

# The Dynamics of Lagrange and Hamilton

Nisheeth K. Vishnoi

September 20, 2018

## Abstract

In 1788, Lagrange presented a set of equations of motion that, unlike Newtonian mechanics, are independent of the choice of coordinates of the physical system and ultimately led to the formulation of general relativity. Hamilton came up with a different set of equations of motion in 1833 that arguably led to the development of quantum mechanics. Remarkably, in classical mechanics, these sets of equations turn out to be equivalent via a beautiful duality due to Legendre. The goal of this note is to present Lagrangian and Hamiltonian dynamics, and the connection between them. This material presented here is well-known in the physics community and is primarily written for a CS/optimization audience.

## 1 The phase space

When studying a physical system, one is typically interested in the behavior of its constituent particles. The following notation is a useful way to think about particles.

**Definition 1.1 (Particle).** A particle in a  $d$ -dimensional Euclidean space consists of a  $d$ -dimensional *position* vector  $q \in \mathbb{R}^d$  and a  $d$ -dimensional *momentum* vector  $p \in \mathbb{R}^d$ .

Thus, a particle is simply a point  $(q, p)$  in  $\mathbb{R}^d \times \mathbb{R}^d$ , called its *phase space*. If there are  $k$  particles, then the phase space consists of  $(q_1, p_1, q_2, p_2, \dots, q_k, p_k)$ , each component lying in  $\mathbb{R}^d$ . As these particles move due to the presence of various forces or fields, the  $p_i$ s and  $q_i$ s vary with time; understanding their dynamics is the essence of classical mechanics.

## 2 Newtonian dynamics

Newtonian dynamics describes the time-evolution of particles in an *inertial frame of reference* – a frame of reference in classical physics in which a body with zero net force acting upon it is not accelerating. The momentum is simply proportional to the time derivative of the position, and Newton’s Second Law of Motion relates the time derivative of momentum to its position. Formally, Newtonian Dynamics for a particle in a potential well described by  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  can be described using the following equations, for a particle  $(q(t), p(t))$  moving in a Cartesian coordinate system

$$\begin{aligned} m \frac{dq}{dt} &= p, \\ \frac{dp}{dt} &= F(q) \\ &= -\nabla V(q), \end{aligned}$$

where  $m$  is the mass of the particle,  $F(q)$  is the force applied at position  $q$  and  $V(q)$  is the potential at position  $q$ .

Newton’s equations give a general and powerful tool to understand how particles move. However, in complex physical systems, they might be hard to apply. In fact, identifying all the

forces in a physical system is often nontrivial. Furthermore, the assumption that the physical system is in an inertial reference frame often causes problems as certain phenomena cannot be explained purely by Newtonian dynamics. As an example one can consider the Coriolis force which arises because of the rotation of earth around its axis. It causes, for instance, rivers (on the northern hemisphere) running due north to erode more strongly along the right bank. This phenomenon is paradoxical when studied with respect to the earth as a reference frame, and to be explained has to be “looked at from a distance” where the rotational movement of earth is apparent.

Another famous example is the Foucault pendulum which is a simple physical system in which a mass is attached to a long cable and swings in a plane. If the pendulum is not at the equator, then after a while one can observe that the plane in which the pendulum swings changes. Newtonian dynamics cannot explain this phenomenon without introducing an external force due to the rotation of the earth.

Furthermore, Newtonian dynamics are written in terms of the Cartesian coordinate system and are not invariant under change of variables. Simply going from Cartesian to polar coordinates can completely change the form of Newton’s equations and requires a rederivation. As we will see, this does not occur when using Lagrangian dynamics.

