THE CINDERELLAS OF ANALYTICAL MECHANICS

i.e.

THE LAGRANGE EQUATIONS, FIRST FORM.

DE

3rd Edition

CENERENTOLA
https://upload.wikimedia.org/wikipedia/commons/8/86/John_Everett_Millais_-_Cinderella.jpg
John Everett Millais [Public domain], via Wikimedia Commons

The Lagrange Equations, first form, are one among the most neglected results and concepts of analytical mechanics. In fact, they are seldom used, and their fame is obscured by their sisters, the Lagrange Equations of the

second form, in which use is made of the so-called generalized or Lagrangian coordinates (generally written q_i) and the concept of Lagrangian function (generally written L) is introduced. (Neither concept is necessary in the Lagrange's equation, first form.)
Modern advanced quantum mechanics still uses the concept of Lagrangian Function. The use of the Lagrange Equation , First form, reappears in analytical mechanics when "non-holonomic" velocity dependent constraints appear in problems, which the Lagrange equations of the Second form cannot solve by themselves. However, the student beware: as soon as he sees the symbols q_i and L, he has abandoned the clean kingdom of the First Form equations and has entered a hybrid kingdom.

The only new symbols in the First form are the Lagrange multipliers (usually written λ_i). The multipliers appear here in a new context, only remotely related to the problem of finding the maxima of functions subjected to constraints, which is the typical problem in which use of the Lagrange multipliers is first met by the student. We will see later whether the relationship between the two problems can be at least vaguely clarified.

In any case, normally, even celebrated textbook avoid the use of Lagrange Equation of the first type. At best, they use them when non-holonomic constraints enter the game.

In this short essay I will try to clarify a number of painful pebbles in my shoes:

1. The various types of constraints.

2. Holonomic and non-holonomic systems.

3. Virtual displacements.

4. Constraint equations. Frictionless constraints. A hint to the principle of virtual works.

(A translated paraphrase of the notes handed out by Prof. T. Zeuli, of the University of Turin, in 1962).

Although we will be dealing with one among the most neglected of all ideas of Lagrange, some introduction is necessary, at least to clarify the language we will use.

What makes the motion of systems of particles interesting, and stimulated the gigantic growth of mechanics in the XVIII-XIX century, making new ideas possible, is the fact that *particles of the systems of interest are not free*. Otherwise, all particles in the system could be dealt with singularly, and Newton's Laws would be amply sufficient to describe the motion of a constraint-free system.

1. For us, a "constraint" will be any *condition* which limits the freedom of the motion of the point masses a mechanical system consists of. To be tractable, such a constraint must have a mathematical expression, that is, it must be expressed by an *equation or an inequality*, either in finite or in differential form, linking the coordinates of the points of the system.

This fact gives us the possibility of classifying the constraints on the basis of the nature of the equations and inequalities which represent them analytically.

i) The simplest case is when the constraint is given by an equation such as:

$$(1) \qquad f(x_1, y_1, z_1, \ x_2, y_2, z_2, \ \ldots . . x_N, y_N, z_N, \ \ t) = 0$$

The constraint thus represented is called a "bilateral positional constraint".

For example, suppose that we have a sphere whose radius R is a function of time. We can constrain a point to move on such a surface by forcing it to move in the interspace between two spheres with the same centre, whose radii differ by an infinitely small amount. The constraint to which our (single) point will be subjected will then be represented by the equation:

$$x_1{}^2 + y_1{}^2 + z_1^2 = R^2(t)$$

Here the distance of the point from the centre is fixed, i.e. the point cannot move neither above nor below the spherical surface, which explains why the constraint is called bilateral.

ii) Equally simple to express but not necessarily equally simple to deal with, is the case in which the constraint equation is given by:

$$(2) \qquad f(x_1, y_1, z_1,\ x_2, y_2, z_2,\ \ldots..x_N, y_N, z_N,\ \ t) \geq 0$$

It is the case, for example, in which we remove one of the two infinitely close spheres of the above example, which will then allow the point to leave the spherical surface, being however constrained not to go to the inside of the sphere (or non to go outside), according to how we write the relationship between R^2 and the sum $x^2 + y^2 + z^2$.

We will call the constraints represented in this form "unilateral positional constraints".

*We thus see that, in simpler words, "positional" means that the constraint can define a relationship of type (1), that is, **in finite terms**, that is, not involving derivatives and infinitesimal calculus terms, between the coordinates of the points of the system; while "bilateral" means that the relationship can be written as an expression **equal to zero**, without "greater than" or "less than"" signs.*

Just for completeness we can add that the configurations in which the $> sign$ is satisfied are called "ordinary configurations", while those in which the = *sign* is satisfied are called "boundary configurations".

If the time t does not appear explicitly in the equations (1) or (2), the constraints will be called "fixed" or "time independent". Another, more high-brow name meaning exactly the same thing is "scleronomic constraint". On the other hand, if time appears in equations (1) or (2), we talk about "time dependent", or "rheonomic" constraints.

I stressed the word "explicitly", because time is underlying, implicitly, all problems of dynamics. No time, no motion – no problem.

Other examples of particular interest:

(a) A system consisting of two points $P_1(x_1,\ y_1, z_1)$ e $P_2(x_2,\ y_2, z_2)$ whose distance d does not vary with time. The analytical expression of such constraint will be:

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = d^2$$

According to the above "glossary", this is a bilateral, positional, fixed (scleronomic) constraint.

(b) A **free rigid system** (including a rigid body, which we can imagine as consisting of a very large number of points rigidly connected to each other, by links as in example (a) above).

If the system consists of N points, there will be N(N-1)/2 distances which must be constant in time (this is nothing but the number of the combinations of N elements taken two at the time).

However, if N>4, *such distances are not all independent*. In fact, if we want that all points are kept at the same distance from each other, it is sufficient to impose that any three points, let's say P_1, P_2, P_3, keep the same distances among them and all other N-3 points keep unchanged their distance from P_1, P_2, P_3. We thus have a total of 3 + 3(N-3) = 3N-6 independent equalities among the 3N coordinates of the N points. A rigid, N-point system is thus subjected to 3N-6 bilateral positional, time independent constraints. (Why must N be greater than 4? Because if N =4 we have that 3N-6 = N(N-1)/2, which means that the number is the same as if the distances between couples of points were all *independent*). Such constraints have been given the name of "rigidity constraints". Please note that the 3 in 3N has *nothing to do* with the number of coordinates of each point, but refers to the number of points which are kept fixed in the system.

(c) A rigid body with a fixed point.

In this case, having called the fixed point $P_c(x_c,\ y_c, z_c)$, we have three more conditions, namely:

$$x_c = const;\ y_c = const;\ z_c = const.$$

The system is thus subjected to 3N-6+3 = 3(N-1) positional, bilateral, time independent constraints.

Each condition, even involving a single coordinate, we recall, represents a constraint.

2. The positional constraints, whether unilateral or bilateral (Equations 1 and 2) limit the configurations of the system at each instant. But they limit also all the possible displacement of all its points dP_i (dx_i, dy_i, dz_j). In fact, taking (1) as a starting point, differentiation gives us:

$$\sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}dx_i + \frac{\partial f}{\partial y_i}dy_i + \frac{\partial f}{\partial z_i}dz_i\right) + \frac{\partial f}{\partial t}dt = 0$$

And, after division by dt,

$$(3') \qquad \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\dot{x}_i + \frac{\partial f}{\partial y_i}\dot{y}_i + \frac{\partial f}{\partial z_i}\dot{z}_i\right) + \frac{\partial f}{\partial t} = 0$$

Here the summation includes 3N terms, extending, as it does, to all coordinates appearing in (1), here grouped in groups of three (a notation which is irrelevant to the calculation). On the other hand, $\frac{\partial f}{\partial t}$ is outside of the summation.

I have also introduced Newton's "dot" notation for time derivatives: $\dot{x}_i \equiv \frac{dx}{dt}$.

There exists, therefore, a "linear" relationship between the velocity components $\dot{x}_i, \dot{y}_i, \dot{z}_i$ of the velocities $\vec{v_i}$, the coordinates x_i, y_i, z_i of the points P_i of the system, and, in case, the time t. If the constraint (1) is time independent, then we say that $(3')$ is "homogeneous" in $\dot{x}_i, \dot{y}_i, \dot{z}_i$.

The derivatives $\frac{\partial f}{\partial x_i}$ are just functions of the coordinates. Accordingly, we might think that constraints could also exist, which could be expressed by more general equations such as:

$$\sum_{i=1}^{N}(\alpha_i dx_i + \beta_i dy_i + \gamma_i dz_i) + \tau\, dt = 0$$

Or, alternatively:

$$(4') \qquad \sum_{i=1}^{N}(\alpha_i \dot{x}_i + \beta_i \dot{y}_i + \gamma_i \dot{z}_i) + \tau = 0$$

Here, the $\alpha_i, \beta_i, \gamma_i$ are functions of the coordinates x_i, y_i, z_i and, if in case, of time t. Such functions are predefined in any possible way. If there exists a function $f(x_i, y_i, z_i, t)$ such that

$$(5) \qquad \alpha_i = \frac{\partial f}{\partial x_i}; \quad \beta_i = \frac{\partial f}{\partial y_i}; \ \gamma_i = \frac{\partial f}{\partial z_i}; \ \tau = \frac{\partial f}{\partial t}$$

then $(4')$ identifies with $(3')$ and does not tell us anything more than the form:

$$f(x_1, y_1, z_1, \ x_2, y_2, z_2, \ \ldots\ldots x_N, y_N, z_N, \ \ t) = 0$$

which can be obtained by integrating $(4')$ with respect to time.

