"Über die Form der Lagrange'schen Gleichungen für nichtholonome, generalisierte Coordinaten," Sitz. Kais. Akad. Wiss. **111** (1902), 1603-14.

On the form of the Lagrange equations for non-holonomic generalized coordinates

By

Ludwig Boltzmann in Vienna
c. M. k. Akad.

(Presented at the session in 18 December 1902)

Translated by D. H. Delphenich

In **Borchartdt**'s Journal for pure and applied mathematics (Bd. 98, Heft 1, pp. 87), I considered the following case as an example to explain a general theorem:

Two pulleys are connected by a belt that can be displaced back and forth parallel to the axes of the pulleys in a manner that is similar to how one displaces a belt from an active pulley to an unused pulley (*Leerscheibe*) or conversely. The one pulley is tapered in one direction, while the other one is tapered in the opposite direction according to a law that makes the same belt always fit when it is displaced in the manner described. Such a displacement of the belt implies a certain variability of the radii r_1 and r_2 of the two pulleys, and therefore the transmission ration $a = r_1 / r_2$, as well.

Therefore, if ω_1 is the angular velocity of the one pulley, while ω_2 is that of the other, then one will have $\omega_2 = a\ \omega_1$. If a is variable then one can choose a and the total angular rotation w_1 for the first pulley during a certain time interval to be the coordinate of a point P that either belongs to the second pulley or is fixed with it. That will be a non-holonomic coordinate, since it does not amount to the same thing when one first changes a and then w_1 , or conversely, first w_1 and then a, by the same amounts.

One would achieve the same effect when one rotates a disc S that cannot slip in any rotating body and which can be displaced parallel to its axis of rotation between two rotating conical bodies that are tapered in the opposite direction.

One might perhaps doubt whether such conditions can be realized without any slipping. However, in any event, the conditions to which we imagine that our mechanical system composed of pulleys is constrained have precisely the same properties that **Hertz** required in his mechanics with non-holonomic constraints (**Hertz**'s *Mechanik*, Book I, Section IV).

The increase in the rectangular coordinates of each of those points can be represented in the form:

$$dx = A\ dw_1 + B\ da,$$

in which $B = 0$, but A depends upon a, such that dx will not be a complete differential. Therefore, e.g., an infinitely-small angle of rotation dw_2 of the second pulley will be equal to $a\,dw_1$.

In the reference that was cited to begin with, I hesitated to not apply the Lagrange equations to this case, but later, I recognized that such an application would not be permissible, as the following considerations will show:

The sum of the *vis vivas* of all of the masses that are linked with the two pulleys can be expressed as a function of a and dw_1 / dt . One has:

$$T = \tfrac{1}{2}(K + La^2)\left(\frac{dw_1}{dt}\right)^2,$$

if t is the time and K and L are the moments of inertia of all of masses that are associated with first (second, respectively) pulley relative to the respective axes of rotation. We think of all of the remaining components as therefore massless, so all of the other belts (the disc S, resp.). Partial differentiation of T with respect to dw_1 / dt will give the correct force that acts on the coordinate w_1 and which we can think of as being applied to a point that is fixed in the first pulley.

a depends upon only the coordinates that determine the position of the belt (disc S, resp.). However, the differential quotients of those coordinates with respect to time do not enter into T, in fact, but those coordinate coordinates probably do enter into T undifferentiated. When one partially differentiates the quantity T, the Lagrange equations will then yield forces along those coordinates that act upon those coordinates; i.e., they try to change them, but they do not exist.

The Lagrange equations are correct for only holonomic coordinates then, and it would probably be worth the effort to ask what equations would appear in place of Lagrange's for non-holonomic coordinates.

In order to answer that question, we next imagine that the positions of the components of an arbitrary mechanical system (an arbitrary number n of material points) are determined by rectangular coordinates. Let x_i be any one of them. Condition equations can exist between them that we can write in the form:

$$\eta_j\,dt + \sum_i \eta_j^i\,dx_i = 0, \qquad j = 1, 2, 3, \dots, \sigma. \tag{1}$$

Some of them can be holonomic, while others can be non-holonomic. Naturally, the left-hand sides of the equations above are integrable when they are holonomic.