### 3 Lagrangian dynamics

In Lagrangian dynamics one typically denotes the position of a particle by  $q$  and its (generalized) velocity by  $\dot{q}$ , as it is normally a time derivative of the position. In most cases one should still think of  $\dot{q}$  as a formal variable independent of  $q$ , but if we consider a trajectory  $(q(t), \dot{q}(t))$  of a particle, then clearly

$$\frac{d}{dt}q(t) = \dot{q}(t).$$

The central object of Lagrangian dynamics is the *Lagrangian* that defined as

$$L(q_1, \dots, q_d, \dot{q}_1, \dots, \dot{q}_d, t) = K(q_1, \dots, q_d, \dot{q}_1, \dots, \dot{q}_d, t) - V(q_1, \dots, q_d, t),$$

where  $K$  is the kinetic energy function and  $V$  is the potential energy function. An example of a Lagrangian in 1-dimension is

$$L = \frac{1}{2}m\dot{q}^2 - V(q). \quad (1)$$

The Lagrangian can also depend on time, for instance if  $\beta < 0$  is a fixed constant, then one can consider the following Lagrangian

$$L = e^{\beta t} \left( \frac{1}{2}m\dot{q}^2 - V(q) \right),$$

which describes a physical system with friction.

We now present the Lagrange dynamics.

**Definition 3.1.** The Lagrange dynamics describing the motion of a particle with coordinates  $(q, \dot{q})$  with respect to a Lagrangian  $L(q, \dot{q}, t)$  are given by

$$\frac{\partial L}{\partial q_i} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \quad i = 1, 2, \dots, d. \quad (2)$$

We can see, for instance, that the Lagrange dynamics for the Lagrangian in (1) are compatible with Newtonian dynamics.

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{d}{dt}(m\dot{q}) = \frac{d}{dt}p = -\nabla V(q) = \frac{\partial L}{\partial q}.$$

The quantity  $\frac{\partial L}{\partial q_i}$  is usually referred to as conjugate momenta or generalized momenta, while  $\frac{\partial L}{\partial \dot{q}_i}$  is called the generalized force.

## 4 Deriving Lagrangian dynamics via calculus of variations

We now derive the Lagrangian dynamics (2) using the calculus of variations, a theory that is attributed to Euler and Lagrange.<sup>1</sup> The key is to define an *action integral* which in this case is the quantity

$$S(q) := \int_a^b L(q, \dot{q}, t) dt,$$

that can be interpreted as the action or work performed by the particle from time  $a$  to  $b$ . Now, the *Principle of Least Action* asserts that the trajectory taken by the particle to go from state  $(q(a), \dot{q}(a))$  to the state  $(q(b), \dot{q}(b))$  should minimize  $S(q)$ . Thus, in order to characterize the motion of such a particle we need to find out which curves minimize  $S(q)$ .

**Theorem 4.1 (Euler-Lagrange Equations).** Let  $q$  and  $\dot{q}$  be functions  $\mathbb{R} \rightarrow \mathbb{R}^d$  (depending on  $t$ ) with  $\dot{q} = \frac{dq}{dt}$ . If  $q$  is a stationary point of

$$S(q) = \int_a^b L(q, \dot{q}, t) dt,$$

then

$$\frac{\partial L}{\partial q_i} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \quad i = 1, 2, \dots, d.$$

*Proof.* Equivalently, we need to show that  $\frac{\partial S}{\partial q}$  and  $\frac{\partial S}{\partial \dot{q}}$  should be 0. Here, the derivatives are taken with respect to curves. In optimization this is tantamount to establishing first order optimality conditions. Let us consider a curve  $y$ , such that  $y(a) = 0$  and  $y(b) = 0$ . Think of this curve as an infinitesimal perturbation to a fixed curve whose optimality we are trying to establish (see Figure 1 for an illustration).

If  $(q, \dot{q})$  is a stationary point of  $S$ , then the directional derivative  $D_y S$  should be 0.

$$0 = D_y S = \lim_{h \rightarrow 0} \frac{1}{h} \int_a^b L(q + hy, \dot{q} + h\dot{y}, t) - L(q, \dot{q}, t) dt.$$

