

Power and beauty of the Lagrange equations

L. de la Peña, A. M. Cetto and A. Valdés-Hernández

*Instituto de Física, Universidad Nacional Autónoma de México,
Apartado Postal 20-364, Ciudad de México, Mexico.*

Received 14 August 2019; accepted 30 October 2019

The Lagrangian formulation of the equations of motion for point particles is usually presented in classical mechanics as the outcome of a series of insightful algebraic transformations or, in more advanced treatments, as the result of applying a variational principle. In this paper we stress two main reasons for considering the Lagrange equations as a fundamental description of the dynamics of classical particles. Firstly, their structure can be naturally disclosed from the existence of integrals of motion, in a way that, though elementary and easy to prove, seems to be less popular—or less frequently made explicit—than others in support of the Lagrange formulation. The second reason is that the Lagrange equations preserve their form in *any* coordinate system—even in moving ones, if required. Their covariant nature makes them particularly suited to deal with dynamical problems in curved spaces or involving (holonomic) constraints. We develop the above and related ideas in clear and simple terms, keeping them throughout at the level of intermediate courses in classical mechanics. This has the advantage of introducing some tools and concepts that are useful at this stage, while they may also serve as a bridge to more advanced courses.

Keywords: Lagrange equations; classical mechanics; covariance; integrals of motion.

PACS: 45.20.Jj; 45.20.D-; 01.40.-d

DOI: <https://doi.org/10.31349/RevMexFisE.17.47>

1. Introduction

Classical dynamics has a mighty range of tools to describe the motion of particles; among them we find in elementary treatments the forces, potentials, constants of motion, systems of reference and, in a prominent place, the equations of motion themselves. The first contact with the latter is undoubtedly through Newton's widely-known second-order differential equation that relates accelerations with forces. It is of interest to recall that the invention of differential calculus by Newton was driven precisely by the need of a tool to describe the motion of bodies. Usually it is only in more advanced courses that the student learns about other approaches serving the same purpose, in particular the Lagrangian formulation (established about a century after Newton's), and the Hamiltonian formulation (developed about half a century later). We thus have at least three different sets of equations, all of them fundamental, and equivalent in that they can be used alternatively to describe the behavior of the same classical system. However, what is considered fundamental depends on the intended descriptive level. From a historical perspective, Newton's Second Law is fundamental, the others having been derived from it with the introduction of more elaborate principles and demands. In particular, the Lagrange equations have a most remarkable and important property, namely that they preserve their form in any system of coordinates, even moving ones [1, 2]. This form-invariant property reflects the fundamental fact that they express a law of nature, which is naturally independent of our description. Also interesting is the fact—expressing another law of nature—that the Lagrange equations result from the single ele-

gant and powerful demand that a function called action attain a minimum value along the trajectory followed by the particle [1,3]. The Hamiltonian formulation, in its turn, represents a further elaboration providing a no less elegant set of differential equations. However, usually these derivations make use of the more advanced calculus of variations, with which we assume the reader-student not to be acquainted.

In this paper we arrive at the Lagrange equations by following a procedure that allows us to expose in clear and simple terms their connection with the integrals of motion—particularly, the energy—as well as their covariance, *i.e.*, their form-invariance with respect to a change of coordinate system. An additional benefit of this alternative route is that it allows us to disclose in a natural way the general structure of the Lagrange equations. Our intention is to develop the above ideas (and other related ones) in simple terms, working all the time at the level of intermediate courses in classical mechanics.

The paper is structured as follows. In Sec. 2 the Lagrange equations for a conservative system are derived, and the mechanical energy is shown to be the integral of motion associated with them. In Sec. 3 the invariance of the Lagrange equations under a transformation of the coordinate system is discussed, and the covariant form of Newton's Second Law is derived. Special attention is paid to the emergence of inertial forces as a result of the curvature of the coordinates (curvilinear coordinates). A couple of illustrative examples are presented in Sec. 3.4. Finally, Sec. 4 establishes the connection between the Lagrangian and the Hamiltonian functions, and briefly introduces Hamilton's equations.

2. The Lagrange equations and integrals of motion

2.1. Integrals of motion

For the description of the motion of a particle it is often convenient to use a system of (generalized) coordinates $\{q_i\}$ suited for the particular problem, instead of Cartesian coordinates $\{x_i\}$. In what follows we shall consider such a system $\{q_i\}$, which is arbitrary except for the condition that the coordinates can be expressed as regular, continuous and invertible functions of the Cartesian coordinates $\{x_i\}$, so that

$$q_i = q_i(\{x_k\}), \quad x_i = x_i(\{q_l\}). \quad (1)$$

This is an instance of a *passive* transformation [1], *i.e.*, one involving a change of coordinate system only, without affecting in any way the physical system. For simplicity we shall omit the appearance of the time as a parameter in the transformation; this restricts our discussion to *point transformations*, that is, to purely *geometric* transformations [1-6].

