(q_k + \delta q_k)$ during the time interval dt should satisfy the constraint equations, so when one switches the indices k and l :

$$\sum_l a_{lr} (q_k + \delta q_k) \frac{d\,(q_l + \delta q_l)}{dt} = 0 ,$$

and therefore, with a **Taylor** development and consideration given to (3), one will have:

$$\sum_k \left(\sum_l \frac{\partial a_{lr}}{\partial q_k} \delta q_k \right) \frac{dq_l}{dt} + \sum_l a_{lr} \frac{d}{dt} \delta q_l = 0 . \tag{5}$$

Upon subtracting that from (4), one will get:

$$\sum_{l,k} \left(\frac{\partial a_{kr}}{\partial q_l} - \frac{\partial a_{lr}}{\partial q_k} \right) \frac{dq_l}{dt} \delta q_k = 0 . \tag{6}$$

since the second summations obviously cancel. Due to the arbitrariness in the δq_i, that relation can be fulfilled only when the term in parentheses vanishes by itself ([1]). That shows that the varied paths are indeed kinematically possible for holonomic systems, but by no means for non-holonomic ones.

One can easily make that result more intuitive with an example. If one lets a blade glide on a plane then the non-holonomic auxiliary condition that it cannot slide laterally will imply that direction of the blade must always coincide with the direction of the tangent to the trajectory. One will then get the neighboring curve when one moves forward along each tangent by a segment that varies continuously with the arc-length along the curve and then combines those points. Obviously, the tangents to the neighboring curve cannot coincide with the directions of the blade.

21. The Lagrange central equation. – One can put d'Alembert's principle into a form that no longer includes the second derivatives with the help of the previous consideration, and which is therefore frequently convenient for the applications. One writes it in the form:

$$\sum_i X_i \, \delta x_i = \sum_i m_i \, \ddot{x}_i \, \delta x_i \ .$$

Now, as has already been said many times, one has:

$$\ddot{x}_i \, \delta x_i = \frac{d}{dt} (\dot{x}_i \, \delta x_i) - \dot{x}_i \frac{d}{dt} (\delta x_i) \ .$$

With the commutation relations (1) of no. **20**, one has:

([1]) That conclusion is not entirely rigorous since the δq_k satisfy the conditions (3), and the $\dot{q}_k$ satisfy the corresponding one $\sum_k a_{kr} \dot{q}_k = 0$, but one can still follow through the proof that (6) will actually vanish only for holonomic constraints by considering just those restrictions.

$$\frac{d}{dt}(\delta x_i) = \delta \frac{dx_i}{dt} = \delta \dot{x}_i \ ,$$

so one will get:

$$\ddot{x}_i \, \delta x_i = \frac{d}{dt}(\dot{x}_i \, \delta x_i) - \dot{x}_i \, \delta \dot{x}_i = \frac{d}{dt}(\dot{x}_i \, \delta x_i) - \tfrac{1}{2} \delta (\dot{x}_i)^2 \, . \tag{1}$$

If one substitutes that in d'Alembert's principle then that will give:

$$\sum_i X_i \, \delta x_i + \sum_i \tfrac{1}{2} m_i \, \delta (\dot{x}_i)^2 = \frac{d}{dt}\left(\sum_i m_i \, \dot{x}_i \, \delta x_i \right) \, . \tag{2}$$

The second term is nothing but the variation of the kinetic energy *T*, and we then get the so-called ***Lagrange*** *central equation:*

$$\delta T + \sum_i X_i \, \delta x_i = \frac{d}{dt}\left(\sum_i m_i \, \dot{x}_i \, \delta x_i \right) \tag{3}$$

as the new form of d'Alembert's principle. Here, we can introduce more-general (holonomic) coordinates: The expression:

$$\sum_i X_i \, \delta x_i = \sum_k Q_k \, \delta q_k$$

is again the work done by the virtual displacement δq_k . We can also denote it symbolically by δA then. With the help of the transformation formulas that take x_i to the generalized coordinates:

$$x_i = x_i \, (q_k, \, t) \ ,$$

so:

$$\dot{x}_i = \sum_k \frac{\partial x_i}{\partial q_k} \dot{q}_k + \frac{\partial x_i}{\partial t} = \sum_k \alpha_{ik} \, \dot{q}_k + \alpha_i \, ,$$

$$\delta x_i = \sum_k \alpha_{ik} \, \delta q_k \, ,$$

we can easily confirm the identity:

$$\sum_i m_i \, \dot{x}_i \, \delta x_i = \sum_k \frac{\partial T}{\partial \dot{q}_k} \delta q_k = \sum_k p_k \, \delta q_k \ ,$$

and the Lagrange central equation will then become:

$$\delta T + \delta A = \frac{d}{dt}\left(\sum_k p_k \, \delta q_k \right) \, . \tag{4}$$

The term in parentheses on the right-hand side is composed from the impulse components p_k and the virtual displacements in the same way as the virtual work is composed of the force vector and the δq_k . Equation (4) then expresses the idea that the sum of the variations of the kinetic energy and the work done by the external forces is equal to change in the virtual impulse-work per unit time.

That result is linked essentially with the commutability of d / dt and δ. Formula (4) will also break down for general variations and quasi-coordinates, for which that is not true. An extension to that case was given by **Hamel** ([1]).

III. – Integral principles.

22. Hamilton's principle. – The defining characteristic of the principles that were considered up to now is that the variations of the system were always examined at a certain moment. Accordingly, they do not give a property of the true motion that distinguishes it from the neighboring motions, but they only allow one to arrive at differential equations for the motion with their help. However, one can also ask what the properties of the true path as a whole might be, and one will then arrive at integral principles.

As in no. **20**, we imagine that we are comparing the true path to neighboring paths. If we assign all possible values to the increases δq_k then we will get a whole family of comparison paths, and the problem becomes that of finding special properties for the true paths. Naturally, one can no longer demand that this can be achieved by an ordinary minimal principle like Gauss's principle, but rather such a result can come about only when we impose some condition on a function of the entire trajectory, which is itself again a function of the coordinates and time, so a function of a function. The calculus of variations will enter in place of the differential calculus accordingly.