The directional derivative with respect to a curve is equivalent to taking an infinitesimal step in the direction of the curve. We can rewrite the right hand side using the first order Taylor approximation as follows:

$$\begin{aligned} D_y S &= \lim_{h \rightarrow 0} \frac{1}{h} \int_a^b L(q, \dot{q}, t) + hy \frac{\partial}{\partial q} L(q + hy, \dot{q} + h\dot{y}, t) + h\dot{y} \frac{\partial}{\partial \dot{q}} L(q + hy, \dot{q} + h\dot{y}, t) \\ &\quad + O(h^2) - L(q, \dot{q}, t) dt \\ &= \int_a^b \lim_{h \rightarrow 0} \frac{1}{h} \left( hy \frac{\partial}{\partial q} L(q + hy, \dot{q} + h\dot{y}, t) + h\dot{y} \frac{\partial}{\partial \dot{q}} L(q + hy, \dot{q} + h\dot{y}, t) + O(h^2) \right) dt \\ &= \int_a^b \lim_{h \rightarrow 0} y \frac{\partial}{\partial q} L(q + hy, \dot{q} + h\dot{y}, t) + \lim_{h \rightarrow 0} \dot{y} \frac{\partial}{\partial \dot{q}} L(q + hy, \dot{q} + h\dot{y}, t) + \lim_{h \rightarrow 0} O(h) dt \\ &= \int_a^b y \frac{\partial}{\partial q} L(q, \dot{q}, t) + \dot{y} \frac{\partial}{\partial \dot{q}} L(q, \dot{q}, t) dt. \end{aligned}$$

By integrating by parts, we obtain

$$D_y S = \int_a^b y \frac{\partial L(q, \dot{q}, t)}{\partial q} dt + \left( y \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}} \right) \Big|_a^b - \int_a^b y \frac{d}{dt} \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}} dt.$$

---

<sup>1</sup>The Lagrangian dynamics are also often referred to as Euler-Lagrange equations.

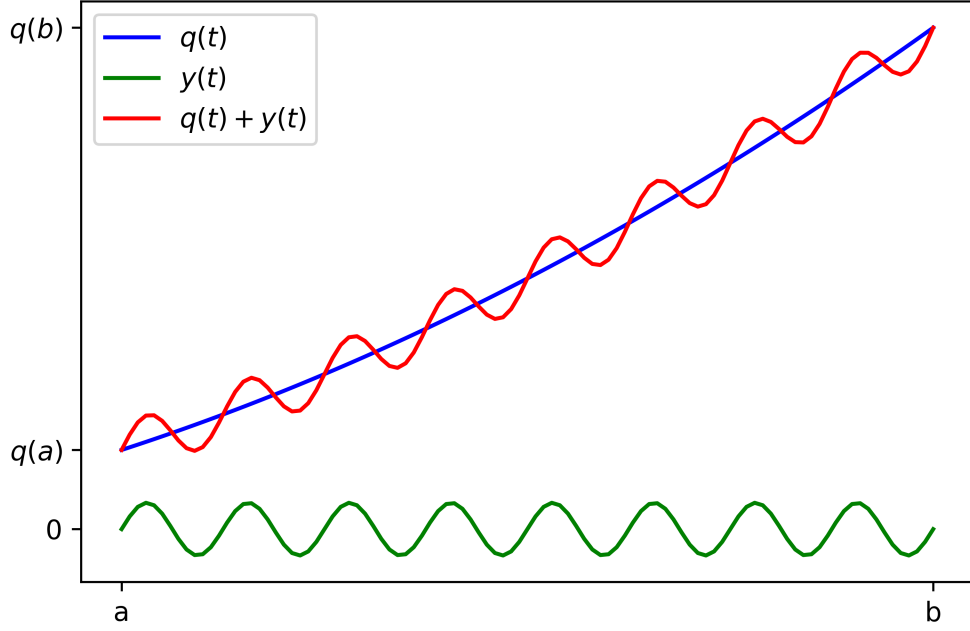


Figure 1: A curve  $q(t)$  and its perturbation  $q(t) + y(t)$ .

