3. Fully nonlinear first order PDE

In this section let $\mathcal{U} \subset_o \mathbb{R}^n$ be an open subset of $\mathbb{R}^n$ and $(x,z,p) \in \bar{\mathcal{U}} \times \mathbb{R}^n \times \mathbb{R} \to F(x,z,p) \in \mathbb{R}$ be a C^2 – function. Actually to simplify notation let us suppose $\mathcal{U} = \mathbb{R}^n$. We are now looking for a solution $u : \mathbb{R}^n \to \mathbb{R}$ to the fully non-linear PDE,

$$F(x, u(x), \nabla u(x)) = 0. \tag{3.1}$$

As above, we "reduce" the problem to solving ODE's. To see how this might be done, suppose u solves (3.1) and $x(s)$ is a curve in $\mathbb{R}^n$ and let

$$z(s) = u(x(s)) \text{ and } p(s) = \nabla u(x(s)).$$

Then

$$z'(s) = \nabla u(x(s)) \cdot x'(s) = p(s) \cdot x'(s) \text{ and} \tag{3.2}$$

$$p'(s) = \partial_{x'(s)} \nabla u(x(s)). \tag{3.3}$$

We would now like to find an equation for $x(s)$ which along with the above system of equations would form and ODE for $(x(s), z(s), p(s))$. The term, $\partial_{x'(s)} \nabla u(x(s))$, which involves two derivative of u is problematic and we would like to replace it by something involving only ∇u and u. In order to get the desired relation, differentiate Eq. (3.1) in x in the direction v to find

$$\begin{aligned} 0 &= F_x \cdot v + F_z \partial_v u + F_p \cdot \partial_v \nabla u = F_x \cdot v + F_z \partial_v u + F_p \cdot \nabla \partial_v u \\ &= F_x \cdot v + F_z \ \nabla u \cdot v + (\partial_{F_p} \nabla u) \cdot v, \end{aligned}$$

wherein we have used the fact that mixed partial derivative commute. This equation is equivalent to

$$\partial_{F_p} \nabla u|_{(x,u(x),\nabla u(x))} = -(F_x + F_z \ \nabla u)|_{(x,u(x),\nabla u(x))}. \tag{3.4}$$

By requiring $x(s)$ to solve $x'(s) = F_p(x(s), z(s), p(s))$, we find, using Eq. (3.4) and Eqs. (3.2) and (3.3) that $(x(s), z(s), p(s))$ solves the **characteristic equations**,

$$\begin{aligned} x'(s) &= F_p(x(s), z(s), p(s)) \\ z'(s) &= p(s) \cdot F_p(x(s), z(s), p(s)) \\ p'(s) &= -F_x(x(s), z(s), p(s)) - F_z(x(s), z(s), p(s))p(s). \end{aligned}$$

We will in the future simply abbreviate these equations by

$$\begin{aligned} x' &= F_p \\ z' &= p \cdot F_p \\ p' &= -F_x - F_z p. \end{aligned} \tag{3.5}$$

The above considerations have proved the following Lemma.

Lemma 3.1. *Let*

$$A(x,z,p) := (F_p(x,z,p), p \cdot F_p(x,z,p), -F_x(x,z,p) - F_z(x,z,p)p),$$

$$\pi_1(x,z,p) = x \text{ and } \pi_2(x,z,p) = z.$$

If u solves Eq. (3.1) and $x_0 \in U$, then

$$e^{sA}(x_0, u(x_0), \nabla u(x_0)) = (x(s), u(x(s)), \nabla u(x(s))) \text{ and}$$

$$u(x(s)) = \pi_2 \circ e^{sA}(x_0, u(x_0), \nabla u(x_0)) \tag{3.6}$$

where $x(s) = \pi_1 \circ e^{sA}(x_0, u(x_0), \nabla u(x_0))$.

We now want to use Eq. (3.6) to produce solutions to Eq. (3.1). As in the quasi-linear case we will suppose $\Sigma : U \subset_o \mathbb{R}^{n-1} \longrightarrow \mathbb{R}^n$ is a surface, $\Sigma(0) = x_0$, $D\Sigma(y)$ is injective for all $y \in U$ and $u_0 : \Sigma \to \mathbb{R}$ is given. We wish to solve Eq. (3.1) for u with the added condition that $u(\Sigma(y)) = u_0(y)$. In order to make use of Eq. (3.6) to do this, we first need to be able to find $\nabla u(\Sigma(y))$. The idea is to use Eq. (3.1) to determine $\nabla u(\Sigma(y))$ as a function of $\Sigma(y)$ and $u_0(y)$ and for this we will invoke the implicit function theorem. If u is a function such that $u(\Sigma(y)) = u_0(y)$ for y near 0 and $p_0 = \nabla u(x_0)$ then

$$\partial_v u_0(0) = \partial_v u(\Sigma(y))|_{y=0} = \nabla u(x_0) \cdot \Sigma'(0)v = p_0 \cdot \Sigma'(0)v.$$