Let us consider a differentiable scalar function $\varphi(\mathbf{q}, \dot{\mathbf{q}})$, with $\dot{\mathbf{q}} = d\mathbf{q}/dt$ and \mathbf{q} the vector with N components $\{q_i\}$, N being the number of degrees of freedom. As a first step we analyze the time derivative of φ , which we obtain applying the chain rule,

$$\begin{aligned} \frac{d\varphi}{dt} &= \sum_i \left[\frac{\partial\varphi}{\partial q_i} \dot{q}_i + \frac{\partial\varphi}{\partial \dot{q}_i} \ddot{q}_i \right] \\ &= \sum_i \left[\frac{\partial\varphi}{\partial q_i} \dot{q}_i + \frac{d}{dt} \left(\frac{\partial\varphi}{\partial \dot{q}_i} \dot{q}_i \right) - \dot{q}_i \frac{d}{dt} \frac{\partial\varphi}{\partial \dot{q}_i} \right]. \end{aligned} \quad (2)$$

A rearrangement of terms leads to the expression

$$\frac{d}{dt} \left(\sum_i \dot{q}_i \frac{\partial\varphi}{\partial \dot{q}_i} - \varphi \right) = \sum_i \dot{q}_i \left(\frac{d}{dt} \frac{\partial\varphi}{\partial \dot{q}_i} - \frac{\partial\varphi}{\partial q_i} \right), \quad (3)$$

which can be written in the more compact form

$$\frac{d\gamma}{dt} = \sum_i \dot{q}_i G_i \quad (4)$$

with

$$\gamma(\mathbf{q}, \dot{\mathbf{q}}) \equiv \sum_i \dot{q}_i \frac{\partial\varphi}{\partial \dot{q}_i} - \varphi, \quad G_i(\mathbf{q}, \dot{\mathbf{q}}) \equiv \frac{d}{dt} \frac{\partial\varphi}{\partial \dot{q}_i} - \frac{\partial\varphi}{\partial q_i}. \quad (5)$$

Assume now that the function φ is selected such that the corresponding γ does not evolve in time, *i.e.*, $d\gamma/dt = 0$. We denote such constant function with ξ , and the corresponding φ with $\varphi^{(\xi)}$. Equations (4) and (5) give then

$$\frac{d\xi}{dt} = 0 = \dot{\mathbf{q}} \cdot \mathbf{G}^{(\xi)}, \quad (6)$$

where $\mathbf{G}^{(\xi)}$ is the vector with components

$$G_i^{(\xi)} = \frac{d}{dt} \frac{\partial\varphi^{(\xi)}}{\partial \dot{q}_i} - \frac{\partial\varphi^{(\xi)}}{\partial q_i}. \quad (7)$$

The above equations establish a correspondence between the constant of motion ξ and the vector $\mathbf{G}^{(\xi)}$, which, according to Eq. (6), is either zero or orthogonal to the velocity $\dot{\mathbf{q}}$ along the trajectory. Equation (7) exhibits in a natural way the characteristic structure of the Lagrange equations. Indeed, in the following section we will see that an analysis of the components $G_i^{(\xi)}$ takes us directly to the Lagrange equations, revealing their equivalence with Newton's Second Law, and to the identification of the mechanical energy as the associated integral of motion.

It is worthwhile to recall at this point that a closed mechanical system with N degrees of freedom may have at most $2N - 1$ nontrivial and functionally independent constants of motion, that is, functions of \mathbf{q} , $\dot{\mathbf{q}}$ and t whose value does not change with time. The complete set of these constants determines the trajectory of the particle in the $2N$ -dimensional phase space. Among the possible constants of motion there are some that have a particular importance, the so-called *integrals of motion*. These are (continuous, single-valued, differentiable) time-independent functions of the generalized coordinates and their corresponding momenta (or velocities), defined over the entire accessible phase space and having a constant value along the trajectory. Thus each one serves in principle to eliminate a degree of freedom from the description.

2.2. The Lagrange equations of motion

Let us consider a particle of mass m under the action of a time-independent, conservative force $\mathbf{f}(\mathbf{x}) = -\nabla V(\mathbf{x})$, with $V(\mathbf{x})$ the potential energy. In Cartesian coordinates the mechanical energy of this particle reads

$$\mathcal{E} = \frac{1}{2} m \dot{\mathbf{x}}^2 + V(\mathbf{x}). \quad (8)$$

According to our previous results, for this $\mathcal{E}(\mathbf{x}, \dot{\mathbf{x}})$ we can construct a $\varphi^{(\xi)}$ such that

$$\mathcal{E} = \sum_i \dot{x}_i \frac{\partial\varphi^{(\xi)}}{\partial \dot{x}_i} - \varphi^{(\xi)}. \quad (9)$$

The general solution $\varphi^{(\xi)}(\mathbf{x}, \dot{\mathbf{x}})$ of this differential equation reads

$$\varphi^{(\xi)}(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} m \sum_i \dot{x}_i^2 - V(\mathbf{x}) + \sum_i Q_i(\mathbf{x}) \dot{x}_i, \quad (10)$$

where the Q_i are arbitrary functions of \mathbf{x} only. From Eqs. (7) and (10) we obtain for the components $G_i^{(\xi)}$

$$G_i^{(\xi)} = m \ddot{x}_i - f_i + \sum_j \left(\frac{\partial Q_i}{\partial x_j} - \frac{\partial Q_j}{\partial x_i} \right) \dot{x}_j \quad (11)$$

$$= m \ddot{x}_i - f_i + \frac{dQ_i}{dt} - \frac{\partial}{\partial x_i} \sum_j Q_j \dot{x}_j. \quad (12)$$

On the other hand, by separating the terms in Eq. (10) that encode the dynamical information from the auxiliary Q_i , $\varphi^{(\mathcal{E})}$ can be rewritten in the form

$$\varphi^{(\mathcal{E})} = \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) + \sum_i Q_i(\mathbf{x}) \dot{x}_i, \quad (13)$$

with

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}m \sum_i \dot{x}_i^2 - V(\mathbf{x}). \quad (14)$$

We thus obtain using Eq. (7)

$$G_i^{(\mathcal{E})} = \frac{d}{dt} \frac{\partial \varphi^{(\mathcal{E})}}{\partial \dot{x}_i} - \frac{\partial \varphi^{(\mathcal{E})}}{\partial x_i} \quad (15)$$

$$= \left(\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \frac{\partial \mathcal{L}}{\partial x_i} \right) + \frac{dQ_i}{dt} - \frac{\partial}{\partial x_i} \sum_j Q_j \dot{x}_j. \quad (16)$$