The most important, and most useful, principle is ***Hamilton**'s principle*, which we would like to focus on in its simplest form for conservative holonomic systems in order to clarify the general character and advantages of integral principles for them.

It reads:

Let T be the kinetic energy and let U be potential energy of the system, so L = T – U is the kinetic potential, which is a function of any sort of Lagrange coordinates q_k, its time derivatives $\dot{q}_k$, and time t. One will then have:

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

for the motion that actually occurs, in which the integral is taken between two given configurations of the system at well-defined times, so:

([1]) **G. Hamel**, Math. Ann. **59** (1904), pp. 416.

$$q_k(t_1) = q_k^{(1)} \qquad \text{and} \qquad q_k(t_2) = q_k^{(2)} \tag{2}$$

are given values, and all paths in the field of competition that emerge from the true path by a variation in the sense of no. **20** *are allowable.*

The δq_k must then be continuous functions of time that satisfy the auxiliary conditions. However, they must also vanish at the limits of the integral.

From the rules of the calculus of variations ([1]), one will get the necessary condition for the occurrence of an extremum in the form of the Lagrange differential equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0 \qquad (k = 1, \ldots, f). \tag{3}$$

A direct derivation of then will also be given in no. **27**.

One can also admit holonomic auxiliary conditions of the form:

$$\varphi_r(q_k, t) = 0 \qquad (r = 1, \ldots, g).$$

As is known, one will then get the differential equations when one adds the auxiliary conditions with undetermined multipliers – λ_r to the integrand in (1), so one then looks for the extremum of:

$$\int_{t_1}^{t_2} (L - \sum_r \lambda_r \varphi_r)\, dt$$

and regards the λ_r as new variables whose derivatives do not, however, appear. One will then get the differential equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} + \sum_r \lambda_r \frac{\partial \varphi_r}{\partial q_k} = 0\,. \tag{3.a}$$

Now, equations (3) [(3.a), resp.] are just the Lagrange equations of the second kind in no. **9** [no. **12**, resp.]. With that, the identity of Hamilton's principle, in its simplest form, with d'Alembert's principle is verified.

From the form of (1), one sees the great significance of integral principles and the advance that was achieved by them. They include nothing that relates to the coordinates in any way. Energy, and therefore the kinetic potential, are mechanical quantities whose meanings are independent of how they are described in terms of special coordinates, so naturally the integral over the trajectory will be like that, as well. The statement of Hamilton's principle is therefore independent of the coordinate system, and its recalculation in any other coordinates will become very convenient with

([1]) See this *Handbuch*, v. III.

its help. Since it includes only the first derivatives of the coordinates, calculating with it will be simpler than with the differential equations themselves, which include second derivatives.

That is also the reason why one always looks for the field equations of the theory of relativity, if not all of the foundations of modern physics, in the form of variational principles: One will always get a formulation of them that is independent of the special choice of representation. The variational principles can be likewise extended to continuous media.

One can also get Hamilton's principle from a direct conversion of d'Alembert's principle (and that is generally the usual way of doing that). In order to do that, one starts from Lagrange's central equation (4) in no. **21**:

$$\delta T + \delta A = \frac{d}{dt}\left(\sum_k \frac{\partial T}{\partial \dot{q}} \, \delta q_k \right) .$$

The form of the right-hand side suggests an integration over time:

$$\int_{t_1}^{t_2} (\delta T + \delta A)\, dt \; = \left[\sum_k \frac{\partial T}{\partial \dot{q}} \, \delta q_k \right]_{t=t_1}^{t=t_2} .$$

If we now demand that the displacements δq_k must vanish at the limits of integration, i.e., that all of the comparison curves must go through the same starting and ending points, then the right-hand side will drop out, and we will get:

$$\int_{t_1}^{t_2} (\delta T + \delta A)\, dt \; = \int_{t_1}^{t_2} (\delta T + \sum Q_k \, \delta q_k)\, dt \; = 0 \, . \tag{4}$$

However, that is nothing but a generalization of Hamilton's principle (1). Namely, if the system has a potential then we will have:

$$\delta A = \sum Q_k \, \delta q_k = - \sum_k \frac{\partial U}{\partial q_k} \delta q_k = - \, \delta U \, ,$$

and as a result, since one can switch the order of variation and integration in this case:

$$\int_{t_1}^{t_2} (\delta T - \delta U)\, dt \; = \delta \int_{t_1}^{t_2} (T - U)\, dt \; = \delta \int_{t_1}^{t_2} L \, dt \; = 0 \, ,$$

which is identical to (1). Equations (4) then represent an extension of Hamilton's principle to forces that do not possess potentials. The character of a variation principle is generally lost in that way. Nevertheless, one can obtain the equations of motion from it by formal processes, and indeed in the general form of no. **13**. We shall return to that concept in no. **27**.

23. The variation of time. – It is possible to give Hamilton's principle the form of a true variational principle only in the case where a kinetic potential exists, which is certainly quite important, but it still does not possess the generality in its applications that the differential principles possess. We have already seen the direction that we must pursue in order to make that generalization in the form of equation (4) of no. **22**. We must go beyond the scope of the calculus of variations, properly speaking, and apply more-general variations and processes of integration to the formulas of d'Alembert's principle or also other differential principles.

Up to now, we have not varied time, i.e., we have allowed points of the varied paths and the true path to correspond only when they belong to the same time t. The virtual displacements were then defined in such a way that they were always performed while time was held constant. We shall now drop that restriction and consider displacements that we shall denote by Δx_i, Δq_k, Δt, to distinguish them, which will then associate the space-time point q_k, t on the true path with the point $q_k + \Delta q_k$, $t + \Delta t$. However, in that way, the Δq_k, Δt shall again be continuous functions of time, such that the set of all varied points will define a continuous varied path. That means that the neighboring paths cannot be traversed with a definite time scale that is coupled with the true path in the spirit of virtual displacements, but with an arbitrary time scale. Naturally, the domain of the allowable functions q_k (t) will be extended considerably in that way.

