



**HAL**  
open science

# From Newton to Hamilton, bypassing Lagrange and Legendre.

Bahram Houchmandzadeh

► **To cite this version:**

Bahram Houchmandzadeh. From Newton to Hamilton, bypassing Lagrange and Legendre.. 2021.  
hal-03259814v1

**HAL Id: hal-03259814**

**<https://hal.science/hal-03259814v1>**

Preprint submitted on 14 Jun 2021 (v1), last revised 18 Jun 2021 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# From Newton to Hamilton, bypassing Lagrange and Legendre.

Bahram Houchmandzadeh

*Univ. Grenoble Alpes, CNRS, LIPhy, F-38000 Grenoble, France*

The Hamilton equations of movement have a deep geometrical meaning and are of prime importance in many areas of physics. Usually, the Hamiltonian approach is taught after the introduction of the Lagrangian formalism through a Legendre Transform. As a complementary approach, I believe that it can be helpful for students to learn the geometrical meaning of the Hamiltonian formulation by deriving it directly from the more familiar Newton dynamics, without using the variational principles and the Legendre transform.

## I. INTRODUCTION.

The Hamiltonian formulation of mechanics, a “*geometrization*” of Lagrangian mechanics, provides the natural framework in many areas of physics such as chaos, statistical and quantum mechanics.

Historically, analytical mechanics emerged at the end of the eighteenth century from the investigations of Euler, d’Alembert and Lagrange[1, 2], 100 years after the publication of *Principia* by Newton and the Johann Bernoulli’s challenge[3]. The field took its more advanced form following the works of Hamilton and Jacobi in the 1830-40’s. A detailed historical development of mechanics can be found in the Dugas’ book[4].

The teaching of Hamiltonian dynamics, as found in nearly all advanced textbooks[5–9], follows closely the historical developments of the field: After introducing the variational formulation of mechanics and the Euler-Lagrange equations obtained through the application of d’Alembert’s principle, the Hamiltonian mechanics is obtained from a Legendre transform of the Lagrangian. The canonical transformations of Hamilton equations are then introduced using the fact that the Lagrangian is defined up to a total differential of a function, which leads to the definition of various generating functions and to the Hamilton-Jacobi equation[10].

This is a very elegant and powerful approach to the field of analytical mechanics : the applications of the variational formulation extends far beyond the classical mechanics and serves as a unifying principle in many areas of physics.

On the other hand, this approach to the Hamiltonian mechanics poses some difficulties to many students. The aim of Lagrange in his founding book was to take the geometry out of mechanics and bring this latter field solely in the area of Analysis: the most quoted sentence of his book is “no figures will be found in this work”[2].

Generalized momenta for example appear as an abstract concept in the standard derivation, associated with the Legendre transform[11]. This fundamental mathematical transformation is usually not taught with all the attention it deserves and many associate it with minimization problems as they occur in thermodynamics or large deviation theory ; its link to differential equations seems rather abstract to many students.

In contrast, for many students geometry is the most

intuitive way of approaching a given field. Indeed, the aim of Hamilton, who began his scientific work in geometric optics, was precisely to put back geometry into mechanics. I propose to introduce Hamiltonian mechanics to students following this geometrical road as a complementary approach to the standard one. The approach I am proposing by no means is intended to replace the standard one, which in my opinion, is the most fundamental way of understanding physics. However, visiting the same subject from different angles can be enriching and shed new lights on our understanding of a given field.

The approach I propose in this article, restricted to conservative classical mechanics, is to stress the fundamental property of Hamiltonian mechanics, *i.e.* volume conservation in phase space, as the main motivation to write the Newton’s equations of motion in this frame. The generalized momenta are obtained by highlighting the role of the metric tensor in defining the underlying geometry of space. The canonical transformation are also introduced as a mean to preserve the volume conservation property of Hamilton’s equation. References to variational principle and Legendre transforms are avoided.

The approach I propose avoids also references to the field of symplectic geometry that is covered in advanced books such as Arnold’s[12]. Symplectic geometry uses the language of differential forms[13, 14] that was formalized by Elie Cartan at the beginning of 20th century[15]. Even though I believe forms are the most natural ways of teaching multi-variable calculus [16], their use has not impregnated the physics community in the same way as for example tensor calculus. Therefore, in this paper, I only use the basic concepts of linear algebra (determinant and trace) without referring explicitly to forms.

This paper is organized as follows. The next section is devoted to a simple introduction of Hamiltonian mechanics directly from Newton’s equations and the relation between momenta and the metric tensor. In section III, I recall the basic concepts of linear algebra, namely “determinant” and “trace” and show how the Liouville’s theorem is a direct consequence of these concepts and the structure of Hamilton’s equation. In section IV, the relation between momenta and the metric tensor is demonstrated in a more rigorous way. Section V is devoted to canonical transformations, using Liouville’s theorem as the guiding line. A final section is devoted to concluding remarks. Two appendices are also devoted to more

technical aspect of the computation.

## II. MOTIVATIONS.

Consider a second order differential equation that describes, for example, the movement of a particle whose coordinate is  $x$  :

$$\ddot{x} + f(\dot{x}, x) = 0. \quad (1)$$

For simplicity throughout this article, we suppose that the independent variable  $t$  does not appear explicitly in the equation. We can always transform this equation into a system of *two* first order equations

$$\dot{x} = y \quad (2)$$

$$\dot{y} = -f(y, x) \quad (3)$$

where  $(x(t), y(t))$  are coordinates of the point in a two dimensional space. We have a much better understanding (visualization) of first order equations than second order one. However, the geometry of equations (2,3) is not yet limpid.

Suppose now that we are able to find a mapping between the space  $(x, y)$  and  $(x, p)$ , *i.e.* we are able to find a function  $p = \phi(x, y)$  such that equations (2,3) are transformed into

$$\dot{x} = \frac{\partial H}{\partial p} \quad (4)$$

$$\dot{p} = -\frac{\partial H}{\partial x} \quad (5)$$

where the function  $H = H(x, p)$  results from our mapping. For a multidimensional system involving  $n$  variables  $x^i$ , equations (4,5) have to be obtained for each variable separately, *i.e.*  $\dot{x}^i = \frac{\partial H}{\partial p_i}$ ,  $\dot{p}_i = -\frac{\partial H}{\partial x^i}$ .

In principle, we can directly write equations (2,3) in this form, *i.e.*  $\dot{x} = \frac{\partial F}{\partial y}$ ,  $\dot{y} = -\frac{\partial F}{\partial x}$ ; however, finding directly the function  $F(\cdot)$  requires solving a first order partial differential equation that in turn necessitate solving equations (2,3), and thus the direct approach is not very useful. On the other hand, as we will see below, finding the mapping  $p = \phi(x, y)$  can be achieved *algebraically* for problems of classical dynamics if we use the geometric properties of our space.

