Notes for Gelfand and Fomin's *Calculus of Variations*

Linus Setiabrata

Todd "Tood" Lensman and I have made a deal: He has challenged me to read Gelfand and Fomin's *Calculus of Variations*, whereas I have challenged him to read Fomin, Williams, and Zelevinsky's *Introduction to Cluster Algebras, Ch 1–3*. Here are my notes, made mostly for my personal use and for proof that I actually did the reading.

Contents

1 Ch I, *Elements of the Theory*

1.1 Sec 1, *Functionals. Some Simple Variational Problems*

One studies functionals and their extrema. A typical example is given by the functional

$$J[y] \stackrel{\text{def}}{=} \int_a^b F(x, y, y')\, dx,$$

where F is some C^2 function of three variables. Examples include functions of the form $F = f(x,y)\sqrt{1+(y')^2}$, corresponding to functionals which are integrating $f(x,y)$ along the curve $y = y(x)$ with respect to arclength $ds = \sqrt{1+(y')^2}\, dx$.

Functionals eat functions and spit out numbers. Hence they are defined on some subset of the set of all functions. They're usually defined on some subset of continuous (or even nicer, e.g. C^r) functions, e.g. on the set of continuous functions with prescribed endpoints. Because the set of continuous functions are normed vector spaces, it helps to think about them a little bit.

1.2 Sec 2, *Function Spaces*

Normed vector spaces are vector spaces equipped with a norm. On the normed vector space $\mathscr{D}_1(a,b)$ of C^1 functions on $[a,b]$, endowed with the norm $\max_{x\in[a,b]} f(x) + \max_{x\in[a,b]} f'(x)$, the functional

$$J[y] \stackrel{\text{def}}{=} \int_a^b F(x, y, y')\, dx$$

is continuous. But on the normed vector space $\mathscr{C}(a,b)$ of continuous functions on $[a,b]$, endowed with the norm $\max_{x\in[a,b]} f(x)$, the arclength functional is not continuous.

A complete normed vector space is called a *Banach space*. The space $\mathscr{C}(a,b)$ is a Banach space. The set $\overline{\mathscr{C}}(a,b)$ of continuous functions on $[a,b]$ endowed with the L^2 norm, however, is a normed vector space that is *not* Banach [Ch I, Ex 3 – 4].

Cute remark: A normed vector space with countable dimension is never complete. This follows from the Baire Category Theorem, apparently. For example, it's easy to see that the set of polynomials on $[0,1]$, with the sup-norm, is a noncomplete normed vector space: the Cauchy sequence (p_n) given by

$$p_n \stackrel{\text{def}}{=} \sum_{k=1}^{n} \frac{x^k}{2^k}$$

wants to converge to $x \mapsto (1 - x/2)^{-1}$, but this is not a polynomial.

1.3 Sec 3, *The Variation of a Functional. A Necessary Condition for an Extremum*

If y is to be an extremum of the functional $J[y]$, then its *variational derivative* $\delta J[y]$ must equal the zero linear functional. (We consider *differentiable* functionals, which are functionals such that for every admissible curve y the difference

$$\Delta J[y] \stackrel{\text{def}}{=} J[y+h] - J[y]$$

can be written as $\Delta J[y] = \varphi[h] + \varepsilon||h||$, where φ is a linear functional and $\varepsilon \to 0$ as $||h|| \to 0$; then δJ is the linear functional φ.)

A series of cute lemmas, beginning with

$$\int \alpha(x) h(x)\, dx = 0$$

for enough $h(x)$ implies $\alpha(x) = 0$, are proven. In light of the $\delta J[y] = 0$ condition, these cute lemmas turn out to be important.

1.4 Sec 4, *The Simplest Variational Problem. Euler's Equation*

Suppose we are considering functionals $J[y]$ defined on the set of continuous functions with fixed endpoints $y(a) = A$ and $y(b) = B$. After spelling everything out, that an extremum y must satisfy $\delta J[y] = 0$ is the same as saying that y solves *Euler's equation*, which says $F_y - \frac{d}{dx}F_{y'} = 0$. This uses structural lemmas from Sec 3.

With more assumptions on F (e.g. that it doesn't depend explicitly on x, or on y, or on y'), Euler's equation becomes a first order differential equation for which we should be confident we are able to solve. Let me single out the case where $F = f(x, y)\sqrt{1 + (y')^2}$, in which case Euler's equation can be transformed into

$$f_y - f_x y' - f\frac{y''}{1 + (y')^2} = 0.$$

Let's also be clear that f is given to us, and we are solving for y; in other words, we have reduced the problem of finding extrema of a functional to the problem of solving a (nasty?) ordinary differential equation.

Suppose we want to find the curve joining two given points (x_0, y_0) and (x_1, y_1) such that the surface obtained by rotating the curve about the x-axis has the minimal area. The relevant functional is

$$J[y] = 2\pi \int_{x_0}^{x_1} y\sqrt{1 + (y')^2}\, dx,$$

which does not depend explicitly on x. Using relevant structural results, we see that the general solution to Euler's equation is

$$y = C \cosh \frac{x + C_1}{C}.$$

Thus, the required curve is a *catenary*. The surface is a *catenoid*.

