

The Legendre Transform

Abstract

This is a deep dive in to the meaning of the Legendre transform.

Index Terms

Calculus of variations, Lagrangian, Hamiltonian, Legendre

I. INTRODUCTION

The Legendre transform appears in various topics in a typical physics curriculum:

- In classical mechanics when transitioning from the Lagrangian description to the Hamiltonian description,
- In statistical thermodynamics when relating various quantities.

It is typically introduced in a single line of equation without providing much of a careful consideration. Here we discuss the meaning of the Legendre transform and its geometric interpretation. We will limit the discussion to classical mechanics, but the observations made here can be transferred to quantum field theory and thermodynamics. In order to derive the equations of motion for classical mechanics in the Lagrangian formalism, we will first use calculus of variations to minimize the action functional. We will then use the Legendre transform to derive Hamiltonian description. We finally circle back to dive deeper into the meaning of the Legendre transform.

II. LAGRANGIAN AND HAMILTONIAN MECHANICS

A functional can be considered as an operation that takes in a function and returns a number. The most familiar functional is integration with fixed limits. It takes in f and returns $\mathcal{S} = \int_a^b f(t)dt$, which is just a number. In a typical mechanics problem, the functional \mathcal{S} will be of the form:

$$\mathcal{S} = \int_{t_0}^{t_1} \mathcal{L}(q, \dot{q}) dt, \quad (1)$$

where \mathcal{L} is the Lagrangian, and $q = q(t)$ is the generalized coordinate with $\dot{q} = \frac{dq}{dt}$. Let's assume that we have a function $q(t)$ that gives the minimum value for \mathcal{S} . If we fiddle q around the optimal function by a small amount $\alpha\eta(t)$, i.e., $q(t) \rightarrow q(t) + \alpha\eta(t)$, where $\eta(t)$ is an arbitrary function and α is a small number, then the change in \mathcal{S} should be 0. This is analogous to requiring that the derivative of a function f should vanish at a local extremum, that is: $\frac{df(t)}{dt}|_{t=t^*} = 0$. Rigorously speaking [1], we can define the following functional

$$\mathcal{S}(\alpha) = \int_{t_0}^{t_1} \mathcal{L}(q + \alpha\eta, \dot{q} + \alpha\dot{\eta}) dt, \quad (2)$$

and require that

$$\left. \frac{d}{d\alpha} \mathcal{S}(\alpha) \right|_{\alpha=0} = 0. \quad (3)$$

Consider a problem where the end points are specified. This implies that we are not free to wiggle q at the end points t_0 and t_1 , i.e.,

$$\eta(t_0) = \eta(t_1) = 0. \quad (4)$$

The variation is illustrated in Fig. 1.

email: quarktetra@gmail.com

Find the interactive HTML-document here.

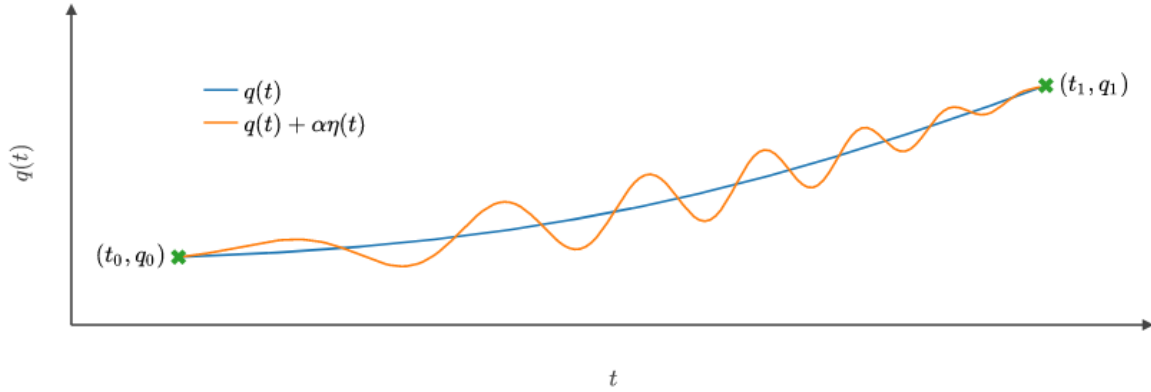


Figure 1: The orange curve $q(t)$, which is unknown at the moment, gives the minimum value for the functional \mathcal{L} . The green curve represents the new curve with random deformations around $q(t)$. The variation $\eta(t)$ must vanish at the end points since the values of q are fixed at these points.