Notation 3.2. Let $\nabla_\Sigma u_0(y)$ denote the unique vector in $\mathbb{R}^n$ which is tangential to Σ at $\Sigma(y)$ and such that

$$\partial_v u_0(y) = \nabla_\Sigma u_0(y) \cdot \Sigma'(0)v \text{ for all } v \in \mathbb{R}^{n-1}.$$

Theorem 3.3. *Let $F : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be a C^2 function, $0 \in U \subset_o \mathbb{R}^{n-1}$, $\Sigma : U \subset_o \mathbb{R}^{n-1} \xrightarrow{C^2} \mathbb{R}^n$ be an embedded submanifold, $(x_0, z_0, p_0) \in \Sigma \times \mathbb{R} \times \mathbb{R}^n$ such that $F(x_0, z_0, p_0) = 0$ and $x_0 = \Sigma(0)$, $u_0 : \Sigma \xrightarrow{C^1} \mathbb{R}$ such that $u_0(x_0) = z_0$, $\mathbf{n}(y)$ be a normal vector to Σ at y. Further assume*

1. $\partial_v u_0(0) = p_0 \cdot \Sigma'(0)v = p_0 \cdot \partial_v \Sigma(0)$ *for all* $v \in \mathbb{R}^{n-1}$.
2. $F_p(x_0, y_0, z_0) \cdot \mathbf{n}(0) \neq 0$.

Then there exists a neighborhood $V \subset \mathbb{R}^n$ of x_0 and a C^2-function $u : V \to \mathbb{R}$ such that $u \circ \Sigma = u_0$ near 0 and Eq. (3.1) holds for all $x \in V$.

Proof. Step 1. There exist a neighborhood $U_0 \subset U$ and a function $p_0 : U_0 \to \mathbb{R}^n$ such that

$$p_0(y)^{\tan} = \nabla_\Sigma u_0(y) \text{ and } F(\Sigma(y), u(\Sigma((y)), p_0(y)) = 0 \tag{3.7}$$

for all $y \in U_0$, where $p_0(y)^{\tan}$ is component of $p_0(y)$ tangential to Σ. This is an

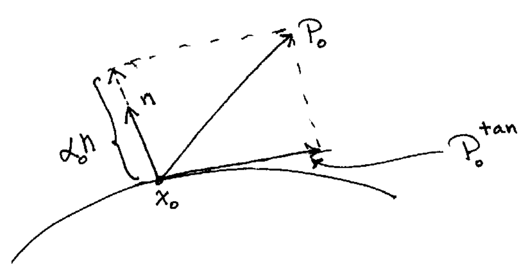

FIGURE 11. Decomposing p into its normal and tangential components.

exercise in the implicit function theorem.

Choose $\alpha_0 \in \mathbb{R}$ such that $\nabla_\Sigma u_0(0) + \alpha_0 \mathbf{n}(0) = p_0$ and define

$$f(\alpha, y) := F(\Sigma(y), u_0(y), \nabla_\Sigma u_0(y) + \alpha\, \mathbf{n}(y)).$$

Then

$$\frac{\partial f}{\partial \alpha}(\alpha, 0) = F_p(x_0, z_0, \nabla_\Sigma u_0(0) + \alpha \mathbf{n}(0)) \cdot \mathbf{n}(0) \neq 0,$$

so by the implicit function theorem there exists $0 \in U_0 \subset U$ and $\alpha : U_0 \to \mathbb{R}$ such that $f(\alpha(y), y) = 0$ for all $y \in U_0$. Now define

$$p_0(y) := \nabla_\Sigma u_0(y) + \alpha(y)\mathbf{n}(y) \text{ for } y \in U_0.$$

To simplify notation in the future we will from now on write U for U_0.

Step 2. Suppose (x, z, p) is a solution to (3.5) such that $F(x(0), z(0), p(0)) = 0$ then

$$F(x(s), z(s), p(s)) = 0 \text{ for all } t \in J \tag{3.8}$$

because

$$\begin{aligned}\frac{d}{ds}F(x(s), z(s), p(s)) &= F_x \cdot x' + F_z z' + F_p \cdot p' \\ &= F_x \cdot F_p + F_z(p \cdot F_p) - F_p \cdot (F_x + F_z p) = 0.\end{aligned} \tag{3.9}$$

Step 3. (Notation). For $y \in U$ let

$$(X(s, y), Z(s, y), P(s, y)) = e^{sA}(\Sigma(y), u_0(y), p_0(y)),$$

ie. $X(s, y)$, $Z(s, y)$ and $P(s, y)$ solve the coupled system of O.D.E.'s:

$$\begin{aligned} X' &= F_p \text{ with } X(0, y) = \Sigma(y) \\ Z' &= P \cdot F_p \text{ with } Z(0, y) = u_0(y) \\ P' &= F_x - F_z P \text{ with } P(0, y) = p_0(y). \end{aligned} \tag{3.10}$$

