March 6, 2013
notes

Legendre transform

Let M be a manifold, $L : TM \to \mathbb{R}$ a Lagrangian.
The Legendre transform

$$\mathcal{L}_L : TM \to T^*M$$

is defined as follows: for $q \in M$, $v \in T_q M$,
$\mathcal{L}_L(q,v)$ is the unique covector in $T_q^* M$ so that

$$(*) \quad \boxed{\mathcal{L}_L(q,v)\, w = \frac{d}{dt}\Big|_0 L(q, v+tw)}$$

for any $w \in T_q M$

Let's see what this means in coordinates.
Let $(q_1 \ldots q_n) : \mathcal{U} \to \mathbb{R}^n$ be coordinates on M.
Then $(q_1 \ldots q_n, v_1 \ldots v_n)$ are the corresponding coordinates on $T\mathcal{U} \subseteq TM$ (recall that for $w \in T_q M$

$$v_i(w) = (dq_i)_q(w) \;)$$

and $(q_1 \ldots q_n, p_1 \ldots p_n)$ the coordinates on $T^*\mathcal{U} \subseteq T^*M$
(for $\eta \in T_q^* M$,

$$p_i(\eta) = \eta\left(\frac{\partial}{\partial q_i}\Big|_q\right) \Big)$$

$$L(q,v) = L(q_1 \ldots q_n, v_1 \ldots v_n)$$

For any $w = (w_1 \ldots w_n) \in T_q \mathcal{U}$,

$$\frac{d}{dt}\Big|_0 L(q, v+tw) = \frac{d}{dt}\Big|_0 L(q_1 \ldots q_n, v_1 + tw_1, \ldots v_n + tw_n)$$

$$= \sum_{i=1}^n \frac{\partial L}{\partial v_i}(q,v)\, w_i = \sum \frac{\partial L}{\partial v_i}(q,v)\, dq_i(w)$$

$$\Rightarrow \mathcal{L}_L(q_1 \ldots q_n, v_1 \ldots v_n) = \mathcal{L}_L\left(q, \sum v_i \frac{\partial}{\partial q_i}\right) = \left(q, \sum \frac{\partial L}{\partial v_i}(q,v) (dq_i)_q\right)$$

$$= \left(q_1 \ldots q_n, \frac{\partial L}{\partial v_1}(q,v), \ldots \frac{\partial L}{\partial v_n}(q,v)\right)$$

If $L : TM \to \mathbb{R}$ is a regular Lagrangian then

__Claim__ $\mathcal{L}_L : TM \to T^*M$ is a local diffeomorphism

__Proof__ We compute the differential of $\mathcal{L}_L$ in coordinates and apply the inverse function theorem.

Since $\mathcal{L}_L(q_1, \dots, q_n, v_1, \dots, v_n) = (q_1 \dots q_n, \frac{\partial L}{\partial v_1}(q,v), \dots, \frac{\partial L}{\partial v_n}(q,v))$

$$(D\mathcal{L}_L)(q,v) = \left(\begin{array}{c|c} \frac{\partial q_i}{\partial q_j}(q,v) & \frac{\partial q_i}{\partial v_j}(q,v) \\ \hline \frac{\partial^2 L}{\partial q_i \partial v_j}(q,v) & \frac{\partial^2 L}{\partial v_i \partial v_j}(q,v) \end{array}\right) = \left(\begin{array}{c|c} \begin{matrix} 1 & & 0 \\ 0 & & 1 \end{matrix} & 0 \\ \hline // & \frac{\partial^2 L}{\partial v_i \partial v_j}(q,V) \end{array}\right)$$

$$\Rightarrow \det(D\mathcal{L}_L(q,v)) \neq 0 \iff \det\left(\frac{\partial^2 L}{\partial v_i \partial v_j}(q,v)\right) \neq 0$$

Recall: $L : TM \to \mathbb{R}$ is regular $\iff \det\left(\frac{\partial^2 L}{\partial v_i \partial v_j}(q,v)\right) \neq 0$.