Further,  $y(a) = y(b) = 0$  implies that

$$\left( y \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}} \right) \Big|_a^b = 0.$$

Therefore, for any curve  $y$  such that  $y(a) = 0$  and  $y(b) = 0$ , we have

$$\int_a^b y \left( \frac{\partial L(q, \dot{q}, t)}{\partial q} - \frac{d}{dt} \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}} \right) dt = 0.$$

This completes the proof.  $\square$

## 5 Invariance of Lagrangian dynamics under coordinate transformations

Now we show that the Lagrangian dynamics are invariant with respect to coordinate changes. Let us first try to explain what this invariance means. Let  $q = (q_1, \dots, q_d) \in \mathbb{R}^d$  be the position coordinates and  $\dot{q} = (\dot{q}_1, \dots, \dot{q}_d) \in \mathbb{R}^d$  be the generalized velocities which we use to write the Lagrangian  $L(q, \dot{q}, t)$ . Consider now a change of coordinates  $x = x(q)$  where  $q$  is a continuous and differentiable map  $q : \mathbb{R}^d \rightarrow \mathbb{R}^d$ . We would like to show that we can write the Lagrangian in terms of the new variables  $(x, \dot{x}) \in \mathbb{R}^{2d}$  (denote it by  $\tilde{L}$ ), so that

$$L(q, \dot{q}, t) = \tilde{L}(x, \dot{x}, t).$$

While it is clear what  $x$  means in the above equation (simply  $q = q(x)$ ) we now explain how  $\dot{x}$  arises.

We would like to map every curve  $(q, \dot{q}) : \mathbb{R} \rightarrow \mathbb{R}^{2d}$  that satisfies

$$\frac{d}{dt}q(t) = \dot{q}(t)$$

to  $(x, \dot{x}) : \mathbb{R} \rightarrow \mathbb{R}^{2d}$  (with  $x(t) = x(q(t))$ ) such that

$$\frac{d}{dt}x(t) = \dot{x}(t).$$

Note that for every  $i \in [d]$

$$\frac{d}{dt}x_i(q) = \sum_{j=1}^d \frac{\partial x_i}{\partial q_j} \frac{dq_j}{dt} = \sum_{j=1}^d \frac{\partial x_i}{\partial q_j} \dot{q}_j.$$

Thus, the coordinate map  $x = x(q)$  uniquely determines the velocity vector  $\dot{x}$  which is given by

$$\dot{x} = J(q)\dot{q}$$

where  $J(q) \in \mathbb{R}^{d \times d}$  is the Jacobian matrix of the transformation  $x = x(q)$  at a point  $q$ . Thus, the transformation of configuration spaces is given by

$$(q, \dot{q}) \mapsto (x, \dot{x}) = (x(q), J(q)\dot{q}). \quad (3)$$

Note that for the above to be well defined we assume in addition that the Jacobian matrix  $J(q)$  is invertible at every point  $q \in \mathbb{R}^d$ , otherwise the above map is not a bijection.

Now we can define the Lagrangian  $\tilde{L}$  in the  $x$  coordinates as:

$$\tilde{L}(x, \dot{x}, t) = L(q, \dot{q}, t), \quad (4)$$

where  $(x, \dot{x})$  and  $(q, \dot{q})$  are related as in (3). The coordinate invariance theorem can now be stated.

**Theorem 5.1 (Invariance of Lagrangian dynamics).** Let  $x = x(q)$  be a coordinate change inducing a reparametrization of the configuration space  $(q, \dot{q}) \mapsto (x, \dot{x})$  as in (3). Let  $L(q, \dot{q}, t)$  be a Lagrangian and  $\tilde{L}$  be the corresponding Lagrangian written in the  $x$  coordinates (as in (4)). Then for every trajectory  $(q, \dot{q}) : \mathbb{R} \rightarrow \mathbb{R}^{2d}$  satisfying  $\frac{d}{dt}q = \dot{q}$  and the Euler-Lagrange equation  $\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}$  (for all  $i \in [d]$ ), the corresponding trajectory  $(x, \dot{x}) : \mathbb{R} \rightarrow \mathbb{R}^{2d}$  satisfies the Euler-Lagrange equation with respect to  $\tilde{L}$ :

$$\text{for all } i \in [d], \quad \frac{d}{dt} \frac{\partial \tilde{L}}{\partial \dot{x}_i} = \frac{\partial \tilde{L}}{\partial x_i}.$$

