Additional material for Chapter 4, Hamilton-Jacobi theory

Chapter 4 concerns the equation

$$I_s + H(s, q, I_q) = 0, \tag{1}$$

which is split as

$$I_s + H(s, q, p) = 0, \quad p = I_q.$$

Differentiating (1) with respect to q we get

$$I_{sq} + H_q + H_p I_{qq} = 0$$

Consider $q = q(s)$ as unknown function of s and define $p(s) = I_q(s, q(s))$ and $z(s) = I(s, q(s))$. If we demand that $\dot{q} = H_p$ along the curve $s \to (q(s), p(s))$, then

$$\dot{p} = \frac{dp}{ds} = \underbrace{I_{qs}(s, q(s))}_{-H_q - H_p I_{qq}} + I_{qq}(s, q(s))\dot{q}(s) = -H_q + (\underbrace{\dot{q} - H_p}_{0})I_{qq} \quad \text{and} \quad \dot{z} = \underbrace{I_s + I_q\dot{q}}_{-H+pH_p}.$$

So

$$\dot{q} = H_p, \quad \dot{p} = -H_q, \quad \dot{z} = -H + pH_p,$$

in which $H = H(s, q(s), p(s))$. Chapter 4 shows, among other things, how solutions of $\dot{q} = H_p, \dot{p} = -H_q$ define solutions of (1) by considering the integral $I(q, s)$ below, starting from the standard Lagrangian integral

$$I = \int L(t, x(t), \dot{x}(t)dt$$

over some bounded time interval which will be renamed later. I am skipping indices for x and $\dot{x}$ throughout these notes. It is a minimal typographical operation to put them in later.

What was F and u before in Chapter 1 is now L and x, and p will be neither $\dot{u}$ nor $\dot{x}$: the letter p will be used for

$$p = \frac{\delta L}{\delta \dot{x}} = L_{\dot{x}}, \quad L = L(t, x, \dot{x}), \quad H(t, x, p) = p\dot{x} - L(t, x, \dot{x}),$$

in which $\dot{x}$ is a symbol for now. Invertibility of the (t, x)-dependent transformation $\dot{x} \to p$ corresponds to

$$L(t, x, \dot{x}) + H(t, x, p) = p\dot{x} \quad \text{with} \quad p = L_{\dot{x}} \iff \dot{x} = H_p \quad \text{and} \quad L_x + H_x = 0.$$

Exercise 1. Assuming that L is C^2 with L_{xx} invertible: prove these equalities without making obvious mistakes, and show that the Lagrangian equations

$$\frac{d}{dt}L_{\dot{x}} = L_x \qquad \text{(i.e.} \quad \frac{d}{dt}L_{\dot{x}}(t, x(t), \dot{x}(t) = L_x(t, x(t), \dot{x}(t))\,)$$

are equivalent to the Hamiltonian equations

$$\dot{x} = H_p;\ \dot{p} = -H_x \qquad \text{(i.e.} \quad \dot{x}(t) = H_p(x(t), p(t));\ \dot{p}(t) = -H_x(x(t), p(t))\,).$$

Distinguish carefully between symbols $x, \dot{x}, p$ and functions $x(t), \dot{x}(t), p(t)$, between partial and total derivatives, and show that $\dot{H} = H_t$. Generalize to $x, \dot{x}, p \in \mathbb{R}^n$.

Exercise 2. In a more general setting, with $L = L(y)$ (think of $y = \dot{x}$), if $\Omega \subset \mathbb{R}^n$ is convex and open, and if $L \in C^k(\Omega)$ with $k \geq 2$ has a matrix of second derivatives which is positive definite throughout Ω (in other words: $L_{yy} > 0$), then the map

$$\phi : y \to p = L_y \in \mathbb{R}^n$$

is locally a C^{k-1} diffeomorphism, as a direct consequence of the inverse function theorem.

- Prove that ϕ is injective. Hint: assume first that $0 \in \Omega$ and $\phi(0) = 0$ and examine $\phi(y) = \int_0^1 \frac{d}{dt}\phi(ty)\,dt$ and $y \cdot \phi(y)$.
- Prove that $\Omega^* = \phi(\Omega)$ is open.
- Thus the inverse map $\psi : p \in \Omega^* \to y \in \Omega$ exists and is C^{k-1}. Define $L^*(p) = p \cdot y - L(y)$. Prove that $L^* : \Omega^* \to \mathbb{R}$ is C^k with $L^*_{pp} > 0$. Hint: show that the first order derivatives of ψ and ϕ have the same smoothness.
- Prove that $y = \psi(p)$ globally maximizes $p \cdot y - L(y)$. Thus

  $$L^*(p) = \max_{y \in \Omega} p \cdot y - L(y),$$

  which is usually called the Legendre transform of L.
- Explain why $p \cdot y \leq L(y) + L^*(p)$ with equality only if $p = \phi(y) = L_y$.
- Assume that Ω^* is convex. Show that $(L^*)^* = L$. Hint: use the symmetry above.
- Example: let $1 < s < \infty$ and $L(y) = \frac{1}{s}|y|^s$. Determine $L^*(p)$. Hint: we wrote s instead of p because p is already in use.
- Suppose $\Omega = \mathbb{R}^n$ and $\alpha > 0$. Express the Legendre transform of $y \to L(\alpha y)$ in terms of L^*.