Comparison of this with Eq. (12) gives

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \frac{\partial \mathcal{L}}{\partial x_i} = m\ddot{x}_i - f_i. \quad (17)$$

From here it follows that Newton's Second Law (for a conservative system)

$$m\ddot{x}_i = -\frac{\partial V}{\partial x_i} = f_i \quad (18)$$

is equivalent to the set of equations

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \frac{\partial \mathcal{L}}{\partial x_i} = 0, \quad (19)$$

which hold *irrespective* of the functions Q_i , and hence of the specific vector $\mathbf{G}^{(\mathcal{E})}$. The equivalence means that any one of the equations (18) and (19) implies the other one, via (17).

The function $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$ is the *Lagrangian* of the system, and Eqs. (19) are the *Lagrange equations of motion*, which, as just seen, are equivalent to Newton's Second Law. Note that \mathcal{L} is a scalar, and as such it remains invariant under a change of (spatial) coordinates.

Introducing Eq. (19) into (16) one obtains

$$G_i^{(\mathcal{E})} = \frac{dQ_i}{dt} - \frac{\partial}{\partial x_i} \sum_j Q_j \dot{x}_j, \quad (20)$$

and hence, from Eq. (4),

$$\frac{d\mathcal{E}}{dt} = \sum_i \dot{x}_i G_i^{(\mathcal{E})} = \sum_i \dot{x}_i \frac{dQ_i}{dt} - \sum_j \dot{x}_j \frac{dQ_j}{dt} = 0. \quad (21)$$

This verifies Eq. (6) (for $\xi = \mathcal{E}$), as expected.

If the vector \mathbf{Q} is selected as irrotational, the terms under the summation sign in Eq. (11) vanish and $\mathbf{G}^{(\mathcal{E})}$ becomes the null vector along the trajectory, $G_i^{(\mathcal{E})} = m\ddot{x}_i - f_i = 0$, which is the trivial solution of Eq. (6). Since in this case one can write $Q_i = \partial K / \partial x_i$ with $K(\mathbf{x})$ an arbitrary scalar function of \mathbf{x} , Eq. (13) becomes

$$\varphi_K^{(\mathcal{E})} = \mathcal{L} + \sum_i \frac{\partial K(\mathbf{x})}{\partial x_i} \dot{x}_i = \mathcal{L} + \frac{dK}{dt}. \quad (22)$$

Further (see Eq. (11)),

$$\frac{d}{dt} \frac{\partial \varphi_K^{(\mathcal{E})}}{\partial \dot{x}_i} - \frac{\partial \varphi_K^{(\mathcal{E})}}{\partial x_i} = \sum_j \left(\frac{\partial Q_i}{\partial x_j} - \frac{\partial Q_j}{\partial x_i} \right) \dot{x}_j = 0, \quad (23)$$

a result that discloses the well-known and very useful property of invariance of the Lagrange equations with respect to the addition of a total time derivative of a scalar function of \mathbf{x} (and t , we may add),

$$\mathcal{L} \rightarrow \mathcal{L}' = \mathcal{L} + \frac{dK}{dt} = \varphi_K^{(\mathcal{E})}, \quad (24)$$

with $K(\mathbf{x}, t)$ freely selected [1, 2, 4, 6]. \mathcal{L} and \mathcal{L}' are *equivalent Lagrangians*, meaning that the presence of K is totally inconsequential, since it does not affect the integrals of motion in any way. Thus there is an infinity of Lagrangians associated with each mechanical system. This gives us the freedom to select one that is appropriate for our purposes, as is sometimes done in the literature. From Eq. (19) we also verify that \mathcal{L} can be multiplied by an arbitrary constant. Thus any Newtonian system accepts an infinity of equivalent Lagrangians.

As mentioned above, in more advanced textbooks it is common to derive the Lagrange equations from a variational principle, known as Hamilton's principle, applied to the *action* of the system, defined as

$$S = \int_{t_1}^{t_2} \mathcal{L}[\mathbf{x}(t), \dot{\mathbf{x}}(t); t] dt. \quad (25)$$

The variational principle becomes then the single one that enters in place of the set of Newton's equations of motion; it states that the deviation δS of the action for any infinitesimal arbitrary deviation from the real trajectory followed by the particle, for initial and final times t_1 and t_2 fixed, is null. Otherwise stated, the action acquires an extremum (usually a minimum) value along the real trajectory from t_1 to t_2 . A relatively immediate application of this demand to Eq. (25) leads to the Lagrange equations of motion [1-6]. Since the variation of \mathcal{L}' given by equation (25) is the same as for \mathcal{L} for t_1 and t_2 fixed, the variational principle holds for all equivalent Lagrangians.

The fact that the whole set of equations of motion follows from a *single* principle of minimum action is of course most appealing, and it is not surprising that it gives rise to extremely valuable tools for the development of theoretical physics. However, some caution is needed to understand its significance. Newton's Second Law is strictly local, meaning that the particle responds *instantaneously* to the force applied just on it, at its position and at precisely that instant. The principle of least action, by contrast, refers to what happens *along the entire trajectory* from t_1 to t_2 ; it says that the particle adjusts its trajectory from the initial to the final time so that the accumulated action will result in a minimum. Of

course the principle applies only because it holds for each infinitesimal displacement, so the particle accumulates minimal actions one after the next along its trajectory. The particle does not know at time t_1 what it will be performing at a future time t_2 , in order to adjust itself to the demand of a global minimum action: the global action is minimal because each microscopic action is minimal. For an elementary derivation of the Lagrange equations from Hamilton's principle see, *e.g.*, [8].