We therefore assume that the Δ-process means a true variation, i.e., in general coordinates, between which no sort of constraints exist, the most-general Δ-operation is applied to a function φ (q_k, t):

$$\Delta\varphi = \sum_k \frac{\partial\varphi}{\partial q_k}\Delta q_k + \frac{\partial\varphi}{\partial t}\Delta t\,, \tag{1}$$

in which the Δq_k, Δt are completely arbitrary, except for the fact that they should vary continuously over time, in the spirit of no. **20**. Naturally, that process also makes good sense for quasi-coordinates. When they are also chosen in such a way that no other sort of constraints exist between them, the Δ-process will once more mean the changes in the relevant quantities when all differentials of the quasi-coordinates and time t are assigned arbitrary increases $\Delta\pi_\rho$, Δt. Similarly, the function φ itself does not need to be representable in integrable form either, such as, e.g., the virtual work done by non-conservative forces. The expressions Q_k will then appear in place of the $\partial\varphi / \partial q_k$, which do not need to satisfy the integrability conditions:

$$\frac{\partial Q_k}{\partial q_l} = \frac{\partial Q_l}{\partial q_k}\,.$$

In other coordinates between which constraint equations still exist (e.g., in rectangular coordinates), the Δq_k, Δt will no longer be completely arbitrary either, but they must fulfill the conditions [e.g., in the general non-holonomic constraints (2.a) of no. **4**]:

$$\sum_i a_{ik}\,\Delta q_i + a_r\,\Delta t = 0\,.$$

Along with the Δ-variation, we also consider the original δ-variation and the advance of the system itself along its path. After a time dt, it will arrive at $x_i + \dot{x}_i\, dt$ (in generalized coordinates, $q_k + \dot{q}_k\, dt$ resp.). Therefore, the corresponding values of Δq_k will also change, such that the expressions:

$$\frac{d}{dt}\Delta q_k\,, \qquad \frac{d}{dt}\Delta t\,, \qquad \Delta\frac{dq_k}{dt} = \Delta\dot{q}_k$$

will also have a well-defined meaning.

We can now couple any Δ-process with a certain δ-process by constructing a δ-process from it by combining it with a d-process. That would be necessary when we wish to employ d'Alembert's principle since it is valid only for the δ-processes. Along with (1), we have:

$$\delta\varphi = \sum_k \frac{\partial\varphi}{\partial q_k}\delta q_k\,, \tag{2}$$

$$\frac{d\varphi}{dt} = \sum_k \frac{\partial\varphi}{\partial q_k}\dot{q}_k + \frac{\partial\varphi}{\partial t} \tag{3}$$

as the definitions of the d and δ-processes. Thus, the operation:

$$\Delta\varphi - \dot{\varphi}\Delta t = \sum_k \frac{\partial\varphi}{\partial q_k}(\Delta q_k - \dot{q}_k\,\Delta t) = \delta\varphi \tag{4}$$

is a virtual displacement, with:

$$\Delta q_k = \Delta q_k - \dot{q}_k\,\Delta t \qquad \text{or} \qquad \delta x_i = \Delta x_i - \dot{x}_i\,\Delta t\,, \text{ resp.} \tag{5}$$

That δ-displacement has the simple geometric meaning that it is the projection of Δ onto a plane t = const. in the general configuration space of q_k, t. Namely, if one moves forward along the true path by Δt and then performs the δ-process then one will arrive at just the point:

$$q_k + \dot{q}_k\,\Delta t + \delta(q_k + \dot{q}_k\,\Delta t)\,,$$

and since the last term can be neglected for being of order two, that is just the point $q_k + \Delta q_k$. The δq_k are the variations at constant time that correspond to the Δq_k then. It emerges from that argument that the manifold of allowable paths in φ-space will not be enlarged by the introduction of time variation, although the rate of variation (*Durchlaufungsgeschwindigskeit*), which was previously established by the true path through the coupling $\delta t = 0$, will be changed now.

A special property of the Δ-variation is that its commutability with differentiation with respect to time for holonomic coordinates, as well, is no longer true. Namely, we have, from (5), that:

$$\frac{d}{dt}\delta q_k = \frac{d}{dt}(\Delta q_k - \dot{q}_k \Delta t) = \frac{d}{dt}\Delta q_k - \ddot{q}_k \Delta t - \dot{q}_k \frac{d\Delta t}{dt} .$$

Now, from (4), one will have, with $\varphi = \dot{q}_k$, that:

$$\Delta \dot{q}_k = \delta \dot{q}_k + \ddot{q}_k \Delta t ,$$

and since d and δ commute, we will get the following commutation relation for d and Δ:

$$\Delta \frac{dq_k}{dt} = \Delta \dot{q}_k = \frac{d}{dt}\Delta q_k - \dot{q}_k \frac{d\Delta t}{dt} , \tag{6}$$

so we will also have:

$$\Delta \dot{x}_i = \frac{d}{dt}\Delta x_i - \dot{x}_i \frac{d\Delta t}{dt} \tag{6.a}$$

in rectangular coordinates.