Keeping the boundary conditions in Eq. (4) in mind, let us calculate Eq. (3):

$$\begin{aligned}
\left. \frac{d}{d\alpha} \mathcal{L}(\alpha) \right|_{\alpha=0} &= \int_{t_0}^{t_1} \left. \frac{d}{d\alpha} \mathcal{L}(q + \alpha\eta, \dot{q} + \alpha\dot{\eta}(t)) \right|_{\alpha=0} dt = \int_{t_0}^{t_1} \left[\frac{\partial}{\partial q} \mathcal{L}(q, \dot{q}) \eta + \frac{\partial}{\partial \dot{q}} \mathcal{L}(q, \dot{q}) \frac{d\eta}{dt} \right] dt \\
&= \int_{t_0}^{t_1} \left[\frac{\partial}{\partial q} \mathcal{L}(q, \dot{q}) \eta + \frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}} \mathcal{L}(q, \dot{q}) \eta \right) - \frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}} \mathcal{L}(q, \dot{q}) \right) \eta \right] dt \\
&= \int_{t_0}^{t_1} \left[\frac{\partial \mathcal{L}(q, \dot{q})}{\partial q} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}(q, \dot{q})}{\partial \dot{q}} \right) \right] \eta dt + \cancel{\frac{\partial \mathcal{L}(q, \dot{q})}{\partial \dot{q}} \eta \Big|_{t_0}^{t_1}} \\
&= \int_{t_0}^{t_1} \left[\frac{\partial \mathcal{L}(q, \dot{q})}{\partial q} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}(q, \dot{q})}{\partial \dot{q}} \right) \right] \eta dt, \tag{5}
\end{aligned}$$

where the boundary terms vanish due to the constraints in Eq. (4). Since η is an arbitrary function, in order to set this equation to 0, we require the following:

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) = 0, \tag{6}$$

which is known as the Euler-Lagrange equation.

As it is typically done in physics classes, we will first pull the definition of Legendre transform out of a hat.¹ In order to do that, we first define the conjugate momenta p as

$$p \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}}, \tag{7}$$

and the Legendre transform as

$$\mathcal{H}(q, p) = p\dot{q} - \mathcal{L}(q, \dot{q}), \tag{8}$$

which will enable us to move from the independent variables $\{q, \dot{q}\}$ to $\{q, p\}$. We can now compute the differential of this new quantity \mathcal{H} by expanding out the right hand side as

$$\begin{aligned}
d\mathcal{H}(q, p) &= dp \dot{q} + p \frac{\partial \dot{q}}{\partial p} dp + p \frac{\partial \dot{q}}{\partial q} dq - \frac{\partial \mathcal{L}}{\partial q} dq - \frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{\partial \dot{q}}{\partial p} dp - \frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{\partial \dot{q}}{\partial q} dq \\
&= dp \left[\dot{q} + \frac{\partial \dot{q}}{\partial p} \left(p - \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \right] + dq \left[-\frac{\partial \mathcal{L}}{\partial q} - \frac{\partial \dot{q}}{\partial q} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} - p \right) \right], \tag{9}
\end{aligned}$$

¹I took several classes in which Legendre transform was used, and every time it was pulled out of thin air with little or no motivation. However, there are certainly exceptions, see Leonard Susskind's class[2]

where the terms in the parenthesis are zero due to the definition in Eq. (7) . Therefore we get:

$$d\mathcal{H}(q, p) = dp\dot{q} - dq\frac{\partial\mathcal{L}}{\partial q} = dp\dot{q} - dq\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}}\right) = dp\dot{q} - dq\dot{p}. \quad (10)$$

We can also write the $d\mathcal{H}(q, p)$ in terms of its functional arguments:

$$d\mathcal{H}(q, p) = dq\frac{\partial\mathcal{H}}{\partial q} + dp\frac{\partial\mathcal{H}}{\partial p}. \quad (11)$$

Matching the coefficients of the differentials in Eqs. (10) and (11), we arrive at the Hamiltonian equations of motions:

$$\dot{q} = \frac{\partial\mathcal{H}}{\partial p}, \text{ and } \dot{p} = -\frac{\partial\mathcal{H}}{\partial q}, \quad (12)$$

and this is how one moves from the Lagrangian equations to Hamiltonian equations via the Legendre transform.