3. Covariance of the equations of motion

In the foregoing section it became clear that for the single-particle, conservative problem treated here, the Lagrange equations of motion are equivalent to Newton's equations of motion, so one might wonder what is the advantage of using the former. One important difference is that the Lagrange equations, unlike Newton's, have the *same form* in *all* coordinate systems, as we shall now see. This property ultimately follows (in the present derivation) from the fact that the scalar product is invariant under geometric transformations. Specifically, Eqs. (6) and (7) hold for all coordinate systems, so that in terms of another set of coordinates $q'_i = q'_i(\{q_l\})$ for instance, we have $\varphi^{(\mathcal{E})} = \varphi^{(\mathcal{E})}[\mathbf{q}(\mathbf{q}')]$, and

$$\frac{d\mathcal{E}}{dt} = 0 = \sum_i \dot{q}_i G'_i{}^{(\mathcal{E})} = \sum_k \dot{q}'_k G'_k{}^{(\mathcal{E})}, \quad (26)$$

with

$$G'_k{}^{(\mathcal{E})} = \frac{d}{dt} \frac{\partial \varphi^{(\mathcal{E})}}{\partial \dot{q}'_k} - \frac{\partial \varphi^{(\mathcal{E})}}{\partial q'_k}. \quad (27)$$

Since the Cartesian coordinates $\{x_i\}$ and the generalized coordinates $\{q_i\}$ are related according to Eq. (1), we have

$$dx_i = \sum_k \frac{\partial x_i}{\partial q_k} dq_k, \quad dq_i = \sum_k \frac{\partial q_i}{\partial x_k} dx_k, \quad (28)$$

and therefore,

$$\begin{aligned} \sum_i \dot{x}_i G'_i{}^{(\mathcal{E})}(\mathbf{x}) &= \sum_i \left(\sum_k \frac{\partial x_i}{\partial q_k} \dot{q}_k \right) G'_i{}^{(\mathcal{E})}[\mathbf{x}(\mathbf{q})] \\ &= \sum_k \dot{q}_k \left(\sum_i \frac{\partial x_i}{\partial q_k} G'_i{}^{(\mathcal{E})}[\mathbf{x}(\mathbf{q})] \right). \end{aligned} \quad (29)$$

Thus, the invariance of the scalar product in (26) gives the law of transformation of $G'_i{}^{(\mathcal{E})}(\mathbf{x})$,

$$G'_k{}^{(\mathcal{E})}(\mathbf{q}) = \sum_i \frac{\partial x_i}{\partial q_k} G'_i{}^{(\mathcal{E})}(\mathbf{x}(\mathbf{q})). \quad (30)$$

From Eqs. (13), (16) and (19) we note that the choice $Q_i = 0$ implies $\varphi^{(\mathcal{E})} = \mathcal{L}$ and $G'_i{}^{(\mathcal{E})} = 0$, hence also $G'_k{}^{(\mathcal{E})} = 0$ according to Eq. (30). Therefore, from Eq. (27) we arrive at

$$\frac{d}{dt} \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})}{\partial q_i} = 0, \quad (31)$$

which confirms that the Lagrange equations look the same in *any* system of coordinates $\{q_i\}$. In more formal terms we say that the Lagrange equations are *covariant*, or that they are written in covariant form.

But then, what is the importance of covariance? Newton's law is a *physical* law, and as such it is totally independent of which system of coordinates is used to express it. The law expresses a fact of nature, whereas our coordinate system is something external to it, freely and arbitrarily selected. Expressing the physical law in such a way that it does not depend on the coordinate system, would of course be most appropriate and satisfactory.

To guarantee the independence of the law from the selected basis vectors, the matrix that transforms the vector components must be the *inverse* of the matrix that transforms the basis vectors. We say that such vector components transform *contravariantly with respect to the basis vectors*. This feature, which is crucial to guarantee that the laws of nature are the same in *any* coordinate system, applied to the second of equations (28), implies that the k -th component of the velocities transforms as

$$\dot{q}_k = \sum_i \frac{\partial q_k}{\partial x_i} \dot{x}_i, \quad (32)$$

whereas the k -th component of the vector \mathbf{G} transforms *inversely*, in terms of the matrix elements $\partial x_i / \partial q_k$ instead of $\partial q_k / \partial x_i$, as follows from Eq. (30). Below we come back to this point.

3.1. Some geometry

In preparation for the following section we develop here some mathematical tools that will be useful to distinguish between the geometric properties and the dynamical properties of a system, when arbitrary coordinates $\{q_i\}$ are employed.

Consider first the arc element ds . We may express this element in both the Cartesian coordinate system and the more general one $\{q_i\}$; we then have

$$(ds)^2 = \sum_{i,j} \eta_{ij} dx_i dx_j = \sum_{k,l} g_{kl} dq_k dq_l, \quad (33)$$

where η_{ij} and g_{kl} are the components of the *metric tensor* g in terms of the Cartesian and the generalized coordinates, respectively. Using the first equation in (28) we are led to

$$(ds)^2 = \sum_{i,j,k,l} \eta_{ij} \frac{\partial x_i}{\partial q_k} \frac{\partial x_j}{\partial q_l} dq_k dq_l = \sum_{k,l} g_{kl} dq_k dq_l; \quad (34)$$

thus $g_{kl} = \sum_{i,j} \eta_{ij} (\partial x_i / \partial q_k) (\partial x_j / \partial q_l)$. For a Euclidean (flat) space, $\eta_{ij} = \delta_{ij}$, and therefore

$$g_{kl} = g_{lk} = \sum_i \frac{\partial x_i}{\partial q_k} \frac{\partial x_i}{\partial q_l}. \quad (35)$$

Thus the components g_{kl} are constant for linear transformations; for curvilinear coordinates they are nontrivial functions of the generalized variables.