For example, suppose we have the equation expressing the relationship between velocities written as

$$x\dot{x} + y\dot{y} + z\dot{z} = 0$$

This can be readily integrated as:

$$x^2 + y^2 + z^2 = C$$

which is the equation we already know of a point constrained to stay on a spherical surface.

In fact,

$$\frac{df}{dt} = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\dot{x}_i + \frac{\partial f}{\partial y_i}\dot{y}_i + \frac{\partial f}{\partial z_i}\dot{z}_i\right) + \frac{\partial f}{\partial t}$$

If, however, such is not the case, i.e. there is no function $f(x_i, y_i, z_i, t)$ for which

$$\alpha_i = \frac{\partial f}{\partial x_i}; \quad \beta_i = \frac{\partial f}{\partial y_i}; \; \gamma_i = \frac{\partial f}{\partial z_i}; \; \tau = \frac{\partial f}{\partial t}$$

and therefore

$$\sum_{i=1}^{N} (\alpha_i dx_i + \beta_i dy_i + \gamma_i dz_i) + \tau \, dt = 0$$

cannot be deduced by differentiation of a function "in finite terms" such as

$$f(x_1, y_1, z_1, \; x_2, y_2, z_2, \; \ldots\ldots x_N, y_N, z_N, \; \mathrm{t}) = 0$$

(which, we know, represents a "positional constraint") we then say that the constraint represented by $(4')$, is a bilateral "mobility" constraint.

Of course, if instead of the symbol " = " we had in $(4')$ the symbol "$\geq$ ", then we would still say that the constraint is unilateral, and we could consider again "ordinary" and "boundary" configurations.

If the mobility constraint does not depend on time, this means that
1) $\alpha_i, \; \beta_i, \; \gamma_i$ do not depend on time;
2) the τ term is absent.
As a consequence, the first member of $(4')$ will be homogeneous in terms of the $\dot{x}_i, \dot{y}_i, \dot{z}_i$.

In conclusion, of all constraints, which can be expressed via $(4'), i.e.$

$$\sum_{i=1}^{N} (\alpha_i \dot{x}_i + \beta_i \dot{y}_i + \gamma_i \dot{z}_i) + \tau = 0$$

those for which one can find a function f such that:

$$(5) \qquad \alpha_i = \frac{\partial f}{\partial x_i}; \quad \beta_i = \frac{\partial f}{\partial y_i}; \; \gamma_i = \frac{\partial f}{\partial z_i}; \; \tau = \frac{\partial f}{\partial t}$$

are "positional " constraints; the others are "mobility" constraints.

2. Holonomic and non-holonomic systems.

(A translated paraphrase of the notes handed out by Prof. T. Zeuli, of the University of Turin, in 1962-63).

2.1 *Definition*: a mechanical system is said to be "holonomic" if all constraints to which it is subjected are positional and bilateral. On the other hand they may be time-dependent.

Let's now suppose that there are m constraint equations

(6) $$f_r(x_1, y_1, z_1,\ x_2, y_2, z_2,\ \ldots..x_N, y_N, z_N,\ \ t) = 0 \quad (r = 1,2,\ldots m)$$

(where we will obviously assume that m < 3N). Then, if n is the difference between *3N* (number of cartesian coordinates of the points of the system) and *m* (number of the independent constraints), we can say that *n* is the "number of degrees of freedom" of the system, or that *n* is "the degree of freedom of the system".

Examples:

i) a point free (unconstrained) in space has three degrees of freedom (3-0);

ii) a point constrained to move on a surface has 3-1 =2 degrees of freedom;

iii) a point constrained to move on a line (which is determined by two equations in space), has 3-2 =1 degrees of freedom;

iv) a system consisting of two points at a constant distance has 3 x 2 -1 = 5 degrees of freedom;

v) a rigid free system consisting of N points has 3N – (3N-6) =6 degrees of freedom. As this number 6 is independent of the number of points N, also a rigid body, which we can treat as if it were a system consisting of a very large number N of mass points still has 6 degrees of freedom. They reduce to 3 if one point is fixed.

We say that a system has n *degrees of freedom*, because in order to fully know at each instant t the configuration of the system, it is sufficient to know the value of only *n* among the *3N* coordinates. Indeed, once they are known, the *m* equations (6), which are independent, give the values of all *3N-n (=m)* remaining coordinates. If you prefer, just think that $n+m=3N$: you know n coordinates, and you derive the remaining m from the m constraint equations.

2.2. Lagrangian coordinates of a holonomic system

The fact that the knowledge of n (= 3N-m) independent cartesian coordinates is sufficient to determine the configuration of the system, suggests a new possibility, that instead of the n independent cartesian coordinates, one uses *n* new parameters ($q_1, q_2, \dots q_n$), also independent from each other (plus the time, if necessary), to determine the 3N Cartesian coordinates of all points of the system as follows:

$$(7) \qquad \begin{cases} x_i = x_i(q_1, q_2 \dots q_n, t) \\ y_i = y_i(q_1, q_2 \dots q_n, t) \\ z_i = z_i(q_1, q_2 \dots q_n, t) \end{cases} \qquad (i = 1,2 \dots . N)$$

That is

$$(7') \qquad P_i = P_i(q_1, q_2 \dots q_n, t) \qquad (i = 1,2 \dots . N)$$

For example, if P is a single point on the surface of a sphere of constant radius r , the constraint equation must be respected

$$x^2 + y^2 + z^2 = r^2$$

and for this reason P has only two degrees of freedom, expressed by two coordinates q_1, q_2. We *can* select for q_1 the co-latitude θ, and for q_2 the longitude φ, and the (7) becomes:

$$x = \sin\theta\, cos\varphi; \quad y = r\, \sin\theta\, sin\varphi; \quad z = r\cos\theta\ .$$

We can see that the (7) represent a holonomic system with n degrees of freedom, i.e a system with 3N-n = m positional bilateral constraints. To do so, we can think of obtaining the n values of the q_i from n of the 3N equations in (7) as function of the x_i, y_i, z_i. If we substitute these values, now in terms of the x_i, y_i, z_i, in the remaining 3N-n = m relationships, we obtain precisely m equations in terms of the same variables, which are of the form

(6) $$f_r(x_1, y_1, z_1,\ x_2, y_2, z_2,\ \ldots\ldots x_N, y_N, z_N,\ \ t) = 0 \quad (r = 1,2,\ldots m)$$

i.e. positional, bilateral (=holonomic) constraints.

Not convinced? Let's study the simpler case of a point constrained to remain on a circumference of radius 1. The system has two Cartesian coordinates and one constraint, and therefore one degree of freedom. The two equations of the two Cartesian coordinates expressed in term of $q_1 = \theta$ are:

$x = \cos\theta$ and $y = \sin\theta$

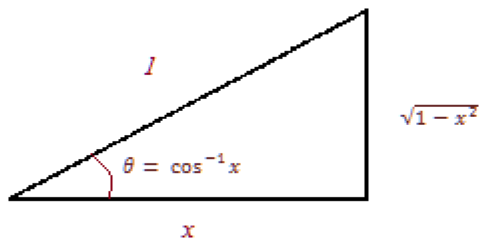
Basically, what we want is to show that with these data one can derive the equation of the constraint, and that the constraint is holonomic (i.e. positional and bilateral). From the first equation we express θ in terms of x. We obtain:

$$\theta = \cos^{-1} x$$

Now we substitute this result in y, and get

$y = \sin(\cos^{-1} x) = \sqrt{1 - x^2}$

I remind the student than one of the simplest ways to reconstruct such relationship is to draw a suitable triangle:



Squaring both terms, we obtain $x^2 + y^2 = 1$, which is the equation expressing our bilateral, positional constraint.