Equation (4,5) have a deep geometrical meaning (figure 1). First, we see that trajectories in the  $(x, p)$  space correspond to level curves of the function  $H$  because

$$\frac{dH}{dt} = \frac{\partial H}{\partial x} \frac{dx}{dt} + \frac{\partial H}{\partial p} \frac{dp}{dt} = 0. \quad (6)$$

by construction. Moreover, we can envisage the direction of the movement  $\vec{u} = (\dot{x}, \dot{p})$  as “*orthogonal*” to the “*gradient*”  $\nabla H$  of the function  $H(x, p)$ . Speaking of gradient and orthogonality is not correct because in the  $(x, p)$  space we have not defined (and don’t need to define) a scalar product. However, this picture carries the analogy

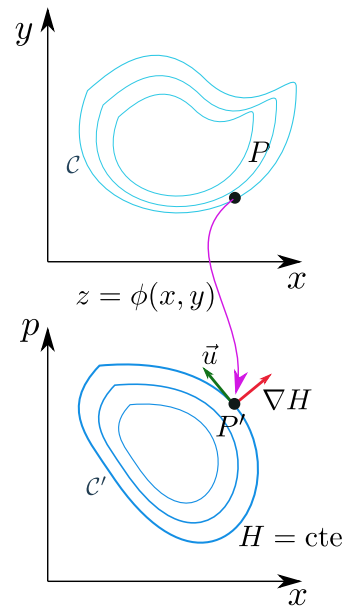


Figure 1. A mapping that transforms the equations of movements (2,3) into a more symmetrical one (4,5) with a geometrical meaning.

with fluid mechanics where the trajectories are similar to current lines : when the level curves of  $H$  are “*closer*” to each other, points on these curves move “*faster*”. We will see below that this “*phase flow*” is indeed incompressible, *i.e.* the volume enclosing  $n$  points moving in  $(x, p)$  space remains constant. This property, called Liouville’s theorem, is at the heart of Hamiltonian formulation of mechanics. We will come back to this fundamental point shortly, but let us first consider few examples.

*Example 1.* As a first example, consider the simple equation

$$m\ddot{x} = -V'(x)$$

that transforms into

$$\dot{x} = y \quad (7)$$

$$\dot{y} = -V'(x)/m \quad (8)$$

These equations are already in the Hamiltonian form where we can check that the function

$$H(x, y) = y^2/2 + V(x)/m$$

has the correct form. Note that we could have made a simple scaling, a change of variable such as  $p = \alpha y$  where  $\alpha$  is a constant. With this variable, the Hamilton function becomes

$$H(x, p) = \frac{1}{2\alpha} p^2 + \frac{\alpha}{m} V(x)$$

Usually, the choice  $\alpha = m$  is made to give  $H$  the dimension of an energy.

*Example 2.* As a slightly more complicated example, consider a two dimensional movement in a central field  $V(r)$ . Using polar coordinates and setting  $\vec{r} = r\vec{u}_r$ , we have

$$\begin{aligned}\dot{\vec{r}} &= \dot{r}\vec{u}_r + r\dot{\theta}\vec{u}_\theta \\ \ddot{\vec{r}} &= (\ddot{r} - r\dot{\theta}^2)\vec{u}_r + (2\dot{r}\dot{\theta} + r\ddot{\theta})\vec{u}_\theta\end{aligned}$$

and therefore, the Newton equations of motion are

$$\ddot{r} - r\dot{\theta}^2 = -V'(r)/m \quad (9)$$

$$2\dot{r}\dot{\theta} + r\ddot{\theta} = 0 \quad (10)$$

It is straightforward to check that the above equations can be put into the Hamiltonian form if we set  $p_r = m\dot{r}$  and  $p_\theta = mr^2\dot{\theta}$  and define

$$H(r, p_r, \theta, p_\theta) = \frac{1}{2m}p_r^2 + \frac{1}{2m}\frac{p_\theta^2}{r^2} + V(r) \quad (11)$$

Note that  $p_r$  and  $p_\theta$  don't have the same dimension.

Finding the mapping  $(\dot{r}, \dot{\theta}) \rightarrow (p_r, p_\theta)$  is not a guess work. In a given system of coordinates  $(q^1, q^2, \dots, q^n)$ , the property of the space is given by the metric tensor  $g_{ij}$  that defines the infinitesimal distance between two close points  $(q^1, \dots, q^n)$  and  $(q^1 + dq^1, \dots, q^n + dq^n)$

$$ds^2 = \sum_{i,j=1}^n g_{ij}dq^i dq^j \quad (12)$$

and by extension, what we call the kinetic energy  $T = (m/2)(ds/dt)^2$ . For problems where the potential  $V$  is only a function of coordinates  $q^i$ , specifying the metric tensor automatically provides the mapping we have to use:

$$p_i = \sum_j g_{ij}\dot{q}^j \quad (13)$$

In other terms, the arc element is written as

$$\left(\frac{ds}{dt}\right)^2 = \sum_i p_i \dot{q}^i \quad (14)$$

For example, for the polar coordinates,  $ds^2 = dr^2 + r^2 d\theta^2$ ,  $g_{11} = 1$ ,  $g_{22} = r^2$  and  $g_{12} = g_{21} = 0$ . The mapping is therefore (up to the scaling factor  $m$ )  $p_r = g_{11}\dot{r} = \dot{r}$  and  $p_\theta = g_{22}\dot{\theta} = r^2\dot{\theta}$ .

We will see in section IV why relation (13) provides the correct mapping for Hamiltonian formulation. Before this derivation however, let us further study the Liouville's theorem and its role in Hamiltonian formulation of mechanics.

### III. VOLUME IN PHASE SPACE.

#### A. Determinant, trace and volume.

Let us recall the fundamental definitions of *determinant* and *trace* of linear applications. Consider a vectorial space  $\mathcal{E}$  of dimension  $n$ , where we don't need to

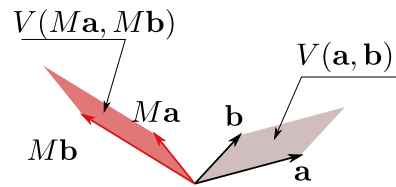


Figure 2. To  $n$  vectors, we can associate a volume. The determinant of an application  $M$  measures the relative change in volume of  $n$  vectors to which  $M$  is applied. Here for a two dimensional example,  $V(M\mathbf{a}, M\mathbf{b}) = |M|V(\mathbf{a}, \mathbf{b})$ .

define any scalar product. Even without the concept of distance, we can define a *volume* in this space. For example, we can speak of Kg.m (a volume in mass-distance space).

To a set of  $n$  vectors  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  in this space, we can associate a number we call *volume* through a  $n$ -linear anti-symmetric application  $V : \mathcal{E}^n \rightarrow \mathbb{R}$ . For example, in a two dimensional space,  $V(\mathbf{e}_1, \mathbf{e}_2) = -V(\mathbf{e}_2, \mathbf{e}_1)$ ,  $V(\lambda\mathbf{e}_1, \mathbf{e}_2) = \lambda V(\mathbf{e}_1, \mathbf{e}_2)$ ,  $V(\mathbf{a} + \mathbf{b}, \mathbf{c}) = V(\mathbf{a}, \mathbf{c}) + V(\mathbf{b}, \mathbf{c})$  and so on. As  $V$  is anti-symmetric,  $V(\mathbf{a}, \mathbf{a}) = 0$ .

If for a given basis of the vector space, we define  $V(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$ , we can find the volume associated to any  $n$ -vectors  $\mathbf{a}_1, \dots, \mathbf{a}_n$  by simply expressing these vectors in the given basis and use the linearity of anti-symmetry of  $V$ .

Now consider a linear application  $M : \mathcal{E} \rightarrow \mathcal{E}$  in this space that associates a vector to a vector. We can measure the *change of volume* induced by the application of  $M$  to  $n$  vectors (figure 2):

$$V(M\mathbf{e}_1, \dots, M\mathbf{e}_n) = |M|V(\mathbf{e}_1, \dots, \mathbf{e}_n)$$

where  $|M|$ , called the *determinant* of  $M$ , is a number that depends only on  $M$ . If we express  $M$  in a given basis and note  $M_j^i$  its matrix elements, a straightforward application of the above concepts allows us to write  $|M|$  in terms of  $M_j^i$ . For example, in two dimensions,

$$\begin{aligned}V(M\mathbf{e}_1, M\mathbf{e}_2) &= V(M_1^1\mathbf{e}_1 + M_2^1\mathbf{e}_2, M_1^2\mathbf{e}_1 + M_2^2\mathbf{e}_2) \\ &= (M_1^1M_2^2 - M_2^1M_1^2)V(\mathbf{e}_1, \mathbf{e}_2)\end{aligned}$$

where to get from the first line to the second line, we have used the linearity and anti-symmetry properties of  $V$ . For the two dimensional case therefore:

$$|M| = M_1^1M_2^2 - M_2^1M_1^2 \quad (15)$$

If we use another basis, the number  $M_j^i$  will change but the quantity  $|M|$  being defined geometrically, remains invariant.