The components of the inverse of the metric tensor $g^{-1} = \bar{g}$ are such thatⁱ

$$\sum_l g_{il} \bar{g}_{ln} = \delta_{in}, \quad (36)$$

and they are given by

$$\bar{g}_{kl} = \bar{g}_{lk} = \sum_i \frac{\partial q_k}{\partial x_i} \frac{\partial q_l}{\partial x_i}. \quad (37)$$

A set of functions that will be of relevance, as we shall see below, are the Christoffel symbols (a kind of affine connections) Γ_{jk}^l , defined as [2, 9, 10]

$$\Gamma_{jk}^l = \Gamma_{kj}^l \equiv \sum_i \frac{\partial q_l}{\partial x_i} \frac{\partial^2 x_i}{\partial q_j \partial q_k}. \quad (38)$$

A comparison with Eq. (35) suggests that the Γ_{jk}^l will be useful to express the derivatives of the metric tensor and viceversa. Indeed, by taking the derivative of g_{kl} and using Eq. (38), one arrives at the expression

$$\frac{\partial g_{kl}}{\partial q_s} = \sum_i (g_{il} \Gamma_{ks}^i + g_{ik} \Gamma_{ls}^i). \quad (39)$$

To invert this relation and express the affine connection in terms of the metric tensor, one combines the above equation and its two permutations, thereby obtaining

$$\frac{\partial g_{kl}}{\partial q_s} + \frac{\partial g_{sl}}{\partial q_k} - \frac{\partial g_{ks}}{\partial q_l} = 2 \sum_i g_{il} \Gamma_{ks}^i. \quad (40)$$

Multiplying the above equation by \bar{g}_{ln} and sum over l leads to a formula for the Christoffel symbols in terms of the derivatives of the metric tensor

$$\Gamma_{ks}^n = \frac{1}{2} \sum_l \bar{g}_{ln} \left(\frac{\partial g_{kl}}{\partial q_s} + \frac{\partial g_{sl}}{\partial q_k} - \frac{\partial g_{ks}}{\partial q_l} \right). \quad (41)$$

Since the affine connections depend on the coordinates only and not on the velocities (see Eq. (38)), they have a geometric meaning and are directly related to the metric of the space in question.

With this we have the geometric tools required to describe Newtonian dynamics in an arbitrary coordinate system, which is the aim of the following section; but before closing this one it is convenient to add a couple of comments regarding the notation.

In the affine connections defined by Eq. (38) we introduced lower and upper indices: these represent different laws of transformation, so their meaning is important and goes beyond a mere convenience in the writing. The upper indices describe the components of *contravariant tensors*, whereas the lower indices refer to the *covariant* ones. The *contravariant* components of a vector are obtained by projecting the vector *onto the coordinate axes*; in their turn, the covariant components are obtained by projecting *onto the lines normal to the coordinate hyperplanes* [10, 12].

An example of a contravariant vector is given in Eq. (32) (another one is the law of transformation for f_i in Eq. (48) below), which transforms with the matrix elements $\partial q_k / \partial x_i$. An example of a covariant vector, transforming with $\partial x_i / \partial q_k$, appears in Eq. (30). In an orthogonal system of coordinates these vectors coincide (except for a possible interchange of coordinates), so there is no need to distinguish between them. However, in general both kinds of indices (of transformations) are necessary; this happens in particular in the general theory of relativity where space-time is curved due to the presence of matter and energy, and is also the case in Eq. (47) below. The scalar product corresponds then to a *contraction*, i.e., the sum over two equal indices, one from a factor that transforms covariantly (written as a subindex), and the other from a factor that transforms contravariantly (written as a superindex), as in $\sum_i q^i G_i = \mathbf{q} \cdot \mathbf{G}$, resulting in an invariant (a scalar). Notice that, in line with classical mechanics terminology, we have avoided the distinction between upper and lower indices when writing covariant and contravariant components. For a more detailed discussion see, e.g., Refs. [1, 2, 4].

3.2. Covariant form of Newton's second law

In this section we go back to Eq. (18) and apply a geometric transformation, to express it in terms of generalized coordinates. The process will lead us to the covariant form of Newton's Second Law, clearly revealing that the *inertial forces* (forces proportional to the mass of the particle) have a geometric origin associated with the curvature of the coordinates.

We first express \ddot{x}_i in terms of the generalized coordinates $\{q_i\}$ and their time derivatives as follows,

$$\begin{aligned} \ddot{x}_i(q) &= \frac{d}{dt} \sum_k \dot{q}_k \frac{\partial x_i}{\partial q_k} = \sum_k \ddot{q}_k \frac{\partial x_i}{\partial q_k} \\ &+ \sum_{j,k} \dot{q}_j \dot{q}_k \frac{\partial^2 x_i}{\partial q_j \partial q_k}. \end{aligned} \quad (42)$$

Since we are looking for an equation for \ddot{q}_k , it is convenient to multiply the above expression by $\partial q_l / \partial x_i$, sum over i , and use the equality

$$\sum_i \frac{\partial x_i}{\partial q_k} \frac{\partial q_l}{\partial x_i} = \frac{\partial q_l}{\partial q_k} = \delta_{kl}, \quad (43)$$

to obtain

$$\ddot{q}_l + \sum_{j,k} \dot{q}_j \dot{q}_k \Gamma_{jk}^l = \sum_i \ddot{x}_i(q) \frac{\partial q_l}{\partial x_i}. \quad (44)$$