With this noration Eq. (3.8) becomes

$$F(X(s, y), Z(s, y), P(s, y)) = 0 \text{ for all } t \in J. \tag{3.11}$$

Step 4. There exists a neighborhood $0 \in U_0 \subset U$ and $0 \in J \subset \mathbb{R}$ such that $X : J \times U_0 \to \mathbb{R}^n$ is a C^1 diffeomorphism onto an open set $V := X(J \times U_0) \subset \mathbb{R}^n$ with $x_0 \in V$. Indeed, $X(0, y) = \Sigma(y)$ so that $D_y X(0, y)|_{y=0} = \Sigma'(0)$ and hence

$$DX(0, 0)(a, v) = \frac{\partial X}{\partial s}(0, 0)a + \Sigma'(0)v = F_p(x_0, z_0, p_0)a + \Sigma'(0)v.$$

By the assumptions, $F_p(x_0, z_0, p_0) \notin \text{Ran } \Sigma'(0)$ and $\Sigma'(0)$ is injective, it follows that $DX(0, 0)$ is invertible So the assertion is a consequence of the inverse function theorem.

Step 5. Define

$$u(x) := Z(X^{-1}(x)),$$

then u is the desired solution. To prove this first notice that u is uniquely characterized by

$$u(X(s, y)) = Z(s, y) \text{ for all } (s, y) \in J_0 \times U_0.$$

Because of Step 2., to finish the proof it suffices to show $\nabla u(X(s, y)) = P(s, y)$.

Step 6. $\nabla u(X(s, y)) = P(s, y)$. From Eq. (3.10),

$$P \cdot X' = P \cdot F_p = Z' = \frac{d}{ds}u(X) = \nabla u(X) \cdot X' \tag{3.12}$$

which shows

$$[P - \nabla u(X)] \cdot X' = 0.$$

So to finish the proof it suffices to show

$$[P - \nabla u(X)] \cdot \partial_v X = 0$$

for all $v \in \mathbb{R}^{n-1}$ or equivalently that

$$P(s,y) \cdot \partial_v X = \nabla u(x) \cdot \partial_v X = \partial_v u(X) = \partial_v Z \tag{3.13}$$

for all $v \in \mathbb{R}^{n-1}$.

To prove Eq. (3.13), fix a y and let

$$r(s) := P(s,y) \cdot \partial_v X(s,y) - \partial_v Z(s,y).$$

Then using Eq. (3.10),

$$\begin{aligned} r' &= P' \cdot \partial_v X + P \cdot \partial_v X' - \partial_v Z' \\ &= (-F_x - F_z P) \cdot \partial_v X + P \cdot \partial_v F_p - \partial_v (P \cdot F_p) \\ &= (-F_x - F_z P) \cdot \partial_v X - (\partial_v P) \cdot F_p. \end{aligned} \tag{3.14}$$

Further, differentiating Eq. (3.11) in y implies for all $v \in \mathbb{R}^{n-1}$ that

$$F_x \cdot \partial_v X + F_z \partial_v Z + F_p \cdot \partial_v P = 0. \tag{3.15}$$

Adding Eqs. (3.14) and (3.15) then shows

$$r' = -F_z P \cdot \partial_v X + F_z \partial_v Z = -F_z r$$

which implies

$$r(s) = e^{-\int_0^s F_z(X,Z,P)(\sigma,y)d\sigma} r(0).$$

This shows $r \equiv 0$ because $p_0(y)^T = (\nabla_\Sigma u_0)(\Sigma(y))$ and hence

$$r(0) = p_0(y) \cdot \partial_v \Sigma(y) - \partial_v u_0(\Sigma(y)) = [p_0(y) - \nabla_\Sigma u_0(\Sigma(y))] \cdot \partial_v \Sigma(y) = 0.$$

■

Example 3.4 (Quasi-Linear Case Revisited)**.** Let us consider the quasi-linear PDE in Eq. (2.39),

$$A(x,z) \cdot \nabla_x u(x) - c(x,u(x)) = 0. \tag{3.16}$$

in light of Theorem 3.3. This may be written in the form of Eq. (3.1) by setting

$$F(x,z,p) = A(x,z) \cdot p - c(x,z).$$

The characteristic equations (3.5) for this F are

$$\begin{aligned} x' &= F_p = A \\ z' &= p \cdot F_p = p \cdot A \\ p' &= -F_x - F_z p = -(A_x \cdot p - c_x) - (A_z \cdot p - c_z)\, p. \end{aligned}$$

Recalling that $p(s) = \nabla u(s,x)$, the z equation above may expressed, by using Eq. (3.16) as

$$z' = p \cdot A = c.$$

Therefore the equations for $(x(s), z(s))$ may be written as

$$x'(s) = A(x,z) \text{ and } z' = c(x,z)$$

and these equations may be solved without regard for the p – equation. This is what makes the quasi-linear PDE simpler than the fully non-linear PDE.