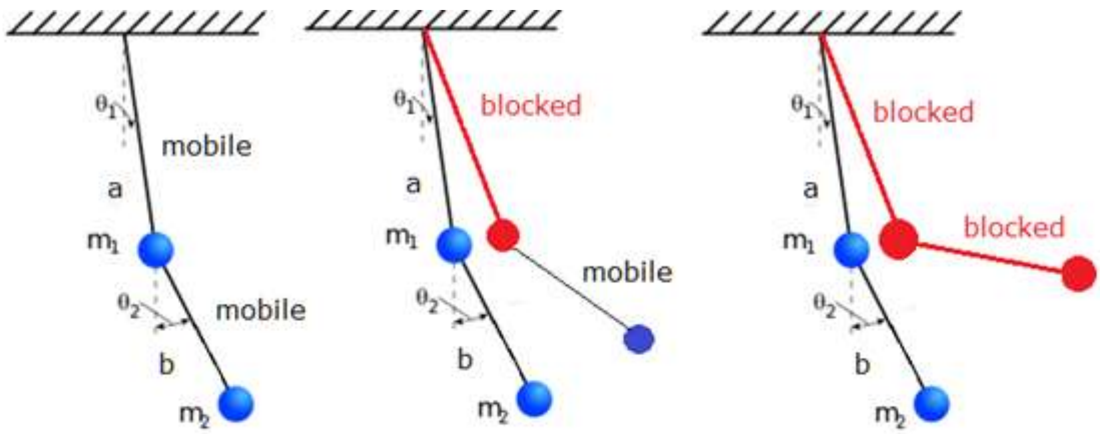
(Naming in a different way the catheti provides other not immediate relationships between the inverse trigonometric functions).

If in equations (7) time does not appear explicitly, the constraints are time independent.

The parameters $q_1, q_2 \dots q_n$ which are *independent among themselves*, and *as many as the degrees of freedom*, in honour of Lagrange are called "Lagrangian coordinates". Their time derivatives ($\dot{q}_1, \dot{q}_2, \dot{q}_3, \dots \dot{q}_n$) are called "Lagrangian velocities" (in case q_i has the dimensions of a length, then it has the dimension of a velocity, L/T), and the second derivatives, such as $\ddot{q}_i$ are called "Lagrangian accelerations".

The number of degrees of freedom coincides with the number n of the independent parameters $q_1, q_2 \dots q_n$. Once such parameters are fixed, the mechanical system is blocked in a given configurations. For example, by fixing θ and φ we block the position of a point P on the surface of a sphere (the system we have considered above). This considerations provides an easy way to determine the degrees of freedom of the system:

The plane double pendulum has two degrees of freedom: we first block the angle $\theta_1$and then the angle θ_1, and the double pendulum is "frozen" in position.



If one wants a more challenging problem, he can calculate the degrees of freedom of a (very essential) bicycle.

Odd as it may seem, many sites on the Web deal with the problem of the number of degrees of freedom of a bicycle. Wikipedia, for the idealized "*Whipple model*" gives a minimum of 7 degrees of freedom.

i) two for the rear wheel to fix the contact point on the road plane;

ii) two (inclination and rotation) of the plane of the rear wheel;

iii) one for the rotation of the front wheel with respect to the bicycle frame (steering angle);

iv) one for each wheel to fix the rotation angle of the wheel with respect to its axis (total two).

So far we already have listed *seven degrees of freedom*, and we know that there are many more mobile parts of the bicycle.

https://en.wikipedia.org/wiki/Bicycle_and_motorcycle_dynamics

As we have seen, the lagrangian coordinates can vary independently from each other, which indicates that there are n (and only n) elementary independent displacements of the system: any other elementary displacement is the resultant of two or more such elementary displacements.

Examples:

> (i)A point on a spherical surface: two elementary displacements, along the meridian ($d\theta$) and along the parallel ($d\varphi$). Any other displacement is the resultant of two such displacements;
>
> (ii) A rigid body: six elementary displacements are possible, three translations and three rotations. Any other displacement is the resultant of two such displacements;
>
> (iii) A rigid, infinitely thin bar: five elementary displacements, three translations and two rotations. The third rigid body rotation, around the line joining the two extremes of the bar, is supposed to leave the bar unchanged.

2.3 Velocity of the points of a holonomic system

Let's take the time derivative of

$$(7') \qquad P_i = P_i(\, q_1, q_2 \dots q_n, t) \qquad\qquad (i = 1,2 \dots . N)$$

We have just to remember the chain rule of derivation, and the fact that the

$(q_1, q_2, \ldots q_n)$, all depend on time. Indicating with $\vec{v}_i$ the velocity of P_i we have:

$$(8) \qquad \vec{v}_i = \sum_{r=1}^{n} \frac{\partial P_i}{\partial q_r} \dot{q}_r + \frac{\partial P_i}{\partial t} \qquad (i = 1,2 \ldots . N)$$

If the constraints are time independent (or fixed, or scleronomic), then

$\frac{\partial P_i}{\partial t} = 0$ and the velocities are simply given by:

$$(9) \qquad \vec{v}_i = \sum_{r=1}^{n} \frac{\partial P_i}{\partial q_r} \dot{q}_r \qquad (i = 1,2 \ldots . N)$$

That is, the velocity of all points of the system are linear homogeneous functions of the lagrangian velocities – but, I repeat, only if the constraints are time-independent.

2.4 Definition of non-holonomic system.

A system is said to be non-holonomic if it is subjected to (i) bilateral constraints, which (ii) cannot be reduced to positional constraints only.

Therefore, the presence of mobility constraints, together with positional constraints, is NOT sufficient to conclude that we are dealing with a non-holonomic system, because it might happen that, using the relationships which represent analytically the positional constraints (with an integration with respect to t), the mobility constraint might be substituted by relationships in finite terms between the coordinates, i.e. with positional constraints. Only if this integration is impossible we can say that we are confronted with non-holonomic constraints.

One could show along this line that the constraint of *rolling without slipping* on a plane, for a **sphere** leads to mobility constraints, which cannot be substituted with positional constraints (and therefore the system is non-holonomic) , while for a **cylinder** it leads to a mobility constraint, which can be substituted by a positional constraint (and therefore such a system is holonomic).

I will now include for completeness an explanation of this statement.

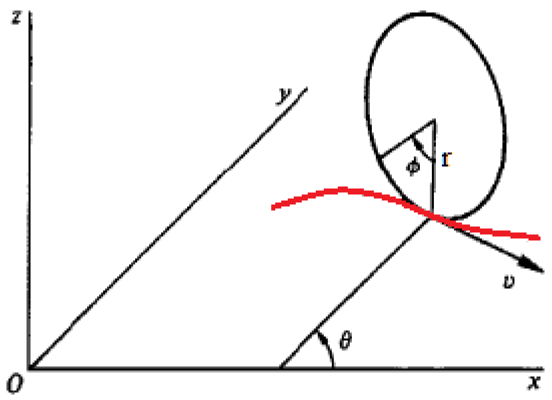
a) The case of the **cylinder rolling without slipping** on a plane is simpler: the cylinder, of radius R, can roll (without slipping) only on a trajectory which is a straight line perpendicular to his axis of rotation. An appropriate (and allowable) choice of coordinates can make the cylinder roll along the x-axis. Calling φ the angle of rotation of the cylinder with respect to the vertical passing through the contact point, we have that

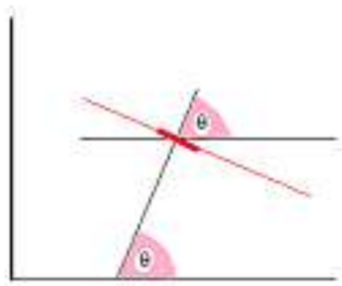
$$dx = R\,d\varphi$$

Or, transforming this relationship into one involving the velocities,

$\dot{x} = R\,\dot{\varphi}$, which can be integrated as Rφ –x = C, a positional constraint, which makes the system holonomic.

b) Following Goldstein (p. 15) let's now take a limiting case of the cylinder: **an extremely thin cylinder of *radius r*, which also rolls without slipping on the plane**. Now, the thinness of the cylinder allows it to spin around the vertical axis going through the point of contact, a case similar to that of a sphere. As is well known, the point of contact is at rest instant by instant, as, contrary to the appearance, *it is the center of rotation of the wheel.*





Still, the velocity of the center of the disk is linked to the speed of rotation of the angle φ, as:

$$v = r\,\dot{\varphi}$$

But now a new variable θ has entered the game, forcing us to introduce also a variable y, and we see that

$$\dot{x} = v\ \cos\left(\frac{\pi}{2} - \ \theta\right) = v \sin\theta$$

$$\dot{y} = -v\ \sin\left(\frac{\pi}{2} - \ \theta\right) = -v\cos\theta$$

(the minus sign is evident from the diagram).

Or, in differential form, using $v = r\ \dot{\varphi}$:

$$(1) \qquad dx - r\sin\theta\ d\varphi\ = 0$$

$$(2) \qquad dy + r\cos\theta\ d\varphi\ = 0$$

Can these two forms be transformed into exact differentials, or, if we prefer, can they be turned into forms like:

$$(3) \quad du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}\ dy + \frac{\partial u}{\partial \varphi}\ d\varphi + \frac{\partial u}{\partial \theta}d\theta \quad ?$$

Let's take eq.(1), and rewrite it as

$$X\ dx + \Phi\ d\varphi + Y\ dy + \Theta d\theta = 0$$

For it to be an exact differential, we should have, as is well known,

X =1 = $\frac{\partial u}{\partial x}$; $\Phi = -r\sin\theta = \frac{\partial u}{\partial \varphi}$; Y = $\frac{\partial u}{\partial y} = 0$; $\Theta = \frac{\partial u}{\partial \theta} = 0$

which entails

$\frac{\partial X}{\partial \varphi} = \frac{\partial \Phi}{\partial x}$ (and all similar pairs, with exchanged coordinates) both sides being equal to $\frac{\partial^2 u}{\partial x\,\partial \varphi}$ and the others likewise.