The dynamics enters by using Eq. (18) in the form

$$\ddot{x}_i = \frac{1}{m} f_i = -\frac{1}{m} \sum_j \frac{\partial V(\mathbf{q})}{\partial q_j} \frac{\partial q_j}{\partial x_i}, \quad (45)$$

so that

$$\begin{aligned}\sum_i \ddot{x}_i \frac{\partial q_l}{\partial x_i} &= -\frac{1}{m} \sum_{i,j} \frac{\partial V(\mathbf{q})}{\partial q_j} \frac{\partial q_j}{\partial x_i} \frac{\partial q_l}{\partial x_i} \\ &= -\frac{1}{m} \sum_j \frac{\partial V(\mathbf{q})}{\partial q_j} \bar{g}_{jl}.\end{aligned}\quad (46)$$

Substitution into Eq. (44) leads to the desired dynamical equation,

$$m\ddot{q}_l = \sum_j -\frac{\partial V(\mathbf{q})}{\partial q_j} \bar{g}_{jl} - m \sum_{j,k} \dot{q}_j \dot{q}_k \Gamma_{jk}^l. \quad (47)$$

This is the explicit covariant form of Newton's Second Law (for the single-particle, conservative and unconstrained problem), written in a general system of coordinates (though without paying due attention to the position of the indices). The geometry of the specific coordinates being encoded in the (inverse of the) metric tensor \mathbf{g} and the Christoffel symbols (see, e.g., Christoffel symbols in *Mathematica*)ⁱⁱ.

Between Eq. (18) and Eq. (47) there is a world of difference. Yet when the coordinates $\{q_i\}$ are the appropriate ones for the visualization of the dynamical problem of interest, Eq. (47) gives us the most transparent description of the behavior of the system, as we shall see below.

3.3. Curvilinear coordinates and geometric forces

Equation (47) shows that there are in general two sources for the acceleration \ddot{q}_l , namely the *generalized force* \tilde{f}_l generated by the external force \mathbf{f} according to

$$\tilde{f}_l = \sum_i f_i \frac{\partial q_l}{\partial x_i} = - \sum_j \frac{\partial V}{\partial q_j} \bar{g}_{jl}, \quad (48)$$

and the additional *inertial force* \tilde{F}_l , always bilinear in the velocities \dot{q}_i and given by

$$\tilde{F}_l = -m \sum_{j,k} \dot{q}_j \dot{q}_k \Gamma_{jk}^l. \quad (49)$$

Now, if the transformation $x_i \rightarrow q_i$ is linear, all coefficients $\partial^2 x_i / \partial q_j \partial q_k$ in Eq. (38) vanish, resulting in a null inertial force. Thus, inertial forces appear only when the transformation $x_i \rightarrow q_i$ is *nonlinear*. A reference system that is nonlinearly related to an inertial one constitutes thus a *noninertial* reference system, in which inertial forces appear as a result of the curvature of the space coordinatesⁱⁱⁱ.

3.4. Two elementary problems as seen from the general theory

The powerful mathematical tools developed above are essential to deal with complex problems, say within general relativity, so the present exposition may be taken as groundwork for those higher matters. However, it seems appropriate to analyze here a couple of elementary problems, to gain familiarity with such tools.

Assume we are interested in the study of a point particle moving on a plane. We may describe its motion using Cartesian coordinates $x_1 = x$, $x_2 = y$, polar coordinates $q_1 = r$, $q_2 = \theta$, or some other set of coordinates; the choice may depend on the symmetry properties of the forces acting on the particle, or on constraints imposed on the motion.

Let us first consider the simple case of a central force, *i.e.*, a force that acts along the line joining the particle and the center of force (taken as the origin), and whose magnitude does not depend on the angle. In such case, because of the circular symmetry, the polar coordinates are the most appropriate ones. The transformation rules between the Cartesian and polar coordinates are^{iv}

$$x(r, \theta) = r \cos \theta, \quad y(r, \theta) = r \sin \theta; \quad (50)$$

$$r(x, y) = \sqrt{x^2 + y^2}, \quad \theta(x, y) = \tan^{-1}(y/x). \quad (51)$$

Upon comparison of the arc element

$$(ds)^2 = (dx)^2 + (dy)^2 = (dr)^2 + r^2(d\theta)^2 \quad (52)$$

with Eq. (34) one gets

$$g_{11} = 1, \quad g_{12} = g_{21} = 0, \quad g_{22} = r^2. \quad (53)$$

The components of the inverse metric tensor are obtained by inverting the matrix

$$g = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}, \quad (54)$$

which gives

$$\bar{g}_{11} = 1, \quad \bar{g}_{12} = \bar{g}_{21} = 0, \quad \bar{g}_{22} = r^{-2}. \quad (55)$$

The metric tensor and its inverse can also be obtained of course by resorting to Eqs. (35) and (50), and to Eqs. (37) and (51), respectively.