Now consider the linear application  $I + \epsilon M$ , where  $I$  is the identity application and  $\epsilon$  a scalar. Using the same concept of linearity, it is straightforward to see that we can write

$$|I + \epsilon M| = 1 + \epsilon a_1(M) + \epsilon^2 a_2(M) + \dots + \epsilon^n a_n(M)$$

where  $a_i(M)$  are numbers that depend only on  $M$ . The function  $a_1(M)$ , noted  $\text{tr}(M)$ , is called the *trace* of  $M$

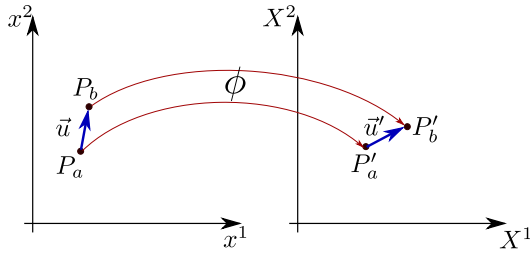


Figure 3. fsdfdf

and  $a_n(M) = |M|$ . A simple computation shows that in a given basis,  $\text{tr}(M)$  is the sum of diagonal elements of the matrix of  $M$  in this basis:

$$a_1(M) = \sum_i M_i^i \quad (16)$$

### B. Change of variable.

Consider two  $n$ -dimensional spaces referenced by coordinates  $x^i$  and  $X^i$  ( $i = 1 \dots n$ ) and a mapping  $\phi$  between them where

$$X^i = X^i(x^1, x^2, \dots, x^n) \quad i = 1, \dots, n$$

Consider two close points  $P_a(x^1, \dots, x^n)$  and  $P_b(x^1 + h^1, \dots, x^n + h^n)$  in the first space and the vector (figure 3)

$$\vec{u} = \overrightarrow{P_a P_b} = (h^1, \dots, h^n)^T$$

Under the mapping  $\phi$ , these points and vector are transformed into  $P'_a(X^1, \dots, X^n)$  and  $P'_b(X^1 + k^1, \dots, X^n + k^n)$   $\vec{u}' = \overrightarrow{P'_a P'_b} = (k^1, \dots, k^n)^T$  where, to the first order in  $h^j$ :

$$k^i = \sum_j \frac{\partial X^i}{\partial x^j} h^j$$

The above relations in matricial form can be written as

$$\vec{u}' = \mathbf{J} \vec{u}$$

where the linear application  $\mathbf{J}$  is called the Jacobian matrix, denoted as  $\partial(X^1, \dots, X^n)/\partial(x^1 \dots x^n)$ ; its matrix elements are

$$(\mathbf{J})_j^i = \frac{\partial X^i}{\partial x^j}$$

The determinant of this application is often simply called the Jacobian and measures the change of volume delimited by  $n$ -vectors at a given point from one space to another.

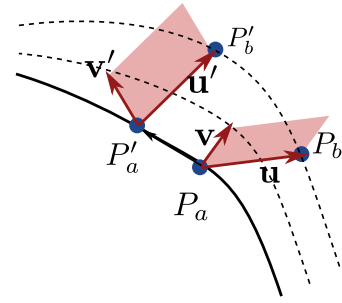


Figure 4. When a point  $P_a$  moves to a point  $P'_a$  during a short time  $dt$  according to the differential equation 17, nearby vectors  $\mathbf{u}, \mathbf{v}$ , are transformed into vectors  $\mathbf{u}', \mathbf{v}'$ ,... according to equation (18). To the first order in  $dt$ , the relative change in the volume enclosed by these vectors is given by  $\text{div}(\mathbf{f}) = \text{tr}(\partial\mathbf{f}/\partial\mathbf{x})$ .

### C. Differential equations.

The above discussion applies directly to system of first order differential equations of the form

$$\dot{x}^i = f^i(x^1, \dots, x^n) \quad i = 1, \dots, n \quad (17)$$

where we can see the numbers  $x^1, \dots, x^n$  as coordinates of a point  $P_a$  in an  $n$ -dimensional space at time  $t$ . During a short (infinitesimal) time  $dt$ , a point  $P_a$  transforms into a point  $P'_a$  with coordinates  $X^i = x^i + dt f^i(x^1, x^2, \dots, x^n)$ . Consider a close point  $P_b$  with coordinates  $x^i + h^i$  and denote  $\mathbf{u} = \overrightarrow{P_a P_b}$ . After the time  $dt$ , the point  $P_b$  transforms into  $P'_b$  and the vector  $\mathbf{u}$  into the vector  $\mathbf{u}'$  (figure 4):

$$\mathbf{u}' = \mathbf{u} + dt \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \mathbf{u} = \left( I + dt \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right) \mathbf{u} \quad (18)$$

where the matrix elements  $f_j^i$  of the linear application  $\partial\mathbf{f}/\partial\mathbf{x}$  are the usual  $\partial f^i/\partial x^j$ . If we consider  $n$  vectors  $\mathbf{u}_k$  around a point  $P$  and the volume  $V$  they enclose and recall the previous discussion of *trace*, we see that the rate of change in this volume during the time  $dt$  is

$$\frac{1}{V} \frac{dV}{dt} = \text{tr} \left( \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right) = \sum_i \partial f^i / \partial x^i \quad (19)$$

The sum is usually called the *divergence* of the field  $\mathbf{f}$ .

### D. Liouville's theorem.

If we look at a Hamiltonian system such as (4,5) we observe that the change in the volume enclosing a point  $P$  in this space, according to equation (19) is

$$\frac{1}{V_H} \frac{dV_H}{dt} = \frac{\partial}{\partial x} \left( \frac{\partial H}{\partial p} \right) + \frac{\partial}{\partial p} \left( -\frac{\partial H}{\partial x} \right) = 0 \quad (20)$$

The Hamiltonian equations conserve *automatically* the volumes in the phase space because by construction, the

linear mappings they induce are trace-less. Relation (20) is known as the Liouville's theorem.

On the other hand, in the original  $(x, y)$  space of equations (2,3), the change in the volume around a point is

$$\frac{1}{V_O} \frac{dV_O}{dt} = \frac{\partial}{\partial x}(y) + \frac{\partial}{\partial y}(f(x, y)) = \frac{\partial f}{\partial y}$$

The change of variable  $y \rightarrow p$  is made precisely to correct for this loss (or gain) of volume in the original space as points move along their trajectories.

Liouville's theorem is at the heart of the Hamiltonian formulation and is used in fundamental approaches to mechanics. For example, the Poincaré's recurrence theorem, stating that for bounded systems, given enough time, a system always comes back arbitrary close to its initial state, is based on Liouville's theorem. Liouville's theorem is also used in discussing the ergodic hypothesis in statistical mechanics.

#### IV. FINDING THE MOMENTS

##### A. Momenta and the metric tensor.

Hamiltonian formulation is interesting only we are able to reformulate easily the Newton's equation in this frame. Here for simplicity, we consider at first one particle of mass  $m$  moving in a  $n$ -dimensional space and then generalize the formulation to a system of  $N$  particles.