III. WHAT IS THE LEGENDRE TRANSFORM?

The Legendre transform can be interpreted as a mapping between two different encodings of a function [3]. Conceptually, it is similar to the Fourier-pair functions that encode a function in the x (length) domain or in the k (wave-number) domain:

$$\{f, x\} \iff \{F, k\}, \quad (13)$$

with the explicit transformation rule as:

$$F(k) = \int dx e^{ikx} f(x). \quad (14)$$

The transformed function F operates in the wave-number domain k , but it still encodes the same information as f .

To discuss the Legendre transform, let us consider a function $\mathcal{L}(v)$, where we labeled the argument as v to make it easier to relate it back to the classical mechanics case we discussed earlier (it will be apparent later that $v = \dot{q}$). The Legendre transform maps the original function $\mathcal{L}(v)$ to a new one which takes $\mathcal{L}'(v) = \frac{d\mathcal{L}}{dv}$ as the argument, instead of v . The original parameter v is traded with the slope of the function, $\mathcal{L}'(v)$. One can rather quickly see that this will be possible only if there is one to one mapping between $\mathcal{L}'(v)$ and v , i.e., given the value of $\mathcal{L}'(v)$, if we can invert it to get v , then we can simply use $\mathcal{L}'(v)$ as the argument of the function. We can do the inversion provided that the function we are dealing with is convex, i.e., the second derivative is always positive and smooth. As we will switch from v to $\mathcal{L}'(v)$, it is convenient to define this derivative as a new function:

$$p(v) \equiv \frac{\partial\mathcal{L}}{\partial v}, \quad (15)$$

where we use partial derivatives for reservations for functions with multiple arguments. The other arguments are not explicitly shown at the moment since they will not enter into the Legendre transform. With this definition, we are equipped to study the actual meaning of the transformation. Before we do that, it is important to emphasize that there is only one independent variable here: you can either choose v to be the independent one, which will completely fix the value of p as $p(v)$, or you can decide to use p as the independent one, which sets $v = v(p)$.

A. Geometric derivation using slopes

Consider the set up in Fig. 2 where we take a function $\mathcal{L}(v)$ and draw a tangent line to it at a value of v , which has a slope $p = \mathcal{L}'(v)$ [3]. The height of the triangle can be computed as the slope multiplied by the base length.

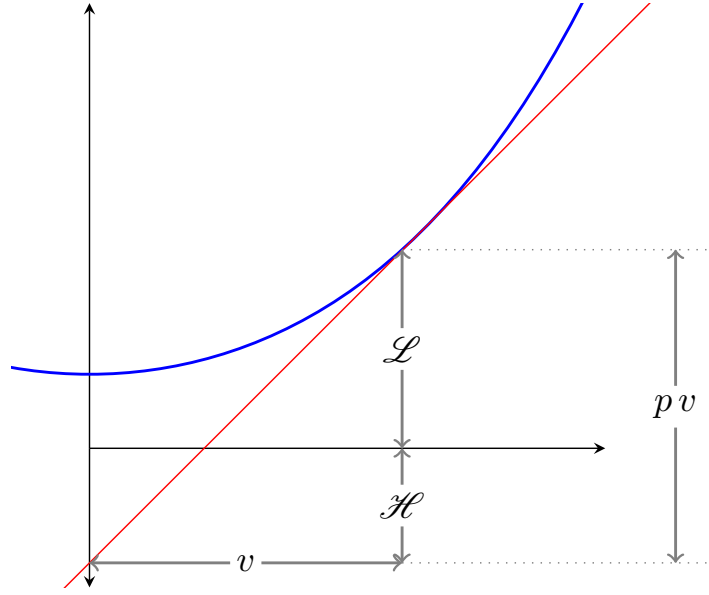


Figure 2: The geometric interpretation of the Legendre transform. The blue curve is the function of interest, $\mathcal{L}(v)$. The red line is the derivative at point v , which we called $p(v)$. From the geometry, we have $\mathcal{L}(v) + \mathcal{H}(p) = pv$.

