

# Variation

Jacob Denson

February 16, 2023

# Chapter 1

## Lagrangian Formulations

### 1.1 The Legendre Transform

In this section, we introduce a duality to the theory of convex functions. The main use for us here occurs in the Lagrangian and Hamiltonian mechanics, in order to convert a convex function  $H : \mathbb{R}^n \rightarrow \mathbb{R}$  representing the *energy* of a system, into a convex function  $L : \mathbb{R}^n \rightarrow \mathbb{R}$ , called the *Lagrangian* of the system.

Let  $V$  be a finite dimensional vector space. Let  $H : V^* \rightarrow \mathbb{R}$  be a convex function. We define the *Legendre transform* of  $H$ , to be the convex function  $L : V \rightarrow \mathbb{R}$  defined by

$$L(v) = \sup_{\xi \in V^*} \xi(v) - H(\xi).$$

We say a convex function  $H$  is *superlinear* if  $\lim_{\xi \rightarrow \infty} H(\xi)/|\xi| = \infty$ . For a superlinear function  $H$ , the resulting function  $L$  will be everywhere finite. As an example, if  $H(\xi) = |\xi|^2/2m$ , then

$$L(v) = \sup_{\xi \in \mathbb{R}^n} \xi \cdot v - |\xi|^2/2m = m|v|^2/2.$$

If  $H$  is superlinear, then  $L$  is superlinear, and  $H$  is the Legendre transform of  $L$ . Thus the Legendre transform is an *order reversing* involution on the space of convex functions. By definition, we have *Fenchel's inequality*

$$H(\xi) + L(v) \geq \xi(v)$$

for all  $v \in V$  and  $\xi \in V^*$ . If  $H$  is  $k$  times continuously differentiable for  $k \geq 2$ , convex, superlinear, and *locally uniformly convex*, in the sense that the Hessian of

$H$  is non-vanishing, then  $L$  is  $k$  times continuously differentiable,  $dH : V^* \rightarrow V$  is a  $C^{k-1}$  diffeomorphism, and its inverse is precisely  $dL : V \rightarrow V^*$ . Moreover, we have

$$L(v) = \langle dL(v), v \rangle - H(dL(v)).$$

In general, we will also deal with Hamiltonian and Lagrangians depending on space. We thus work on a manifold  $M$ , and a function  $H : T^*M \rightarrow \mathbb{R}$ , which is convex when restricted to each point in  $M$ . We can then take the Legendre transform at each point in  $M$ , which gives a function  $L : TM \rightarrow \mathbb{R}$ .

## 1.2 Minimizing Action

We now consider a Lagrangian  $L : \mathbb{R}_v^n \rightarrow \mathbb{R}$  which is locally uniformly convex, uniformly superlinear, and  $k$  times continuously differentiable, for  $k \geq 2$ . Our goal is to study the *action functional*

$$I[\gamma] = \int_a^b L(\gamma(t), \dot{\gamma}(t)) dt,$$

where  $a < b$  are fixed, and  $\gamma$  takes values in a certain class of curves  $\mathcal{A}$ , with fixed endpoints  $x_0$  and  $x_1$ . Our goal is to show that *minimizers* of  $I$  exist, and that they have certain regularity properties.

Let us start with a classical framework, where  $\mathcal{A}$  is the class of all continuous, piecewise  $C^1$  curves  $\gamma$  defined on  $[a, b]$ , with  $\gamma(a) = x_0$  and  $\gamma(b) = x_1$ .

**Lemma 1.1.** *If  $\gamma \in \mathcal{A} \cap C^2[a, b]$  is a minimizer of  $I$  in  $\mathcal{A}$ , then*

$$\frac{d}{dt} \{(\nabla_v L)(\gamma(t), \dot{\gamma}(t))\} = (\nabla_x L)(\gamma(t), \dot{\gamma}(t)).$$

*Proof.* We apply a classical variational method. Fix a smooth function  $\eta : [a, b] \rightarrow \mathbb{R}^n$  with  $\eta(a) = \eta(b) = 0$ . Then for any  $s \in \mathbb{R}$ ,  $\gamma + s\eta \in \mathcal{A}$ . Define

$$f(s) = I[\gamma + s\eta].$$

Then because  $\gamma$  is a minimizer,  $f'(0) = 0$ , and differentiating under the integral sign yields that

$$\begin{aligned} 0 &= \int_a^b (\nabla_x L)(\gamma(t), \dot{\gamma}(t)) \cdot \eta + (\nabla_v L)(\gamma(t), \dot{\gamma}(t)) \cdot \dot{\eta} \\ &= \int_a^b \left( (\nabla_x L)(\gamma(t), \dot{\gamma}(t)) - \frac{d}{dt} (\nabla_v L)(\gamma(t), \dot{\gamma}(t)) \right) \cdot \eta. \end{aligned}$$