1.5 Sec 5, *The Case of Several Variables*

One may also consider, instead of functionals $J[y]$ eating curves $y = y(x)$ and spitting out numbers, functionals $J[z]$ eating surfaces $z = z(x, y)$ and spitting out numbers. A typical example is given by the functional

$$J[z] = \iint_R F(x, y, z, z_x, z_y)\, dxdy.$$

Rebuilding the analogous theory, every extremum $z = z(x, y)$ of J satisfies $\delta J[z] = 0$ and hence *Euler's equation*

$$F_z - \frac{\partial}{\partial x}F_{z_x} - \frac{\partial}{\partial y}F_{z_y} = 0.$$

As in Sec 4, concluding that Euler's equation is satisfied requires some structural lemmas: in this case, a lemma asserting that

$$\iint \alpha(x, y)h(x, y)\, dxdy = 0$$

for enough $h(x, y)$ implies $\alpha(x, y) = 0$ is used crucially.

Suppose we want to find the surface of least area spanned by a given contour. The relevant functional is

$$J[z] = \iint_R \sqrt{1 + z_x^2 + z_y^2}\, dxdy.$$

After a hairy computation, the condition that Euler's equation is satisfied precisely says that the mean curvature of $z = z(x, y)$ is zero everywhere. These are called *minimal surfaces*.

1.6 Sec 6, *A Simple Variable Endpoint Problem*

Instead of functionals merely defined on the set of continuous functions with fixed endpoints $y(a) = A$ and $y(b) = B$, one may consider functionals $J[y]$ defined on the set of continuous functions with endpoints on the vertical lines $x = a$ and $x = b$. An extremum y of J, say with $y(a) = A$ and $y(b) = B$, is simultaneously an extremum of the functional obtained by restricting J to continuous functions satisfying $y(a) = A$ and $y(b) = B$; in particular, an extremum y of J must still satisfy Euler's equation

$$F_y - \frac{d}{dx}F_{y'} = 0.$$

This condition actually reflects the fact that restricting the linear functional $\delta J[y] = 0$ to the set of continuous functions with $y(a) = A$ and $y(b) = B$ implies Euler's equation must hold. But $\delta J[y] = 0$ on all continuous functions with endpoints at $x = a$ and $x = b$; this says equivalently that

$$F_{y'}|_{x=b}h(b) - F_{y'}|_{x=a}h(a) = 0.$$

Since h is arbitrary this implies both

$$F_{y'}|_{x=b} = 0 \quad \text{and} \quad F_{y'}|_{x=a} = 0.$$

1.7 Sec 7, *The Variational Derivative*

The variational derivative of a functional is the analogue of a partial derivative. It is denoted and defined by

$$\left.\frac{\delta J}{\delta y}\right|_{x=x_0} \stackrel{\text{def}}{=} \lim_{\Delta\sigma\to 0} \frac{J[y+h] - J[y]}{\Delta\sigma},$$

where $h(x)$ is nonzero in a neighborhood around x_0, and $\Delta\sigma$ is the area between $y = y(x)$ and $y = y(x) + h(x)$; the limit is taken in such a way that both $\max|h(x)|$ and the length of the interval in which $h(x)$ is nonvanishing go to 0. Right now, it's not clear to me what the significance of this construction is. It seems that nothing in Ch I uses this concept.

[$\star$But many things in Ch II use this!$\star$]

1.8 Sec 8, *Invariance of Euler's Equation*

Finally, one can prove that solutions to Euler's equation are preserved under curvilinear changes of variables. Although showing this is a straightforward task, it seems to have powerful consequences (as is the case in usual calculus, I guess).

2 Ch II, *Further Generalizations*

[$\star$I am a little disturbed by the quantity $F_{y'}$, for a function $F(x, y, y')$. If I understand correctly, for a function $F(x, y, z)$, we denote by $F_{y'} = F_{y'}(x, y, y')$ the quantity $F_z(x, y, z)|_{z=y'}$. I presume, then, when we have more complicated computations that they are all to be done first with $F(x, y, z)$ and then substitute $z = y'$ at the very end. But at times in the book it seems we are taking the actual quantities $F(x, y, y')$ and doing the computations to them. Then there are questions about whether substitution commutes with whatever we are doing. I can't really phrase the question in a precise way, though...$\star$]

[$\star$I guess the variables y and z are independent quantities in the algebraic sense, i.e., there is no polynomial relation between them; similarly the quantities y and y' are algebraically independent, hence substitution $z = y'$ and algebraic manipulation commute.$\star$]

2.9 Sec 9, *The Fixed End Point Problem for* n *Unknown Functions*

Let me fix notation $\mathbf{y} = (y_1, \dots, y_n)\colon \mathbb{R} \to \mathbb{R}^n$ and $\mathbf{y}' = (y_1', \dots, y_n')$. Consider functionals of the form

$$J[\mathbf{y}] = \int_a^b F(x, \mathbf{y}, \mathbf{y}')\, dx,$$

and let us try to find extremals $\mathbf{y}$ subject to $\mathbf{y}(a) = \mathbf{A}$ and $\mathbf{y}(b) = \mathbf{B}$. I found it useful to think about x as a "time variable" and $\mathbf{y}$ as a curve in $\mathbb{R}^n$, but the book uses the language of n unknown functions y_i treated separately. In any case, an extremum $\mathbf{y}$ satisfies $\delta J[\mathbf{y}] = 0$. By varying each y_i separately we see that Euler's equation must hold coordinatewise, i.e. we have

$$F_{y_i} - \frac{d}{dx} F_{y_i'} = 0 \qquad \text{for all } i \in [n].$$

An example of a problem that can be solved is that of finding geodesics on a manifold. Let $\mathbf{r} = \mathbf{r}(u, v)$ be a surface. Then a curve on this surface can be parametrized with $u = u(t)$ and $v = v(t)$. The arclength is given by the functional

$$J[u, v] = \int_{t_1}^{t_2} \sqrt{E(u')^2 + 2Fu'v' + G(v')^2}\, dt,$$

where E, F, and G are coefficients of the first fundamental form of the surface.