The positions of all parts of the mechanical system shall also be determined by generalized coordinates p_k , and indeed we next consider the case in which the number of generalized coordinates is equal to the number s of degrees of freedom of the system, so it is equal to the difference between the number $3n$ of rectangular coordinates and the number σ of condition equations that exist between them.

However, the p_k (or at least one of them) shall be non-holonomic, so the differentials of the rectangular coordinates shall be given by the generalized coordinates with the equations that have the form:

$$dx_i = \xi_i\, dt + \sum_h \xi_i^h\, dp_h\,. \tag{2}$$

The right-hand sides of arbitrarily-many of these equations can be integrable, such that equations in question can be brought into the form:

$$s_i = f\,(t, p_1, p_2, \ldots),$$

however, some of them shall not be integrable, such that generalized coordinates are not holonomic. The quantities ξ and η are functions of the coordinates that might also possibly include the time t explicitly.

Let the forces X_i be given that act upon the various material points of the system in the directions of the rectangular coordinates x_i. As is known, one will then have:

$$\sum_{i=1}^{3n} (X_i - m_i\, \dot{x}_i)\, \delta x_i = 0, \tag{3}$$

from the Lagrange-d'Alembert principle.

In the following, we shall not consider the case of one-sided constraints, for which an inequality sign will enter in place of the equality sign.

The symbols have the usual meanings here, namely, $m_1 = m_2 = m_3$ is the mass of the first material point, $m_4 = m_5 = m_6$ is that of the second, an overhead dot means the first differential quotient with respect to time, two means the second, and δx_i are the so-called virtual displacements, which fulfill the conditions (1) at a certain unvarying moment in time at which $dt = 0$, but δx_i is set equal to dx_i, such that one will have:

$$\sum_{i=1}^{3n} \eta_j^i\, \delta x_i = 0, \qquad j = 1, 2, \ldots, \sigma. \tag{4}$$

We call the differential quotients of x_i with respect to t that follow from equations (2) when we hold all p_i constant the *partial* derivatives with respect to t and denote them by $\partial x_i / \partial t$; $\partial x_i / \partial p_k$ has an analogous meaning, such that one has:

$$\frac{\partial x_i}{\partial t} = \xi_i\,, \qquad \frac{\partial x_i}{\partial p_k} = \xi_i^k\,. \tag{5}$$

It follows from the same equations:

$$\dot{x}_i = \xi_i + \sum_{h=1}^{s} \xi_i^h\, \dot{p}_h\,. \tag{6}$$

By contrast, time is to be kept constant when one defines the δx_i such that one has:

$$\delta x_i = \sum_{h=1}^{s} \xi_i^h \, \delta p_h \,. \tag{7}$$

If one then understands a partial differential quotients with respect to $\dot{x}_i$ to mean one for which the variables t, x_i , and $\dot{x}_i$ are kept constant, except for one, with respect to which one differentiates, so:

$$\frac{\partial \dot{x}_i}{\partial p_h} = \frac{\partial \xi_i}{\partial p_h} + \sum_{h=1}^{s} \frac{\partial \xi_i^h}{\partial p_h} \dot{p}_h \,, \tag{8}$$

$$\frac{\partial \dot{x}_i}{\partial \dot{p}_h} = \xi_i^h = \frac{\partial x_i}{\partial p_h} \,, \tag{9}$$

the last of which follows from equations (5). When one differentiates those equations, it will follow that:

$$\frac{d}{dt}\left(\frac{\partial x_i}{\partial p_h}\right) = \frac{\partial \xi_i^k}{\partial t} + \sum_{k=1}^{s} \frac{\partial \xi_i^k}{\partial p_k} \dot{p}_k \,. \tag{10}$$