In the form given, some such equations are not satisfied because , for example $\frac{\partial \Phi}{\partial \theta} = -r\ cos\theta$, while $\frac{\partial \Theta}{\partial \varphi}$ =0, besides the example already given $\frac{\partial X}{\partial \varphi} \neq \frac{\partial \Phi}{\partial x}$.

Thus, in the form we have them, neither of the two equations is an exact differential (actually we have shown such a result only for one of the two equations, but the other has exactly the same problem).

However, we could think of finding an "integrating factor" z, i.e, in this case, a function of x,y, φ, θ which, multiplied by the whole equation (3), will turn at least one of the two first member functions into an exact differential. Let's examine Eq.(1)

Here it is impossible to find an integrating factor, however, because, for example, zΘ = 0 while zΦ cannot become 0 unless z is zero itself, which, as an integrating factor, makes it useless.

We conclude that such mobility constraint s cannot be substituted with positional constraints (and therefore the system is non-holonomic).

2.5 Lagrangian coordinates, and degree of freedom of a non holonomic system.

Let's now consider a non-holonomic system, and let the positional bilateral constraints be:

(10) $f_r(x_1, y_1, z_1,\ x_2, y_2, z_2,\ \ldots..x_N, y_N, z_N,\ t) = 0$ (r= 1,2,…μ)

And the bilateral mobility constraints

(11) $\sum_{i=1}^{N}\left(\alpha_{ji}dx_i + \beta_{ji}dy_i + \gamma_{ji}dz_i\right) + \tau_j\, dt = 0$ (j= 1,2,…ν)

Also in this case the (10) will allow us to introduce n (= 3N-μ) parameters q_1, q_2…q_n, as we have seen for fully holonomic systems. We will still call such parameters "lagrangian coordinates" of the non-holonomic system, with which we will express all 3N Cartesian coordinates of the system points, in the form:

(12) $$\begin{cases} x_i = x_i(q_1, q_2 \ldots q_n, t) \\ y_i = y_i(q_1, q_2 \ldots q_n, t) \\ z_i = z_i(q_1, q_2 \ldots q_n, t) \end{cases} \qquad (i = 1,2 \ldots . N)$$

However, the existence of the (11), tells us that the q_1, q_2, …q_n will no longer be independent. The (11), instead, tell us that between the increments dq_1, dq_2…dq_n , which the lagrangian coordinates receive while going from one configuration of the system to another, infinitely close the initial one, only $n - \nu$ can be arbitrarily assigned, while the remaining ones, ν, result from them, being determined by the (11).

Thus one can see that the number of lagrangian coordinates which can be made to vary arbitrarily to define a displacement of the system is not equal to n, but to $n - v$; and it is this number (instead of n) which is called "degree of freedom" of the system.

It remains to be said that, for what concerns the non-holonomic systems, the velocities of the system points will still be given by the (8) or (9), where one must remember that the lagrangian velocities $\dot{q}_i$ are not independent, but are now linked together by the relationships (11).

3. Virtual displacements.

> The obstinate reader which is reading the introduction to analytical mechanics on this essay or other more important books, is now about to meet one of the most extraordinary creations of human mind in the field of mechanics, comparable to the Newton Laws. Of course, one can deal with the subject with a certain light-heartedness, giving the definition of virtual displacements as one concept like many, which, in addition, is now all but out of fashion. We are more or less accustomed to it, and might not appreciate the original, revolutionary idea, of building the whole of mechanics on an abstract concept. Because virtual displacements, in principle, do not exist.

3.1 Given an arbitrary mechanical system, consisting of points P_1, P_2, …P_n, we call (elementary) "effective displacement of the system", corresponding to given applied forces, in the infinitesimal time interval (t, t+dt), the totality of all displacements dP_i, which the points P_i perform in the time interval dt, consistent with the constraints applied to the system.

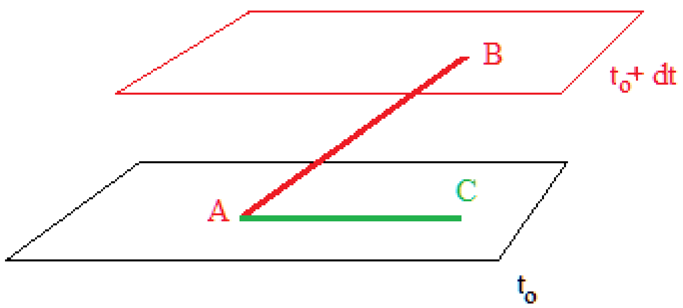
Together with effective or "actual" displacement of the system, one can also consider all other infinite displacements, which the system could perform under different applied forces: these are called "possible displacements". If the constraints are time-independent (or scleronomic), one does not consider any other types of displacements. However, if the constraints are time-dependent (rheonomic), *together with the "possible" displacements (which include the "effective" displacements),* we consider also another type of displacements, the so-called "virtual displacements".

More precisely, having fixed a time t_0, and having taken into consideration the infinitesimal time interval (t_0, t_0 + dt), we suppose that during this time interval the constraints do not vary, but remain "frozen" in the configuration they had at the time t_0. The elementary displacements which the system could perform in this case in the time interval dt, are called "virtual displacements corresponding to the instant t_0".

In other words, “the virtual displacements of the system corresponding to the instant t_0 are the elementary displacements consistent with the constraints “frozen” at the instant t_0”.

Simple examples of virtual displacement can be figured out by thinking of a rigid surface (such as a plane) uniformly moving in space, or a surface changing its shape (such as a sphere with a fixed centre, O, and a radius r(t) , varying in time according to a given law).

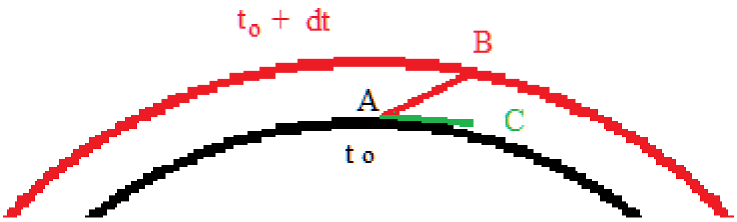
In the first case, the virtual displacements corresponding to the time t_0, are all the elementary displacements of the point on the plane “fixed” in the position it had at the instant t_0;



AB: an effective displacement compatible with the motion of the constraining plane in the time interval (t_0, t_0 + dt);

AC: a virtual displacement compatible with the constraining plane frozen at the instant t_0 in the interval dt.

in the second case the virtual displacements are all the elementary displacements on the sphere of center O, and radius r(t_0).



AB: an effective displacement compatible with the changing shape of the constraining sphere in the time interval (t_0, t_0 + dt). Here we have assigned an expansion, which is linear in time.

AC: a virtual displacement compatible with the shape of the constraining sphere frozen at the instant t_0 in the interval dt.

3.2 Whatever be the system, to indicate the elementary variation of any physical quantity which occurs in the course of a virtual displacement, the symbol **δ** is adopted (followed by the symbol proper to the quantity being considered), while to indicate the variation which occurs during an effective displacement, the symbol **d** is used (followed by the symbol proper to the quantity being considered). Thus dP_i and δP_i respectively will be the effective elementary displacement and the virtual displacement of point P_i; dq_i and δq_i, the elementary effective and the virtual variations of the lagrangian coordinate q_i. *Be it clear, however, that, independently of the symbols used, in both cases we are dealing with differentials.*

In the case of a holonomic system with n degrees of freedom, taking account of (8) we have for the effective and for the virtual elementary displacements dP_i and δP_i the remarkable expressions:

$$(13) \qquad dP_i = \sum_{r=1}^{n} \frac{\partial P_i}{\partial q_r} dq_r + \frac{\partial P_i}{\partial t}\, dt$$

And

$$(14) \qquad \delta P_i = \sum_{r=1}^{n} \frac{\partial P_i}{\partial q_r} \, \delta q_r$$

If, in addition, we are dealing with a rigid system, recalling that for the elementary displacement of one of his points, P_i, we have the expression

$$dP_i = dO + \vec{\omega} \, dt \wedge (P_i - O)$$

I remind the student that this equation descends from the formula:

$$\frac{d(P-O)}{dt} = \frac{dP}{dt} - \frac{dO}{dt} = x\frac{d\vec{\imath}}{dt} + y\frac{d\vec{\jmath}}{dt} + z\frac{d\vec{k}}{dt} = \vec{\omega} \wedge (x\vec{\imath} + y\vec{\jmath} + z\vec{k}) = \vec{\omega} \wedge (P - O)$$

In which, while the coordinates x,y,z of the P fixed in the rigid system are constant, the time derivatives of $\vec{\imath}, \vec{\jmath}, \vec{k}$ are given by Poisson formulae: $\frac{d\vec{\imath}}{dt} = \vec{\omega} \wedge \vec{\imath}\ etc.$ e the vectors $\vec{\imath}, \vec{\jmath}, \vec{k}$, are the unit vectors along the principal Cartesian axes.