Hence the $(D\mathcal{L}_L)(q,v)$ is invertible for all $(q,v) \in TM$

$\Rightarrow$ $\mathcal{L}_L : TM \to T^*M$ is a <u>local</u> diffeo. $\square$

__Proposition__ If $\left(\frac{\partial^2 L}{\partial v_i \partial v_j}(q,v)\right)$ is positive definite for all $(q,v) \in TM$, then

$$\mathcal{L}_L : TM \to T^*M$$

is <u>globally</u> 1-1

__Proof__ Note that since $\forall q \in M$

$$\mathcal{L}_L(T_qM) \subseteq T_q^*M$$

we only need to show that

$$\mathcal{L}_L : T_qM \to T_q^*M$$

is 1-1 for every $q \in M$.

Therefore it is enough to prove: suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a function such that $\forall x \in \mathbb{R}^n$ the matrix of second partials $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)$ is positive definite.

Then $\mathcal{L}_f : \mathbb{R}^n \to (\mathbb{R}^n)^*$, defined by

$$\boxed{\mathcal{L}_f(x)(w) = \frac{d}{dt}\Big|_0 f(x+tw) \quad \text{for all } w \in \mathbb{R}^n \quad \text{is 1-1}}$$

We first consider the case $n=1$. Then $f: R \to \mathbb{R}$ is just a function of 1 variable and $\mathcal{L}_f(x): T_x\mathbb{R} \to \mathbb{R}$ is the differential df_x. In fact

$$\mathcal{L}_f(x)\, w = \frac{d}{dt}\Big|_0 f(x+tw) = f'(x)\, w \qquad (= df_x(w))$$

We can think of $\mathcal{L}_f : \mathbb{R} \to (\mathbb{R}^*)$ as sending $x \in \mathbb{R}$ to the linear functional $\ell : \mathbb{R} \to \mathbb{R}$,

$$\ell(w) = f'(x)\, w.$$

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right) = \left(f''(x)\right) \qquad [\text{a } 1\times 1 \text{ matrix}]$$

So "$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)$ is positive definite" means:

"$f''(x) > 0$ for all x."

We now identify $\ell \in (\mathbb{R})^*$ with its slope.

With this identification, $\mathcal{L}_f : \mathbb{R} \to \mathbb{R}$, $\mathcal{L}_f(x) = f'(x)$.

Now, for $\ell \in \mathbb{R}$,

$$\ell = \mathcal{L}_f(x_0) \Leftrightarrow f'(x_0) = \ell \Leftrightarrow \left(f(x) - \ell\cdot x\right)'(x_0) = 0$$

i.e. $\ell = \mathcal{L}_f(x_0) \Longleftrightarrow x_0$ is a critical point of

$$f_\ell(x) := f(x) - \ell \cdot x$$

Since $\frac{d^2}{dx^2} f_\ell = \frac{d}{dx}\left(f'(x) - \ell\right) = f''(x) > 0$ for all x, the function $f_\ell(x)$ has <u>only one critical point</u>; This point is a global minimum of $f_\ell(x)$.

It corresponds to the point $x \in \mathbb{R}$ where the tangent line to the graph $y = f(x)$ has slope ℓ.

Consequently $\mathcal{L}_f : \mathbb{R} \to \mathbb{R}$ is 1-1.

Remark $\mathcal{L}_f$ need not be onto:
consider $f(x) = e^x$.
Convince yourself that $\mathcal{L}_f(\mathbb{R})$ = lines with <u>positive</u> slope

---

We now tackle the case of $n > 1$; ie we consider $f \in C^\infty(\mathbb{R}^n)$ with $d^2 f(x) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right) > 0$ ("> 0" here means positive definite. Richard should think of Sylvester's criterion and explain it to Ven, Bryce and Bill).

Lemma For $h \in C^\infty(\mathbb{R}^n)$ with $d^2 h(x) > 0$ there is <u>at most one</u> critical point.

Proof If $y, z \in \mathbb{R}^n$ are two critical points of f with $y \neq z$, consider

$$g(t) = f(ty + (1-t)z) \in C^\infty(\mathbb{R})$$

By the chain rule,

$$g'(t) = df(ty+(1-t)z) \cdot (y-z)$$
$$= \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(q) y_i - \frac{\partial f}{\partial x_i}(q) z_i \right)$$

where $q = ty + (1-t)z$.

$$\Rightarrow g''(t) = \sum_{i,j} \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(q) y_i y_j - \frac{\partial^2 f}{\partial x_j \partial x_i}(q) y_i z_j - \frac{\partial^2 f}{\partial x_j \partial x_i} z_i y_j + \frac{\partial^2 f}{\partial x_j \partial x_i}(q) z_i z_j \right)$$

$$= \sum_{i,j} \frac{\partial^2 f}{\partial x_j \partial x_i}(q)(y_i - z_i)(y_j - z_j)$$

$$= (y-z)^T \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(q) \right)(y-z) > 0$$

since $y - z \neq 0$ and $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(q) \right) > 0$.