In a generalized system of coordinates  $q^1, \dots, q^n$ , the fundamental quantity that defines the underlying geometry of the physical space is the distance  $ds$  between two close points whose coordinates differences are  $dq^i$ :

$$ds^2 = g_{ij} dq^i dq^j \quad (21)$$

Note that from now on, we use the repeated index summation convention: if an index is repeated in an expression, *once in the lower position and once in the higher one*, we assume that it is summed over. Hence, instead of writing for example  $\sum_{i,j} \Gamma_{jk}^i p_i q^j$  (whatever the meaning of this expression), we just write  $\Gamma_{jk}^i p_i q^j$ . The expression (21) therefore assumes summation over the two dummy indexes  $i, j$ .

the quantity  $(g_{ij})$  (which can be seen as an  $n \times n$  matrix) is called the metric tensor and without loss a generality, we suppose it to be symmetric:  $g_{ij} = g_{ji}$ . We note  $\tilde{g}^{ij}$  its inverse with the usual property of the inverse matrix:

$$g_{ij} \tilde{g}^{jk} = \delta_i^k$$

where  $\delta_i^k$  is the Kronecker symbol. In general,  $g_{ij}$  is a function of coordinates.

The Newton equations of motion in generalized coordinates (see appendix 1) are

$$\frac{d}{dt}(g_{ij} \dot{q}^j) - \frac{1}{2} \frac{\partial g_{kl}}{\partial q^i} \dot{q}^k \dot{q}^\ell = -\frac{1}{m} \frac{\partial V}{\partial q^i} \quad i = 1, \dots, n \quad (22)$$

Where in the above expression, all indexes except  $i$  are repeated and therefore summed over. One can check the above formula for the example of the 2 dimensional movement in a central field (equations 9,10).

The left part of relation (22) is the equation of "geodesics" when no external potential exists. In a "flat" space, geodesics are straight lines, which are the trajectories of particles moving in the absence of external forces. Geodesics generalize the concept of straight lines to curved spaces. For example, on the unit sphere where spherical coordinates  $\theta, \phi$  are used and  $ds^2 = (d\theta)^2 + \sin^2 \theta (d\phi)^2$ , solving equations (22) for  $V = 0$  shows that geodesics are great circles.

When an outside potential exists, trajectories deviate from the geodesics. Einstein realized that for gravitational problems, one can "include" the potential term into the metric tensor: particles still move on geodesics, but the metric tensor is modified compared to the metric of a flat manifold.

Coming back to the question of finding a correct mapping for Hamiltonian formulation of conservative classical dynamics, we see in the expression (22) that the mapping  $\dot{q} \rightarrow p$  seems obvious:

$$p_i = m g_{ij} \dot{q}^j \quad (23)$$

which we can invert as

$$\dot{q}^i = \frac{1}{m} \tilde{g}^{ij} p_j \quad (24)$$

Expression (22) thus transforms into

$$\frac{d}{dt} p_i - \frac{1}{2m} \frac{\partial g_{kl}}{\partial q^i} \tilde{g}^{kj} \tilde{g}^{\ell m} p_j p_m = -\frac{\partial V}{\partial q^i} \quad i = 1, \dots, n \quad (25)$$

In order to put the above expression into its final Hamiltonian form, we have to arrange the middle term:

$$\frac{\partial g_{kl}}{\partial q^i} \tilde{g}^{kj} \tilde{g}^{\ell m} = \frac{\partial (g_{kl} \tilde{g}^{kj})}{\partial q^i} \tilde{g}^{\ell m} - \frac{\partial \tilde{g}^{kj}}{\partial q^i} g_{kl} \tilde{g}^{\ell m}$$

but

$$g_{kl} \tilde{g}^{kj} = \delta_l^j ; \quad g_{kl} \tilde{g}^{\ell m} = \delta_k^m$$

so

$$\frac{\partial g_{kl}}{\partial q^i} \tilde{g}^{kj} \tilde{g}^{\ell m} p_j p_m = -\frac{\partial \tilde{g}^{kj}}{\partial q^i} \delta_k^m p_j p_m = -\frac{\partial \tilde{g}^{kj}}{\partial q^i} p_j p_k$$

and expression (25) can finally be written as

$$\dot{p}_i = -\left\{ \frac{1}{2m} \frac{\partial \tilde{g}^{kj}}{\partial q^i} p_j p_k + \frac{\partial V}{\partial q^i} \right\} = -\frac{\partial H}{\partial q^i}$$

where we have defined

$$H = \frac{1}{2m} \tilde{g}^{kj} p_j p_k + V(q) \quad (26)$$

Note that for  $H = H(q, p)$ , coordinates  $q$  and momenta  $p$  are considered *independent* variables. Therefore, for

derivations  $\partial/\partial q^i$ , the momenta  $p_j$  are considered as constants. By the same token,

$$\frac{\partial H}{\partial p_j} = \tilde{g}^{kj} p_k = \dot{q}^j$$

according to expression (24). We see that the form of the Hamiltonian (relation 26) and the mapping  $\dot{q}^i \rightarrow p_i$  are automatically obtained from the metric tensor.

The generalization to a system of  $N$  particles is trivial and is obtained by repeating the above arguments for each particle. In particular, the momenta are defined as

$$p_{\alpha,i} = m_{\alpha} g_{ij} \dot{q}_{\alpha}^i \quad i = 1 \dots n, \alpha = 1 \dots N$$

and the Hamiltonian is defined as

$$H = \sum_{\alpha=1}^N \frac{1}{2m_{\alpha}} \tilde{g}^{kj} p_{\alpha,j} p_{\alpha,k} + V(q_1^1, \dots, q_N^n)$$

The generalization is even broader if we consider the movement of  $N$  particle in an  $n$  dimensional space as the movement of a single particle in an  $N \times n$ -dimensional space.

### B. Gauge functions.

There is some indeterminacy in the definition of the momenta. To see this, consider the Newton equation for a one dimensional system

$$m \frac{d}{dt} (g(q)\dot{q}) - \frac{m}{2} g'(q)\dot{q}^2 = -V'(q) \quad (27)$$

which we can write, by adding the *same* quantity to both side of the equation as

$$m \frac{d}{dt} \left( g(q)\dot{q} + \frac{1}{m} f(q) \right) - \frac{m}{2} g'(q)\dot{q}^2 = -V'(q) + \dot{q} f'(q)$$

where  $f(q)$  is an *arbitrary* smooth function of the position  $q$ . Following our previous discussion, we can set

$$p = mg(q)\dot{q} + f(q) ; \quad \dot{q} = \frac{1}{mg(q)} p - f(q), \quad (28)$$

write the Hamiltonian as

$$H(q, p) = \frac{1}{2mg(q)} (p - f(q))^2 + V(q)$$

and still get the canonical equations

$$\dot{q} = \partial_p H ; \quad \dot{p} = -\partial_q H$$

This generalization of the momenta can be easily extended to  $n$  spatial dimensions when dealing with the Newton equation (22). Writing the Newton equation as

$$\frac{d}{dt} (m g_{ij} \dot{q}^j + f_i) - \frac{1}{2} \frac{\partial g_{k\ell}}{\partial q^i} \dot{q}^k \dot{q}^{\ell} = -\frac{\partial V}{\partial q^i} + \frac{\partial f_i}{\partial q^j} \dot{q}^j$$

by adding the arbitrary functions  $f_i$ , we can define the momenta as

$$p_i = m g_{ij} \dot{q}^j + f_i(q) \quad (29)$$

Following the same line of arguments as above, we see that the Hamiltonian is now:

$$H = \frac{1}{2m} \tilde{g}^{kj} (p_j - f_j) (p_k - f_k) + V(q) \quad (30)$$

provided that the functions  $f_i(q)$  obey the following constraints:

$$\frac{\partial f_j}{\partial q^k} - \frac{\partial f_k}{\partial q^j} = 0 \quad (31)$$

In Lagrangian dynamics, this indeterminacy is reflected in the fact that the Lagrangian is defined up to a total differential.