Or, setting $\vec{\omega} = \frac{d\theta}{dt} \vec{k}$

$$dP_i = dO + d\theta\vec{k} \wedge (P_i - O)$$

We have, for the virtual displacement, the expression:

$$\delta P_i = dO + \delta\theta\vec{k} \wedge (P_i - O)$$

3.3 Ordinarily, virtual displacements are not possible displacements. It is obvious, however, that if we are dealing with time-independent constraints, the virtual displacements identify themselves with the possible displacements , and, on the other hand, the effective displacement is one particular virtual displacement.

Virtual displacements can be either invertible/ reversible or not invertible/irreversible, depending on whether they can happen in both directions or only in one of them. For example, for a body standing on a supporting surface, a displacement is invertible if it is tangent to the surface, because it can occur either toward the left or toward the right. On the other hand, a displacement which tends to detach the body from the surface is not invertible, because the opposite displacement (through the surface) is obviously impossible.

We conclude this section by noting that, if the system is non-holonomic, nothing needs to be added to what we just said: virtual displacements corresponding to the time t_o are those displacements which are compatible with all the constraints applied to the system, considered as fixed in the configuration they have at the instant t_o.

3.4. If $\vec{F}(t)$ is the force acting on point P_i, we define "**Virtual work**" which the force $\vec{F}(t)$ performs for the virtual displacement δP_i corresponding to the instant t_o, the internal product

$$\delta L = \vec{F}(t_o) \cdot \delta P_i$$

For "virtual work of a system of forces $\vec{F_i}(t)$ (I =1,2 … N) " acting respectively on the points P_i of a material system we define the sum of the virtual works of each of them:

$$\delta L = \sum_{i=1}^{N} \vec{F_i}(t_o) \cdot \delta P_i$$

4. Constraint equations. Frictionless constraints.

4.1 Given a constrained system (or system subjected to constraints), we call “active forces” the forces acting on the points of the system, independently of the existence of the constraints, while we call “reactive forces” or “constraint reactions” the forces which the constraints exert on the system. We will indicate the first ones with the symbol $\vec{F_\iota}$ and the second ones with $\vec{R_\iota}$.

While the active forces are known a priori, this is not the case with the constraint reactions. They are therefore supplementary unknowns. In general, mechanical problems do not require the calculation of such forces, and therefore one attempts to eliminate them from the calculations. In fact, the problem of their determination, if required, is usually tackled aside from the main problem, after the (statical or dynamical) configuration of the system under the action of the “active forces” has been determined.

To arrive at the elimination of the reactive, constraint forces, one start from observing that , if the constraints are *time independent* and *frictionless,* the constraint reactions , even if variable, perform zero work during the motion (as it happens, for example, in the case of a heavy mass point moving on a frictionless horizontal plane). This is not the case if the constraint varies with time: consider the same case of a heavy mass point on a plane, when the point is immobile on the plane, but the plane moves parallel to itself.

However, the notion of constraint is an abstract, schematic concept, which was introduced to facilitate the solution of mechanical problems: in general, a constraint is nothing but the simplified image of a real, more complex phenomenon. Thus, the motion we mentioned above, of a point mass on a rigid horizontal plane, is nothing but the schematization of the real motion of the body, which moves on the surface of an elastic body, slightly deforming it under the action of its own weight. When we substitute this “physical” surface with the abstract entity represented by a rigid plane, we, *de facto*, substitute to real and continuous reactive forces some abstract and discontinuous constraint reactions. For example, the discontinuity becomes evident when we detach the point from a rigid plane, the constraint being unilateral: here, the constraint reaction instantaneously stops acting on the point.

Such forces, discontinuous functions of the position coordinates, are not considered as such in Vectorial Mechanics (the first step in Mechanics) and therefore it is necessary to introduce a **new postulate** to include them in an acceptable mathematical description: such a postulate is accepting as the **definition of “frictionless constraints”** those for which “the virtual work of reactive forces is zero if all constraints are bilateral, and either positive or zero if at least one of the constraints is unilateral”, in one word those constraints for which

$$(1) \qquad \sum_{i=1}^{N} \overrightarrow{R_\iota} \cdot \delta P_i \geq 0$$

Here, the sign “ =” is applied to reversible displacements, and the sign “≥ " is applied to irreversible displacements. Simple examples like the one quoted above (the mass point moving on a plane and subjected either to a bilateral or to a unilateral constraint) verify the plausibility of the postulate/definition.

We are not going to discuss the principle of virtual works, which is generally well illustrated in texts of vectorial mechanics. We only make two remarks:

i) The definition of the principle:

The necessary ad sufficient condition for a material system with frictionless constraints to be in equilibrium is that the virtual work of the active forces, $\overrightarrow{F_\iota}$, acting on the system be ***zero*** *if all constraints are* *bilateral,* ***negative or zero*** *if at least one of the constraints is unilateral, namely*

(2) $$\sum_{i=1}^{N} \overrightarrow{F_i} \cdot \delta P_i \le 0\,.$$

2) A comment:

With the above postulate/definition of frictionless constraints, the principle of virtual works becomes the "theorem of virtual works" , because the relationship (2) can be demonstrated by means of (1).

This statement descends from the fact that for a system in equilibrium , for each point P_i it must be true that the resultant of all forces acting on it ($\vec{F}_i$) is equal and opposite to the resultant of all constraint reactions ($\vec{R}_i$), or:

$$\vec{F}_i + \vec{R}_i = 0$$

Therefore, the total virtual work of active forces,

$$\delta L = \sum_{i=1}^{N} \overrightarrow{F_i} \cdot \delta P_i = -\sum_{i=1}^{N} \overrightarrow{R_i} \cdot \delta P_i \le 0$$

Note that at this point, somewhat magically, the constraint forces are gone, and one can work with the active forces only.

The fundamental importance of the principle of virtual works is well known. Lagrange put it at the basis of all of Rational Mechanics and Equation (2) is known as "symbolic relationship of Statics". But, as we will shortly show, a slight modification makes of it the basis of both Statics and Dynamics.

Conclusion.

In any serious science, there is no way to escape the need for non-demonstrable principles. The Laws of Newton were examples of such principles, and their success (especially in solving some celestial mechanics problems) put them on a firm basis. However, they were justified by the results of their application. In the above section we have seen that the "*principle* of virtual works*", therefore not demonstrable*, can become a theorem (and therefore is demonstrated) at the cost of introducing another *principle*, disguised as a definition of "frictionless constraints".

The purpose of the principle (in any of the two forms given) is that of "simplifying the laws of Newton", by eliminating the constraint reactions from our equations. This can be done if :

a) the constraints are frictionless;

b) the constraints are time independent or, more in general, we deal with virtual displacements (which assume that the constraints are fixed in time).

The principle cannot be proved, but can be verified, and this is what most textbooks do. We leave the interested student to such developments and we go straight to an exciting amplification of the principle.

A simple verification for a point on a plane.

If the point is constrained to stay on the plane (bilateral constraint), the fact that the constraint is frictionless means that the constraint reaction is normal to the plane, and therefore its internal product with the virtual displacements is zero.

5. D'Alembert Principle and its consequences

It would be interesting to know the mental procedures of the great mathematicians of the past. In the case of *D'Alembert's Principle* I think that he must have noticed that finding the solution of all mechanical problems, regardless whether of Statics or Dynamics, amounts to solving certain finite or differential equations, showing that a certain left hand element is equal to an appropriate right hand element. If we are dealing with forces, it means that certain forces are equal to other forces. In other words, an equilibrium of forces comes into play. In Statics, for example, we look for equilibrium between active forces and constraint reaction (forces).

Equilibrium in general means that the system is not moving, but after all also Newton's Law , written as

$$\vec{F} = m\,\vec{a}$$

can be interpreted as the equilibrium of two elements, and its solution provides the motion of the system being considered. Let's see how, through D'Alembert's principle, together with the principle of virtual works, any problem of general dynamics can be treated as problem of statics.

1. The fundamental problem of dynamics consists of the determination of the motion of any given mechanical system, acted upon by a given system of forces. Such a definition is, in a sense, a major contribution of Newton, who set the problem in this clear way, even giving a semi-tautological definition of what a force is (*Def. IV: An impress'd force is an action exerted upon a body, in order to change its state, either of rest, or of moving uniformly forward in a right line).*

A mechanical system, for us, will be a system of N mass points P_i (i=1,2,…N), each endowed with mass m_i. Each point will be acted upon by various forces, whose resultant will be called $\vec{F_i}$. We will thus consider a system of N forces $\vec{F_i}$ applied to N points P_i respectively. *We consider such forces as* ***known*** *if they are expressed as functions of the configuration of the system, the simultaneous velocities of the N points and time.*

If the N points are free, as is the case of N bodies in celestial mechanics, and the system of forces acting on the system is known, the problem is immediately translated into equations. In fact, by applying to each of the N points the fundamental equation of dynamics, one has N equations:

$$(1) \qquad m_i \vec{a_i} = \vec{F_i} \qquad (i = 1,2, \dots N)$$

Here, $\vec{a_i}$ is the acceleration of the generic point P_i, with reference to a Galileian system (i.e. and inertial system, either at rest with respect to the fixed stars, or in rectilinear uniform motion).

