Calculus of Variations
Summer Term 2016

Lecture 7

Universität des Saarlandes

20. Mai 2016

Purpose of Lesson:

- To discuss why does the Lagrange multiplier approach work.
- Consider problems with non-integral constraints (holonomic and non-holonomic).
- Study general geodesic problem.

Why the Lagrange multiplier approach works?

- Consider the approximation of the functional

$$J[y] = \int_a^b F(x, y, y')dx \simeq \sum_{i=1}^{n} F\left(x_i, y_i, \frac{\Delta y_i}{\Delta x_i}\right) \Delta x = F(y_1, \ldots, y_n)$$

where $\Delta x = \frac{(b-a)}{n}$, and $\Delta y_i = y_i - y_{i-1}$.

- The problem of finding an extremal curve now becomes one of finding stationary points of the function $F(y_1, \ldots, y_n)$.

- We solve this by looking for

$$\frac{\partial F}{\partial y_i} = 0 \quad \text{for all} \quad i = 1, 2, \ldots, n.$$

- The constraint can be likewise approximted to give

$$\mathcal{G}[y] \simeq \sum_{i=1}^{n} G\left(x_i, y_i, \frac{\Delta y_i}{\Delta x_i}\right) \Delta x = G(y_1, \dots, y_n) = L.$$

- Under our usual conditions on J and $\mathcal{G}$, the limit as $n \to \infty$ gives

$$F(y_1, \dots, y_n) \to J[y]$$
$$G(y_1, \dots, y_n) \to \mathcal{G}[y]$$

- That is, the functions of the approximation $y_1, \dots, y_n$ converge to the functionals of the curve $y(x)$.

- In the finite dimensional case the constraint is

$$G(y_1, \dots, y_n) - L = 0$$

and we use a standard Lagrange multiplier

$$H(y_1, \dots, y_n, \lambda) = F(y_1, \dots, y_n) + \lambda \left[G(y_1, \dots, y_n) - L\right]$$

- We solve this by looking for

$$\frac{\partial H}{\partial y_i} = 0, \quad \forall i = 1, 2, \dots, n, \quad \text{and} \quad \frac{\partial H}{\partial \lambda} = 0.$$

- The last equation just gives you back your constraint.

- In our formulation of the isoperimetric problem we take

$$\mathcal{H}[y] = J[y] + \lambda \mathcal{G}[y]$$

and we also have

$$H(y_1, \ldots, y_n, \lambda) = F(y_1, \ldots, y_n) + \lambda \left[G(y_1, \ldots, y_n) - L \right].$$

- In the limit as $n \to \infty$ we find that

$$H(y_1, \ldots, y_n, \lambda) \to \mathcal{H}[y] - \lambda L.$$

- The EL equations for $\mathcal{H}[y] - \lambda L$ and $\mathcal{H}[y]$ are the same, so they have the same extremals.

Remarks about multiple constraints

- We can also handle multiple constraints via multiple Lagrange multipliers.
- For instance, if we wish to find extremals of $J[y] = \int\limits_{x_0}^{x_1} F(x, y, y')dx$ with the *m* constraints

$$\mathcal{G}_k[y] = \int\limits_{x_0}^{x_1} G_k(x, y, y')dx = L_k$$

we would look for extremals of

$$\mathcal{H}[y] = \int\limits_{x_0}^{x_1} H(x, y, y')dx = \int\limits_{x_0}^{x_1} \left[F(x, y, y') + \sum_{k=1}^{m} \lambda_k G_k(x, y, y') \right] dx$$

§6. Problems with non-integral constraints

It is relatively easy to adapt the Lagrange multiplier technique to the case with non-integral constraints.

- Holonomic constraints are of the form

$$G(x, y) = 0$$

- Non-Holonomic constraints are of the form

$$G(x, y, y') = 0$$

- ***"Holonomic"*** comes from the greek ***"holos"***, for ***"whole"***. In this context it refers to integrability of the constraint.
- The ***non-holonomic*** constraints are really ***DEs***.

Problem 7-1

Consider the problem of finding extremals of

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

subject to the constraint

$$G(x, y) = 0.$$

- In this case we introduce a function $\lambda(x)$ (also called a Lagrange multiplier), and look for the extremals of

$$\mathcal{H}[y] = J[y] + \int_{x_0}^{x_1} \lambda(x)G(x, y)dx.$$

Remarks

- Constraints of the form $G(x, y) = 0$ which don't involve derivatives of $y(x)$ can also be handled using a Lagrange multiplier technique.
- But we have to introduce a Lagrange multiplier function $\lambda(x)$, not just a single value λ.
- Effectively we introduce one Lagrange multiplier at each point where the constraint is enforced.

Why the Lagrange multiplier approach works here?