One can also associate any Δ-displacement with a δ-displacement in a different way, e.g., by the definition:

$$\delta^* x_i = \Delta x_i - \frac{\delta x_i}{\delta t}\Delta t , \tag{7}$$

in which x_i is thought of as a function of the generalized configuration parameters q_k and t. Due to the fact that:

$$\Delta x_i = \sum_k \frac{\partial x_i}{\partial q_k}\Delta q_k + \frac{\partial x_i}{\partial t}\Delta t , \quad \text{so} \quad \delta^* x_i = \sum_k \frac{\partial x_i}{\partial q_k}\Delta q_k ,$$

δ^* will be a virtual displacement, with $\delta^* q_k = \Delta q_k$, but from (1) and (2), Δ and δ^* cannot simultaneously represent true variations, which would probably give priority to the definition (5) that was employed here. Naturally, one can also operate further with (7) or any other definition of the virtual displacements and go on to what will be done in the next subsection by converting d'Alembert's expression into a general integral principle, as long as one always observes the proper limit conditions. Thus, **Hölder** ([1]) employed the association (7) in his ground-breaking work, while (5) was introduced by **Voss** ([2]). We have discussed that point so thoroughly here because many misconceptions about it seem to exist in the literature ([3]).

([1]) **O. Hölder**, Göttinger Nachr. (1896), pp. 122.
([2]) **A. Voss**, Göttinger Nachr. (1900), pp. 322.
([3]) See, in particular, the discussion between **M. Rethy** [Math. Ann. **58** (1905), pp. 169 and *ibid.* **64** (1906), pp. 156] and **Ph. E. B. Jourdain** [*ibidem*, **62**, pp, 413 and **65** (1904), pp. 513]. The latter paper gave the complete explanation; cf., also **H. Brell**, Wiener Ber. (II.a) **122** (1913), pp. 122.

24. General transformation of d'Alembert's principle. – With the help of the general Δ-process and integration over the time, we shall now look for a conversion of d'Alembert's principle into an integral formula that is as general as possible. That will then allow us, on the one hand, to derive the integral principle from d'Alembert's principle, and on the other, to clarify the connection between the two. In order to do that, we first apply the Δ-process to the kinetic energy:

$$T = \sum_i \tfrac{1}{2} m_i \dot{x}_i^2 .$$

From formula (6.a) of no. **23**, we will have:

$$\Delta T = \Delta \sum_i \tfrac{1}{2} m_i \dot{x}_i^2 = \sum_i m_i \dot{x}_i \Delta \dot{x}_i = \sum_i m_i \dot{x}_i \frac{d\,\Delta x_i}{dt} - \sum_i m_i \dot{x}_i^2 \frac{d\,\Delta t}{dt}$$

for the true coordinates x_i, so:

$$\Delta T + 2T \frac{d\,\Delta t}{dt} = \sum_i m_i \dot{x}_i \frac{d\,\Delta x_i}{dt} . \tag{1}$$

Now:

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

and therefore:

$$-\sum_i m_i \ddot{x}_i \Delta x_i = \Delta T + 2T \frac{d\,\Delta t}{dt} - \frac{d}{dt} \sum m_i \dot{x}_i \Delta x_i . \tag{2}$$

Moreover, from (5) in no. **23**, we have:

$$\left.\begin{aligned} \sum_i m_i \ddot{x}_i \Delta x_i &= \sum_i m_i \ddot{x}_i \delta x_i + \sum_i m_i \ddot{x}_i \dot{x}_i \Delta t \\ &= \sum_i m_i \ddot{x}_i \delta x_i + \frac{dT}{dt} \Delta t . \end{aligned}\right\} \tag{3}$$

However, we must still free ourselves of our reference to the special choice of rectangular coordinates.

We will get a conversion of $\sum m_i \dot{x}_i \Delta x_i$ into general coordinates by returning to the explicit form of the kinetic energy. In completely-free coordinates, which can be quasi-coordinates and rheonomic, as well, let:

$$\dot{x}_i = \sum_k \alpha_{ik} \dot{q}_k + \alpha_i ,$$

so, from our definition:

$$\Delta x_i = \sum_k \alpha_{ik} \Delta q_k + \alpha_i \Delta t ,$$

in which the Δq_k, Δt are completely free. Moreover, we will then have:

$$\left.\begin{aligned} 2T &= \sum_{i,j,k} m_i\, \alpha_{ik}\, \alpha_{il}\, \dot{q}_k\, \dot{q}_l + 2\sum_{i,k} m_i\, \alpha_{ik}\, \alpha_{il}\, \dot{q}_k + \sum_i m_i\, \alpha_i^2 \\ &= 2T_2 + 2T_1 + 2T_0\,, \end{aligned}\right\} \tag{4}$$

which will reduce to a homogeneous function of degree two for scleronomic systems, i.e., $\alpha_i = 0$. Upon substituting those expressions, we will easily verify the identity:

$$\sum_i m_i\, \dot{x}_i\, \Delta x_i \;=\; 2T\,\Delta t + \sum_k \frac{\partial T}{\partial \dot{q}_k}(\Delta q_k - \dot{q}_k\,\Delta t) \;=\; 2T\,\Delta t + \sum_k \frac{\partial T}{\partial \dot{q}_k}\delta q_k\ . \tag{5}$$

From (2), (3), and (5), we will finally have:

$$-\sum_i m_i\, \ddot{x}_i\, \Delta x_i + \frac{d}{dt}\left(2T\,\Delta t + \sum_k \frac{\partial T}{\partial \dot{q}_k}\delta q_k\right) = \Delta T + 2T\,\frac{d\,\Delta t}{dt} + \frac{dT}{dt}\Delta t$$

then. If we add the work done by the external forces under the virtual displacement δq_k to both side of that, namely:

$$\sum_i X_i\,\delta x_i \;=\; \sum_k Q_k\,\delta q_k \;=\delta A\ ,$$

and integrate over time between the limits t_1 and t_2 then we will get the following general (*so it is also true for non-holonomic-rheonomic systems*) identity:

$$\int_{t_1}^{t_2}\left\{\Delta T + 2T\,\frac{d\,\Delta t}{dt} + \frac{dT}{dt}\Delta t + \delta A\right\}dt \;=\; \int_{t_i}^{t_2}\left\{\sum_i (X_i - m_i\,\ddot{x}_i)\,\delta x_i\right\}dt + \left[2T\,\Delta t + \sum_k \frac{\partial T}{\partial \dot{q}_k}\delta q_k\right]_{t=t_1}^{t=t_2}, \tag{6}$$

which can be called the *master formula for all integral principles*.