If we indicate with x_i, y_i, z_i the coordinates of P_i with respect to this inertial system, and with X_i, Y_i, Z_i the Cartesian components of the force $\vec{F_i}$, then the N vector equations (1), projected on the coordinate axes O(x,y,z) produce the 3N scalar equations:

$$(1') \qquad m_i \ddot{x_i} = X_i, \quad m_i \ddot{y_i} = Y_i, \quad m_i \ddot{z_i} = Z_i \qquad (i = 1,2,\dots N),$$

where, in general, the X_i, Y_i, Z_i are functions of the 6N +1 arguments x_i, y_i, z_i, $\dot{x_i}, \dot{y_i}, \dot{z_i}, t$ (I = 1, 2…N).

We thus have a system $(1')$ of 3N ordinary differential equations of the second order, in the 3N unknown functions x_i, y_i, z_i, of the unique independent variable, t.

As is well known, in general, system $(1')$ cannot be solved in finite terms. However, well known theorems of existence make us sure that, under very broad conditions for the functions X_i, Y_i, Z_i, system $(1')$ admits a general integral depending on 6N arbitrary constant. We can thus say that for the N free points P_i, acted upon by the given system of forces, there are ∞^{6N} different possible motions, and each of them can be identified for example by giving the initial position of the system and the initial velocities of all single points.

The case we have just considered, practically, applies only to celestial mechanics. The forces being considered in celestial mechanics are of the "positional" type.

In the majority of cases, on the other hand, one is brought to consider constrained material systems. Now, as is well known, in a system of N point masses P_i, under any constraints, the actions of the latter on each point of the system, is equivalent to **a fictitious force,** which is called reaction or constraint force. It follows that the given system can be considered as a system of N free points, each of them being subjected to the simultaneous action of the resultant of the active forces directly applied to it, and at the same time to the resultant of the reactions which express the action of the constraint on it.

> I note *en passant* that this is a remarkable generalization of the concept of force, which we are accustomed to accept uncritically. If we consider the so called "active" forces, such as gravitation, and others, both known at the times of Newton and later, the reactions exerted by the constraints have no similarity to them. For example, reactions come into existence only when active forces start acting on constrained points. In contrast, an active force such as gravitation is "always there". However, we are helped by the fact that the constraint reactions can be represented by vectors, and, in the case of Statics, they contribute to the equilibrium of the point by opposing the active forces. Still, as we wrote above, "*…forces [are] known if they are expressed as functions of the configuration of the system, the simultaneous velocities of the N points and time*" and, as far as Classical Mechanics is concerned, this is not the case of constraint reactions.

It follows that also in the more general case of constrained systems equation (1) is valid, provided we interpret each $\overrightarrow{F_\iota}$ as the resultant of both active and reactive forces acting on P_i. Unfortunately, while we know the active forces and the ways the constraints act on the system, the corresponding reactive forces are unknown, and assume the character of auxiliary unknowns. Equations (1), therefore, for the motion of a constrained system, have the character of a provisional approach.

If one is only interested in the motion of the system, the need arises to eliminate from equation (1) the constraint reactions, in order to have, for the determination of the motion, differential equations uniquely depending of the direct data of the problem.

It can be shown that on the basis of (i) the classification of active and constraint forces, (ii) under the hypothesis of frictionless constraints, the elimination of the unknown reactions from the differential equations of the motion can be carried out in general and in a so-to-speak automatic way, on the basis of the Principle of Virtual Works, and of D'Alembert Principle, thus arriving at the classical Lagrange Equation of the Second form. These are dealt with at large in any textbook on Classical Mechanics.

On the other hand, *if one is interested in calculating also the constraint reactions besides the motion of the system,* the same principles of Virtual Work and D'Alembert Principle allowed Lagrange to formulate his equation in the first form, which, although much more cumbersome than the second form equations, completely solve the problem. It is to these equations that I shall dedicate, in a very introductory way, the remaining sections of this essay.

2. By classifying the forces acting on point P_i in active (F_i) and constraint forces (R_i) , equation (1) becomes:

$$(2) \qquad m_i\, \overrightarrow{a_i} = \overrightarrow{F_i} + \overrightarrow{R_i} \qquad (i = 1,2, \dots N)$$

which can be written as

$$(3) \qquad \overrightarrow{F_i} - m_i\, \overrightarrow{a_i} + \overrightarrow{R_i} = 0 \qquad (i = 1,2, \dots N)$$

We can interpret each $m_i\, \overrightarrow{a_i}$, which is a vector and has the dimensions of a force, as a fictitious force, which we shall call "inertial force concerning point P_i" . Equation (3) thus tells us that, in the course of the motion of a material system, acted upon by whatever forces and constraints, active forces, inertial forces and constraint reactions are in equilibrium at each instant.

If we do not want to promote the action of the constraints to the rank of forces, we can also say that (I.) *in the course of the motion of a material system, acted upon by whatever forces and constraints, active forces, and* *inertial forces are in equilibrium at each instant* ***thanks to the action of the constraints.***

This is quite a remarkable shift in focus. We can perform another easy trick by writing:

$$(4) \qquad \vec{F_i} = m_i\,\vec{a_i} + (\,\vec{F_i} - m_i\,\vec{a_i}\,) \qquad (i = 1,2,\dots N)$$

By so doing we split the active force $\vec{F_i}$ into two components-

The first component, $m_i\,\vec{a_i}$, represents the "force" which could impress on point P_i, if it were free, the same motion which it acquires under the combined action of the whole force $\vec{F_i}$ and of the constraints. We can call it the "effective component of the force $\vec{F_i}$".

The second component, $(\,\vec{F_i} - m_i\,\vec{a_i}\,)$, the sum of the active force and the force of inertia, represents the part of $\vec{F_i}$, which is, so to speak, lost because of the constraints. Therefore traditionally it received the name of "lost force".

We can thus enunciate D'Alembert's Principle:

(II) "in the course of the motion of a material system, acted upon by whatever forces and constraints, in each instant the lost forces are in equilibrium thanks to the action of the constraints".

*We have just changed statement (I) into (II) substituting the wording "active forces and inertial forces " with the wording "Lost forces". Still, t*he importance of D'Alembert's Principle is remarkable, because it reduces a problem of general dynamics to a problem of statics, the problem of the "equilibrium of the lost forces". But in Statics we have another principle ready to be applied to systems subjected to frictionless constraints, admitting virtual displacements, *which are all invertible*: the Principle of Virtual Works.

In the case of statics, the equilibrium conditions are all included in the equation:

$$\delta L = \sum_{i=1}^{N} \vec{F_i} \cdot \delta P_i = 0$$

And here, by applying D'Alembert's principle, we can characterize the motion of the system through the equation:

$$\delta L = \sum_{i=1}^{N} (\vec{F_i} - m_i \, \vec{a_i}) \cdot \delta P_i = 0$$

the so-called "**Symbolic equation** of Dynamics".

We remind the persistent reader that if the constraints were not invertible, we would have a $\leq$ sign , instead of the much more comfortable = sign (see Sec. 4.1).

The Symbolic Equation of Dynamics, which is valid for all dynamical systems with frictionless constraints and admitting only invertible virtual displacements, is a direct consequence of the Principle of Virtual Works, written as

$$\delta L = \sum_{i=1}^{N} \vec{R_i} \cdot \delta P_i = 0$$

which is valid both in equilibrium and in motion conditions, because of the relationship (3), $\vec{F_i} - m_i \, \vec{a_i} + \vec{R_i} = 0$. The latter, immediately produces the symbolic equation.

If the constraints were not invertible, one would have instead

$$\delta L = \sum_{i=1}^{N} \vec{R_i} \cdot \delta P_i \geq 0$$

And, consequently,

$$\delta L = \sum_{i=1}^{N} (\vec{F_i} - m_i \, \vec{a_i}) \cdot \delta P_i \leq 0$$

Which is called "the **symbolic relationship** of Dynamics".