As seen from the geometry, the original function \mathcal{L} and the transformed function \mathcal{H} add up to pv . Therefore we can define the transformation mapping as

$$\{\mathcal{L}, v\} \iff \{\mathcal{H}, p\}, \quad (16)$$

with the explicit transformation rule as:

$$\mathcal{L}(v) + \mathcal{H}(p) = pv. \quad (17)$$

B. Geometric derivation using areas

Since there is a one to one map between v and p , we can plot the functions $p(v)$ or $v(p)$, and construct Fig. 3 [2].

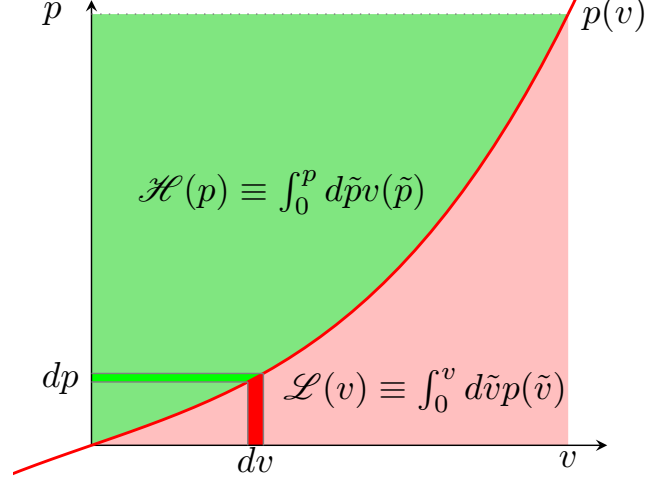


Figure 3: The geometric interpretation of the Legendre transform in terms of the areas. The red line is $p(v)$. Since there is a one to one map between v and p , we can compute the areas using horizontal slices or vertical slices. Since the areas add up to the area of the rectangle, we have $\mathcal{L}(v) + \mathcal{H}(p) = pv$.

The shaded areas can be calculated using horizontal slices or vertical slices as follows:

$$\mathcal{L}(v) \equiv \int_0^v d\tilde{v}p(\tilde{v}), \quad \text{and} \quad \mathcal{H}(p) \equiv \int_0^p d\tilde{p}v(\tilde{p}). \quad (18)$$

As the areas add up to the area of the rectangle, we have

$$\mathcal{L}(v) + \mathcal{H}(p) = pv, \quad (19)$$

which is nothing but the definition of the Legendre transform. Furthermore, we can take the derivatives of the functions in Eq. (18) with respect to their arguments to get:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial v} &= \frac{\partial}{\partial v} \left(\int_0^v d\tilde{v}p(\tilde{v}) \right) = p(v), \\ \frac{\partial \mathcal{H}}{\partial p} &= \frac{\partial}{\partial p} \left(\int_0^p d\tilde{p}v(\tilde{p}) \right) = v(p) \end{aligned} \quad (20)$$

which recovers the original definition of p in Eq. (15) and the first part of the Hamiltonian equations of motion in Eq. (12) with $v = \dot{q}$.

If we want to recover the other part of the Hamiltonian equations, we need to expose the other argument of \mathcal{L} and \mathcal{H} , that is $\mathcal{L} = \mathcal{L}(v, q)$ and $\mathcal{H} = \mathcal{H}(p, q)$. But note that q is a totally independent variable. Taking the derivative of Eq. (19) with respect to q , we get:

$$\frac{\partial \mathcal{L}(v, q)}{\partial q} + \frac{\partial \mathcal{H}(p, q)}{\partial q} = \frac{\partial (pv)}{\partial q} = 0 \implies \frac{\partial \mathcal{H}(p, q)}{\partial q} = -\frac{\partial \mathcal{L}(v, q)}{\partial q} = -\dot{p}, \quad (21)$$

where we used Euler's equation of motion in Eq. (6) to convert $\frac{\partial \mathcal{L}(v, q)}{\partial q}$ to \dot{p} .