Under the integral on the right-hand side, one finds the expression (2) in no. **8** for d'Alembert's principle precisely, so it will vanish for all motions, and one will then get a statement that is equivalent to d'Alembert's principle in:

$$\left.\begin{aligned} \int_{t_1}^{t_2}\left\{\Delta T + 2T\,\frac{d\,\Delta t}{dt} + \frac{dT}{dt}\Delta t + \delta A\right\}dt \;&=\left[2T\,\Delta t + \sum_k \frac{\partial T}{\partial \dot{q}_k}\delta q_k\right]_{t=t_1}^{t=t_2} \\ &=\left[2T\,\Delta t + \sum_k \frac{\partial T}{\partial \dot{q}_k}(\Delta q_k - \dot{q}_k\,\Delta t)\right]_{t=t_1}^{t=t_2}. \end{aligned}\right\} \tag{7}$$

As will be shown in the next subsection, by specializing that, one will get a series of properties of mechanical trajectories that serve to characterize them uniquely in comparison to certain families of neighboring paths and would then be suitable as principles of dynamics.

Naturally, we can also convert the expression for d'Alembert's principle with the methods that were used for the differential principles themselves. We learned about the general form in generalized coordinates in equation (8) of no. **19**, such that we can also write it:

$$\int_{t_i}^{t_2} \left\{ \sum_i (X_i - m_i \ddot{x}_i)\, \delta x_i \right\} dt = \int_{t_i}^{t_2} \left\{ \sum_k \left(Q_k - \frac{\partial S}{\partial \ddot{q}_k} \right) \delta q_k \right\} dt \,.$$

Brell ([1]) achieved that form by direct calculation. On the other hand, we can also start from the Lagrange equations of the second kind in the best form for them [(2) in no. **12**]. They are equivalent to the identity:

$$\sum_i (m_i \ddot{x}_i - X_i)\, \delta x_i = \sum_k \left\{ \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} - \sum_r \lambda_r a_{rk} - Q_k \right\} \delta q_k \,.$$

The term with the multipliers will drop out as a result of the auxiliary condition:

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

and that will give the identity:

$$\int_{t_i}^{t_2} \left\{ \sum_i (m_i \ddot{x}_i - X_i)\, \delta x_i \right\} dt = \int_{t_1}^{t_2} \left\{ \sum_k \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} - Q_k \right] \delta q_k \right\} dt \qquad (8)$$

when the q_k are true coordinates, and even in the non-holonomic case. Upon substituting (8) in (6), one will get an integral conversion of the Lagrange equations of the second kind that was first exhibited by Voss.

25. General form of Hamilton's principle and the principle of least action ([2]). – With those preliminaries, we shall now go on to exhibit the integral principle itself. If we choose the displacements Δq_k, Δt [δq_k, resp.] in (7) of no. **24** in such a way that the terms on the right-hand side that are free of the integral sign vanish at the limits of integration then we will get the most-general *principle of least action*:

([1]) **H. Brell**, Wiener Ber. (II.a) **122** (1913), pp. 933.

([2]) In addition to the fundamental articles that were cited in no. **23**, see al **H. Brell** and **E. Schenkl**, Verh. d. D. Phys. Ges. **15** (1913), pp. 1082 and *ibid*. **16** (1914), pp. 479.

$$\int_{t_1}^{t_2} \left\{ \Delta T + 2T \frac{d\,\Delta t}{dt} + \frac{dT}{dt} \Delta t + \delta A \right\} dt = 0 , \tag{1}$$

which was first exhibited by Voss, and in which the allowable motions must only satisfy the boundary conditions:

$$\left[2T\Delta t + \sum_k \frac{\partial T}{\partial \dot{q}_k} (\Delta q_k - \dot{q}_k \, \Delta t) \right]_{t=t_1}^{t=t_2} = 0 . \tag{2}$$

Their validity will be guaranteed, e.g., when all Δq_k, Δt vanish for $t = t_1$ and $t = t_2$, so all comparison trajectories will go through the same starting and ending points at the same time.

If we now write the left-hand side of equation (7) in no. **24** in the form:

$$\int_{t_1}^{t_2} \left\{ \Delta T + 2T \, \frac{d\,\Delta t}{dt} + \frac{dT}{dt} \Delta t + \delta A \right\} dt = \int_{t_1}^{t_2} \left\{ \Delta T - \frac{dT}{dt} \Delta t + \delta A \right\} dt + \left[2T \, \Delta t \right]_{t=t_1}^{t=t_2}$$

then the integrated term $2T$ will drop out in comparison to corresponding one on the right-hand side of that equation, and when we consider the fact that:

$$\Delta T - \frac{dT}{dt} \Delta t = \delta T ,$$

which follows from (4) in no. **23**, we will get:

$$\int_{t_1}^{t_2} (\delta T + \delta A)\, dt = \left[\sum_k \frac{\partial T}{\partial \dot{q}_k} \delta q_k \right]_{t=t_1}^{t=t_2} . \tag{3}$$

However, that is ***Hamilton**'s principle* in its most-general form, which is then completely identical to the principle (7) in no. **24**. The integral will vanish under the limit conditions:

$$\left[\sum_k \frac{\partial T}{\partial \dot{q}_k} \delta q_k \right]_{t=t_1}^{t=t_2} = 0 , \tag{4}$$

so, e.g., for $\delta q_k^{(1)} = 0$, $\delta q_k^{(2)} = 0$, i.e., when all curves go through the starting and ending points. It is very remarkable that the variation of time drops out automatically here, so its introduction would not really imply any generalization. Naturally, that is based upon the fact that time is not varied in d'Alembert's principle.