2.10 Sec 10, *Variational Problems in Parametric Form*

Taking the previous perspective a bit further, suppose that rather than considering functionals of curves $y = y(x)$, we considered functionals of parametrized curves $(x, y) = (x(t), y(t))$. In this case,

$$\int_{t_0}^{t_1} F\left(x(t), y(t), \frac{\dot{y}(t)}{\dot{x}(t)}\right) \dot{x}(t)\, dt = \int_{t_0}^{t_1} \Phi(x, y, \dot{x}, \dot{y})\, dt$$

becomes a functional Φ in terms of x, y, $\dot{x}$, and $\dot{y}$ (here, the overdot is a derivative with respect to t). In particular, this is of the form $F(x, y_1, y_2, y_1', y_2')$ as in the previous section, with the additional structure that Φ does not depend on x, and further that Φ is positive-homogeneous in $\dot{x}$ and $\dot{y}$, i.e.,

$$\Phi(x, y, \lambda\dot{x}, \lambda\dot{y}) = \lambda\Phi(x, y, \dot{x}, \dot{y}).$$

The functional obtained by integrating such a Φ is independent of parametrization of the curve $(x(t), y(t))$. A certain converse of this is proven, namely, a necessary and sufficient condition for

$$\int_{t_0}^{t_1} \Phi(t, x, y, \dot{x}, \dot{y})\, dt$$

to depend only on the curve in the xy-plane defined by the parametric equations $x = x(t)$ and $y = y(t)$, and not on the parametrization, is that Φ should not depend on t explicitly, and should be positive-homogeneous of degree 1 in t.

2.11 Sec 11, *Functionals Depending on Higher-Order Derivatives*

Suppose we were considering not just

$$\int_a^b F(x, y, y')\, dx,$$

but rather, more generally,

$$\int_a^b F(x, y, y', \dots, y^{(n)})\, dx.$$

Thus we are considering curves $y = y(x) \in \mathscr{D}_n(a, b)$ satisfying the conditions

$$\begin{aligned} y(a) &= A_0, & y'(a) &= A_1, & \dots, && y^{(n-1)}(a) &= A_{n-1}, \\ y(b) &= B_0, & y'(b) &= B_1, & \dots, && y^{(n-1)}(b) &= B_{n-1}. \end{aligned}$$

The usual yoga, setting $\delta J = 0$ and applying the relevant (generalization of the) structural lemma from Section 3, says that

$$F_y - \frac{d}{dx}F_{y'} + \frac{d^2}{dx^2}F_{y''} - \cdots + (-1)^n \frac{d^n}{dx^n}F_{y^{(n)}} = 0.$$

This is called Euler's equation as well.

[⋆I presume one can combine Sections 9 and 11 in the obvious way.⋆]

2.12 Sec 12, *Variational Problems with Subsidiary Conditions*

Two (related) problems of great importance are solved in this section. One of them is the isoperimetric problem, which can be solved using "Lagrange multipliers": Consider the functional

$$J[y] = \int_a^b F(x, y, y')\, dx,$$

defined on curves $y = y(x)$ with $y(a) = A$, $y(b) = B$, and

$$K[y] = \int_a^b G(x, y, y')\, dx = \ell$$

for some G. Suppose $J[y]$ has an extremum for $y = y(x)$. If y is not an extremal of $K[y]$, then there exists a constant λ such that $y = y(x)$ is an extremal of the functional

$$\int_a^b (F + \lambda G)\, dx,$$

so y satisfies the differential equation

$$F_y - \frac{d}{dx}F_{y'} + \lambda\left(G_y - \frac{d}{dx}G_{y'}\right) = 0.$$

Thus, finding an extremum subject to the condition that $K[y]$ is held constant translates to finding an extremum subject to a "mysterious" constant λ (but now without the subsidiary condition). The latter is easy; just do the Euler equation.

The proof uses variational derivatives quite crucially; indeed, the constant λ turns out to be

$$\lambda = -\frac{\left.\dfrac{\delta F}{\delta y}\right|_{x=x_2}}{\left.\dfrac{\delta G}{\delta y}\right|_{x=x_2}},$$

where x_2 is an arbitrary point subject to the condition that the denominator above is nonzero. [⋆It's not obvious to me that this λ is independent of choice of x_2, but the proof says it is.⋆]

The more general problem of functionals depending on several functions $\mathbf{y} = (y_1, \dots, y_n)$ and constraints

$$K_j[\mathbf{y}] = \int_a^b G_j(x, \mathbf{y}, \mathbf{y}')\, dx = \ell_j, \qquad \text{for all } j \in [k]$$

can also be resolved; a necessary condition is that

$$\frac{\partial}{\partial y_i}\left(F + \sum_{j=1}^k \lambda_j G_j\right) - \frac{d}{dx}\left(\frac{\partial}{\partial y_i'}\left(F + \sum_{j=1}^k \lambda_j G_j\right)\right) = 0 \qquad \text{for all } i \in [n].$$