6. The first form of Lagrange's Equation.

(A translated paraphrase of notes handed over by Prof. T. Zeuli, of the University of Turin, anno 1962-63)

The study of the motion of a holonomic system consisting of the points P_i (x_i, y_i, z_i) (i = 1,2,…N), is now the study of the "general equation of dynamics", namely:

(1) $$\sum_{i=1}^{N}(\vec{F}_i - m_i\vec{a}_i)\cdot\delta\vec{P}_i = 0$$

subjected to the m conditions:

(2) $$f_r(x_1, y_1, z_1, \ldots . x_N, y_N, z_N) = 0 \qquad (r = 1,2, \ldots m)$$

which represent the constraints to which the system is subjected. The virtual displacements $\overrightarrow{\delta P_i}$ are vectors whose components are δx_i , δy_i , δz_i. As a consequence of (2), the $\overrightarrow{\delta P_i}$ (= $\delta x_i, \delta y_i, \delta z_i$) are connected by the equations

(3) $$\sum_{i=1}^{N}\left(\frac{\partial f_r}{\partial x_i}\delta x_i + \frac{\partial f_r}{\partial y_i}\delta y_i + \frac{\partial f_r}{\partial z_i}\delta z_i\right) = 0 \qquad (r = 1,2, \ldots m)$$

Or, in a more compact form:

(3') $$\sum_{i=1}^{N}(grad_{P_i} f_r)\cdot\delta\vec{P}_i = 0 \qquad (r = 1,2, \ldots m).$$

Certainly I would not need writing these equations, which can be found on many analytical mechanics textbooks, were it not because of my personal experience based on the difficulties I experienced at the appropriate time, and the observations I made of students who had the same difficulties, which convinced me that equations (1) and (3) are frequently misinterpreted by a hurried reader.

As we said earlier, equation (1) represents the total virtual work done by the lost forces on all points the system consists of, and it is a single equation , which – as is -cannot be disentangled, because the displacements $\overrightarrow{\delta P_i}$ are linked together by equation (3) and $(3')$. The equation given in the form (1) can lead the reader into error because he could absentmindedly consider that this is not a single equation, but a set of N equations, more or less the same set which will immediately appear thanks to the stroke of genius of Lagrange. In fact the idea of Lagrange was that of introducing *m* undetermined coefficients λ_r, which are multiplied "in an orderly way "(to be exemplified below), by the $(3')$ and added to (1), thus obtaining

(4) $$\sum_{i=1}^{N}(\vec{F}_i - m_i\vec{a}_i + \sum_{r=1}^{m}\lambda_r\, grad_{P_i} f_r) \cdot \delta\vec{P}_i = 0$$

As the λ_r are undetermined, we can use (or hope to use) them to make equal to zero each parenthesis appearing in (4), which at this point still is a single equation. In other words, each parenthesis becomes an equation with λ_r as unknown and, as we shall see in the minimal example of section 7, what one does is to attempt to solve each of the N "parenthesis" independently (which generally is possible in textbook exercises). Once the solutions are found and all parentheses are equal to zero, whether the $\overrightarrow{\delta P_i}$ are linked together or not becomes irrelevant, because the effect of the constraints is already taken into account in the summation of the gradients.

In other words, we can transform the single, generally unmanageable equation (1) into the N vector equations

(5) $$\vec{F}_i - m_i\vec{a}_i + \sum_{r=1}^{m}\lambda_r\, grad_{P_i} f_r = 0$$

Or, equivalently, into the 3N scalar equations:

$$(5') \qquad m_i \frac{d^2 x_i}{dt^2} = X_i + \lambda_1 \frac{\partial f_1}{\partial x_i} + \lambda_2 \frac{\partial f_2}{\partial x_i} + \cdots \lambda_m \frac{\partial f_m}{\partial x_i}$$

$$m_i \frac{d^2 y_i}{dt^2} = Y_i + \lambda_1 \frac{\partial f_1}{\partial y_i} + \lambda_2 \frac{\partial f_2}{\partial y_i} + \cdots \lambda_m \frac{\partial f_m}{\partial y_i}$$
$$m_i \frac{d^2 z_i}{dt^2} = Z_i + \lambda_1 \frac{\partial f_1}{\partial z_i} + \lambda_2 \frac{\partial f_2}{\partial z_i} + \cdots \lambda_m \frac{\partial f_m}{\partial z_i}$$

which, together with the *m* equations expressing the constraints

$$(2) \qquad f_r(\mathrm{x}_1, \mathrm{y}_1, \mathrm{z}_1, \ldots . \mathrm{x}_\mathrm{N}, \mathrm{y}_\mathrm{N}, \mathrm{z}_\mathrm{N}) = 0 \qquad (\mathrm{r} = 1,2, \ldots \mathrm{m})$$

Form a system in the *3N + m* unknowns
$(\mathrm{x}_1, \mathrm{y}_1, \mathrm{z}_1, \ldots . \mathrm{x}_\mathrm{N}, \mathrm{y}_\mathrm{N}, \mathrm{z}_\mathrm{N};\ \lambda_1, \lambda_2 \ldots \lambda_m)$.

(Use has been made of the relationship $grad_{P_i} f_r = \vec{\imath}\,\frac{\partial f_r}{\partial x_i} + \vec{\jmath}\,\frac{\partial f_r}{\partial y_i} + \vec{k}\,\frac{\partial f_r}{\partial z_i}$)

Equations $(5')$ *and* (2) together are the **first form of Lagrange's equations.** And they can characterize all possible motions of the holonomic system under exam. Their beauty resides in the fact that all variables are treated in a completely symmetrical way. On the down side we must admit that we must now cope with a larger number of equations (3N+m), which most of the time are quite difficult, if not impossible to solve.

But there is more to the First Form Equation. First of all, they can be written directly and are formulated in the natural system of coordinates, i.e. Cartesian coordinates. No efforts of imagination are necessary to discover the degrees of freedom and any set of generalized coordinates. Moreover, they afford more information than the celebrated Equations of the second type, because the so called " Lagrange multipliers" λ_i have a remarkable mechanical meaning. If we compare the equations (5) with the equation of motion of the given system, i.e.

$$m_i\vec{a}_i = \overrightarrow{F}_i + \overrightarrow{R}_i$$

where $\overrightarrow{R}_i$ is the resultant of the constraint reactions acting on P_i , we obtain the expression:

$$\overrightarrow{R}_i = \lambda_1\, grad_{P_i} f_1 + \lambda_2\, grad_{P_i} f_2 + \cdots \lambda_m\, grad_{P_i} f_m.$$

This means that the reactions of the constraints will be determined once we will know the quantities $\lambda_1, \lambda_2 \ldots \lambda_m$, resulting from the solution of our system of 3N+m equations.

A simple example is the equation of the motion of a point mass on a surface whose equation is *f(x,y,z,t)=0* , which is allowed to vary with time. It is:

$$m\vec{a} = \overrightarrow{F} + \lambda\, grad\, f, \text{ where } f(x,y,z,t)=0.$$

No doubt, the results thus obtained afford a complete solution of the problem of the motion of the system, giving us, in addition, also the values of the reactions of the constraints. However, now we must deal with 3N+m equations, far more than the number of the degrees of freedom of the mechanical system, which is 3N-m = n.

If we just wish to determine the motion of the system, and not , in addition, the reactions of the constraints, it is possible to extract form D'Alembert's Principle a system of n equations, the Second form of the Lagrange equations. These are the major glory of Lagrange, but I will not deal with them in this page: they are not Cinderellas at all.

Yet, I would like to add at this point a remark made by Sommerfeld, which stresses one more advantage of our true Cinderellas.

"So far, we have assumed that the constraint are holonomic; we can easily convince ourselves that all the preceding can be carried over to the case of non-holonomic constraints, with only slight modifications. The only difference is that the components of the gradients must now be replaced by functions of the coordinates, which cannot be written in the form of partial derivatives of functions" (Mechanics, p.67).

Thus, problems which are not tractable by means of the Lagrange equations of the second form *alone*, can become tractable by using the Equations of the first form. In fact, to reduce the number of equations to be solved, various authors propose examples and exercises in a hybrid form for non-holonomic systems, in which the Newton equations (totaling 3N equations) appearing in the First Form equations are substituted by Lagrange's equations of the second form (totaling 3N-m equations), which take care of the m holonomic constraints, to which ν Lagrange multipliers are added to take care of the ν non holonomic constraints – thus going partially back to the first form. This is more or less directly suggested by the author, Prof. Zeuli, to which this essay is dedicated, in his section 2.5, given above.

7. The simplest problem which can be solved via the Lagrange Equation of the first form.

(In my opinion) the simplest problem is that of a point mass constrained to move on a circumference under the action of no forces.

One might think that the problem of a point mass constrained to move on a *straight line* under the action of no forces is even simpler. That's true. However, I think than such a problem is really too simple, and does not open the window on a broader and more attractive landscape. In fact, an accurate choice of coordinates immediately kills the lagrangian multiplier, the window is so to speak slammed closed, and we go straight back to Newton's equation, whose solution is a particle moving with constant speed.

Leaving this remark aside, to simplify as much as possible the formalism of the problem I have chosen , we set the mass of the mass point m = 1. The constraint is represented by the equation

$$x^2 + y^2 - R^2 = 0$$

We thus have our single first form equation written as:

$$\big((X - \ddot{x}) + 2\,\lambda x\big)\,\delta x + \big((Y - \ddot{y}) + 2\,\lambda y\big)\,\delta y = 0$$

The term $2\lambda x$ comes from the expression $\lambda \frac{\partial(x^2+y^2-R^2=0)}{\partial x}$ and the term $2\,\lambda y$ likewise. Note also that X = Y = 0, as no forces are acting on the mass point. The stroke of genius of Lagrange, as we said, is that the undetermined multiplier λ can be determined in such a way as to allow us to separate the two term multiplying δx and δy respectively in the equation. To do so we assume that the task can be accomplished, we separate the equations and we see whether we can solve them, thus setting both equal to zero, which achieves the separation.