3.1. **An Introduction to Hamilton Jacobi Equations.** A Hamilton Jacobi Equation is a first order PDE of the form

$$\frac{\partial S}{\partial t}(t,x) + H(x, \nabla_x S(t,x)) = 0 \text{ with } S(0,x) = g(x) \tag{3.17}$$

where $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are given functions. In this section we are going to study the connections of this equation to the Euler Lagrange equations of classical mechanics.

3.1.1. *Solving the Hamilton Jacobi Equation (3.17) by characteristics.* Now let us solve the Hamilton Jacobi Equation (3.17) using the method of characteristics. In order to do this let

$$(p_0, p) = (\frac{\partial S}{\partial t}, \nabla_x S(t,x)) \text{ and } F(t,x,z,p) := p_0 + H(x,p).$$

Then Eq. (3.17) becomes

$$0 = F(t, x, S, \frac{\partial S}{\partial t}, \nabla_x S).$$

Hence the characteristic equations are given by

$$\frac{d}{ds}(t(s), x(s)) = F_{(p_0,p)} = (1, \nabla_p H(x(s), p(s))$$

$$\frac{d}{ds}(p_0, p)(s) = -F_{(t,x)} - F_z(p_0, p) = -F_{(t,x)} = (0, -\nabla_x H(x(s), p(s)))$$

and

$$z'(s) = (p_0, p) \cdot F_{(p_0,p)} = p_0(s) + p(s) \cdot \nabla_p H(x(s), p(s)).$$

Solving the t equation with $t(0) = 0$ gives $t = s$ and so we identify t and s and our equations become

$$\dot{x}(t) = \nabla_p H(x(t), p(t)) \tag{3.18}$$

$$\dot{p}(t) = -\nabla_{\vec{x}} H(x(t), p(t)) \tag{3.19}$$

$$\frac{d}{dt}\left[\frac{\partial S}{\partial t}(t, x(t))\right] = \frac{d}{dt} p_0(t) = 0 \text{ and}$$

$$\begin{aligned} \frac{d}{dt}[S(t, x(t))] &= \frac{d}{dt} z(t) = \frac{\partial S}{\partial t}(t, x(t)) + p(t) \cdot \nabla_p H(x(t), p(t)) \\ &= -H(x(t), p(t)) + p(t) \cdot \nabla_p H(x(t), p(t)). \end{aligned}$$

Hence we have proved the following proposition.

Proposition 3.5. *If S solves the Hamilton Jacobi Equation Eq. 3.17 and $(x(t), p(t))$ are solutions to the Hamilton Equations (3.18) and* **??** *(see also Eq. (3.29) below) with $p(0) = (\nabla_x g)(x(0))$ then*

$$S(T, x(T)) = g(x(0)) + \int_0^T [p(t) \cdot \nabla_p H(x(t), p(t)) - H(x(t), p(t))]\, dt.$$

In particular if $(T, x) \in \mathbb{R} \times \mathbb{R}^n$ then

$$S(T, x) = g(x(0)) + \int_0^T [p(t) \cdot \nabla_p H(x(t), p(t)) - H(x(t), p(t))]\, dt. \tag{3.20}$$

provided (x, p) is a solution to Hamilton Equations (3.18) and (3.19) satisfying the boundary condition $x(T) = x$ and $p(0) = (\nabla_x g)(x(0))$.

Remark 3.6. Let $X(t,x_0,p_0)=x(t)$ and $P(t,x_0,p_0)=p(t)$ where $(x(t),p(t))$ satisfies Hamilton Equations (3.18) and (3.19) with $(x(0),p(0))=(x_0,p_0)$ and let $\Psi(t,x):=(t,X(t,x,\nabla g(x))$. Then $\Psi(0,x)=(0,x)$ so

$$\partial_v\Psi(0,0)=(0,v) \text{ and } \partial_t\Psi(0,0)=(1,\nabla_p H(x,\nabla g(x)))$$

from which it follows that $\Psi'(0,0)$ is invertible. Therefore given $a\in\mathbb{R}^n$, the exists $\epsilon>0$ such that $\Psi^{-1}(t,x)$ is well defined for $|t|<\epsilon$ and $|x-a|<\epsilon$. Writing $\Psi^{-1}(T,x)=(T,x_0(T,x))$ we then have that

$$(x(t),p(t)):=(X(t,x_0(T,x),\nabla g(x_0(T,x)),P(t,x_0,\nabla g(x_0(T,x)))$$

solves Hamilton Equations (3.18) and (3.19) satisfies the boundary condition $x(T)=x$ and $p(0)=(\nabla_x g)(x(0))$.

3.1.2. *The connection with the Euler Lagrange Equations.* Our next goal is to express the solution $S(T,x)$ in Eq. (3.20) solely in terms of the path $x(t)$. For this we digress a bit to Lagrangian mechanics and the notion of the "classical action."