However, we will obtain an essentially new form when we narrow down the field of the paths that are allowed to compete by demanding that:

$$\delta A = \delta T = \Delta T - \frac{dT}{dt}\Delta t \tag{5}$$

at each moment. That means that energy should remain constant under the transition to the neighboring path since the change in the energy is indeed equal to the difference between the kinetic energy and the work done under that displacement. That is not to say that the energy should remain constant during the entire motion, so the law of energy would apply, but only that the true path is always associated with neighboring paths that have corresponding energies at each point, so the law of energy applies to only the transitional motions. If we express the work δA in equation (7) of no. **24** in terms of ΔT by means of the aforementioned transition conditions (5) then we will get simply:

$$\int_{t_1}^{t_2} \left\{ \Delta T + 2T \frac{d\,\Delta t}{dt} \right\} dt = \left[2T\Delta t + \sum_k \frac{\partial T}{\partial \dot{q}_k} (\Delta q_k - \dot{q}_k\, \Delta t) \right]_{t=t_1}^{t=t_2}. \tag{6}$$

We can write the left-hand side symbolically:

$$\int_{t_1}^{t_2} \left\{ \Delta T + 2T \frac{d\,\Delta t}{dt} \right\} dt = 2\Delta \int_{t_1}^{t_2} T\, dt\,, \tag{6.a}$$

in which the symbol Δ is supposed to mean the difference between the integrals over the true and varied trajectories. In fact, from no. **23**, we will have:

$$\Delta \int T\, dt = \int T\,(q_k + \Delta q_k, t + \Delta t) dt - \int T\, dt = \Delta \int T\, dt\,,$$

i.e., when we consider the auxiliary condition (5), the Δ-variation of the integral of the kinetic energy will likewise vanish with the limit condition (2).

Since the integral of the kinetic energy is referred to as the *action*, we will then have the most-general form of the ***Euler-Maupertuis*** *principle of least action*, which is just as applicable to non-holonomic-rheonomic systems. it reads:

The integral of the kinetic energy assumes an extremal value for the true motion in comparison to all Δ*-variations of the trajectory* (see no. **23**) *that satisfy the condition* (5) *for every point of the path and the boundary conditions* (2) *at the endpoints.*

In general, the principle is usually only stated in the scleronomic case since it will then simplify considerably. Namely, if T is a homogeneous quadratic function of the velocities. T_0 and T_1 will then drop out, and from Euler's theorem on homogeneous functions:

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

the boundary condition (2) will become simply:

$$\left[\sum_k \frac{\partial T}{\partial \dot{q}_k}\Delta q_k\right]^{t=t_2}_{t=t_1} = 0\,, \tag{7}$$

i.e., no boundary condition is imposed upon Δt at all, while (7) will certainly be fulfilled when one demands that $\Delta q_k^{(1)} = 0$, $\Delta q_k^{(2)} = 0$. The comparison curves are then the ones that lead from the starting point $q_k = q_k^{(1)}$ to the end point $q_k = q_k^{(2)}$ at any time. The principle of least action was first proved rigorously by **Helmholtz** ([1]) in that form. Thus, the variation of time by the δ-variation in Hamilton's principle is essential to the principle of least action in that form.

The calculation of the time integral over the kinetic energy as the action finds its justification in the fact that it can represented as an integral of the impulse. Namely, let s_n be the arc-length along the path of the n^{th} mass-point and its velocity, so:

$$v_n = \frac{ds_n}{dt}\,,$$

and one will have:

$$\int T\,dt = \int \sum_n m_n v_n \frac{ds_n}{dt}\,dt = \int \sum_n m_n v_n\,ds_n = \int \sum_n p_n\,ds_n\,.$$

Let it be remarked as an aside that one can attempt to perform corresponding conversions of Jourdain's and Gauss's principles, as in no. **24**. One would then arrive at new integral principles in which it would correspondingly not be the coordinates, but the velocities (accelerations, resp.) that are varied, which is naturally not as intuitive as when one considers the varied path as a whole. **Schenkl** ([2]) took a step in that direction when he derived the principle that is analogous to Hamilton's:

$$\int_{t_1}^{t_2}\left\{\delta\frac{d^2T}{dt^2} - \frac{d^2\delta A}{dt^2}\right\}dt = 0$$

from Gauss's principle, in which the symbol δ now refers to the Gaussian variation of the accelerations for fixed configurations and velocities.

26. The Jacobi principle and Hertz's principle. – One can now go further and assume that the law of energy is also valid for the true path, which means that one must naturally restrict oneself to conservative systems. One will then have the energy integral:

([1]) **H. von Helmholtz**, Ber. Berlin Akad. (1887), pp. 225.
([2]) **E. Schenkl**, Wiener Ber. (II.a) **122** (1913), pp. 721.

$$T + U = \text{const.} = E.$$

One can eliminate time from equation (6) in no. **25** with its help, in such a way that a function of the trajectory will now be under the integral sign. The kinetic energy is once more assumed to be a homogeneous quadratic function of the velocities:

$$2\,T = \sum_{k,l} \alpha_{kl}\,\frac{dq_k}{dt}\,\frac{dq_l}{dt}\ . \tag{1}$$

Thus:

$$dt^2 = \sum \frac{1}{2T}\,\alpha_{kl}\,dq_k\,dq_l\,.$$

If we use that in the energy equation $T = E - U$, and substitute that in equation (6) in no. **25** then we will get the ***Jacobi*** *form of the principle of least action:*

$$\Delta \int_{t_1}^{t_2} \sqrt{E-U}\,\sqrt{\sum\ \alpha_{kl}\,dq_k\,dq_l}\ = 0\ . \tag{2}$$

For practical use, it is convenient to introduce a new parameter τ over which we integrate. We can then write:

$$\Delta \int_{t_1}^{t_2} \sqrt{E-U}\,\sqrt{\sum\ \alpha_{kl}\,\frac{dq_k}{d\tau}\,\frac{dq_l}{d\tau}}\ d\tau\ = 0\ . \tag{3}$$