- Go back to the approximation of the functional

$$J[y] \simeq \sum_{i=1}^{n} F\left(x_i, y_i, \frac{\Delta y_i}{\Delta x_i}\right) \Delta x = F(y_1, \ldots, y_n).$$

- The constraint applies a condition on each (x_i, y_i).
- So, in the approximation there are n constraints

$$G(x_i, y_i) = 0 \quad \text{for} \quad i = 1, \ldots, n.$$

- There are n constraints,

$$G(x_i, y_i) = 0 \quad \text{for} \quad i = 1, \dots, n.$$

- For optimization problems with n constraints, we introduce n Lagrange multipliers, and maximize

$$H(y_1, \dots, y_n) = F(y_1, \dots, y_n) + \sum_{k=1}^{n} \lambda_k G(x_k, y_k).$$

- In the limit as $n \to \infty$

$$\Delta x \sum_{k=1}^{n} \lambda_k G(x_k, y_k) \to \int_{x_0}^{x_1} \lambda(x) G(x, y) dx$$

and hence the choice of

$$\mathcal{H}[y, \lambda] = J[y] + \int_{x_0}^{x_1} \lambda(x) G(x, y) dx.$$

$$\mathcal{H}[y,\lambda] = J[y] + \int\limits_{x_0}^{x_1} \lambda(x) G(x,y) dx$$
$$= \int\limits_{x_0}^{x_1} \left(F(x,y,y') + \lambda(x)G(x,y)\right) dx$$

- So, we can apply out standard arguments to the integrand

$$H(x,y,y',\lambda) = F(x,y,y') + \lambda(x)G(x,y)$$

and get the Euler-Lagrange equation

$$\boxed{\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \lambda(x)\frac{\partial G}{\partial y} = 0}.$$

With multiple dependent variables holonomic constraints are of the form

$$G(t, \mathbf{q}) = 0$$

and they don't involve derivatives.

Example 7.1

To minimize the functional

$$J[x, y, z] = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} dt$$

subject to the constraint

$$x^2 + y^2 - r^2 = 0.$$

Remarks

- In Example 7.1 we have to find geodesics on a right circular cylinder with radius r.
- ***Geodesic*** is the shortest line between two points on a mathematically defined surface (as a straight line on a plane or an arc of a great circle (like the equator) on a sphere).
- ***Geodesic*** is a curve whose tangent vectors remain parallel is they are transported along it.

$$\mathcal{H}[\mathbf{q}, \lambda] = J[\mathbf{q}] + \int_{t_0}^{t_1} \lambda(t) G(t, \mathbf{q}) dt$$

- So, we can again apply our standard arguments to the integrand

$$H(t, \mathbf{q}, \dot{\mathbf{q}}, \lambda) = F(t, \mathbf{q}, \dot{\mathbf{q}}) + \lambda(t) G(t, \mathbf{q})$$

and get the system of the Euler-Lagrange equations

$$\boxed{\frac{\partial F}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{q}_k}\right) + \lambda(t)\frac{\partial G}{\partial q_k} = 0}$$

for all k.

General geodesic problem can be written as

Problem 7-2 (general geodesic problem)

To minimize

$$J[x,y,z] = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\,dt$$

sublect to

$$G[x,y,z] = 0,$$

where $G[x,y,z] = 0$ is the equation describing the surface of interest.

- As usual instead of $J[x,y,z]$ we minimize

$$\mathcal{H}[x,y,z,\lambda] = \int_{t_0}^{t_1} \left(\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} + \lambda(t) G(x,y,z) \right) dt.$$

Given this formulation of the geodesic problem, the Euler-Lagrange equations become

$$\frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}\right)-\lambda(t)\frac{\partial G}{\partial x}=0$$
$$\frac{d}{dt}\left(\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}\right)-\lambda(t)\frac{\partial G}{\partial y}=0$$
$$\frac{d}{dt}\left(\frac{\dot{z}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}\right)-\lambda(t)\frac{\partial G}{\partial z}=0$$

which may be easier to solve in some cases.

Example 7.2 (Geodesics on the sphere)

Find the geodesics on the sphere: e.g., we wish to find a parametric curve $(x(t), y(t), z(t))$ to minimize distance

$$J[x, y, z] = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} dt$$

subject to being on the surface of a sphere

$$x^2 + y^2 + z^2 = a^2.$$

- We get

$$H(t, x, y, z, \dot{x}, \dot{y}, \dot{z}, \lambda) = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} + \lambda(t)\left(x^2 + y^2 + z^2 - a^2\right)$$

  and there are three dependent variables (x, y, z).

Example 7.2 (cont.)

- The simple calculation shows that

$$\frac{\partial H}{\partial x} = 2\lambda x \qquad \frac{\partial H}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}$$

$$\frac{\partial H}{\partial y} = 2\lambda y \qquad \frac{\partial H}{\partial \dot{y}} = \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}$$

$$\frac{\partial H}{\partial z} = 2\lambda z \qquad \frac{\partial H}{\partial \dot{z}} = \frac{\dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}$$

Example 7.2 (cont.)