This is only possible for all  $\eta$  if the required formula holds.  $\square$

The result continues to hold for  $\gamma \in \mathcal{A}$ , provided we interpret the formula distributionally. This means that if  $\gamma \in \mathcal{A} \cap C^1[a, b]$ , then the right hand side is then  $C^1$ . But this means that  $\gamma$  satisfies a distributional ordinary differential equation with  $C^1$  values. Thus  $\gamma$  is actually  $C^2$ . But iterating this process actually guarantees that the equation is  $C^k$ . But now if  $\gamma$  is only piecewise  $C^1$ , then, on each interval, we can conclude that  $\gamma$  is  $C^k$ , and that the derivative formula above holds. But this is sufficient to conclude the left and right hand derivatives at each of the ends of these intervals agrees with one another, so that  $\gamma$  is actually  $C^1$ , and thus  $C^k$ .

To prove the *existence* of minimizers, it is natural to work with the smaller class  $\mathcal{A}_0 \subset \mathcal{A}$  of *absolutely continuous* curves, since this class has a useful *compactness property*, namely, that any sequence  $\{\gamma_k\}$  such that:

- The functions  $\dot{\gamma}_k$  are uniformly integrable.
- There exists  $t_0 \in [a, b]$  such that  $\{\gamma_k(t_0)\}$  is bounded.

Then the sequence  $\{\gamma_k\}$  has a subsequence  $\{\gamma_{k_j}\}$  which converges uniformly on  $[a, b]$ , and the derivatives  $\dot{\gamma}_{k_j}$  converge weakly in  $L^1$ .

To prove the existence of minimizers, we apply the *direct method* for the calculus of variations. The general framework goes as follows:

- There is a constant  $C$  such that for all  $\gamma \in \mathcal{A}_0$ ,  $I[\gamma] \geq C$ . Thus an infimum  $C_* = \inf I[\gamma]$  exists.
- Consider a sequence  $\{\gamma_k\}$  such that  $I[\gamma_k] \rightarrow C_*$ . Applying a compactness property, we can conclude a subsequence converges to some  $\gamma$ .
- We argue that  $I$  is weakly lower semicontinuous, i.e.

$$I[\gamma] \leq \liminf_k I[\gamma_k].$$

But this means  $I[\gamma] = C_*$ , completing the argument.

Let's now prove the existence of minimizers.

**Theorem 1.2.** *Minimizers exist in  $\mathcal{A}_0$ .*

*Proof.* Since  $L$  is superlinear, for each  $\theta > 0$ , there is  $C_\theta > 0$  such that

$$L(x, v) \geq \theta|v| - C_\theta.$$

In particular,

$$L(x, v) \geq |v| - C_1.$$

Thus for any  $\gamma \in \mathcal{A}_0$ ,

$$I[\gamma] = \int_a^b L(\gamma(t), \dot{\gamma}(t)) dt \geq \left( \int_a^b |\dot{\gamma}(t)| dt - C_1(b-a) \right) \geq -(b-a)C_1.$$

In particular, setting  $\theta = 1$  gives the existence of the  $C_0$  required in the first step of the argument.

Now take a sequence  $\{\gamma_k\}$  such that  $I[\gamma_k]$  converges to  $C_*$ . Let  $C^*$  denote the maximum value of  $I[\gamma_k]$ . We must show that this sequence satisfies the tightness condition. To show uniform integrability, we calculate that for any Borel set  $E$  with  $|E| < \delta$ ,

$$\begin{aligned} \int_E L(\gamma_k(t), \dot{\gamma}_k(t)) dt &= I[\gamma_k] - \int_{[a,b]-E} L(\gamma_k(t), \dot{\gamma}_k(t)) dt \\ &\leq C^* + |[a,b] - E|C_1 \leq C^* + (b-a)C_1 \lesssim 1. \end{aligned}$$

But this means that

$$\int_E |\dot{\gamma}_k(t)| dt \leq \int_E (C_\theta + \theta^{-1} L(\gamma_k(t), \dot{\gamma}_k(t))) dt \lesssim \delta C_\theta + \theta^{-1}.$$

Choosing  $\theta$  sufficiently large, and  $\delta$  sufficiently small, yields the required claim. Thus  $\{\gamma_k\}$  converges uniformly to an absolutely continuous function  $\gamma$ , such that  $\dot{\gamma}_k$  converges weakly in  $L^1$  to  $\dot{\gamma}$ .