Until now, we have considered conservative systems where the potential is defined only as a function of positions. The above discussion shows how to deal with some type of potentials that contain also the velocities. A case in point is the time-independent electromagnetic force where the right hand side of Newton equation is of the form

$$-\frac{\partial V}{\partial q^i} + \left( \frac{\partial A_j}{\partial q^i} - \frac{\partial A_i}{\partial q^j} \right) \dot{q}^j$$

and summation over the dummy index  $j$  is assumed. Following the same procedure as discussed above, we see that we can write the momenta and the Hamiltonian as in relations (29-30) where  $f_i = A_i + C_i$  and the arbitrary functions  $C_i$  (usually called gauge functions) obey the constraint (31).

### V. CANONICAL TRANSFORMATIONS.

The Hamiltonian formulation is of fundamental importance for a deep understanding of mechanics. The Hamiltonian approach can also be useful to find an exact solution if we make an adequate change of variables. In the Newtonian approach, making a change of variable to simplify the potential term usually makes the kinetic energy term complicated. In Hamiltonian mechanics on the other hand, coordinates and momenta are considered as independent variables and we have more freedom to find an interesting change of variables. We are interested in the kind of change that *preserves* the structure of the Hamiltonian. In one dimension, if we call  $X, P$  the new variables, we wish to have

$$\dot{X} = \partial_P H ; \quad \dot{P} = -\partial_X H \quad (32)$$

The change of variables that preserve the Hamiltonian structure are called *canonical*. We will see below the general principles guiding these transformations. Canonical transformations are valid for all Hamiltonians and

are not restricted to a specific problem. However, a specific problem can be greatly simplified if the transformed Hamiltonian does not contain one or more position coordinate  $X^i$ . Then for this coordinate,  $\dot{P}_i = 0$  and the transformed momenta  $P_i$  is a conserved quantity of our mechanical system, associated to a symmetry of the problem through the Noether theorem[17]. A basic example is the angular momentum of the movement in a central field (equation 11). I don't develop the systematic search for conserved quantities in this article. Let us just mention that this investigation is usually performed by using the Hamilton-Jacobi equation, which, in the words of Arnold[12], "[...] is the most powerful method known for the exact integration [of Hamilton equations]".

### A. The structure of canonical transformations.

For simplicity, let us first consider a one-dimensional system in the physical space. The mapping  $(x, p)$  we have found to transform the Newton equation into its Hamiltonian equivalent is by no mean unique. Once we have found a correct mapping  $(x, p)$ , we can find infinitely many others  $(X, P)$  that have the same property.

Consider a change of variable from  $(x, p)$  to  $(X, P)$  space, *i.e.*  $X = X(x, p)$  and  $P = P(x, p)$ . Using the usual tools of change of variables (see below), we find that

$$\frac{dX}{dt} = J(X, P) \frac{\partial H}{\partial P} \quad (33)$$

$$\frac{dP}{dt} = -J(X, P) \frac{\partial H}{\partial X} \quad (34)$$

where the proportionality factor  $J(X, P)$  is precisely the Jacobian of the transformation

$$J = \left| \frac{\partial(X, P)}{\partial(x, p)} \right| = \begin{vmatrix} \partial_x X & \partial_p X \\ \partial_x P & \partial_p P \end{vmatrix}$$

discussed in subsection III B. When a point moves in  $(x, p)$  space, the volume  $V_{xp}$  around it is conserved. If  $J = 1$  (or more generally any constant) throughout the whole phase space, at each time, the volume around the transformed point  $(X, P)$  equals the volume  $V_{xp}$  and hence it is also conserved as the point moves in the  $(X, P)$  space.

Therefore, in one dimension, any transformation for which  $J = 1$  is a canonical one and conserves the form of the Hamilton equations as in relation (32).

As an example, consider the change of variable

$$X = x \cos s + p \sin s \quad (35)$$

$$P = -x \sin s + p \cos s \quad (36)$$

where  $s$  is an arbitrary constant. It is trivial to check that this is indeed a canonical transformation. More over, if the Hamiltonian for the original equation were  $H = p^2 + x^2$ , in the new variables it becomes  $H = P^2 + X^2$ .

A slightly less trivial canonical transformation is

$$x = \sqrt{2P} \sin X ; p = \sqrt{2P} \cos X \quad (37)$$

For this canonical transformation, the Hamiltonian  $H = p^2 + x^2$  transforms into  $H = 4P$ , effectively eliminating the variable  $X$  from the Hamiltonian.

It is worthwhile to detail the transformation. In order to shorten the notations, we use the (column) vector  $\vec{\eta} = (x, p)^T$  for the old variables and  $\vec{\zeta} = (X, P)^T$  for the new variables. The Hamilton equation for the old variables is

$$\frac{d}{dt} \begin{pmatrix} \eta^1 \\ \eta^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\partial_{\eta^1} H, \partial_{\eta^2} H)^T \quad (38)$$

where we stress that derivatives of  $H$  constitute a row vector (a linear form). Noting  $\Omega$  the anti-symmetric matrix of the above relation, we can write this equation as

$$\frac{d}{dt} \vec{\eta} = \Omega.(\partial_{\vec{\eta}} H)^T \quad (39)$$

Consider now a change of variable from  $\eta^i$  to  $\zeta^i$  that we write in differential form and using the summation convention as

$$d\eta^i = f_j^i d\zeta^j \quad (40)$$

$$d\zeta^j = \tilde{f}_i^j d\eta^i \quad (41)$$

The matrix  $(f_j^i)$  and  $(\tilde{f}_j^i)$  are the usual transition matrix where

$$f_j^i = \partial \eta^i / \partial \zeta^j ; \tilde{f}_j^i = \partial \zeta^i / \partial \eta^j$$

The chain rule of derivation applied to  $H$  is

$$\frac{\partial H}{\partial \eta^i} = \tilde{f}_i^j \frac{\partial H}{\partial \zeta^j} \quad (42)$$

or in matricial notation,

$$(\partial_{\vec{\eta}} H) = (\partial_{\vec{\zeta}} H). \tilde{f}$$

where again distinction between column and row vectors is stressed. Relation (42) is just a shorthand for expressions like

$$\frac{\partial H}{\partial x} = \frac{\partial H}{\partial X} \frac{\partial X}{\partial x} + \frac{\partial H}{\partial P} \frac{\partial P}{\partial x}$$

Now, making the actual change of variables in relation (39), we have

$$f.\vec{\zeta} = \Omega.\tilde{f}^T.(\partial_{\vec{\zeta}} H)^T$$

Left-multiplying both side by  $\tilde{f}$ , we see that the transformation is canonical if (and only if)

$$\tilde{f}\Omega\tilde{f}^T = \Omega \quad (43)$$



or equivalently

$$f\Omega = \Omega \tilde{f}^T \quad (44)$$

Multiplication by the  $2 \times 2$   $\Omega$  matrix is straightforward. Performing this multiplication in relation (43) results in our claim

$$J = 1 \quad (45)$$

Performing this multiplication in relation (44) results in a more detailed relation :

$$\frac{\partial x}{\partial X} = \frac{\partial P}{\partial p}; \frac{\partial p}{\partial P} = \frac{\partial X}{\partial x} \quad (46)$$

$$\frac{\partial x}{\partial P} = -\frac{\partial X}{\partial p}; \frac{\partial p}{\partial X} = -\frac{\partial P}{\partial x} \quad (47)$$

The above relations (45) or (46-47) give the sufficient and necessarily conditions for a transformations to be canonical.