Definition 3.7. Let $T>0$, $L:\mathbb{R}^n\times\mathbb{R}^n\longrightarrow\mathbb{R}$ be a smooth "Lagrangian" and $g:\mathbb{R}^n\to\mathbb{R}$ be a smooth function. The $g-$ **weighted action** $I_T^g(q)$ of a function $q\in C^2([0,T],\mathbb{R}^n)$ is defined to be

$$I_T^g(q)=g(q(0))+\int_0^T L(q(t),\dot q(t))dt.$$

When $g=0$ we will simply write I_T for $I_{0,T}$.

We are now going to study the function $S(T,x)$ of "least action,"

(3.21)

$$\begin{aligned}S(T,x)&:=\inf\left\{I_T^g(q):q\in C^2([0,T])\text{ with }q(T)=x\right\}\\&=\inf\left\{g(q(0))+\int_0^T L(q(t),\dot q(t))dt:q\in C^2([0,T])\text{ with }q(T)=x\right\}.\end{aligned}$$

The next proposition records the differential of $I_T^g(q)$.

Proposition 3.8. *Let $L\in C^\infty(\mathbb{R}^n\times\mathbb{R}^n,\mathbb{R})$ be a smooth Lagrangian, then for $q\in C^2([0,T],\mathbb{R}^n)$ and $h\in C^1([0,T],\mathbb{R}^n)$*

$$\begin{aligned}DI_T^g(q)h&=\left[(\nabla g(q)-D_2L(q,\dot q))\cdot h\right]_{t=0}+\left[D_2L(q,\dot q)\cdot h\right]_{t=T}\\&\quad+\int_0^T(D_1L(q,\dot q)-\frac{d}{dt}D_2L(q,\dot q))h\ dt\end{aligned}\tag{3.22}$$

Proof. By differentiating past the integral,

$$\begin{aligned}\partial_h I_T(q)&=\frac{d}{ds}|_0 I_T(q+sh)=\int_0^T\frac{d}{ds}|_0 L(q(t)+sh(t),\dot q(t)+s\dot h(t))dt\\&=\int_0^T(D_1L(q,\dot q)h+D_2L(q,\dot q)\dot h)dt\\&=\int_0^T(D_1L(q,\dot q)-\frac{d}{dt}D_2L(q,\dot q))h\ dt+D_2L(q,\dot q)h\Big|_0^T.\end{aligned}$$

This completes the proof since $I_T^g(q)=g(q(0))+I_T(q)$ and $\partial_h\left[g(q(0))\right]=\nabla g(q(0))\cdot h(0)$. ■

Definition 3.9. A function $q \in C^2([0,T],\mathbb{R}^n)$ is said to solve the **Euler Lagrange equation for** L if q solves

$$D_1L(q,\dot{q}) - \frac{d}{dt}[D_2L(q,\dot{q})] = 0. \tag{3.23}$$

This is equivalently to q satisfying $DI_T(q)h = 0$ for all $h \in C^1([0,T],\mathbb{R}^n)$ which vanish on $\partial[0,T] = \{0,T\}$.

Let us note that the Euler Lagrange equations may be written as:

$$D_1L(q,\dot{q}) = D_1D_2L(q,\dot{q})\dot{q} + D_2^2L(q,\dot{q})\ddot{q}.$$

Corollary 3.10. *Any minimizer q (or more generally critical point) of $I_T^g(\cdot)$ must satisfy the Euler Lagrange Eq. (3.23) with the boundary conditions*

$$q(T) = x \text{ and } \nabla g(q(0)) = \nabla_{\dot{q}}L(q(0),q(0)) = D_2L(q(0),q(0)). \tag{3.24}$$

Proof. The corollary is a consequence Proposition 3.8 and the first derivative test which implies $DI_T^g(q)h = 0$ for all $h \in C^1([0,T],\mathbb{R}^n)$ such that $h(T) = 0$. ∎

Example 3.11. Let $U \in C^\infty(\mathbb{R}^n,\mathbb{R})$, $m > 0$ and $L(q,v) = \frac{1}{2}m\,|v|^2 - U(q)$. Then

$$D_1L(q,v) = -\nabla U(q) \text{ and } D_2L(q,v) = mv$$

and the Euler Lagrange equations become

$$-\nabla U(q) = \frac{d}{dt}\,[m\dot{q}] = m\ddot{q}$$

which are Newton's equations of motion for a particle of mass m subject to a force $-\nabla U$. In particular if $U = 0$, then $q(t) = q(0) + t\dot{q}(0)$.

The following assumption on L will be assumed for the rest of this section.

Assumption 1. We assume $\left[D_2^2L(q,v)\right]^{-1}$ exists for all $(q,v) \in \mathbb{R}^n \times \mathbb{R}^n$ and $v \to D_2L(q,v)$ is invertible for all $q \in \mathbb{R}^n$.