IV. INVERSE LEGENDRE TRANSFORM

Ever heard of the inverse Legendre transform? Probably not. That is because there is no need to define the inverse transform since it is its own inverse! The definition of the transform in Eq. (19) manifestly shows this: the original function and the transformed one are added to give pv . In other words, we can swap \mathcal{L} with \mathcal{H} and v with p , and nothing will change. Equivalently, we can try to transform $\mathcal{H}(p)$ one more time. Let's do that. Remember that we trade the argument of the function with the derivative of the function:

$$u(p) = \frac{\partial \mathcal{H}}{\partial p}, \quad (22)$$

and define the transformed function

$$\mathcal{H}_2(u) = up - \mathcal{H}(p). \quad (23)$$

However, from the Hamiltonian equations of motion, we already know that $u(p) = \frac{\partial \mathcal{H}}{\partial p} = v$. Plugging this back in gives

$$\mathcal{H}_2(v) = vp - \mathcal{H}(p), \quad (24)$$

and we explicitly see that $\mathcal{H}_2(v) = \mathcal{L}(v)$, i.e., transforming for the second time undoes the first one and returns back the original function \mathcal{L} .

V. EXTENDED LAGRANGIAN AND HAMILTONIAN

We have carefully stressed that there is only one independent variable in the game: p or v (ignoring the spectator variable q). The relation between p and v is dictated by physics, such as $p = mv$ for a non-relativistic particle or $p = \frac{mv}{\sqrt{1-v^2/c^2}}$ for the relativistic one, where m is the mass of the particle, and c is the speed of light. However, we can still let v and p go *off-shell* and be independent of each other [4]. This formalism extends the Lagrangian and Hamiltonian into a larger phase-space. The extended functions are defined as [5]

$$\begin{aligned} \mathcal{H}_e(p, \dot{q}, q) &\equiv p\dot{q} - \mathcal{L}(\dot{q}, q) \\ \mathcal{L}_e(\dot{q}, p, q) &\equiv p\dot{q} - \mathcal{H}(p, q), \end{aligned} \quad (25)$$

where we switched the notation a bit by defining $v = \dot{q}$. Now we treat p and \dot{q} as independent variables, and therefore \mathcal{H}_e will be a surface. Let us consider a non-relativistic particle of unit mass with with no potential: $\mathcal{L}(\dot{q}, q) = \mathcal{L}(\dot{q}) = \frac{\dot{q}^2}{2}$. Then the corresponding \mathcal{H}_e becomes

$$\mathcal{H}_e(\dot{q}, p, q) \equiv p\dot{q} - \mathcal{L}(\dot{q}) = p\dot{q} - \frac{\dot{q}^2}{2}, \quad (26)$$

which is shown in Fig. 4.