Generalization of these relations to  $n$ -dimensional systems (and hence  $2n$ -dimensional phase space) is straightforward (see appendix B for details). Hamilton equations have a fine structure: not only the linear mapping they induce is trace-less (relation 20), but it is so on each set of variables  $(x^i, p_i)$ . Writing  $dH = \sum_{i=1}^n dH_i$  where

$$dH_i = \frac{\partial H}{\partial x^i} dx^i + \frac{\partial H}{\partial p_i} dp_i$$

we see that *each*  $dH_i = 0$  if the variables  $(\vec{x}, \vec{p})$  follow the Hamilton equations.

For a transformation  $\vec{X}(\vec{x}, \vec{p}), P(\vec{x}, \vec{p})$  to be canonical, it has to preserve this fine structure, *i.e.* for each couple  $i, j$ , we must have (see appendix B)

$$\frac{\partial x^i}{\partial X^j} = \frac{\partial P_j}{\partial p_i}; \frac{\partial p_i}{\partial P_j} = \frac{\partial X^j}{\partial x^i} \quad (48)$$

$$\frac{\partial x^i}{\partial P_j} = -\frac{\partial X^j}{\partial p_i}; \frac{\partial p_i}{\partial X^j} = -\frac{\partial P_j}{\partial x^i} \quad (49)$$

The meaning of these relations is close to what we discussed above (figure 5): Consider an areas (2-volumes) in a given  $(x^i, p_i)$  plane and the *projection* of its transformed into all planes  $(X^j, P_j)$ . Relations (48-49) imply that the sum of the projected area is conserved, *i.e.*

$$\sum_{j=1}^n \left| \frac{\partial(X^j, P_j)}{\partial(x^i, p_i)} \right| = 1 \quad (50)$$

Details of the explanation is given in appendix B. The conservation can be generalized to  $2k$ -volumes ( $k = 1, \dots, n$ ).

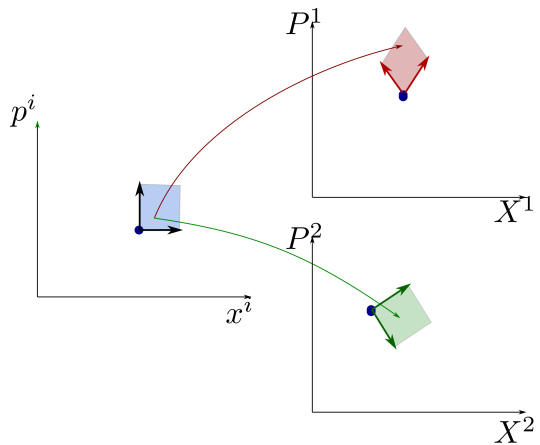


Figure 5. For a canonical transformation, the sum of a projected area in a given plane  $(x^i, p_i)$  into all planes  $(X^j, P_j)$  ( $j = 1 \dots n$ ) is conserved.

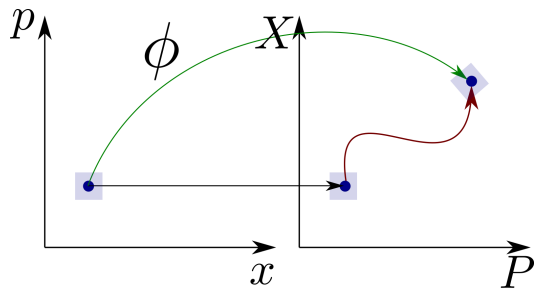


Figure 6. Consider in  $(X, P)$  space points transforming according to relation (A12) with the initial condition  $X(0) = x$ ;  $P(0) = p$ . For a given value of  $s$ ,  $(X(s), P(s))$  is a mapping  $\phi : (x, p) \rightarrow (X, P)$  that is canonical, *i.e.* conserves volumes.

## B. Finding canonical transformations.

Relations (45) or (46-47) are not only necessary conditions, they also provide for an efficient method to *find* canonical transformations.

Consider an arbitrary function  $K(X, P)$ , an arbitrary variable  $s$  and set

$$\frac{dX}{ds} = \frac{\partial K}{\partial P}; \frac{dP}{ds} = -\frac{\partial K}{\partial X} \quad (51)$$

with the initial condition  $X(0) = x$  and  $P(0) = p$ . For a given value of  $s$ , we can see  $(X(s), P(s))$  as a mapping from  $(x, p)$  space into  $(X, P)$  space (figure 6). Because equation (51) has an Hamiltonian form, we know that it automatically conserves the volumes and therefore, whatever the value of  $s$ , we have  $J(s) = 1$ . The example (35-36) of canonical transformation, which corresponds to a rotation in  $(x, p)$  space, is found by setting

$$K = P^2 + X^2$$

In continuous group theory (Lie groups), relation (51) is called an infinitesimal generator of the group of canonical

transformations. The Original Hamilton equations are themselves such a generator if we set  $s = t$  and  $K = H$ . It is straightforward to generalize the Lie approach to  $n$ -dimensional systems.

The Lie approach is very elegant. However, we are only able to solve analytically simple differential equations and this approach does not seem to provide for an easy way to obtain all canonical transformations.

There exists however an algebraic way of finding canonical transformations. Consider an arbitrary function  $F(x, X)$  (called a generating function) and set

$$p = \frac{\partial F}{\partial x} ; P = -\frac{\partial F}{\partial X} \quad (52)$$

Mixing old and new variables may seem at first strange. Relations (52) are however *two* equations relating *four* variables, so in principle and at least locally (provided some conditions on  $F$ ), we can always solve it to express  $(X, P)$  as a function of  $(x, p)$ . Transformations (37) for example are obtained by choosing  $F(x, X) = (1/2)x^2 \cot X$ .

Differentiating relations (52) and rearranging the terms, we have

$$\begin{pmatrix} \frac{\partial^2 F}{\partial x \partial X} & 0 \\ -\frac{\partial^2 F}{\partial X^2} & -1 \end{pmatrix} \begin{pmatrix} dX \\ dP \end{pmatrix} = \begin{pmatrix} -\frac{\partial^2 F}{\partial x^2} & 1 \\ \frac{\partial^2 F}{\partial x \partial X} & 0 \end{pmatrix} \begin{pmatrix} dx \\ dp \end{pmatrix} \quad (53)$$

or in shorthand matricial notations,

$$A.d\vec{\zeta} = B.d\vec{\eta}$$

Provided that  $A$  is an invertible matrix, *i.e.*  $\partial^2 F / \partial x \partial X \neq 0$ , the Jacobian of the transformation is

$$J = |A^{-1}B| = |B| / |A| = 1$$

Some care must be taken if  $\partial^2 F / \partial x \partial X = 0$  at some points of null measure.

Four generating functions can be obtained by mixing differently the old and new variable. For example, by choosing the generating function  $F(p, X)$  and setting

$$x = -\frac{\partial F}{\partial p} ; P = -\frac{\partial F}{\partial X} \quad (54)$$

we also obtain a canonical transformation.

The method of generating function is generalized to  $n$ -dimensional systems. For example, choosing the generating function  $F(x^1, \dots, x^n; X^1, \dots, X^n)$  and setting

$$p_i = \frac{\partial F}{\partial x^i} ; P_i = -\frac{\partial F}{\partial X^i}$$

leads to a canonical transformation.

## VI. CONCLUSION.

The main intention of this article was to stress the geometry of the Hamilton's equation in classical dynamics,

using volume conservation (Liouville's theorem) as the guiding line. In particular, the usual road of Lagrangians and Legendre transforms have been avoided.

I believe this approach can simplify some of the concepts of Hamiltonian dynamics for some undergraduate students that discover this field for the first time, by presenting the material from a different angle. It can also serve to initiate students to powerful tools such as tensor calculus or symplectic geometry that they will discover later in their studies.

As I stressed in the introduction, the approach I propose is only complementary to the powerful standard one. The materials developed in this short article, which does not contain the usual mathematical complexity found in most textbooks, can be covered in one or two lectures and I hope help students to get a better understanding of the Hamiltonian formalism.