The chapter is concerned about I as a function of the boundary conditions

$$x(\sigma) = \kappa \quad \text{and} \quad x(s) = q \quad \text{leading to} \quad I = I(s, q) = I(s, \sigma, q, \kappa), \tag{2}$$

and finding stationary points for $I(s, \sigma, q, \kappa)$ when s and q are fixed and σ and κ are varied over a manifold of the form $T(\sigma, \kappa) = 0$. First however we consider the case that σ and κ are fixed, say $\sigma = 0 \in \mathbb{R}$ and $\kappa = 0 \in \mathbb{R}^n$ (with notationally $n = 1$). Thus

$$I(s, q) = \int_0^s L(t, x(t), \dot{x}(t)dt$$

in which $x(t)$ solves

$$\frac{d}{dt}L_{\dot{x}} = L_x, \quad x(0) = 0, \quad x(s) = q.$$

Exercise 3. Take a simple example, say

$$L(t, x, \dot{x}) = L(x, \dot{x}) = \frac{1}{2}\dot{x}^2 - V(x),$$

with V a smooth function. Assume that $x = x(t; s, q)$ is a solution which depends smoothly on s and q. Find differential equations and boundary conditions for x_s and x_q as we did in the course and derive a first order partial differential equation for $I(s, q)$.

Exercise 4. We did not discuss the solvability of the boundary value problems for x_s and x_q. Consider the general case, i.e. with (2). Suppose that $I(s_0, q_0)$ is realised by a solution $x = x(t)$. Explain why you can compute I_s, I_q, I_σ and I_κ if (s, q) and (σ, κ) are not conjugate along the solution.

Exercise 5. (continued) Keep σ and κ fixed. Derive that $I_s + H = 0$ and explain which arguments you should have in H.

Background: characteristics

In relation to the first order equation for $I = I(s, q)$ we encountered you may have seen the following. I am using the notation that Evans uses in his PDE book.

Let $F = F(x, z, p)$ be a function of (x, z, p). A general first order equation ordinary differential equation

$$F(x, u(x), u_x(x)) = 0$$

for $u = u(x)$ can be solved by putting

$$x = x(\tau), \quad z = z(\tau) = u(x(\tau)), \quad p = p(\tau) = u_x(x(\tau)).$$

Differentiating $F(x(\tau), z(\tau), p(\tau)) = 0$ with respect to τ we get, omitting the arguments,

$$F_x\dot{x} + F_z\dot{z} + F_p\dot{p} = 0.$$

Here dots denoting differentiation with respect to the artificial time variable τ. Since

$$\dot{z} = u_x\dot{x} = p\dot{x},$$

we must have

$$(F_x + pF_z)\dot{x} + F_p\dot{p} = 0,$$

which is certainly the case if we put

$$\dot{x} = F_p, \quad \dot{p} = -F_x - pF_z,$$

whence

$$\dot{z} = p\dot{x} = pF_p.$$

We have thus "reduced" a simple first order ODE to a 3-dimensional system of first ODE's for x, p, z. The right hand side of $\dot{z} = \cdots$ has to be rewritten using $F = 0$, depending on the exact form of the equation.

With $p_i = u_{x_i}$ you easily see that first order PDE's

$$F(x_1, \dots, x_n, u, u_{x_1}, \dots, u_{x_n}) = 0$$

in n variables lead to (numbering $i = 1, \dots, n$)

$$\dot{x}_i = F_{p_i}, \quad \dot{p}_i = -F_{x_i} - p_i F_z, \quad \dot{z} = \sum_{j=1}^{n} p_j F_{p_j}.$$

This is called the method of characteristics. Note it may happen that the right hand sides F_{p_i} in the equations

$$\dot{x}_i = F_{p_i}$$

depend only on the independent variables $x_1, \dots, x_n$. Solution curves of this system in n-dimensional x-space are called characteristics.

Of course you may also treat equations of the form

$$u_t + H(x, t, u, u_x) = 0,$$

Exercise 6. Show that you get

$$\dot{x} = H_p, \quad \dot{p} = -H_x - pH_z, \quad \dot{z} = pH_p - H$$

(in which H depends on t, x, z, p).

Exercise 7. Show that equations of the form

$$u_t + H(x_1, \dots, x_n, t, u, u_{x_1}, \dots, u_{x_n}) = 0,$$

lead to

$$\dot{x}_i = H_{p_i}, \quad \dot{p}_i = -H_{x_i} - p_i H_z, \quad \dot{z} = \sum_{j=1}^{n} p_j H_{p_j} - H$$