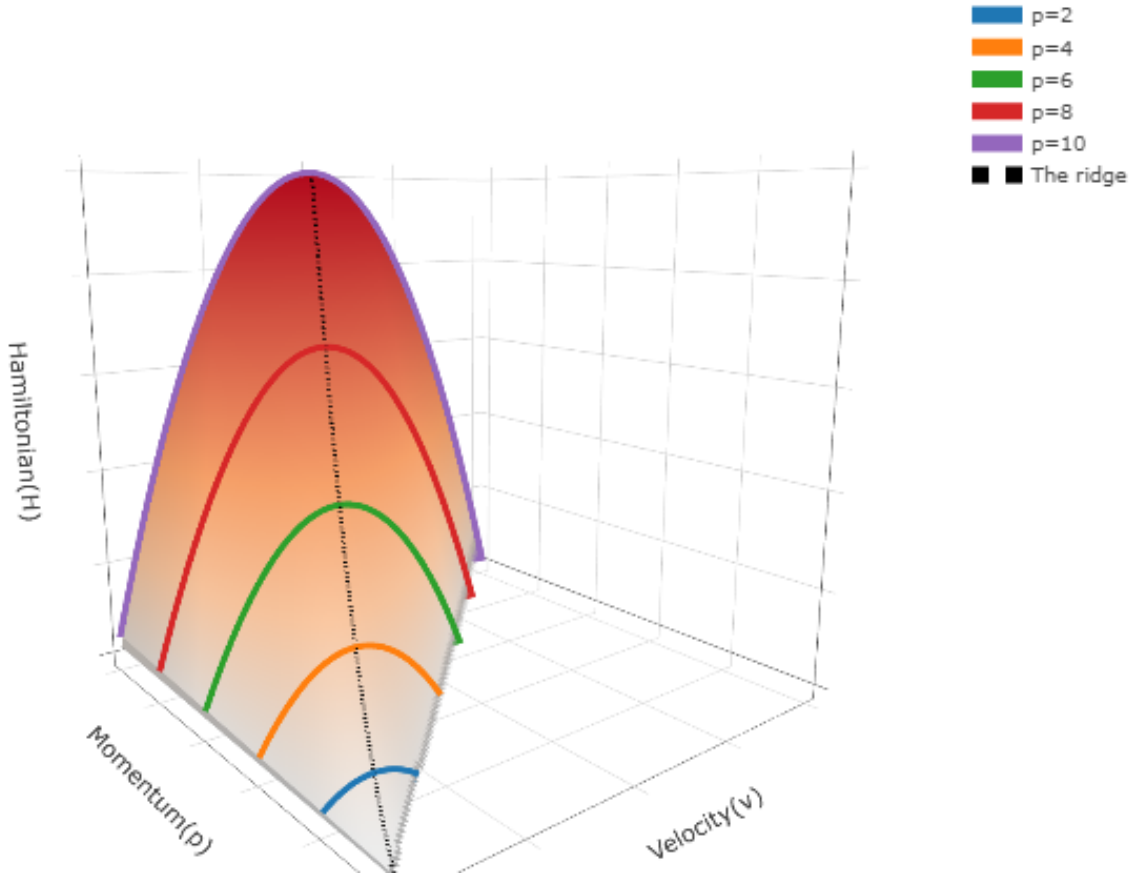


Figure 4: Extended Hamiltonian for a non-relativistic particle . The value of the ordinary Hamiltonian is a slice of the surface, which is the ridge shown in the dashed black line.

We can still recover the original \mathcal{H} since it is the value of \mathcal{H}_e when $p = \dot{q}$ and it is the ridge of the surface shown with the dashed black line. Mathematically, we have $\mathcal{H}(p) = \mathcal{H}_e(\dot{q}(p), p)$. More generically we can write the following:

$$\mathcal{H}(p, q) \equiv \max_{\dot{q}} (p\dot{q} - \mathcal{L}(\dot{q}, q)), \quad (27)$$

where $\max_{\dot{q}}$ means that the equation is evaluated at the value of \dot{q} which maximizes the result. Note that this can be used as the definition of the Legendre transform, and in fact it is what mathematicians do. The Legendre transform of a function $f(x)$ is defined as follows [6]:

$$f^*(x^*) = \sup_{x \in I} (x^*x - f(x)), \quad x^* \in I^*, \quad (28)$$

where sup denotes the supremum, I and I^* are the domains of the functions f and f^* , respectively. The mapping between the quantities we used earlier and the ones in this formal definition is as follows: $x \sim \dot{q}$ and $x^* \sim p$, $f \sim \mathcal{L}$, and $f^* \sim \mathcal{H}$.

This completes our deep dive into the Legendre transform!

- [1] L. D. Elsgolc, *Calculus of variations*. Dover Publications, 2007.
- [2] S. Leonard, “Modern physics: Classical mechanics (stanford, lecture 6, 2007).” 2007 [Online]. Available: <https://www.youtube.com/watch?v=14Yhzbn96Bc&t=1345s>. [Accessed: 12-Nov-2021]
- [3] R. K. P. Zia, E. F. Redish, and S. R. McKay, “Making sense of the legendre transform,” *American Journal of Physics*, vol. 77, no. 7, pp. 614–622, Jul. 2009, doi: 10.1119/1.3119512. [Online]. Available: <http://dx.doi.org/10.1119/1.3119512>

- [4] D. Cline, “Lorentz-invariant formulations of Hamiltonian Mechanics.” University of Rochester, 2021 [Online]. Available: [https://phys.libretexts.org/Bookshelves/Classical_Mechanics/Variational_Principles_in_Classical_Mechanics_\(Cline\)/17%3A_Relativistic_Mechanics/17.07%3A_Lorentz-invariant_formulations_of_Hamiltonian_Mechanics](https://phys.libretexts.org/Bookshelves/Classical_Mechanics/Variational_Principles_in_Classical_Mechanics_(Cline)/17%3A_Relativistic_Mechanics/17.07%3A_Lorentz-invariant_formulations_of_Hamiltonian_Mechanics). [Accessed: 12-Nov-2021]
- [5] M. Alford, “Legendre transforms.” 2019 [Online]. Available: https://web.physics.wustl.edu/alford/physics/Legendre_introduction.pdf. [Accessed: 12-Nov-2021]
- [6] Wikipedia, “Legendre transformation — Wikipedia, the free encyclopedia.” <http://en.wikipedia.org/w/index.php?title=Legendre%20transformation&oldid=1054288669>, 2021.