Using now Eq. (48) we obtain for the components of the generalized force

$$\tilde{f}_l = -\bar{g}_{l1} \frac{\partial V}{\partial q_1} - \bar{g}_{l2} \frac{\partial V}{\partial q_2}, \quad (56)$$

whence

$$\tilde{f}_r \equiv \tilde{f}_1 = -\bar{g}_{11} \frac{\partial V}{\partial q_1} = -\frac{\partial V}{\partial r}, \quad (57)$$

$$\tilde{f}_\theta \equiv \tilde{f}_2 = -\bar{g}_{22} \frac{\partial V}{\partial q_2} = -\frac{1}{r^2} \frac{\partial V}{\partial \theta}. \quad (58)$$

The affine connections follow from Eq. (41), the only elements different from zero being

$$\Gamma_{22}^1 = -\frac{1}{2} \bar{g}_{11} \frac{\partial g_{22}}{\partial q_1} = -r, \quad (59)$$

$$\Gamma_{12}^2 = \Gamma_{21}^2 = \frac{1}{2} \bar{g}_{22} \frac{\partial g_{22}}{\partial q_1} = \frac{1}{r}. \quad (60)$$

From Eq. (49) applied to this case,

$$\tilde{F}_l = -m (2\Gamma_{12}^l \dot{q}_1 \dot{q}_2 + \Gamma_{22}^l \dot{q}_2^2), \quad (61)$$

we obtain thus

$$\tilde{F}_r = -m\Gamma_{22}^1 \dot{q}_2^2 = mr\dot{\theta}^2, \quad (62)$$

$$\tilde{F}_\theta = -2m\Gamma_{12}^2 \dot{q}_1 \dot{q}_2 = -2mr^{-1} \dot{r} \dot{\theta}. \quad (63)$$

Substitution of these results in Eq. (47) gives the equations of motion in polar coordinates:

$$m\ddot{r} = -\frac{\partial V}{\partial r} + mr\dot{\theta}^2, \quad (64)$$

$$m\ddot{\theta} = -\frac{1}{r^2} \frac{\partial V}{\partial \theta} - \frac{2m}{r} \dot{r} \dot{\theta}. \quad (65)$$

Notice that these equations contain each an inertial force term. Let us recast them both in a more familiar form. Firstly, multiplication of (65) by r^2 gives

$$mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta} = \frac{d}{dt}(mr^2\dot{\theta}) = \frac{dL}{dt} = -\frac{\partial V}{\partial \theta}, \quad (66)$$

where $mr^2\dot{\theta}$ has been identified as the angular momentum L . This is the dynamical equation for the angular momentum, driven just by the external torque. In particular, it shows that for radial forces ($V = V(r)$) the angular momentum is conserved (it is an integral of motion),

$$L = mr^2\dot{\theta} = \text{constant}. \quad (67)$$

On the other hand, in terms of L (whether conserved or not), Eq. (64) takes the form

$$m\ddot{r} = -\frac{\partial V}{\partial r} + \frac{L^2}{mr^3}, \quad (68)$$

which contains the centrifugal force $L^2/mr^2 = mr\dot{\theta}^2$, an inertial force that propels the particle radially away from the origin.

Let us now take as a second example the case of a particle that, in addition to being subject to a conservative force $\mathbf{f}(x, y) = -\nabla V(x, y)$, is constrained to move along a prescribed trajectory on the plane; we may think of a frictionless rail that enforces this dynamics, by exerting at all times a force on the particle perpendicular to its trajectory. The geometrical shape of the rail can be expressed as a functional relation that must hold between the two position coordinates (with a constant)

$$g(x, y) = a, \quad (69)$$

which reduces the number of degrees of freedom by one. We may conveniently choose the generalized coordinates as $q_1 = q_1(x, y)$, and $q_2 = a$ (q_2 will be fixed to a constant value through the constriction (69)), which can be inverted using (69) to solve for the Cartesian coordinates,

$$x = x(q_1, q_2), \quad y = y(q_1, q_2). \quad (70)$$

The corresponding Lagrange equation for the unconstrained generalized coordinate reads

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_1} - \frac{\partial T}{\partial q_1} = -\frac{\partial V}{\partial q_1}, \quad (71)$$

where we wrote $\mathcal{L} = T - V$, with T the kinetic energy (see Eqs. (74) and (77) below). Equation (71) can be solved *without knowing the forces of constraint*. This is a great advantage, for the forces of constraint depend on the motion of the particle, and therefore cannot be determined in general until the motion is known. In some instances it is important to know such forces; these can then be calculated from the Lagrange equations for the constrained coordinate (in our case, the coordinate q_2):

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_2} - \frac{\partial T}{\partial q_2} = -\frac{\partial V}{\partial q_2}. \quad (72)$$

Indeed, by substituting here the solution for $q_1(t)$ obtained from solving Eq. (71), we find the constraining force ($-\partial V/\partial q_2$). This shows that the Lagrange formalism is particularly suited to deal with systems subject to holonomic constraints^v.

4. The relation between momenta and velocities. The Hamiltonian

In concluding, it seems convenient to briefly recall the connection between the two fundamental functions that serve to define, each one in their own capacity, the dynamics of a mechanical system, namely the Lagrangian and the Hamiltonian. For this purpose we first introduce the (generalized) momentum p_i , *defined* via the Lagrangian as

$$p_i = \frac{\partial \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{q}_i}. \quad (73)$$

The variables q_i, p_i are thus said to be canonically conjugate.

In order to calculate p_i , we transform $\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})$ to $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$ by resorting to Eqs. (14), (28) and (35),

$$\begin{aligned} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) &= \frac{1}{2}m \sum_{i,j,k} \frac{\partial x_i}{\partial q_k} \frac{\partial x_i}{\partial q_j} \dot{q}_k \dot{q}_j - V(\mathbf{x}(\mathbf{q})) \\ &= \frac{1}{2}m \sum_{j,k} g_{kj} \dot{q}_k \dot{q}_j - V(\mathbf{x}(\mathbf{q})). \end{aligned} \quad (74)$$

Equation (73) gives therefore a linear relation between the momenta and the velocities

$$p_i = m \sum_k g_{ik} \dot{q}_k, \quad (75)$$

which can be inverted to obtain

$$\dot{q}_i = \frac{1}{m} \sum_k \bar{g}_{ik} p_k. \quad (76)$$

Only for rectilinear coordinates (when all the g_{ik} are constant) this relation is constant along the trajectory. Further, Eq. (75) allows us to write the kinetic energy as

$$T = \frac{1}{2m} \sum_{i,j} \bar{g}_{ij} p_i p_j = \frac{1}{2} m \sum_{k,l} g_{lk} \dot{q}_l \dot{q}_k. \quad (77)$$

Note that the kinetic energy is always a bilinear function of either the velocities or the momenta.