The other problem of great importance concerns finite subsidiary conditions: Fix functions $\mathbf{g} = (g_1, \dots, g_k) \colon \mathbb{R} \to \mathbb{R}^k$ with $k < n$, and consider the functional

$$J[\mathbf{y}] = \int_a^b F(x, \mathbf{y}, \mathbf{y}')\, dx$$

defined on curves with $\mathbf{y}(a) = \mathbf{A}$, $\mathbf{y}(b) = \mathbf{B}$, and

$$\mathbf{g}(x, \mathbf{y}) = 0,$$

i.e., on curves living in the manifold defined by $\mathbf{g}$. In the case $n = 2$ and $k = 1$, the precise result is the following: if g_y and g_z do not vanish simultaneously at any point of the surface defined by $g(x, y, z) = 0$, then there exists a function $\lambda(x)$ such that $(y, z) = (y(x), z(x))$ is an extremal of the functional

$$\int_a^b (F + \lambda(x) g)\, dx,$$

i.e., the system of differential equations

$$F_y + \lambda g_y - \frac{d}{dx} F_{y'} = 0,$$
$$F_z + \lambda g_z - \frac{d}{dx} F_{z'} = 0$$

is satisfied.

3 Ch III, *The General Variation of a Functional*

3.13 Sec 13, *Derivation of the Basic Formula*

The most general functionals considered so far have been those defined on $\mathscr{D}_n(a, b)$. The aim is to think about functionals defined on

$$\mathscr{D}_n \overset{\text{def}}{=} \bigsqcup_{(a,b) \in \mathbb{R}^2} \mathscr{D}_n(a, b),$$

i.e., to be able to account for curves with endpoints wherever we want. The first order of business is to put a norm on $\mathscr{D}_n$. For two curves $y = y(x)$ and $y^* = y^*(x)$ in $\mathscr{D}_n$, we define

$$\rho(y, y^*) = \max |y - y^*| + \max |y' - (y^*)'| + \rho(P_0, P_0^*) + \rho(P_1, P_1^*),$$

where the left and right endpoints of y and y^* are denoted P_0, P_0^* and P_1, P_1^* respectively respectively [$\star$what a phrase!$\star$] (here ρ of two points in Euclidean space is the usual norm).

The second order of business is to write down the linear functional $\delta J[y]$, (to be clear: δJ depends on a choice of $y = y(x) \in \mathscr{D}_n$, that eats increments h and spits out numbers), which can be done completely explicitly. In the event that y is an extremum, y is automatically an extremal ($\delta J[y] = 0$ for enough h, cf. Sec 3, so it satisfies Euler's equation) and furthermore

$$F_{y'} \delta y \Big|_{x=x_0}^{x=x_1} - (F - F_{y'} y') \delta x \Big|_{x=x_0}^{x=x_1} = 0;$$

here δy and δx are the perturbations coming from $h = y^* - y$ (in the fixed endpoint problem, all the δ's are zero; in the simple variable endpoint problem of Sec 6, the δx are zero). For curves $\mathbf{y} = (y_1, \dots, y_n) \colon \mathbb{R} \to \mathbb{R}^n$, all the Euler equations hold as usual, and furthermore

$$\left(\sum_{i=1}^n F_{y_i'} \delta y_i - \left(F - \sum_{i=1}^n F_{y_i'} y_i' \right) \delta x \right) \Bigg|_{x=x_0}^{x=x_1} = 0.$$

[$\star$By varying the increments, and hence the δy_i and δx, appropriately, I believe the quantities

$$F_{y_i'}|_{x=x_0}, \quad F_{y_i'}|_{x=x_1}, \quad F|_{x=x_0}, \quad F|_{x=x_1}$$

should all equal zero. I presume these act as boundary conditions for y_i, but I'm a little suspicious because the book seems to say nothing about this...$\star$]

There is a change of variables that one can attempt to do, namely, replace the variables y_i' with the variables $p_i \stackrel{\text{def}}{=} F_{y_i'}$. For the change of variables to work we need a nonzero Jacobian, i.e.

$$\det M \neq 0; \qquad M = (M_{i,j})_{i,j\in[n]} \text{ with } M_{i,j} = F_{y_i',y_j'}.$$

We also write

$$H = H(x, \mathbf{y}, \mathbf{p}) \stackrel{\text{def}}{=} -F + \sum_{i=1}^{n} y_i' p_i,$$

where y_i' should be thought of as a function in $(x, \mathbf{y}, \mathbf{p})$; such a representation of y_i' exists because the Jacobian is nonzero. These are called *canonical variables*. In this language, extrema of linear functionals $J[y]$ defined on $\mathscr{D}_n$ satisfy the Euler equations and furthermore

$$\left(\sum_{i=1}^{n} p_i \delta y_i - H\delta x\right)\Bigg|_{x=x_0}^{x=x_1} = 0.$$