One then takes the fixed starting and ending points of the path to be the limits of integration and assumes that the variations of the coordinates must vanish at them. Since time does not appear at all, one will get only the trajectory itself from Jacobi's principle, which once more represents a true variational principle, while the time evolution of the motion will then be determined by energy equation. If one denotes the derivative with respect to τ by a prime and the integrand by $F\,(q_k, q_k', \tau)$:

$$F = \sqrt{E-U}\,\sqrt{\sum_{k,l} \alpha_{kl}\,q_k'\,q_l'} \tag{4}$$

then from the rules of the calculus of variations, the equations of the trajectory will become:

$$\frac{d}{d\tau}\left(\frac{\partial F}{\partial q_k'}\right) - \frac{\partial F}{\partial q_k} = 0\ . \tag{5}$$

One easily arrives at an *integral form of Hertz's principle of the straightest path* from Jacobi's principle. Namely, if one chooses the parameter to be the arc-length that was defined in no. **17** [cf., (1) of no. **17**, equation (3)]:

$$ds = \sqrt{\sum_{k,l} \alpha_{kl}\, dq_k\, dq_l}$$

then (3) will go to:

$$\Delta \int_{t_1}^{t_2} (E - U)\, ds = 0\,. \tag{6}$$

In particular, if no external forces are present then one will have:

$$U = 0, \qquad E = \text{const.},$$

and (6) will reduce to:

$$\Delta \int_{t_1}^{t_2} ds = 0\,,$$

i.e., the length of the path between the starting and ending point is an extremum for the true path, just as it can also be shown that it is even a minimum for sufficiently-small distances. Naturally, in so doing, the extremum is to be sought among all trajectories that are compatible with the auxiliary kinematical conditions, e.g., for a mass-point constrained to a surface, that would be all curves that lie on the surface. In that way, we will come back to Hertz's principle of the straightest path since from known principles of geometry, the shortest lines are just the geodetic lines, which simultaneously have the property that they are curves of least curvature, so they are straightest lines in the spirit of **Hertz**.

27. Deriving the equations of motion from Hamilton's principle. – Now that we have also shown that, in general, the integral principles are equivalent to d'Alembert's principle and must then be compatible with the equations of motion, all that remains is for us to take the opposite step, namely, to derive the equations of motion once more from the integral principles for the most-general case and only then make them truly fruitful. In so doing, we shall restrict ourselves to Hamilton's principle as the simplest and most-general.

That problem was solved before in no. **22** in the case where a kinetic potential exists by appealing to the rules of the calculus of variations, but not under more-general assumptions. In its most-general form [(3), no. **25**], the principle reads:

$$\int_{t_1}^{t_2} (\delta T + \delta A)\, dt = \left[\sum_k \frac{\partial T}{\partial \dot{q}_k} \delta q_k \right]_{t=t_1}^{t=t_2}. \tag{1}$$

The expression for δA is then known to be regarded as:

$$\delta A = \sum_k Q_k\, \delta q_k\ .$$

One then addresses the problem of finding a corresponding form for δT, namely, a homogeneous linear expression $\sum_k P_k \, \delta q_k$ in the δq_k from which one can infer conclusions about the degree of arbitrariness in the δq_k.

One initially has:

$$\delta T = \sum_k \frac{\partial T}{\partial \dot{q}_k} \delta \dot{q}_k + \sum_k \frac{\partial T}{\partial q_k} \delta q_k \, ,$$

so

$$\int_{t_1}^{t_2} \left\{ \sum_k \frac{\partial T}{\partial \dot{q}_k} \delta \dot{q}_k + \sum_k \frac{\partial T}{\partial q_k} \delta q_k + \sum_k \frac{\partial T}{\partial \dot{q}_k} \delta \dot{q}_k + \sum_k Q_k \, \delta q_k \right\} dt = \left[\sum_k \frac{\partial T}{\partial \dot{q}_k} \delta q_k \right]_{t=t_1}^{t=t_2} . \qquad (2)$$

Now, if we have a holonomic system then the first term can be converted by partial integration since the operations δ and d / dt will commute in that case. One will then get:

$$\int_{t_1}^{t_2} \frac{\partial T}{\partial \dot{q}_k} \delta \dot{q}_k \, dt = \int_{t_1}^{t_2} \frac{\partial T}{\partial \dot{q}_k} \frac{d \, \delta q_k}{dt} dt = \left[\sum_k \frac{\partial T}{\partial \dot{q}_k} \delta q_k \right]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) \delta q_k \, dt \, . \qquad (3)$$

If we substitute that in (2) then the terms that are free of the integral sign will drop out without any sort of restrictions being imposed upon the δq_k at the limits of integration, and we will get:

$$\int_{t_1}^{t_2} \left\{ \sum_k \left[\frac{\partial T}{\partial q_k} - \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) + Q_k \right] \right\} \delta q_k \, dt = 0 \, . \qquad (4)$$

Now, since the δq_k are completely-arbitrary quantities, and that relation will also be true for an arbitrary choice of integration interval, the term in square brackets must vanish by itself, so:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} - Q_k = 0 \, . \qquad (5)$$

However, those are the Lagrange equations of the second kind.

That argument will not suffice for non-holonomic constraints since one must still consider the auxiliary conditions. Therefore, the q_k will no longer be coordinates, but one imposes the auxiliary kinematical conditions upon them that:

$$\sum_k a_{rk} \, \dot{q}_k + a_r = 0 \, , \qquad (6)$$

so the virtual displacements δq_k will be subject to the constraints that:

$$\sum_k a_{rk}\,\delta q_k = 0\ . \tag{7}$$

Naturally, those constraints can also be holonomic, so the differential forms (6) will be integrable, but they do not have to be. However, the q_k should be true coordinates in all cases. (4) will still be correct, except that the δq_k will no longer be free, and one can therefore no longer infer (5) from (4), but only that:

$$\sum_k \left[\frac{\partial T}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) + Q_k\right]\delta q_k = 0\ . \tag{8}$$

One can now consider the auxiliary conditions (6.a) by multiplying them by the undetermined factors λ_r in the known way, adding that to (7), and then treating the δq_k as free. Therefore, in place of (7), one will have:

$$\sum_k \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} - Q_k + \sum_r \lambda_r\, a_{rk}\right]\delta q_k = 0\ ,$$

with which (4) will also be fulfilled since from (6.a) the λ-terms will mutually cancel each other. Now, it will then follow from the reasoning in no. **5** that one will have the Lagrange equations with the auxiliary conditions (2) of no. **12**:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} - Q_k + \sum_k \lambda_k\, a_{rk} = 0\ ,$$

to which one must add the kinematical constraint (6) on the $\dot{q}_k$, from which one will then get sufficiently-many equations for determining the λ_r, as well.