Notation 3.12. For $q,p \in \mathbb{R}^n$ let

$$V(q,p) := [D_2L(q,\cdot)]^{-1}(p). \tag{3.25}$$

Equivalently, $V(q,p)$ is the unique element of $\mathbb{R}^n$ such that

$$D_2L(q,V(q,p)) = p. \tag{3.26}$$

Remark 3.13. The function $V : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is smooth in (q,p). This is a consequence of the implicit function theorem applied to $\Psi(q,v) := (q, D_2L(q,v))$.

Under Assumption 1, Eq. (3.23) may be written as

$$\ddot{q} = F(q,\dot{q}) \tag{3.27}$$

where

$$F(q,\dot{q}) = D_2^2L(q,\dot{q})^{-1}\{D_1L(q,\dot{q}) - D(D_2L(q,\dot{q})\dot{q}\}.$$

Definition 3.14 (Legendre Transform)**.** Let $L \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n,\mathbb{R})$ be a function satisfying Assumption 1. The Legendre transform $L^* \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n,\mathbb{R})$ is defined by

$$L^*(x,p) := p \cdot v - L(x,v) \text{ where } p = \nabla_v L(x,v),$$

i.e.

$$L^*(x,p) = p \cdot V(x,p) - L(x,V(x,p)). \tag{3.28}$$

Proposition 3.15. *Let $H(x,p) := L^*(x,p)$, $q \in C^2([0,T],\mathbb{R}^n)$ and $p(t) := L_v(q(t),\dot{q}(t))$. Then*

(1) *$H \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ and*
$$H_x(x,p) = -L_x(x, V(x,p)) \text{ and } H_p(x,p) = V(x,p)..$$
(2) *H satisfies Assumption 1 and $H^* = L$, i.e. $(L^*)^* = L$.*
(3) *The path $q \in C^2([0,T],\mathbb{R}^n)$ solves the Euler Lagrange Eq. (3.23) then $(q(t),p(t))$ satisfies **Hamilton's Equations**:*
$$\begin{aligned}\dot{q}(t) &= H_p(q(t),p(t)) \\ \dot{p}(t) &= -H_x(q(t),p(t)).\end{aligned} \tag{3.29}$$
(4) *Conversely if (q,p) solves Hamilton's equations (3.29) then q solves the Euler Lagrange Eq. (3.23) and*
$$\frac{d}{dt}H(q(t),p(t)) = 0. \tag{3.30}$$

Proof. The smoothness of H follows by Remark 3.13.

(1) Using Eq. (3.28) and Eq. (3.26)
$$\begin{aligned}H_x(x,p) &= p \cdot V_x(x,p) - L_x(x,V(x,p)) - L_v(x,V(x,p))V_x(x,p) \\ &= p \cdot V_x(x,p) - L_x(x,V(x,p)) - p \cdot V_x(x,p) = -L_x(x,V(x,p)).\end{aligned}$$
and similarly,
$$\begin{aligned}H_p(x,p) &= V(x,p) + p \cdot V_p(x,p) - L_v(x,V(x,p))V_p(x,p) \\ &= V(x,p) + p \cdot V_p(x,p) - p \cdot V_p(x,p) = V(x,p).\end{aligned}$$
(2) Since $H_p(x,p) = V(x,p) = [L_v(x,\cdot)]^{-1}(p)$ and by Remark 3.13, $p \to V(x,p)$ is smooth with a smooth inverse $L_v(x,\cdot)$, it follows that H satisfies Assumption 1. Letting $p = L_v(x,v)$ in Eq. (3.28) shows
$$\begin{aligned}H(x,L_v(x,v)) &= L_v(x,v) \cdot V(x,L_v(x,v)) - L(x,V(x,L_v(x,v))) \\ &= L_v(x,v) \cdot v - L(x,v)\end{aligned}$$
and using this and the definition of H^* we find
$$\begin{aligned}H^*(x,v) &= v \cdot [H_p(x,\cdot)]^{-1}(v) - H(x,[H_p(x,\cdot)]^{-1}(v)) \\ &= v \cdot L_v(x,v) - H(x,L_v(x,v)) = L(x,v).\end{aligned}$$
(3) Now suppose that q solves the Euler Lagrange Eq. (3.23) and $p(t) = L_v(q(t),\dot{q}(t))$, then
$$\dot{p} = \frac{d}{dt}L_v(q,\dot{q}) = L_q(q,\dot{q}) = L_q(q,V(q,p)) = -H_q(q,p)$$
and
$$\dot{q} = [L_v(q,\cdot)]^{-1}(p) = V(q,p) = H_p(q,p).$$
(4) Conversely if (q,p) solves Eq. (3.29), then
$$\dot{q} = H_p(q,p) = V(q,p).$$
Therefore
$$L_v(q,\dot{q}) = L_v(q,V(q,p)) = p$$

and

$$\frac{d}{dt}L_v(q,\dot{q}) = \dot{p} = -H_q(q,p) = L_q(q,V(q,p)) = L_q(q,\dot{q}).$$