Let us now combine Eqs. (9) and (14), to get

$$\mathcal{E}(\mathbf{q}, \dot{\mathbf{q}}) = \sum_i \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L} = \sum_i \dot{q}_i p_i - \mathcal{L} \mathcal{H}(\mathbf{q}, \mathbf{p}). \quad (78)$$

The right-hand side of this equation corresponds to the *definition* of the Hamiltonian $\mathcal{H}(\mathbf{q}, \mathbf{p})$, which, as is clear from the first equality, for a closed, conservative system coincides with the mechanical energy^{vi}.

By using (76) we recover a most important relationship between the Lagrangian and the Hamiltonian,

$$\mathcal{H} = \frac{1}{m} \sum_{i,k} \bar{g}_{ik} p_k p_i - \mathcal{L}, \quad (79)$$

which gives

$$\mathcal{H}(\mathbf{q}, \mathbf{p}) = \frac{1}{2m} \sum_{i,j} \bar{g}_{ij} p_i p_j + V(\mathbf{q}). \quad (80)$$

It is a simple, illustrative exercise to use this result for the derivation of the set of equations of motion [2-4]

$$\frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i, \quad \frac{\partial \mathcal{H}}{\partial q_i} = -\dot{p}_i. \quad (81)$$

These are the notorious Hamilton equations of motion that completely describe the evolution of any system defined by the Hamiltonian $\mathcal{H}(\mathbf{q}, \mathbf{p}, t)$. We see that there are $2N$ differential Hamilton equations of first order for a system with N degrees of freedom, while according to Eq. (19) the dynamics of the same system is described by N Lagrange differential equations of second order. For this reason, in some cases the description of the dynamics of a system in terms of the Hamilton equations of motion is simpler than in Lagrangian terms.

Acknowledgments

The authors gratefully acknowledge financial support from DGAPA-UNAM through project PAPIIT IA101918.

-
- i.* The transition from the x -system to the q -system of coordinates can be viewed as a mere change of variables, where the coefficients $J_{kl} = (\partial x_l / \partial q_k)$ are the elements of the Jacobian of the transformation (for the transition from x_k to q_k) and, similarly, the elements of the inverse matrix $J^{-1} = \bar{J}$ are $\bar{J}_{lk} = (\partial q_k / \partial x_l)$ (for the inverse transition from q_k to x_k).
 - ii.* Though this form is not common in classical mechanics, it can be found from time to time [2, 11].
 - iii.* The observation that forces can be associated with curvatures of the coordinate system—so fundamental for the general theory of relativity—and that a covariant description is needed, is older than one would think; it was mentioned already by the brilliant German mathematician Bernhard Riemann (1826-1866), followed by another brilliant mathematician, the British William K. Clifford (1845-1879), as discussed in Ref. [13], Chapter 5.
 - iv.* In Eq. (51) it should be understood that $\tan^{-1}(y/x)$ is the two-argument inverse tangent, which takes into account the signs of both x and y .
 - v.* Here we have dealt with *holonomic* constraints, which depend on the position coordinates. In the more general case, constraints may include restrictions on the velocity. If these constraints can be integrated so as to lead to an expression of the form (69), they are still holonomic. There are cases, however, in which the equations of constraint cannot be integrated (take for instance, a circular wire rolling on a plane without slipping); we then speak of *nonholonomic* constraints. For a more detailed discussion and additional examples see Ref. [6], Chapter 9.
 - vi.* The function $\mathcal{E}(\mathbf{q}, \dot{\mathbf{q}})$, whose value is just that of the Hamiltonian $\mathcal{H}(\mathbf{q}, \mathbf{p})$ but expressed in terms of velocities instead of momenta, is called the *energy function*; see Ref. [1], Sec. 2.7.

1. H. Goldstein, Ch. Poole and J. Saiko, *Classical Mechanics* 3rd ed. (Addison-Wesley, MA, 1980).
2. H. C. Corben and P. Stehle, *Classical Mechanics* (John Wiley, NY, 1950), Chap. V.
3. N. Rasband, *Dynamics* (John Wiley, NY, 1983), Chap. III.
4. J. V. José and E. J. Saletan, *Classical Dynamics. A contemporary approach* (Cambridge University Press, Cambridge, 1998).
5. J. R. Taylor, *Classical Mechanics* (University Science Books, CA, 2005).
6. K. R. Symon, *Mechanics* 2nd. ed. (Addison-Wesley, MA, 1960).
7. R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman Lectures on Physics*, **Vol. II** (Addison-Wesley, Reading, MA, 1964).
8. J. Hanc, E. F. Taylor and S. Tuleja, *Am. J. Phys.* **72** (2004) 510.
9. See, e.g., R. Price, *Am. J. Phys.* **50** (1982) 300.
10. S. Weinberg, *Gravitation and Cosmology. Principles and Applications of the General Theory of Relativity* (John Wiley, NY, 1972).
11. T. M. Kalotas and B. G. Wybourne, *J. Phys. A* **15** (1982) 2077.
12. See, e.g., R. Courant, *Differential and Integral Calculus*, **Vol. II**, (John Wiley, NY, 1962), Chapt. III.
13. M. Jammer, *Concepts of Space: The History of Theories of Space in Physics* (Dover, N.Y., 1954).