*Proof.* Start by writing (using the chain rule)

$$\frac{\partial \tilde{L}}{\partial \dot{x}_i} = \sum_{j=1}^d \frac{\partial L}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial \dot{x}_i} + \sum_{j=1}^d \frac{\partial L}{\partial q_j} \frac{\partial q_j}{\partial \dot{x}_i} = \sum_{j=1}^d \frac{\partial L}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial \dot{x}_i}. \quad (5)$$

The last equality follows from the fact that  $q$  does not depend on  $\dot{x}$  (it only depends on  $x$ ) and, hence,  $\frac{\partial q_j}{\partial \dot{x}_i} = 0$  for any  $i, j \in [d]$ . Moreover, as  $\dot{q}_j = \sum_{k=1}^d \frac{\partial q_j}{\partial x_k} \dot{x}_k$  this simplifies to

$$\frac{\partial \tilde{L}}{\partial \dot{x}_i} = \sum_{j=1}^d \frac{\partial L}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial \dot{x}_i}.$$

Further, by taking the time derivative of the above we obtain

$$\begin{aligned} \frac{d}{dt} \left( \sum_{j=1}^d \frac{\partial L}{\partial \dot{q}_j} \frac{\partial q_j}{\partial x_i} \right) &= \sum_{j=1}^d \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} \right) \frac{\partial q_j}{\partial x_i} + \sum_{j=1}^d \sum_{k=1}^d \frac{\partial L}{\partial \dot{q}_j} \frac{\partial^2 q_j}{\partial x_i \partial x_k} \frac{\partial x_k}{\partial t} \\ &= \sum_{j=1}^d \frac{\partial L}{\partial q_j} \frac{\partial q_j}{\partial x_i} + \sum_{j=1}^d \sum_{k=1}^d \frac{\partial L}{\partial \dot{q}_j} \frac{\partial^2 q_j}{\partial x_i \partial x_k} \frac{\partial x_k}{\partial t} \end{aligned} \quad (6)$$

Here, we have used the Leibniz rule, the chain rule, and the Lagrangian dynamics for  $q, \dot{q}$ . Similarly, we compute the derivative of  $\tilde{L}$  with respect to  $x_i$ :

$$\frac{\partial \tilde{L}}{\partial x_i} = \sum_{j=1}^d \frac{\partial L}{\partial q_j} \frac{\partial q_j}{\partial x_i} + \sum_{j=1}^d \frac{\partial L}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial x_i}. \quad (7)$$

Here, none of the terms simplify as both  $q$  and  $\dot{q}$  depend on  $x$ . To arrive at a similar form as in (6) we consider the term  $\frac{\partial \dot{q}_j}{\partial x_i}$  for  $i, j \in [d]$ , and using the Leibniz rule, obtain that:

$$\frac{\partial \dot{q}_j}{\partial x_i} = \frac{\partial}{\partial x_i} \left[ \sum_{k=1}^d \frac{\partial q_j}{\partial x_k} \dot{x}_k \right] = \sum_{k=1}^d \frac{\partial^2 q_j}{\partial x_i \partial x_k} \dot{x}_k. \quad (8)$$

Combining this with (7) we obtain

$$\frac{\partial \tilde{L}}{\partial x_i} = \sum_{j=1}^d \frac{\partial L}{\partial q_j} \frac{\partial q_j}{\partial x_i} + \sum_{j=1}^d \sum_{k=1}^d \frac{\partial L}{\partial \dot{q}_j} \frac{\partial^2 q_j}{\partial x_i \partial x_k} \dot{x}_k. \quad (9)$$