One then has:

$$\frac{\partial \dot{x}_i}{\partial \dot{p}_h} - \frac{d}{dt}\left(\frac{\partial x_i}{\partial p_h}\right) = \frac{\partial \xi_i}{\partial p_h} - \frac{\partial \xi_i^h}{\partial t} + \sum_{k=1}^{s} \dot{p}_k \left(\frac{\partial \xi_i^k}{\partial p_h} - \frac{\partial \xi_i^h}{\partial p_k}\right).$$

We now set:

$$\frac{\partial \xi_i}{\partial p_h} - \frac{\partial \xi_i^h}{\partial t} = \mathfrak{r}_i^h \,, \qquad \frac{\partial \xi_i^k}{\partial p_h} - \frac{\partial \xi_i^h}{\partial p_k} = \mathfrak{r}_i^{hk} \,, \tag{11}$$

for brevity, such that we can write:

$$\frac{\partial \dot{x}_i}{\partial \dot{p}_h} - \frac{d}{dt}\left(\frac{\partial x_i}{\partial p_h}\right) = \mathfrak{r}_i^h + \sum_{k=1}^{s} \mathfrak{r}_i^{hk} \, \dot{p}_k \,. \tag{12}$$

The geometric meanings of the quantities that were introduced here are implied by the following considerations:

If one first lets p_h increase by dp_h and then increases p_k by dp_k then the quantities x_i will first increase by $\xi_i^h dp_h$, and then by $\left(\xi_i^k + \frac{\partial \xi_i^k}{\partial p_h} dp_h\right) dp_k$. Those locations in space through which the material point whose mass is m_r is displaced from its initial location during the entire process will be called B_r^{hk}. Now, conversely, let p_k be increased by dp_k and then p_h by dp_h , such that the x_i will first increase by $\xi_i^k dp_k$ and then by $\left(\xi_i^h + \frac{\partial \xi_i^h}{\partial p_k} dp_k\right) dp_h$. In that way, the same material point whose mass is m_r will move from its undisplaced position by A_r^{hk}. One sees directly that when $i = r$, $r + 1$, or $r + 2$, the quantities:

$$\left(\frac{\partial \xi_i^k}{\partial p_h} - \frac{\partial \xi_i^h}{\partial p_k}\right) dp_h \, dp_k = \mathfrak{r}_i^{hk} \, dp_h \, dp_k$$

will be nothing but the projection of the straight connecting line $A_r^{hk} B_r^{hk}$ between the two points A_r^{hk} and B_r^{hk} onto the coordinate axis along which the x_i is measured. If one denotes that projection by $C_r^{hk} D_r^{hk}$ then one will have:

$$\mathfrak{r}_i^{hk} = \lim \frac{C_r^{hk} D_r^{hk}}{dp_h \, dp_k}.$$

Similarly, let E_r^h and F_r^h be the two points in space to which the mass-particle m_i displaces when one first increases t by dt and then p_h by dp_h, and then in the other case, one first increases p_k by dp_k and then t by dt. Furthermore, let $G_r^h H_r^h$ be the projection of $E_r^h F_r^h$ onto the coordinate axes along which the x_i is measured. In the same way that one obtained the geometric meaning of the $\mathfrak{r}_i^{hk}$ before, one will now find that:

$$\mathfrak{r}_i^h = \lim \frac{G_r^h H_r^h}{dt \, dp_h}.$$

If one denotes the factors that **Lagrange** multiplied the condition equations (4) by λ_1, λ_2, …, λ_σ then it will follow from (3) and (4) in the known way that:

$$X_i = m \ddot{x}_i + \sum_{j=1}^{\sigma} \lambda_j \eta_j^i , \qquad i = 1, 2, 3 \ldots, n. \tag{13}$$

In one introduce the δp_h, in place of the δx_i, in the expression by using equations (7) then one will get:

$$\sum_{i=1}^{3n} X_i \, \delta x_i = \sum_{k=1}^{s} \sum_{i=1}^{3n} X_i \, \xi_i^h \, \delta p_h . \tag{14}$$