- There are 3 dependent variables (x, y, z), and, so 3 Euler-Lagrange equations, e.g.,

$$2\lambda x = \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}\right) = \frac{\ddot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}} - \frac{\dot{x}\left[\dot{x}\ddot{x}+\dot{y}\ddot{y}+\dot{z}\ddot{z}\right]}{\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)^{3/2}}$$

- Due to symmetry, the equation

$$2\lambda u = \frac{\ddot{u}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}} - \frac{\dot{u}\left[\dot{x}\ddot{x}+\dot{y}\ddot{y}+\dot{z}\ddot{z}\right]}{\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)^{3/2}}$$

holds for $u = x, y$ and z.

Example 7.2 (cont.)

- Observe that

$$2\lambda u = \frac{\ddot{u}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} - \frac{\dot{u}\left[\dot{x}\ddot{x} + \dot{y}\ddot{y} + \dot{z}\ddot{z}\right]}{\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right)^{3/2}}$$

  is a second order linear DE in u, and so it has **only 2 linearly independent solutions**, but the DE holds for $u = x, y$ and z.

- Therefore, x, y and z are linearly dependent, and so we can write them as

$$Ax + By + Cz = 0$$

  but this is the equation of a plane through the origin.

- We have shown that geodesics on the sphere are great circles.

Remarks

- ***Non-Holonomic constraints*** are constraints of the form

  $$G(x, y, y') = 0 \quad \text{or} \quad G(t, \mathbf{q}, \dot{\mathbf{q}}),$$

  which involve derivatives.

- Non-Holonomic constraints are effectively additional DEs which we need to solve, but we can once again use Lagrange multipliers.

- Sometimes a constraint involving derivatives may be integrated to get a holonomic constraint. So, we refer to these constraints as integrable.

- In general, we will also need to deal with constraints involving derivatives as these may describe an entire systems behaviour, and be very difficult to integrate out of the problem.

Example 7.3 (Non-Holonomic constraints)

Example non-holonomic constraints:

$$G(x, y, y') = 0 \quad \text{or} \quad G(t, \mathbf{q}, \dot{\mathbf{q}}),$$

Instances:

- $y = \dot{x}$
- $y'^2 = \log x$.

- Solution technique for the non-holonomic constraints is just as for holonomic constraints, e.g.,

$$\mathcal{H}[y, \lambda] = J[y] + \int_{x_0}^{x_1} \lambda(x) G(x, y, y') dx$$

and the argument for why it works is almost identical.

Remark

- Non-Holonomic constraints can be used to avoid higher derivatives.

Example 7.4

Minimizing the functional

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

we derive a new form of the Euler-Lagrange (Euler-Poisson) equation for this case, e.g.,

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial F}{\partial y''}\right) = 0 \tag{7.1}$$

Example 7.4 (cont.)

- Non-Holonomic constraints give us an alternative approach to problem (7-1).
- Introduce the new variable $z = y'$, and rewrite the functional as

$$J[y, z] = \int_a^b F(x, y, z, z') dx. \tag{7.2}$$

- Now there is more thatn one dependent variable, but no second order derivatives. However, we must also introduce the constraint that

$$z - y' = 0.$$

- So, we look for stationary curves of the functional

$$\mathcal{H}[y, z, \lambda] = \int_a^b \left(F(x, y, z, z') + \lambda(x)(z - y') \right) dx.$$

Example 7.4 (cont.)

- The Euler-Lagrange equations for y and z are

$$\frac{\partial H}{\partial y} - \frac{d}{dx}\left(\frac{\partial H}{\partial y'}\right) = 0$$
$$\frac{\partial H}{\partial z} - \frac{d}{dx}\left(\frac{\partial H}{\partial z'}\right) = 0$$

- Note that $H(x, y, y', z, z') = F(x, y, z, z') + \lambda(x)\,(z - y')$. So, the Euler-Lagrange equations become

$$\frac{\partial F}{\partial y} + \frac{d}{dx}\left(\lambda(x)\right) = 0$$
$$\frac{\partial F}{\partial z} + \lambda(x) - \frac{d}{dx}\left(\frac{\partial F}{\partial z'}\right) = 0$$

Example 7.4 (cont.)

- The first Euler-Lagrange equation can be rewritten

$$\frac{d\lambda}{dx} = -\frac{\partial F}{\partial y}$$

- Differentiating the second Euler-Lagrange equation w.r.t. x we get

$$\frac{d}{dx}\left(\frac{\partial F}{\partial z}\right) + \frac{d\lambda}{dx} - \frac{d^2}{dx^2}\left(\frac{\partial F}{\partial z'}\right) = 0$$

- Note from above that $\lambda' = -F_y$ and that $z = y'$ and $z' = y''$ we get (as before) the Euler-Poisson equation:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \frac{d^2}{dx^2}\left(\frac{\partial F}{\partial y''}\right) = 0.$$

Remarks

- Earlier we derived the Euler-Lagrange equation assuming treating y and y' as if they were independent variables.
- In reality they are related along the extremal.
- Lets get some motivation for this. Start by taking a new variable $u(x) = y'(x)$, and put this into our minimization problem

$$\mathcal{H}[y, u, \lambda] = \int_a^b \left(F(x, y, u) + \lambda(x)\left[u - y'\right]\right) dx.$$

- We can use the same trick as in previous slides to get the Euler-Lagrange equations.