3.14 Sec 14, *End Points Lying on Two Given Curves or Surfaces*

Because we went through the extra work of writing down $\delta J[y]$ in general, it becomes a matter of specializing the results of the previous section to solve the problem of finding extrema of

$$J[y] = \int_{x_0}^{x_1} F(x, y, y')\, dx$$

among all smooth curves $y = y(x)$ with endpoints on fixed curves $y = \varphi(x)$ and $y = \psi(x)$. Such extrema clearly satisfy Euler's equation; they furthermore satisfy the *transversality conditions*

$$(F + (\varphi' - y')F_{y'})|_{x=x_0} = 0$$
$$(F + (\psi' - y')F_{y'})|_{x=x_1} = 0.$$

In case the functional $F = F(x, y, y')$ is of the form $F = f(x, y)\sqrt{1 + (y')^2}$, the transversality conditions become

$$\frac{(1 + y'\varphi')F}{1 + (y')^2} = 0,$$
$$\frac{(1 + y'\psi')F}{1 + (y')^2} = 0.$$

It follows that $y' = -\frac{1}{\varphi'}$ at the left endpoint and $y' = -\frac{1}{\psi'}$ at the right endpoint.

The analogous problem for curves whose endpoints lie on two given surfaces $\varphi(y, z)$ and $\psi(y, z)$ has an analogous answer: an extremum would satisfy the Euler equations and four transversality conditions.

3.15 Sec 15, *Broken Extremals. The Weierstrass-Erdmann Conditions*

Let's also understand functionals defined on *piecewise* smooth curves. For simplicity let us assume the curve y can have at most one corner, say at $c \in (a, b)$ (where c varies). Let us fix an extremum y and study its properties; this effectively fixes c as well. It's clear that Euler's equation holds on (a, c) and (c, b). We may also write

$$J[y] = \int_a^c F(x, y, y')\, dx + \int_c^b F(x, y, y')\, dx = J_1[y] + J_2[y]$$

and compute the differentials $\delta J_1[y]$ and $\delta J_2[y]$ separately [$\star$Note that J_1 and J_2 depend on y, but y has been fixed$\star$]. Setting $\delta J_1[y] + \delta J_2[y] = 0$ results in the conditions

$$F_{y'}|_{x=c-0} = F_{y'}|_{x=c+0}$$
$$(F - y'F_{y'})|_{x=c-0} = (F - y'F_{y'})|_{x=c+0},$$

i.e. the canonical variables $p = F_{y'}$ and $H = F - y'F_{y'}$ must be continuous at c. These are the *Weierstrass-Erdmann corner conditions*.

4 Ch IV, *The Canonical Form of the Euler Equations and Related Topics*

4.16 Sec 16, *The Canonical Form of the Euler Equations*

Euler's equations are nice when written in terms of canonical variables. We want to rewrite

$$\frac{dy_i}{dx} = y_i', \qquad F_{y_i} - \frac{d}{dx} F_{y_i'} = 0;$$

the y_i' and F need to be replaced with $p_i = F_{y_i'}$ and $H = -F + \sum y_i' p_i$. Along the way, a miracle occurs: taking the differential of H, we obtain

$$dH = \cdots - \sum_{i=1}^{n} \frac{\partial F}{\partial y_i'} \, dy_i' + \sum_{i=1}^{n} p_i \, dy_i' + \ldots,$$

where the terms in $\ldots$ involve dx, dy_i, and dp_i (i.e., non-problematic variables). By definition of p_i, these terms cancel (!), which apparently is an important feature of canonical variables. Thanks to this, it becomes easy to prove that Euler's equations transform to

$$\frac{dy_i}{dx} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dx} = -\frac{\partial H}{\partial y_i}.$$

[$\star$It looks a little like Cauchy-Riemann...$\star$]

4.17 Sec 17, *First Integrals of the Euler Equations*

Consider functionals obtained by integrating a function $F = F(\mathbf{y}, \mathbf{y}')$ that doesn't depend on x explicitly. Then H doesn't depend on x explicitly either. Euler's equations then says $\frac{dH}{dx} = 0$, i.e., H is a first integral of Euler's equation (A *first integral* is a function which has a constant value along integral curves.)

In fact, the same yoga, without the assumption that F doesn't depend on x explicitly, applied to an arbitrary function $\Phi = \Phi(x, \mathbf{y}, \mathbf{y}')$ says that

$$\frac{d\Phi}{dx} = [\Phi, H] + \frac{\partial \Phi}{\partial x}.$$

Here $[\cdot, \cdot]$ is the usual Poisson bracket, defined by

$$[\Phi, H] \stackrel{\text{def}}{=} \sum_{i=1}^{n} \frac{\partial \Phi}{\partial y_i}\frac{\partial H}{\partial p_i} - \frac{\partial \Phi}{\partial p_i}\frac{\partial H}{\partial y_i}.$$

4.18 Sec 18, *The Legendre Transformation*

Although more technical [$\star$and in my opinion, less illuminating$\star$], it will pay off to realize the canonical form of Euler's equation associated to $J[\mathbf{y}] = \int_a^b F$, i.e. to realize the equations

$$\frac{dy_i}{dx} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dx} = -\frac{\partial H}{\partial y_i},$$

as the Euler equations of an equivalent functional $\widetilde{J}[\mathbf{y},\mathbf{p}]$ (so the task is: from J, construct $\widetilde{J}$ and prove that their extrema are the same).