Equation (3.30) is easily verified as well:

$$\begin{aligned}\frac{d}{dt}H(q,p) &= H_q(q,p)\cdot\dot{q} + H_p(q,p)\cdot\dot{p} \\ &= H_q(q,p)\cdot H_p(q,p) - H_p(q,p)\cdot H_q(q,p) = 0.\end{aligned}$$

■

Example 3.16. Letting $L(q,v) = \frac{1}{2}m|v|^2 - U(q)$ as in Example 3.11, L satisfies Assumption 1,

$$V(x,p) = \left[\nabla_v L(x,\cdot)\right]^{-1}(p) = p/m$$

$$H(x,p) = L^*(x,p) = p\cdot\frac{p}{m} - L(x,p/m)) = \frac{1}{2m}|p|^2 + U(q)$$

which is the conserved energy for this classical mechanical system. Hamilton's equations for this system are,

$$\dot{q} = p/m \text{ and } \dot{p} = -\nabla U(q).$$

Notation 3.17. Let $\phi_t(x,v) = q(t)$ where q is the unique maximal solution to Eq. (3.27) (or equivalently 3.23)) with $q(0) = x$ and $\dot{q}(0) = v$.

Theorem 3.18. *Suppose $L \in C^\infty(\mathbb{R}^n\times\mathbb{R}^n,\mathbb{R})$ satisfies Assumption 1 and let $H = L^*$ denote the Legendre transform of L. Assume there exists an open interval $J\subset\mathbb{R}$ with $0\in J$ and $U\subset_o\mathbb{R}^n$ such that there exists a smooth function $x_0 : J\times U\to\mathbb{R}^n$ such that*

$$\phi_T(x_0(T,x), V(x_0(T,x), \nabla g(x_0(T,x))) = x. \tag{3.31}$$

Let

$$q_{x,T}(t) := \phi_t(x_0(T,x), V(x_0(T,x), \nabla g(x_0(T,x))) \tag{3.32}$$

so that $q_{x,T}$ solves the Euler Lagrange equations, $q_{x,T}(T) = x$, $q_{x,T}(0) = x_0(T,x)$ and $\dot{q}_{x,T}(0) = V(x_0(T,x), \nabla g(x_0(T,x))$ or equivalently

$$\partial_v L(q_{x,T}(0), \dot{q}_{x,T}(0)) = \nabla g(x_0(T,x)).$$

Then the function

$$S(T,x) := I_T^g(q_{x,T}) = g(q_{x,T}(0)) + \int_0^T L(q_{x,T}(t), \dot{q}_{x,T}(t))dt. \tag{3.33}$$

solves the Hamilton Jacobi Equation (3.17).

Conjecture 3.19. *For general g and L convex in v, the function*

$$S(t,x) = \inf\{g(q(0)) + \int_0^t L(q(\tau),\dot{q}(\tau))d\tau : q\in C^2([0,t],\mathbb{R}^n) \text{ with } q(t) = x\}$$

is a distributional solution to the Hamilton Jacobi Equation Eq. 3.17. See Evans to learn more about this conjecture.

Proof. We will give two proofs of this Theorem.

First Proof. One need only observe that the theorem is a consequence of Definition 3.14 and Proposition 3.15 and 3.5.

Second Direct Proof. By the fundamental theorem of calculus and differentiating past the integral,

$$\begin{aligned}\frac{\partial S(T,x)}{\partial T} &= \nabla g(x_0(T,x))\cdot\frac{\partial}{\partial T}x_0(T,x) + L(q_{x,T}(T),\dot q_{x,T}(T)) + \int_0^T \frac{\partial}{\partial T}L(q_{x,T}(t),\dot q_{x,T}(t))dt\\ &= \nabla g(x_0(T,x))\cdot\frac{\partial}{\partial T}x_0(T,x) + L(q_{x,T}(T),\dot q_{x,T}(T)) + DI_T(q_{x,T})\left[\frac{\partial}{\partial T}q_{x,T}\right]\\ &= L(q_{x,T}(T),\dot q_{x,T}(T)) + DI_T^g(q_{x,T})\left[\frac{\partial}{\partial T}q_{x,T}\right].\end{aligned} \tag{3.34}$$

Using Proposition 3.8 and the fact that $q_{x,T}$ satisfies the Euler Lagrange equations and the boundary conditions in Corollary 3.10 we find

$$DI_T^g(q_{x,T})\left[\frac{\partial}{\partial T}q_{x,T}\right] = \left(D_2L(q_{x,T}(t),\dot q_{x,T}(t))\frac{\partial}{\partial T}q_{x,T}(t)\right)\Big|_{t=T}. \tag{3.35}$$

Furthermore differentiating the identity, $q_{x,T}(T)=x$, in T implies

$$0 = \frac{d}{dT}x = \frac{d}{dT}q_{x,T}(T) = \dot q_{x,T}(T) + \frac{d}{dT}q_{x,T}(t)|_{t=T} \tag{3.36}$$