*Acknowledgment.* I'm grateful to Marcel Vallade and Lado Samushia for fruitful discussions.

## Appendix A: Newton equation in generalized coordinates

There are various methods to deduce the Newton equations of motion in generalized coordinates. The most natural way is to observe that in Cartesian coordinates, the Newton equations are of variational type, deduce the Lagrangian in these coordinates, make the change of variables directly in the Lagrangian and deduce the Euler-Lagrange equations in the new set of coordinates. Another method is to begin with the Alembert's principle, make the adequate virtual displacement in generalized coordinates and obtain the dynamical equations. In this article, I have avoided the Lagrangian formalism. The d'Alembert's method, to be rigorous, needs the concept of parallel transport of vectors and covariant derivation. Here I use a middle approach, making the change of variables directly in Newton's equations of Cartesian coordinates. Even though the computation can seem of some length, it involves only basic algebraic manipulation such as multiplying and summing over repeated indexes.

Consider a change of coordinates from Cartesian coordinates  $x^i$  to generalized coordinates  $q^i$  :

$$x^i = f^i(q^1, \dots, q^n) \quad i = 1, \dots, n$$

which we can write in differential form as

$$dx^i = f_j^i(q) dq^j \quad (A1)$$

where

$$f_j^i(q) = \frac{\partial f^i}{\partial q^j} = \frac{\partial x^i}{\partial q^j} \quad (A2)$$

As we suppose the matrix  $(f_j^i)$  to be invertible, we can reverse relation (A1) :

$$dq^j = \tilde{f}_i^j dx^i \quad (A3)$$

where the matrix  $(\tilde{f}_i^k) = \partial q^k / \partial x^i$  is the inverse of  $(f_j^i)$ , or in other terms

$$f_j^i \tilde{f}_i^k = \delta_j^k \quad (\text{A4})$$

Furthermore, because of the equality of the cross derivations, we have

$$\frac{\partial f_j^i}{\partial q^k} = \frac{\partial f_k^i}{\partial q^j} \quad (\text{A5})$$

The fundamental quantity in a physical space is the infinitesimal arc-length :

$$ds^2 = \delta_{ij} dx^i dx^j = g_{ij} dq^i dq^j \quad (\text{A6})$$

where  $\delta_{ij}$  is the Kronecker symbol and  $g_{ij}$  is called the metric tensor in generalized coordinates. Comparing (A1,A6) we obtain

$$g_{\alpha\beta} = \delta_{ij} f_\alpha^i f_\beta^j = f_{j\alpha} f_\beta^j \quad (\text{A7})$$

where the operation of rising or lowering of indexes is defined by

$$f_{j\alpha} = f_\alpha^j = \delta_{ij} f_\alpha^i \quad (\text{A8})$$

In Cartesian coordinates  $(x^1 \dots x^n)$ , the Newton's equations  $\ddot{\mathbf{r}} = -\nabla V/m$  can be written as

$$\frac{d}{dt} \delta_{ij} \dot{x}^j = -\frac{1}{m} \frac{\partial V}{\partial x^i} \quad i = 1, \dots, n \quad (\text{A9})$$

The reason to write  $\delta_{ij} \dot{x}^j$  instead of  $\dot{x}^i$  is that the latter is the  $i$ -th element of a (column or contravariant) *vector* while the former is the  $i$ -th element of a *linear form* (a row or covariant vector). As  $\partial_i V$  is also a linear form, both side of the relation (A9) are of the same type. As

$$\begin{aligned} \frac{d}{dt} \dot{x}^j &= \frac{d}{dt} f_k^j \dot{q}^k \\ \frac{\partial V}{\partial x^i} &= \frac{\partial V}{\partial q^k} \frac{\partial q^k}{\partial x^i} = \frac{\partial V}{\partial q^k} \tilde{f}_i^k \end{aligned}$$

We can express the two sides of relation (A9) in the new coordinates:

$$\frac{d}{dt} (f_{ik} \dot{q}^k) = -\frac{1}{m} \frac{\partial V}{\partial q^k} \tilde{f}_i^k \quad i = 1, \dots, n \quad (\text{A10})$$

This form of the Newton equation is not very usable and terms have to be rearranged. Let us multiply both side of relation (A10) by  $f_\ell^i$  and sum over the repeated index  $i$  :

$$f_\ell^i \frac{d}{dt} f_{ik} \dot{q}^k = -\frac{1}{m} \frac{\partial V}{\partial q^k} f_\ell^i \tilde{f}_i^k = -\frac{1}{m} \frac{\partial V}{\partial q^\ell} \quad \ell = 1, \dots, n$$

by virtue of relation (A4). We can also rearrange the left hand side:

$$\begin{aligned} f_\ell^i \frac{d}{dt} (f_{ik} \dot{q}^k) &= \frac{d}{dt} (f_\ell^i f_{ik} \dot{q}^k) - f_{ik} \dot{q}^k \frac{d}{dt} f_\ell^i \\ &= \frac{d}{dt} (g_{\ell k} \dot{q}^k) - f_{ik} \dot{q}^k \frac{\partial f_\ell^i}{\partial q^\alpha} \dot{q}^\alpha \quad \ell = 1, \dots, n \quad (\text{A11}) \\ &= \frac{d}{dt} (g_{\ell k} \dot{q}^k) - f_{ik} \dot{q}^k \frac{\partial f_\alpha^i}{\partial q^\ell} \dot{q}^\alpha \quad \ell = 1, \dots, n \quad (\text{A12}) \end{aligned}$$

where in the last line, we have used the property (A5) to change the index of derivation. To rearrange the last term of relation (A12), let us note that the derivative of the metric tensor defined in (A7) is given by

$$\frac{\partial g_{k\alpha}}{\partial q^\ell} = \frac{\partial f_{ik}}{\partial q^\ell} f_\alpha^i + f_{ik} \frac{\partial f_\alpha^i}{\partial q^\ell} \quad (\text{A13})$$

Multiplying both side of the above expression by  $\dot{q}^k \dot{q}^\alpha$ , summing over the repeated indexes  $k$  and  $\alpha$  and using the symmetry of the product, we deduce

$$2f_{ik} \dot{q}^k \frac{\partial f_\alpha^i}{\partial q^\ell} \dot{q}^\alpha = \frac{\partial g_{k\alpha}}{\partial q^\ell} \dot{q}^k \dot{q}^\alpha$$

So finally, the Newton equation in generalized coordinates reads:

$$\frac{d}{dt} (g_{\ell k} \dot{q}^k) - \frac{1}{2} \frac{\partial g_{k\alpha}}{\partial q^\ell} \dot{q}^k \dot{q}^\alpha = -\frac{\partial V}{\partial q^\ell} \quad \ell = 1, \dots, n$$

The details of coordinate changes  $f_j^i$  have disappeared from the equation, and only the metric tensor  $g_{\ell k}$  remains in the equation.