The process of turning F into H (with the intermediary step of turning y' into p) is an instance of a more general construction known as *Legendre transform*. It is an involution. Define

$$\widetilde{J}[y,p] \stackrel{\text{def}}{=} \int_a^b (-H(x,y,p)+py')\,dx,$$

where p is now independent of y. Setting $p=F_{y'}(x,y,y')$ recovers $J[y]$, but now p is allowed to vary. Well, it turns out that when y is held fixed and p is allowed to vary, the extremum of $J[y,p]$ is $J[y]$ (!). It follows that J and $\widetilde{J}$ are equivalent functionals.

4.19 Sec 19, *Canonical Transformations*

Let us consider changes of variables $Y_i = Y_i(x,\mathbf{y},\mathbf{p})$ and $P_i = P_i(x,\mathbf{y},\mathbf{p})$ preserving the canonical Euler equations, i.e., let us consider functions satisfying

$$\frac{dY_i}{dx} = \frac{\partial H^*}{\partial P_i}, \qquad \frac{dP_i}{dx} = -\frac{\partial H^*}{\partial Y_i}.$$

(Here $H^* = H^*(x,\mathbf{Y},\mathbf{P})$ is some new function.) Such transformations are called *canonical transformations*. Thanks to the work in the last chapter we know already that the canonical Euler equations are the usual Euler equations of

$$\widetilde{J}[\mathbf{y},\mathbf{p}] = \int_a^b \left(\sum_{i=1}^n p_i y_i' - H\right) dx$$

and furthermore that adding the total differential $\frac{d\Phi}{dx}$ to the integrand (here Φ can be anything) leaves the (usual) Euler equations fixed: one can compute that the Euler equations of $\int_a^b \frac{d\Phi}{dx}\,dx$ are identically zero. We get a relation between $\sum P_i Y_i' - H^*$ and $\sum p_i y_i' - H$; spelling everything out we obtain

$$p_i = \frac{\partial \Phi}{\partial y_i}, \qquad P_i = \frac{\partial \Phi}{\partial Y_i}, \qquad H^* = H + \frac{\partial \Phi}{\partial x}.$$

Alternatively, writing $\Psi = \Phi + \sum P_i Y_i$, we may write the above as

$$p_i = \frac{\partial \Psi}{\partial y_i}, \qquad Y_i = \frac{\partial \Psi}{\partial P_i}, \qquad H^* = H + \frac{\partial \Psi}{\partial x}.$$

4.20 Sec 20, *Noether's Theorem*

Noether's Theorem generalizes the fact that H is a first integral of the Euler equations when F doesn't vary on x explicitly. It says the following:

Suppose $\Phi(x,\mathbf{y},\mathbf{y}';\varepsilon)$ and $\Psi_i(x,\mathbf{y},\mathbf{y}';\varepsilon)$ are transformations differentiable with respect to ε satisfying $\Phi(x,\mathbf{y},\mathbf{y}';0)=x$ and $\Psi_i(x,\mathbf{y},\mathbf{y}';0)=y_i$. Suppose further that the functional

$$J[\mathbf{y}] = \int_{x_0}^{x_1} F(x,\mathbf{y},\mathbf{y}')\,dx$$

is invariant under the transformations Φ and Ψ_i for every x_0, x_1. Then

$$\sum_{i=1}^n p_i\psi_i - H\varphi$$

is a first integral of the Euler equations of J; here we write

$$\varphi(x,\mathbf{y},\mathbf{y}') = \left.\frac{\partial \Phi}{\partial \varepsilon}\right|_{\varepsilon=0} \qquad \psi_i(x,\mathbf{y},\mathbf{y}') = \left.\frac{\partial \Psi_i}{\partial \varepsilon}\right|_{\varepsilon=0}.$$

The case where F doesn't depend on x, and $\Phi = x+\varepsilon$ and $\Psi_i = y_i$ gives the result from Sec 17.

4.21 Sec 21, *The Principle of Least Action*

[⋆You know, basic high school physics words (e.g. *momentum*(!)) have become such buzzwords for me that reading a nonironic (= down to earth) treatment of mathematical physics was an amazing experience.⋆]

Suppose we have some particles with mass m_i and position coordinates (x_i, y_i, z_i); the position data depends on a time variable t. Using overdots to denote derivatives with respect to t, we define

$$T \stackrel{\text{def}}{=} \frac{1}{2}\sum_{i=1}^{n} m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2)$$

to be the *kinetic energy*, and suppose that the system has *potential energy* $U = U(t, \mathbf{x}, \mathbf{y}, \mathbf{z})$ so that the force acting on the ith component is

$$X_i = -\frac{\partial U_i}{\partial x_i}, \qquad Y_i = -\frac{\partial U_i}{\partial y_i}, \qquad Z_i = -\frac{\partial U_i}{\partial z_i}.$$

Denote by $L \stackrel{\text{def}}{=} T - U$ the *Lagrangian*; the *Principle of Least Action* says that the x_i, y_i, z_i evolve according to the curve for which the *action*, given by the integral

$$\int_{t_0}^{t_1} L\, dt,$$

is minimal. This is because the Euler equations for this functional are precisely Newton's equations of motion.

4.22 Sec 22, *Conservation Laws*

The canonical variables for the action functional happen to be the momentum ($p_{ix} = m_i\dot{x}_i$, for example) and $H = T + U$ the total energy of the system. Thus, if the Lagrangian L (or equivalently the potential energy U) does not depend on t, then the fact that H is a first integral implies that *total energy is conserved*.