The coefficient of δp_h in the expression on the right shall be called the force that acts upon the coordinate p_h, as it is for holonomic generalized coordinates, and denote it by P_h, such that one will have:

$$P_h = \sum_{i=1}^{3n} X_i \, \xi_i^h = \sum_{i=1}^{3n} \left(m \ddot{x}_i + \sum_{j=1}^{\sigma} \lambda_j \eta_j^i \right) \frac{\partial x_i}{\partial p_h}; \tag{15}$$

the latter is true because of equation (13).

We would now like to differentiate the expression $\dot{x}_i \frac{\partial \dot{x}_i}{\partial p_h}$ with respect to time, as one cares to do in the derivation of Lagrange's equations. It follows that:

$$\frac{d}{dt}\left(\dot{x}_i \frac{\partial x_i}{\partial p_h}\right) = \ddot{x}_i \frac{\partial x_i}{\partial p_h} + \dot{x}_i \frac{d}{dt}\left(\frac{\partial x_i}{\partial p_h}\right). \tag{16}$$

For holonomic coordinates, one can obviously set:

$$\frac{d}{dt}\left(\frac{\partial x_i}{\partial p_h}\right) = \frac{\partial \dot{x}_i}{\partial p_h}.$$

That will no longer be the case only for non-holonomic ones. One will then have:

$$\dot{x}_i = \xi_i + \sum_{k=1}^{s} \xi_i^k \dot{p}_k ,$$

so:

$$\frac{\partial \dot{x}_i}{\partial p_h} = \frac{\partial \xi_i}{\partial p_h} + \sum_{k=1}^{s} \frac{\partial \xi_i^k}{\partial p_h} \dot{p}_k ,$$

while:

$$\frac{\partial \dot{x}_i}{\partial p_h} = \xi_i^h ;$$

hence:

$$\frac{d}{dt}\left(\frac{\partial x_i}{\partial p_h}\right) = \frac{d\xi_i^k}{dt} + \sum_{k=1}^{s} \frac{\partial \xi_i^h}{\partial p_k} \dot{p}_k .$$

One then has:

$$\frac{d}{dt}\left(\dot{x}_i \frac{\partial x_i}{\partial p_h}\right) = \ddot{x}_i \frac{\partial x_i}{\partial p_h} + \dot{x}_i \frac{\partial \dot{x}_i}{\partial p_h} + \dot{x}_i \left(\frac{d}{dt}\frac{\partial x_i}{\partial p_h} - \frac{\partial \dot{x}_i}{\partial p_h}\right), \tag{17}$$

or, from equation (12):

$$\frac{d}{dt}\left(\dot{x}_i \frac{\partial x_i}{\partial p_h}\right) = \ddot{x}_i \frac{\partial x_i}{\partial p_h} + \dot{x}_i \frac{\partial \dot{x}_i}{\partial p_h} - \dot{x}_i \left(\mathfrak{r}_i^k + \sum_{i=1}^{s} \mathfrak{r}_i^{kk} \dot{p}_k\right). \tag{18}$$

We now multiply that equation by m_i, add $\sum_{j=1}^{\sigma} \lambda_j \eta_j^i \frac{\partial x_i}{\partial p_h}$ to both sides, and finally sum over i from 1 to $3n$.

If we begin with the far left and proceed to the terms in equation (18) that lie increasingly to the right then:

1. From equation (9):

$$\sum_{i=1}^{3n} m_i \frac{d}{dt}\left(\dot{x}_i \frac{\partial \dot{x}_i}{\partial p_h} \right) = \frac{d}{dt}\sum_{i=1}^{3n} m_i \, \dot{x}_i \frac{\partial \dot{x}_i}{\partial \dot{p}_h} = \frac{dq_h}{dt} . \tag{19}$$

q_h has the known meaning. It is the momentum that relates to the h^{th} coordinate, and will be defined when one expresses the *vis viva*:

$$T = \sum_{i=1}^{3n} m_i \, \ddot{x}_i^2 \tag{19}$$

as a function of the p_h and $\dot{p}_h$, and then partially differentiates with respect to $\dot{p}_h$.