But $g'(0) = df(z) \cdot (y-z) = 0$ since $df(z) = 0$
and $g'(1) = df(y) \cdot (y-z) = 0$ since $df(y) = 0$

Since $g''(t) > 0$ for all t, $g'(t)$ is strictly increasing.
This contradicts $g'(0) = g'(1)$.

$\Rightarrow \quad y = z.$

This prove the lemma. $\square$

We now argue that $\mathcal{L}_f : \mathbb{R}^n \to (\mathbb{R}^n)^*$ is 1-1:

$l = \mathcal{L}_f(x)$ for some $x \in \mathbb{R}^n$

$\Leftrightarrow \quad l = df(x)$

$\Leftrightarrow \quad df_l(x) = 0 \qquad$ where $f_l(x) = f(x) - l(x)$

But $f_l(x)$ has only one critical point by Lemma.

$\Rightarrow \quad \mathcal{L}_f$ is 1-1.

$\square$

<u>Exercise 1</u> Let $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be an inner product; i.e. a symmetric positive definite bilinear form.
($g(x,y) = x^T A y$ for a positive definite matrix A)
Let $f(x) = \frac{1}{2} g(x,x)$; the corresponding quadratic form. Show that for any $x \in \mathbb{R}^n$

$$df(x) \in (\mathbb{R}^n)^*$$

is given by $\quad df(x)\,w = g(x, w) \qquad$ for all $w \in \mathbb{R}^n$.
Conclude that

$$\mathcal{L}_f : \mathbb{R}^n \to (\mathbb{R}^n)^*, \quad x \mapsto df(x)$$

is given by

$$\mathcal{L}_f(x) = g(x, \cdot).$$

Explain why it follows that $\mathcal{L}_f$ is a diffeomorphism.

Exercise 2 Consider the Lagrangian $L : T\mathbb{R}^n \to \mathbb{R}$ of the form

$$L(q,v) = \frac{1}{2} g(v,v) - V(q)$$

where $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a positive definite inner product and $V \in C^\infty(\mathbb{R}^n)$ a smooth function.
Show that the Legandre transform

$$\mathcal{L}_L : T\mathbb{R}^n \to T^*\mathbb{R}^n \cong \mathbb{R}^n \times (\mathbb{R}^n)^*$$

is given by

$$\mathcal{L}_L(q,v) = (q, \; g(v, \cdot))$$

Exercise 3 Let $(x_1 \ldots x_n) : \mathcal{U} \to \mathbb{R}^n$ and $(y_1 \ldots y_n) : \mathcal{U} \to \mathbb{R}^n$ be two sets of coordinates on a manifold M.
Let $(x_1 \ldots x_n, \eta_1 \ldots \eta_n) : T^*\mathcal{U} \to \mathbb{R}^n \times \mathbb{R}^n$
and $(y_1 \ldots y_n, \xi_1 \ldots \xi_n) : T^*\mathcal{U} \to \mathbb{R}^n \times \mathbb{R}^n$
be the corresponding coordinates on $T^*\mathcal{U} \subseteq T^*M$.
Show that

$$\sum_{i=1}^{n} \eta_i \, dx_i = \sum_{j=1}^{n} \xi_j \, dy_j$$

Conclude that there is a 1-form $\alpha \in \Omega^1(T^*M)$ with

$$\alpha|_{T^*\mathcal{U}} = \sum \eta_i \, dx_i$$

for any coordinate chart $(x_1 \ldots x_n)$ on M.

[More traditionally α is written as $\sum p_i dq_i$.
Note that $\omega = d\alpha$ is a symplectic form:

$$d\omega = d(d\alpha) = d\left(\sum dp_i \wedge dq_i\right) = 0$$

and $\omega = \sum dp_i \wedge dq_i$ is nondegenerate in any local coordinate chart.