One could object: "But why to use the multipliers? We can set to zero both equations separately also without including the λ-terms!" Correct, but by so doing we solve a different problem, the problem of a free particle (in this case in two dimensions), which, as we know, in the absence of forces moves at uniform speed on a straight line, as we would immediately find out by solving the two resulting equations. This is not our original problem. Our two particles must stay on a circle.

We have instead:

$$\ddot{x} = 2\,\lambda x$$

$$\ddot{y} = 2\,\lambda y$$

$$x^2 + y^2 - 1 = 0$$

Just for fun, if we had a particle constrained to move on a straight line, we could chose the constraint as y = 0 (particle moving on the x axis).

The single first form equation would have been:

$$\big((0 - \ddot{x}) + 2\,\lambda 0\big) \cdot \delta x + \big((0 - \ddot{y}) + 2\,\lambda 0\big) \cdot \delta y = 0$$

That is:

$$\ddot{x} = 0$$

$$\ddot{y} = 0$$

$$y = 0$$

Whose solution is x = $\dot{x}_0 t + x_0$,

y = 0

i.e. a uniform motion on a straight line, the x axis. Here λ disappears: indeed no constraint is necessary to keep a mass point on a straight line at constant speed in the absence of forces: it is Newton's First Law!

In our case, the two equations

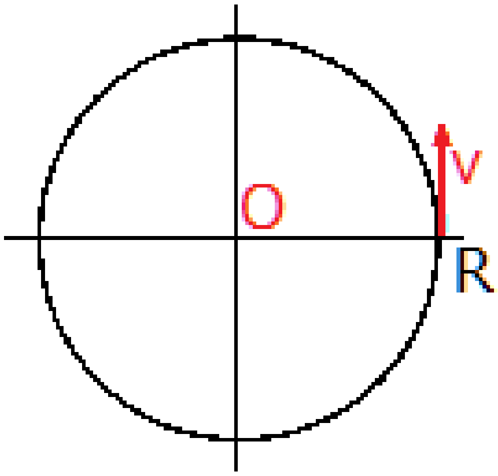
$$\ddot{x} = 2\,\lambda x$$

$$\ddot{y} = 2\,\lambda y$$

can be easily solved, giving:

$$x = A\cos\left(\sqrt{2\lambda}t\right) + B\sin\left(\sqrt{2\lambda}t\right)$$

$$y = C\cos\left(\sqrt{2\lambda}t\right) + D\sin\left(\sqrt{2\lambda}t\right)$$



For our initial conditions we are free to select any point on the circumference

$$x^2 + y^2 - R^2 = 0$$

such as (x_0 = R, y_0 = 0), and an arbitrary initial speed v along the y axis, such as ($\dot{x}_0 = 0,\ \dot{y}_0 = v$), in which case A= R, that is:

$$x_0 = A = R$$

And $y_0 = 0$, and therefore C = 0.

The components of the initial velocity are thus

$$\dot{x}_0 = \sqrt{2\lambda}B = 0, \text{ or } B=0$$

$$\dot{y}_0 = \sqrt{2\lambda}D = v$$

We thus have:

(I)

$$x = R\cos(\sqrt{2\lambda}t)$$

$$y = \frac{v}{\sqrt{2\lambda}}\sin(\sqrt{2\lambda}t)$$

From which, and from the equation of the constraint,

$$x^2 + y^2 = R^2 cos^2(\sqrt{2\lambda}t) + \left(\frac{v}{\sqrt{2\lambda}}\right)^2 sin^2(\sqrt{2\lambda}t) = R^2$$

The equation of the constraint is thus verified only if $\left(\frac{v}{\sqrt{2\lambda}}\right)^2 = R^2$

Which gives: $2\lambda = \frac{v^2}{R^2}$. We can also define an "angular velocity" $\omega = \sqrt{2\lambda} = \frac{v}{R}$.

Thus, the resulting motion is a rotation at constant speed, or with constant angular velocity v/R . The suggested generalized coordinate is $\sqrt{2\lambda}t$, which we can call, for example, θ.

What is interesting to me is that we start with two Cartesian coordinates, one of which is redundant, and we end up with a clear suggestion of a single non-Cartesian coordinate, which could be adopted as our single generalized q.

In other words, had we a much more complex problem, we might even hope that our "Cinderellas" could suggest to us the most appropriate generalized coordinates to treat the problem.

8. Some considerations

The pebbles I hope I have helped to extract from the shoes of some interested students are quite important and often overlooked. As we have seen, the principle of virtual works (PVW), normally used to solve a few equilibrium problems, and today quickly discarded, is a remarkable conceptual step. It was put by Lagrange at the basis of Statics and Dynamics, and can be said to be the door to Analytical mechanics. Yet, its importance is downplayed by most advanced mechanics tests. The principle is not even mentioned in the "Mechanics" textbook of the monumental work by Landau. In fact, the much more celebrated Lagrange Equation of the second form can be obtained without invoking the PVW at all, and Landau, more practically oriented, simply writes (§2) that "*the most general formulation of the law governing the motion of mechanical systems is the principle of least action, or Hamilton's principle*" etc. From this statement follows an Euler-like derivation of the equation of the second form, as the solution of a "variational problem", which leads to the so called "Euler –Lagrange " equations. It is amazing that the same equations can be reached in two totally different ways, but this point is seldom elaborated. However, it was not lost on Hamilton, and others.

Like Landau, Corben and Stehle never mention the principle of virtual works, but in equation 2.9 they introduce in a sneaky way the virtual displacements, without naming them: "we consider a small displacement of the particle defined by changes δx_i in the x_i, *with t held fixed* etc. As we have seen, the four word sentence "*with t held fixed"* conceals a world of reasoning.

The PVW is mentioned and its application to dynamics is clarified in the work by Goldstein. But here again no mention is made of the Lagrange Equation of the first form. On the other hand they are used when non-holonomic constraints are introduced. Our λ_i are called by Goldstein "undetermined multipliers", and no attempt is made to identify them in any way with the multipliers appearing in other problems of constrained maxima, i.e., as a typical example, the problem to maximize (in 2 dimension) a function f(x,y) subjected to the restriction that the maximum is constrained to belong to the curve g(x,y). This is how most student meet the Lagrange multipliers the first time in their life.

The question is whether Lagrange gave the same name to all undetermined coefficients he needed, no matter how different were the conditions of the applications, or he saw any connection between the different cases.

After a certain amount of brain searching I venture the hypothesis that the relationship between the problem of constrained *extrema* and the First form of Lagrange's equations arises from the equations of Statics for the case of a frictionless surface. In this case Statics produces equations which are quite similar to the First Form of Lagrange Equation. In fact, in the ideal case of a such a frictionless surface, the tangential component of the constraint reaction is zero and the reaction is fully normal to the surface, i.e. along the components of the gradient, albeit with an unknown value, which we can call λ. In this case, the equilibrium equation

$$\vec{F} + \vec{R} = 0$$

Becomes

$$(a) \qquad \vec{F} + \lambda\, grad\ f = 0,$$

where $f\ (x, y) =$ C is the equation of the frictionless surface.

But now, if $\vec{F}$ were derived from a potential U(x,y,z), then we would be indeed looking for the constrained extremum (whether maximum or minimum, depends on the convention we adopt that the Force is the gradient of U or the opposite of the gradient of U) of the potential, and we would be back to the original use of the Lagrange multipliers, we first met in analysis.

$$(b) \qquad \pm\, grad\ U + \lambda\, grad\ f = 0,$$

The equipotential surface U(x,y,z) = constant going through the equilibrium position is, in general, tangent to the surface f(x,y) =C going through that point. Indeed, in that point, the vectors *grad U* and *grad f* are parallel, as equation (b) tells us.

D'Alembert's Principle will put $\vec{F} - m\vec{a}$ in the place of $\overrightarrow{F}$, and the Lagrange multiplier λ will appear in a formally similar context:

$$\left(\overrightarrow{F} - m\vec{a}\right) + \lambda \, grad \; f = 0$$

but having assumed a different meaning.

This is as far as I can go without too much hand waving.

9. Conclusions.

"What is the use of Analytic Mechanics?" , asked a young lady student to Prof. Zeuli. And he answered:"Nothing is of use, as long as you don't know it".

I owe much to Prof. T. Zeuli, an excellent teacher who seemed to take a particular pleasure in explaining the most abstruse concepts of many fields of mathematics and classical mechanics in the most interesting possible way. I also owe much to him as a kind gentleman, and I hope that this essay will render a sincere, if inadequate homage to his memory, of a worthy successor of the many great mathematicians who illustrated my city of Turin.