2. The p_h shall fulfill the condition equations identically. For constant time, one will then have:

$$\sum_{i=1}^{3n} \eta_j^i \, \delta x_i = 0 \tag{20}$$

for each $j = 1, 2, \ldots, \sigma$ when the p_h change arbitrarily, so also when all of the other ones remain constant along with time, except for one, which we would like to call p_h. In other words, we will have:

$$\sum_{i=1}^{3n} \eta_j^i \frac{\partial x_i}{\partial p_h} = 0, \qquad j = 1, 2, 3, \ldots, \sigma \tag{21}$$

for each value of j and k.

When the time also increases by δt, the x will become somewhat different functions of p, in addition, and one will have:

$$\eta_j \, dt + \sum_{i=1}^{3n} \eta_j^i \, \delta x_i = 0$$

for every j.

However, that equation is not true for our present considerations, since no variation of time was coupled with any of the variations that were denoted by the sign δ up to now.

3. From equation (15), one has:

$$\sum_{i=1}^{3n} (m_i \, \ddot{x}_i + \lambda_j \, \eta_j^i) \frac{\partial x_i}{\partial p_h} = P_h . \tag{22}$$

4. If follows from equation (19) that:

$$\sum_{i=1}^{3n} m_i \, \dot{x}_i \frac{\partial \dot{x}_i}{\partial p_h} = \frac{\partial T}{\partial p_h} . \tag{23}$$

That will then imply that:

$$\frac{dq_h}{dt} = P_h + \frac{\partial T}{\partial p_h} - \sum_{i=1}^{3n} m_i\,\dot{x}_i \left(\mathfrak{r}_i^h + \sum_{k=1}^{s} \mathfrak{r}_i^{hk} \right). \tag{24}$$

Now let $\mathfrak{v}_r$ be the magnitude, as well as the direction, of the velocity of the r^{th} material point, which has mass $m_r = m_{r+1} = m_{r+2}$, such that $\dot{x}_r$, $\dot{x}_{r+1}$, $\dot{x}_{r+2}$ are the components of $\mathfrak{v}_r$ along the three coordinate directions. Moreover, let $\mathfrak{u}_r^h$ and $\mathfrak{u}_r^{hk}$ be the magnitudes and directions of the lines that were previously denoted by $E_r^h F_r^h$ and $A_r^{hk} B_r^{hk}$, resp. One can also write equation (20) in the form:

$$\frac{dq_h}{dt} = P_h + \frac{\partial T}{\partial p_h} + \sum_r m_r\,\mathfrak{v}_r \left[\mathfrak{u}_r^h \cos(\mathfrak{v}_r, \mathfrak{u}_r^h) + \sum_{k=1}^{s} \mathfrak{u}_i^{hk} \cos(\mathfrak{v}_r, \mathfrak{u}_r^{hk}) \right], \tag{25}$$

in which r runs through only the values 1, 4, 7, …, $3n - 2$ in the first sum.

1. then proves that the Lagrange equations, in their unaltered form, are not valid applied to a non-holonomic coordinate, and 2. gives the correction term that one must add in order for them to be valid again.

The proof will suffer only an inessential modification when the number n of generalized coordinates is greater than the number $s = 3n - \sigma$ of degrees of freedom in the system. $\nu - s = \tau$ condition equations will then remain between the generalized coordinates, some of which can be holonomic, while other can be non-holonomic. Of the σ condition equations that exist between the rectangular coordinates, only $\sigma - \tau$ of them will be fulfilled identically by the generalized coordinates then.