If the action functional is invariant under parallel displacements, then Noether's theorem implies that *momentum is conserved*. Similarly if it is invariant under rotations, then Noether's theorem implies that *angular momentum is conserved*.

4.23 Sec 23, *The Hamilton-Jacobi Equation. Jacobi's Theorem*

Fix a functional

$$J[\mathbf{y}] = \int_{x_0}^{x_1} F(x, \mathbf{y}, \mathbf{y}')\, dx$$

with the property that given any starting conditions $\mathbf{y}(x_0) = \mathbf{A}$ and $\mathbf{y}(x_1) = \mathbf{B}$, there is precisely one extremal (examples include arclength functional, or the action functional). Fix the initial starting condition $\mathbf{y}(a) = \mathbf{A}$ and consider the function

$$S(x_1, \mathbf{B}) \stackrel{\text{def}}{=} \operatorname*{ext}_{\mathbf{y}\colon \mathbf{y}(x_1)=\mathbf{B}} J[\mathbf{y}]$$

The function S satisfies the Hamilton-Jacobi equation, which says

$$\frac{\partial S}{\partial x} + H\left(x, y_1, \dots, y_n, \frac{\partial S}{\partial y_1}, \dots, \frac{\partial S}{\partial y_n}\right) = 0.$$

Jacobi's theorem says that if we have a complete integral $S = S(x, y_1, \dots, y_n, \alpha_1, \dots, \alpha_n)$ (here α_i are parameters) of the Hamilton-Jacobi equation, then the functions $y_i = y_i(x, \alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n)$ defined by $\frac{\partial S}{\partial \alpha_i} = \beta_i$, and $p_i = \frac{\partial S}{\partial y_i}$, are a general solution of the canonical system

$$\frac{dy_i}{dx} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dx} = -\frac{\partial H}{\partial y_i}.$$

5 Ch V, *The Second Variation. Sufficient Conditions for a Weak Extremum*

5.24 Sec 24, *Quadratic Functionals. The Second Variation of a Functional*

A quadratic functional $A[x]$ is something of the form $A[x] = B[x,x]$ for some bilinear functional B. A functional $J[y]$ is said to be *twice differentiable* if $\Delta[y+h] - \Delta[y] = \varphi_1[h] + \varphi_2[h] + \varepsilon||h||^2$, where φ_1 is linear, φ_2 is quadratic, and $\varepsilon \to 0$ as $||h|| \to 0$. The quadratic functional φ_2 is the *second variation* and will be denoted $\delta^2 J[y]$.

A necessary condition for the functional J to have a minimum at y is that $\delta J[y]$ is positive semidefinite, i.e. $(\delta^2 J[y])([h]) \geq 0$ for all admissible h. Similarly, a necessary condition for J to have a maximum at y is that $\delta J[y]$ is negative semidefinite.

We say a quadratic functional $\varphi_2[h]$ is *strongly positive* if there is $k > 0$ with $\varphi_2[h] \geq k||h||^2$. A sufficient condition for J to have a minimum at y is that $\delta J[y]$ is strongly positive. Similarly, a sufficient condition for J to have a minimum at y is that $\delta J[y]$ is strongly negative.

[$\star$(I presume we are looking for conditions on F for which the functional $J = \int F$ has extrema, as opposed to conditions on J.)$\star$]

5.25 Sec 25, *The Formula for the Second Variation. Legendre's Condition*

Although we have

$$(\delta^2 J[y])[h] = \frac{1}{2}\int_a^b (F_{yy}h^2 + 2F_{yy'}hh' + F_{y'y'}(h')^2)\,dx,$$

when considering increments h with $h(a) = h(b) = 0$ then we may integrate by parts and obtain

$$(\delta^2 J[y])[h] = \int_a^b (P(h')^2 + Qh^2)\,dx$$

where

$$P = \frac{1}{2}F_{y'y'}, \qquad Q = \frac{1}{2}\left(F_{yy'} - \frac{d}{dx}F_{yy'}\right).$$

From the nonnegativity of the second variation at an extremum, one obtains the necessary condition that $P \geq 0$.This is Legendre's condition: $F_{y'y'} \geq 0$ must be satisfied at every point of the curve $y = y(x)$.

5.26 Sec 26, *Analysis of the Quadratic Functional* $\int_a^b (P(h')^2 + Qh^2)\,dx$

We consider abstractly the quadratic functional

$$J[h] = \int_a^b (P(h')^2 + Qh^2)\,dx,$$

i.e. forgetting quadratic functionals of this form arise as second variations of $\int_a^b F\,dx$. The Euler equation is

$$-\frac{d}{dx}(Ph') + Qh = 0.$$

A point $\tilde{a} \neq a$ is said to be *conjugate* to a if there is a nonzero function h with $h(a) = h(\tilde{a})$ satisfying the above differential equation. The section culminates in the following result about $J[h]$:

Suppose that $P > 0$. Then $J[h]$, defined on $h(x)$ with $h(a) = h(b) = 0$, is positive definite if and only if the interval $[a,b]$ contains no points conjugate to a.

(The proof of the forwards direction also proves that if you are positive semidefinite, then (a,b) contains no points conjugate to a.)