Hence, using (5), (6) and (9) the theorem follows.  $\square$

## 6 Hamiltonian dynamics

Imagine a particle of unit mass in a potential well  $V : \mathbb{R}^d \rightarrow \mathbb{R}$ . If the particle has position  $q \in \mathbb{R}^d$  and momentum  $p \in \mathbb{R}^d$ , its total energy is given by the function  $H : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ , where

$$H(q, p) = V(q) + \frac{1}{2} \|p\|^2.$$

This function is called the Hamiltonian of the particle. Hamiltonian dynamics for this particle are:

$$\frac{dq}{dt} = p \quad \text{and} \quad \frac{dp}{dt} = -\nabla V(q).$$

We denote the solutions to these equations by  $(q(t), p(t))$  for  $t \geq 0$ . Interestingly, these solutions satisfy a number of conservation properties. In particular, they conserve the Hamiltonian  $H(q(t), p(t)) = H(q(0), p(0))$  for all  $t$ . The underlying geometry and conservation laws have been generalized significantly in physics and have led to the area of symplectic geometry in mathematics.

## 7 The duality between Lagrangian and Hamiltonian dynamics

Finally, we show that Hamiltonian dynamics are related to the Lagrangian dynamics via Legendre duality.<sup>2</sup> Let us compute the Legendre dual  $H(p, q, t)$  of the Lagrangian  $L(q, \dot{q}, t)$  and denote it by

$$\tilde{H}(q, p, t) := \max_{\dot{q}} \langle \dot{q}, p \rangle - L(q, \dot{q}, t).$$

---

<sup>2</sup>This duality is also often referred to as Legendre-Fenchel duality.

If we compute the partial derivative of the right hand side with respect to  $\dot{q}$ , we obtain

$$p - \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}}.$$

Consequently,  $\dot{q}$  maximizes the right hand side when

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad \text{for } i = 1, \dots, d.$$

Although  $\tilde{H}$  function does not directly depend on  $\dot{q}$ , still  $\dot{q}$  appears in it implicitly:

$$\tilde{H}(q, p, t) = \sum_{i=1}^d \dot{q}_i p_i - L(q, \dot{q}, t), \quad (10)$$

where  $\dot{q}$  is the solution of

$$p_i = \frac{\partial L}{\partial \dot{q}_i}.$$

Let us compute total derivative of  $\tilde{H}$ . We can compute it in two ways: either using its definition or the alternative formulation in (10). The first way leads to,

$$d\tilde{H} = \sum_{i=1}^d \frac{\partial \tilde{H}}{\partial p_i} dp_i + \sum_{i=1}^d \frac{\partial \tilde{H}}{\partial q_i} dq_i + \frac{\partial \tilde{H}}{\partial t}, \quad (11)$$

and the second way leads to

$$\begin{aligned} d\tilde{H} &= \sum_{i=1}^d p_i d\dot{q}_i + \sum_{i=1}^d \dot{q}_i dp_i - \sum_{i=1}^d \frac{\partial L}{\partial q_i} dq_i - \sum_{i=1}^d \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial t} \\ &= \sum_{i=1}^d \left( \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \dot{q}_i dp_i - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} dq_i - \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i \right) - \frac{\partial L}{\partial t} \\ &= \sum_{i=1}^d \left( \dot{q}_i dp_i - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} dq_i \right) - \frac{\partial L}{\partial t} \\ &= \sum_{i=1}^d \left( \dot{q}_i dp_i - \frac{d}{dt} p_i dq_i \right) - \frac{\partial L}{\partial t} \\ &= \sum_{i=1}^d (\dot{q}_i dp_i - \dot{p}_i dq_i) - \frac{\partial L}{\partial t}. \end{aligned} \quad (12)$$

By comparing the two equations ((11) and (12)), we obtain

$$\begin{aligned} \frac{\partial \tilde{H}}{\partial p_i} &= \dot{q}_i, \\ -\frac{\partial \tilde{H}}{\partial q_i} &= \dot{p}_i, \\ \frac{\partial \tilde{H}}{\partial t} &= -\frac{\partial L}{\partial t}. \end{aligned} \quad (13)$$

These equations are the same as Hamiltonian dynamics if we identify  $H$  and  $\tilde{H}$ . Thus, the Hamiltonian is nothing but the Legendre dual of the Lagrangian and using this duality we can derive Hamiltonian dynamics from Lagrangian dynamics and vice-versa.