The essential aspect of this argument is the fact that *one can first make use of the auxiliary conditions after the variation.* That is based upon what was established in no. **20**, namely, that the neighboring curves for non-holonomic systems cannot satisfy the auxiliary conditions in their own right, so they will not represent kinematically-possible paths. By contrast, if one includes the auxiliary conditions as one did before in the calculus of variations then one would restrict the field of competition for paths in a way that is not allowable, and Hamilton's principle would no longer be valid since it would imply the false equations:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k} + \sum_k \lambda_k\, a_{rk}\right) - \frac{\partial}{\partial q_k}\left(T + \sum_k \lambda_k\, a_{rk}\,\dot{q}_k\right) - Q_k = 0\ ,$$

as one easily sees.

One can also introduce the quasi-coordinates in no. **13** into Hamilton's principle directly ([1]), although in so doing one must observe only that variation and differentiation will no longer commute for them. Let them be coupled with the true coordinates by the non-integrable relations (see no. **13**):

$$dq_k = \sum_k \beta_{k\rho}\, d\pi_\rho , \qquad \text{so} \qquad \delta q_k = \sum_k \beta_{k\rho}\, \delta\pi_\rho \ .$$

In Hamilton's principle, one will have:

$$\delta A = \sum_\rho \Pi_\rho\, \delta\pi_\rho \ .$$

Moreover, when we once more denote the kinetic energy as a function of $\overset{*}{\pi}_\rho$ by $\mathfrak{T}$, we will have:

$$\delta \mathfrak{T} = \sum_\rho \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho}\, \delta \overset{*}{\pi}_\rho + \sum_\rho \frac{\partial \mathfrak{T}}{\partial \pi_\rho}\, \delta\pi_\rho ,$$

in which we generally introduce:

$$\frac{\partial \mathfrak{T}}{\partial \pi_\rho} = \sum_k \beta_{k\rho} \frac{\partial \mathfrak{T}}{\partial q_k}$$

as only a symbolic abbreviation since the π_ρ themselves have no meaning. The expression for Hamilton's principle will then be precisely as in (2):

$$\int_{t_1}^{t_2} \left\{ \sum_\rho \left(\frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho}\, \delta \overset{*}{\pi}_\rho + \sum_\rho \frac{\partial \mathfrak{T}}{\partial \pi_\rho}\, \delta\pi_\rho + \sum_\rho \Pi_\rho \right) \delta\pi_\rho \right\} dt \ = \left[\sum_\rho \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho}\, \delta\pi_\rho \right]_{t=t_1}^{t=t_2} . \tag{8}$$

In order to convert the first term, we now appeal to the commutation relation (2) of no. **20**:

$$\delta \overset{*}{\pi}_\rho \ = \frac{d}{dt} \delta\pi_\rho - \sum_{\sigma,\tau} \gamma_{\sigma\rho\tau}\, \overset{*}{\pi}_\sigma\, \delta\pi_\tau .$$

We will then have:

$$\int_{t_1}^{t_2} \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho}\, \delta \overset{*}{\pi}_\rho\, dt = \int_{t_1}^{t_2} \left(\frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \frac{d\, \delta\pi_\rho}{dt} - \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \sum_{\sigma,\tau} \gamma_{\sigma\rho\tau}\, \overset{*}{\pi}_\sigma\, \delta\pi_\tau \right) dt \ .$$

The first term on the right-hand side can be converted by partial integration again and will imply that:

([1]) **Cl. Schaefer**, Phys. Zeit. **19** (1918), pp. 406.

$$\int_{t_1}^{t_2} \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho} \frac{d\,\delta\pi_\rho}{dt}\,dt = \left[\frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho}\,\delta\pi_\rho\right] - \int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho}\right)\delta\pi_\rho\,dt\,.$$

If we substitute everything in (8) then upon combining all terms with the same $\delta\pi_\rho$ and making a corresponding re-indexing of the summation sign, we will finally get:

$$\int_{t_1}^{t_2} \left\{\sum_\rho \left[\frac{\partial \mathfrak{T}}{\partial \pi_\rho} - \frac{d}{dt}\left(\frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho}\right) - \sum_{\sigma,\tau} \gamma_{\sigma\rho\tau} \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\sigma} \overset{*}{\pi}_\tau + \Pi_\rho\right]\delta\pi_\tau\right\} dt = 0\,,$$

and therefore since the term in square brackets must again vanish by itself due to the arbitrariness in the $\delta\pi_\rho$, and from the definition (6) in no. **13**, we will have $\gamma_{\tau\sigma\rho} = -\gamma_{\sigma\rho\tau}$, we will get the Lagrange equations of second kind (8) in no. **13**:

$$\frac{d}{dt}\left(\frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\rho}\right) - \sum_{\sigma,\tau} \gamma_{\sigma\rho\tau} \frac{\partial \mathfrak{T}}{\partial \overset{*}{\pi}_\sigma} \overset{*}{\pi}_\tau - \frac{\partial \mathfrak{T}}{\partial \pi_\rho} - \Pi_\rho = 0\,.$$

With that, the equations of motion are derived completely from Hamilton's principle.

Naturally, one can also go through that derivation in the opposite direction and then conclude the Lagrange equations from Hamilton's principle directly.