## Appendix B: General canonical transformations.

The generalization of canonical transformation to more than one spatial dimension is a little more cumbersome, but the procedure is technically the same. We want to find a change of variable from  $(x^i, p_i)$  into  $(X^i, P_i)$  ( $i = 1, \dots, n$ ) such that the structure of Hamilton equation is conserved, *i.e.*

$$\begin{aligned} \frac{dX^i}{dt} &= \frac{\partial H}{\partial P_i} \\ \frac{dP_i}{dt} &= -\frac{\partial H}{\partial X^i} \end{aligned}$$

In order to avoid manipulating  $x^i$  and  $p_i$  separately, we can, as in subsection V A, group them together into a vector  $\vec{\eta}$  of  $2n$  elements where the first  $n$  elements are the  $x^i$  and the last  $n$  elements are the  $p_i$ . By the same token, we can group the partial differential of  $H$  also into a  $2n$  linear form (row vector)  $\partial_\eta H$ :

$$\vec{\eta} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \\ p_1 \\ \vdots \\ p_n \end{pmatrix}$$

$$\partial_\eta H = (\partial H / \partial x^1, \dots, \partial H / \partial x^n | \partial H / \partial p_1, \dots, \partial H / \partial p_n)$$

With these notations, we can write the Hamilton equations as:

$$\frac{d\vec{\eta}}{dt} = \Omega (\partial_\eta H)^T$$

where  $\Omega$  is a *symplectic* matrix which can be written as blocks

$$\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

Let us now make a change of variable from  $\vec{\eta}$  to  $\vec{\zeta}$  that we write in differential form as

$$d\eta^i = \tilde{f}_j^i d\zeta^j ; d\zeta^j = f_i^j d\eta^i$$

where

$$\tilde{f}_j^i = \frac{\partial \eta^i}{\partial \zeta^j}$$

are the matrix element of the passage matrix from old  $(\vec{x}, \vec{p})$  coordinates to the new  $(\vec{X}, \vec{P})$  ones and  $f$  is the inverse matrix:

$$f_i^j = \frac{\partial \zeta^j}{\partial \eta^i} ; f_i^j \tilde{f}_k^i = \delta_k^j \quad (\text{B1})$$

Upon this change of variable,

$$\frac{\partial H}{\partial \eta^i} = \frac{\partial H}{\partial \zeta^j} \frac{\partial \zeta^j}{\partial \eta^i} = f_i^j \frac{\partial H}{\partial \zeta^j}$$

and the Hamilton equation transforms into

$$\tilde{f} \frac{d\vec{\zeta}}{dt} = \Omega f^T (\partial_{\zeta} H)^T$$

where the subscript  $T$  denotes the transposition operation.

If we wish to keep the structure of Hamilton equation, we must have

$$f \Omega f^T = \Omega$$

or in other terms,

$$\Omega f^T = \tilde{f} \Omega \quad (\text{B2})$$

The  $\Omega$  matrix has a very simple block structure. In order to put this structure into use, we also divide the matrices  $f$  and  $\tilde{f}$  into four  $n \times n$  blocks form:

$$f = \begin{pmatrix} f_a^a & f_b^a \\ f_a^b & f_b^b \end{pmatrix}$$

where

$$\begin{aligned} (f_a^a)_j^i &= \frac{\partial X^i}{\partial x^j} ; (f_b^a)_j^i = \frac{\partial X^i}{\partial p_j} \\ (f_a^b)_j^i &= \frac{\partial P_i}{\partial x^j} ; (f_b^b)_j^i = \frac{\partial P_i}{\partial p_j} \end{aligned}$$

and similar relations for the  $\tilde{f}$  matrix. Now, using the product between blocks, the matricial relation (B2) can be written as

$$\begin{pmatrix} (f_b^a)^T & (f_b^b)^T \\ -(f_a^a)^T & -(f_a^b)^T \end{pmatrix} = \begin{pmatrix} -\tilde{f}_b^a & \tilde{f}_a^a \\ -\tilde{f}_b^b & \tilde{f}_a^b \end{pmatrix}$$

Detailing the above expression, we can explicitly write the relations binding the new and old variables :

$$\frac{\partial X^j}{\partial p_i} = -\frac{\partial x^i}{\partial P_j} \quad (\text{B3})$$

$$\frac{\partial P_j}{\partial p_i} = \frac{\partial x^i}{\partial X^j} \quad (\text{B4})$$

$$\frac{\partial X^j}{\partial x^i} = \frac{\partial p_i}{\partial P_j} \quad (\text{B5})$$

$$\frac{\partial P_j}{\partial x^i} = -\frac{\partial p_i}{\partial X^j} \quad (\text{B6})$$

equations (B3-B6) are the necessary and sufficient conditions for a change of variable to be canonical.

Relation (B2) leads automatically to volume conservation : computing the determinant of both side the above relation, and using the fact that  $|AB| = |A||B|$ ,  $|A^T| = |A|$  and  $|\Omega| \neq 0$ , we observe that we must have  $|f| = |\tilde{f}|$ , which implies that

$$J = |f| = 1$$

As we mentioned (relation 50), canonical transformations impose a more restrictive condition than the total volume transformation : the sum of projected areas from a given  $(x^i, p_i)$  into all  $(X^j, P_j)$  must be conserved:

$$\begin{aligned} \sum_{j=1}^n J_{j,i} &= \sum_{j=1}^n \left| \frac{\partial(X^j, P_j)}{\partial(x^i, p_i)} \right| \\ &= \sum_{j=1}^n \frac{\partial X^j}{\partial x^i} \frac{\partial P_j}{\partial p_i} - \frac{\partial X^j}{\partial p_i} \frac{\partial P_j}{\partial x^i} \\ &= \sum_{j=1}^n \frac{\partial X^j}{\partial x^i} \frac{\partial x^i}{\partial X^j} + \frac{\partial x^i}{\partial P_j} \frac{\partial P_j}{\partial x^i} \\ &= 1 \end{aligned}$$

Where to go from the second to the third one, we have used relations (B3-B6). The third line itself is the relation (B1) written explicitly and therefore equals to one.

- 
- [1] Joseph-Louis Lagrange. *Mécanique analytique, volume 1*. Cambridge University Press, Cambridge, 1er édition édition, July 2009.
- [2] Sandro Caparrini. The history of the Mécanique analytique. *Lettera Matematica*, 2(1):47–54, June 2014.
- [3] Herman Erlichson. Johann Bernoulli's brachistochrone solution .using Fermat's principle of least time. *European Journal of Physics*, 20(5):299–304, July 1999.
- [4] René Dugas. *A History of Mechanics*. Dover Publications, November 2012.
- [5] Cornelius Lanczos. *The Variational Principles of Mechanics*. Dover Publications, 4th revised ed. édition édition, April 2012.
- [6] L. D. Landau and E. M. Lifshitz. *Mechanics: Volume 1*. Butterworth-Heinemann, Amsterdam u.a, 3e édition édition, January 1976.
- [7] Herbert Goldstein. *Classical Mechanics*. Pearson, Harlow, 3e édition édition, August 2013.
- [8] M G Calkin. *Lagrangian and Hamiltonian Mechanics*. WORLD SCIENTIFIC, July 1996.
- [9] Louis N. Hand and Janet D. Finch. *Analytical Mechanics*. Cambridge University Press, Cambridge, 1998.
- [10] Bahram Houchmandzadeh. The Hamilton–Jacobi equation: An alternative approach. *American Journal of Physics*, 88(5):353–359, April 2020.
- [11] R. K. P. Zia, Edward F. Redish, and Susan R. McKay. Making sense of the Legendre transform. *American Journal of Physics*, 77(7):614–622, June 2009.
- [12] V. I. Arnol'd, K. Vogtmann, and A. Weinstein. *Mathematical Methods of Classical Mechanics*. Springer-Verlag New York Inc., New York, 2nd ed. 1989. corr. 4th printing 1997 édition édition, September 1997.
- [13] Henri Cartan. *Differential Forms*. Dover Publications, illustrated édition édition, July 2012.
- [14] Harley Flanders. *Differential Forms with Applications to the Physical Sciences*. Dover Publications, revised ed. édition édition, April 2012.
- [15] Victor J. Katz. The history of differential forms from Clairaut to Poincaré. *Historia Mathematica*, 8(2):161–188, May 1981.
- [16] Harold M. Edwards. *Advanced Calculus: A Differential Forms Approach*. Birkhauser Verlag AG, Boston, 3e édition édition, November 1993.
- [17] Edward A. Desloge and Robert I. Karch. Noether's theorem in classical mechanics. *American Journal of Physics*, 45(4):336–339, April 1977.