The σ equations (4) will be true for the variations δx_i of the rectangular coordinates at constant time, as before. We can combine all of those equations into a single one when we multiply each of them by an arbitrary factor λ and add them together. In that way, we will obtain the resulting equation:

$$\sum_{j=1}^{\sigma} \sum_{i=1}^{3n} \lambda_j\,\eta_j^i\,\delta x_i = 0. \tag{26}$$

Establishing that this equation should be true for arbitrary values of the λ comes down to the σ equations (4) completely.

If we now replace the δx in equation (26) with the δp then the number of arbitrary factors λ must be reduced from σ to $\mathfrak{r}$, since indeed only τ equations exist between the δp, which we would like to write in the form:

$$\sum_{i=1}^{r} \zeta_j^i\,\delta p_i = 0, \qquad j = 1, 2, \ldots, \mathfrak{r}. \tag{27}$$

Equation (26) must then reduce to the following one when one introduces the δp :

$$\sum_{j=1}^{\tau}\sum_{i=1}^{r}\mu_j^i\,\zeta_j^i\,\delta p_i = 0,$$

in which the μ are likewise τ linear mutually-independent functions of the λ.

The factors $\lambda_1, \lambda_2, \ldots, \lambda_\sigma$ are now chosen in such a way that the expression:

$$\sum_{i=1}^{3n}\left(X_i - m_i\frac{d^2x_i}{dt^2} + \sum_{j=1}^{\sigma}\lambda_j\,\eta_j^i\right)\delta x_i$$

vanishes for all values of the δx_i. After one introduces the generalized coordinates, that expression will be converted into:

$$\sum_{h=1}^{\nu}\left\{P_h - \frac{dq_h}{dt} + \frac{\partial T}{\partial p_h} - \sum_{i=1}^{3n} m_i\,\dot{x}_i\left(\mathfrak{r}_i^h + \sum_{k=1}^{\nu}\mathfrak{r}_i^{hk}\right) + \sum_{j=1}^{\tau}\mu_j\,\zeta_j^i\right\}\delta p_i = 0,$$

or

$$\sum_{h=1}^{\nu}\left\{P_h - \frac{dq_h}{dt} + \frac{\partial T}{\partial p_h} + \sum_r m_r\,\mathfrak{v}_r\left[\mathfrak{u}_i^h\cos(\mathfrak{v}_r\mathfrak{u}_r^h) + \sum_{k=1}^{\nu}\mathfrak{u}_i^{hk}\cos(\mathfrak{v}_r\mathfrak{u}_r^{hk})\right] + \sum_{j=1}^{\tau}\mu_j\,\zeta_j^i\right\}\delta p_i = 0,$$

in which r again runs through the values 1, 4, 7, …, $3n - 2$. Due to the choice of λ, from which analogous properties will result for the μ, the left-hand side of the last two equations will vanish for all possible values of the δp_i, and one will get the equations of motion:

$$\frac{dq_h}{dt} = P_h + \frac{\partial T}{\partial p_h} + \sum_{i=1}^{3n} m_r\,\dot{x}_r\left(\mathfrak{r}_i^k + \sum_{k=1}^{\nu}\mathfrak{r}_i^{hk}\right) + \sum_{j=1}^{\tau}\mu_j\,\zeta_j^i\ ,$$

or

$$\frac{dq_h}{dt} = P_h + \frac{\partial T}{\partial p_h} + \sum_r m_r\,\mathfrak{v}_r\left[\mathfrak{u}_i^h\cos(\mathfrak{v}_r\mathfrak{u}_r^h) + \sum_{k=1}^{\nu}\mathfrak{u}_i^{hk}\cos(\mathfrak{v}_r\mathfrak{u}_r^{hk})\right] + \sum_{j=1}^{\tau}\mu_j\,\zeta_j^i = 0.$$

The additional terms in the Lagrange equations that are required by the lack of holonomity in the coordinates are therefore entirely the same as in the case where the number of generalized coordinates is equal to the number of degrees of freedom in the system, such that no equations between the generalized coordinates remain any longer, and the problem that was posed is then solved in full generality.

____