Now let's show  $\gamma$  is a minimizer. To begin with, we show that for any  $C_1, C_2 > 0$ , there exists  $\delta > 0$  such that if  $|v_1| \leq C_1$ , and  $|x_1|, |x_2| \leq C_2$ , with  $|x_1 - x_2| \leq \delta$ . Then for all  $v_2 \in \mathbb{R}^n$ ,

$$L(x_2, v_2) \geq L(x_1, v_1) + (\nabla_v L)(x_1, v_1) \cdot (v_2 - v_1) - \varepsilon.$$

How does this Lemma imply the result? Let

$$U_\lambda = \{t \in [a, b] : |\dot{\gamma}(t)| \leq \lambda\}.$$

Then  $[a, b] - \bigcup U_\lambda$  is a set of measure zero. For each  $\lambda$ , set  $C_1 = \lambda$ , set  $C_2 = \sup_k \|\gamma_k\|_{L^\infty}$ , and for each  $\varepsilon > 0$ , fix  $\delta$  as above. We conclude that if  $\|\gamma_k - \gamma\|_{L^\infty} \leq \delta$ , then

$$L(\gamma_k(t), \dot{\gamma}_k(t)) \geq L(\gamma(t), \dot{\gamma}(t)) + (\nabla_v L)(\gamma(t), \dot{\gamma}(t)) \cdot (\dot{\gamma}_k(t) - \dot{\gamma}(t)) - \varepsilon.$$

But this means that

$$\begin{aligned}
I[\gamma_k] &= \left\{ \int_{U_\lambda} + \int_{U_\lambda^c} \right\} L(\gamma_k(t), \dot{\gamma}_k(t)) dt \\
&\geq \int_{U_\lambda} L(\gamma_k(t), \dot{\gamma}_k(t)) dt - C^*|[a, b] - U_\lambda| \\
&\geq \int_{U_\lambda} L(\gamma(t), \dot{\gamma}(t)) dt + (\nabla_v L)(\gamma(t), \dot{\gamma}(t)) \cdot (\dot{\gamma}_k(t) - \dot{\gamma}(t)) dt - (b-a)\varepsilon - C^*|[a, b] - U_\lambda|.
\end{aligned}$$

As  $k \rightarrow \infty$ , we conclude that

$$C_* \geq \int_{U_\lambda} L(\gamma(t), \dot{\gamma}(t)) dt - (b-a)\varepsilon - C^*|[a, b] - U_\lambda|.$$

Taking  $\lambda \rightarrow \infty$  yields that

$$C_* \geq \int_a^b L(\gamma(t), \dot{\gamma}(t)) dt - (b-a)\varepsilon.$$

Taking  $\varepsilon \rightarrow 0$  yields that  $I[\gamma] \leq C_*$ , completing the proof of the claim.

It remains to prove the required inequality. By assumption, we know that for  $|v_1| \leq C_1$  and  $|x_1| \leq C_2$ ,  $|L(x_1, v_1)| \lesssim_{C_1, C_2} 1$  and  $|(\nabla_v L)(x_1, v_1)| \lesssim_{C_1, C_2} 1$ . But this means that

$$L(x_1, v_1) + (\nabla_v L)(x_1, v_1) \cdot (v_2 - v_1) - \varepsilon \leq A_{C_1, C_2}(1 + |v_2 - v_1|),$$

for some constant  $A_{C_1, C_2}$ . Since  $L$  is superlinear, there exists  $R > 0$  such that if  $|v_2| \geq R$ ,

$$L(x_2, v_2) \geq A_{C_1, C_2}(2 + |v_2|) \geq L(x_1, v_1) + (\nabla_v L)(x_1, v_1) \cdot (v_2 - v_1) - \varepsilon,$$

so the proof is trivial for large  $v_2$ . But now apply the uniform continuity of  $L$ , there exists  $\delta$  such that if  $|x_1 - x_2| \leq \delta$ , then  $|L(x_1, v_1) - L(x_2, v_2)| \leq \varepsilon$ , which means that

$$L(x_2, v_2) \geq L(x_1, v_1) - \varepsilon \geq L(x_1, v_1) + (\nabla_v L)(x_1, v_1) \cdot (v_2 - v_1) - \varepsilon,$$

which completes the proof of the inequality in general.  $\square$

Thus minimizers exist. Let's show they're regular. The method is somewhat involved, using the theory of characteristics and Hamilton-Jacobi equations.

**Lemma 1.3.** *Let  $\pi : \mathbb{R}_x^n \times \mathbb{R}_v^n \rightarrow \mathbb{R}_x^n$  be the projection map, and let  $\{\phi_t\}$  be the Lagrangian flow. For any  $C > 0$ , there exists  $\delta > 0$  such that for  $|t| \leq \delta$ , and  $|x| \leq C|s|$ ,*

$$\pi \circ \phi_t(\{x\})$$

TODO: Check Appendix C + Method of Characteristics of Evans.