5.27 Sec 27, *Jacobi's Necessary Condition. More on Conjugate Points*

Putting Sec 24 and 26 together, we obtain that yet another a necessary condition for

$$J[y] = \int_a^b F(x, y, y')\, dx,$$

defined on y with $y(a) = A$ and $y(b) = B$, to have an extremum at y is that y satisfies Euler's equation

$$F_y - \frac{d}{dx} F_{y'},$$

that $F_{y'y'} \geq 0$ or $F_{y'y'} \leq 0$ along this extremum (corresponding to y being a minimum or maximum respectively), and that (a, b) has no points conjugate to a.

(There are also some equivalent definitions for conjugate points, [$\star$which seem technical and unimportant at least right now$\star$])

5.28 Sec 28, *Sufficient Conditions for a Weak Extremum*

Contrast the necessary conditions of the previous section with the following sufficient conditions:
A sufficient condition for

$$J[y] = \int_a^b F(x, y, y')\, dx,$$

defined on y with $y(a) = A$ and $y(b) = B$, to have an extremum at y is that y satisfies Euler's equation

$$F_y - \frac{d}{dx} F_{y'},$$

that $F_{y'y'} > 0$ or $F_{y'y'} < 0$ along this extremum (corresponding to y being a minimum or maximum respectively), and that $[a, b]$ contains no points conjugate to a.

[$\star$Perhaps one thing to note here, which is not mentioned explicitly in the book, is that these sufficiency results would not hold if we were considering functionals defined on more general y, e.g. on variable end point problems.$\star$]

5.29 Sec 29, *Generalization to n Unknown Functions*

You can redo the chapter with n many functions $\mathbf{y}\colon \mathbb{R} \to \mathbb{R}^n$. Specifically, Legendre's condition now becomes that

$$P = \frac{1}{2} F_{\mathbf{y}'\mathbf{y}'} = \frac{1}{2} \begin{bmatrix} F_{y_1'y_1'} & F_{y_1'y_2'} & \cdots & F_{y_1'y_n'} \\ F_{y_2'y_1'} & F_{y_2'y_2'} & \cdots & F_{y_2'y_n'} \\ \vdots & \vdots & \ddots & \vdots \\ F_{y_n'y_1'} & F_{y_n'y_2'} & \cdots & F_{y_n'y_n'} \end{bmatrix}$$

is positive semidefinite. (This is required so that the second variation

$$(\delta^2 J[\mathbf{y}])[\mathbf{h}] = \int_a^b (\langle P\mathbf{h}', \mathbf{h}' \rangle + \langle Q\mathbf{h}, \mathbf{h} \rangle)\, dx$$

is positive semidefinite.)

In considering abstractly the functional

$$J[\mathbf{h}] = \int_a^b (\langle P\mathbf{h}', \mathbf{h}' \rangle + \langle Q\mathbf{h}, \mathbf{h} \rangle)\, dx,$$

forgetting that it arises as a second variation, one is led to study conjugate points again. This time, a *conjugate point* of a is a point $\widetilde{a}$ for which the determinant

$$\begin{bmatrix} h_{11}(x) & h_{12}(x) & \dots & h_{1n}(x) \\ h_{21}(x) & h_{22}(x) & \dots & h_{2n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1}(x) & h_{n2}(x) & \dots & h_{nn}(x) \end{bmatrix}$$

vanishes at $x = \widetilde{a}$; here $\mathbf{h}^{(i)} = (h_{i1}, \dots, h_{in})$ are solutions to the Euler equations

$$-\frac{d}{dx}(P\mathbf{h}') + Q\mathbf{h} = 0.$$

With basically the same proof as in Sec 26, it is proven that for a positive definite matrix P, the quadratic functional

$$J[\mathbf{h}] = \int_a^b (\langle P\mathbf{h}', \mathbf{h}' \rangle + \langle Q\mathbf{h}, \mathbf{h} \rangle)\, dx,$$

is positive definite if and only if $[a, b]$ contains no point conjugate to a; furthermore if it is nonnegative then (a, b) contains no point conjugate to a.

It follows now that if $\mathbf{y}$ is an extremal of

$$J[\mathbf{y}] = \int_a^b F(x, \mathbf{y}, \mathbf{y}')\, dx,$$

defined on curves $\mathbf{y}$ with fixed endpoints $\mathbf{y}(a) = \mathbf{A}$ and $\mathbf{y}(b) = \mathbf{B}$, and if $F_{\mathbf{y}'\mathbf{y}'}$ is positive definite along the extremal, then the open interval contains no points conjugate to a.

Contrast this to the sufficiency conditions: if $\mathbf{y}$ satisfies Euler's equation

$$F_{y_i} - \frac{d}{dx} F_{y_i'} = 0,$$

and P is positive definite or negative definite along this extremum (corresponding to $\mathbf{y}$ being a minimum or maximum respectively), and that $[a, b]$ contains no points conjugate to a, then $\mathbf{y}$ is an extremum.

5.30 Sec 30, *Connection Between Jacobi's Condition and the Theory of Quadratic Forms*

[⋆in progress⋆]

6 Ch VI, *Fields. Sufficient Conditions for a Strong Extremum*

6.31 Sec 31, *Consistent Boundary Conditions. General Definition of a Field*

6.32 Sec 32, *The Field of a Functional*

6.33 Sec 33, *Hilbert's Invariant Integral*

6.34 Sec 34, *The Weierstrass E-function. Sufficient Conditions for a Strong Extremum*