Combining Eqs. (3.34) – (3.36) gives

$$\frac{\partial S(T,x)}{\partial T} = L(x,\dot q_{x,T}(T))) - D_2L(x,\dot q_{x,T}(T))\dot q_{x,T}(T). \tag{3.37}$$

Similarly for $v\in\mathbb{R}^n$,

$$\begin{aligned}\partial_v S(T,x) &= \partial_v I_T^g(q_{x,T}) = DI_T^g((q_{x,T}))\left[\partial_v q_{x,T}\right]\\ &= D_2L(q_{x,T}(T),\dot q_{x,T}(T))\partial_v q_{x,T}(T) = D_2L(x,\dot q_{x,T}(T))v\end{aligned}$$

wherein the last equality we have use $q_{x,T}(T)=x$. This last equation is equivalent to

$$D_2L(x,\dot q_{x,T}(T)) = \nabla_x S(T,x)$$

from which it follows that

$$\dot q_{x,T}(T) = V(x,\nabla_x S(T,x)). \tag{3.38}$$

Combining Eqs. (3.37) and (3.38) and the definition of H, shows

$$\begin{aligned}\frac{\partial S(T,x)}{\partial T} &= L(x,V(x,\nabla_x S(T,x))) - D_2L(x,\dot q_{x,T}(T))V(x,\nabla_x S(T,x))\\ &= -H(x,\nabla_x S(T,x)).\end{aligned}$$

■

Remark 3.20. The hypothesis of Theorem 3.18 may always be satisfied locally, for let $\psi:\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}\times\mathbb{R}^n$ be given by $\psi(t,y):=(t,\phi_t(y,V(y,\nabla g(y)))$. Then $\psi(0,y):=(0,y)$ and so

$$\dot\psi(0,y) = (1,*) \text{ and } \psi_y(0,y) = id_{\mathbb{R}^n}$$

from which it follows that $\psi'(0,y)^{-1}$ exists for all $y \in \mathbb{R}^n$. So the inverse function theorem guarantees for each $a \in \mathbb{R}^n$ that there exists an open interval $J \subset \mathbb{R}$ with $0 \in J$ and $a \in U \subset_o \mathbb{R}^n$ and a smooth function $x_0 : J \times U \to \mathbb{R}^n$ such that

$$\psi(T, x_0(T,x)) = (T, x_0(T,x)) \text{ for } T \in J \text{ and } x \in U,$$

i.e.

$$\phi_T(x_0(T,x), V(x_0(T,x), \nabla g(x_0(T,x))) = x.$$

3.2. **Geometric meaning of the Legendre Transform.** Let V be a finite dimensional real vector space and $f : V \to \mathbb{R}$ be a strictly convex function. The graph of an $\alpha \in V^*$. defines a hyperplane which if translate by some amount called $-f^*(\alpha)$ just touches the graph of f at one point, say v, see Figure 12. That is to

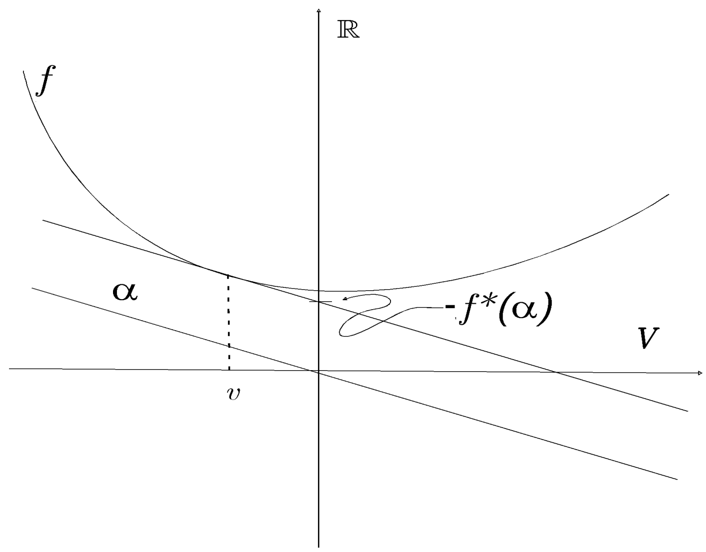

FIGURE 12. Legendre Transform of f.

say, $-f^*(\alpha) + \alpha(v) = f(v)$.

Now suppose further that f is smooth and $f'' > 0$. At the point of contact, v, α and f have the same tangent plane and since α is linear this means that $f'(v) = \alpha$. The function $f^* : V^* \to \mathbb{R}$ defined by this means is called the Legendre transform of f and is given explicitly by

$$f^*(\alpha) = \alpha(v) - f(v) \text{ with } v \text{ such that } f'(v) = \alpha, \text{ i.e. } \alpha = (f')^{-1}(v).$$

The Legendre transform above applied to $L(x, v)$ is this transform applied to $f(v) := L(x, v)$.