Calculus of Variations
Summer Term 2016

Lecture 13

Universität des Saarlandes

21. Juni 2016

Purpose of Lesson:

- The Ritz method applied to the catenary gives additional insights.
- Kantorovich's method generalizes Ritz to $2D$ functions.
- To consider optimal control examples
- To introduce a terminology.

Ritz and Catenary:

Example 13.1 (the catenary, again)

The functional of interest (the potential energy) is

$$J_p[y] = mg \int_{x_0}^{x_1} y\sqrt{1 + y'^2}\,dx.$$

- Take symmetric problem with fixed end points $y(-1) = a$ and $y(1) = a$.
  We know that the solution looks like

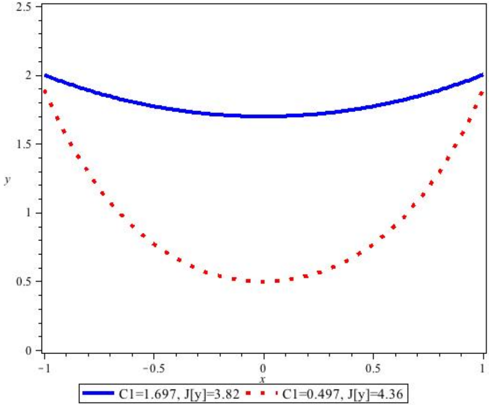
  $$y(x) = c_1 \cosh\left(\frac{x}{c_1}\right)$$

  where c_1 is chosen to match the end points.

Example 13.1 (the catenary, again)

$y(1) = 2$ gives $c_1 = 0.47$ or $c_1 = 1.697$

- Are they both local minima?



Example 13.1 (Ritz and the Catenary)

- Lets try approximating the curve by a polynomial

$$y(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \dots$$

- Note that symmetry of problem implies y is an even function, and hence the odd terms

$$a_1 = a_3 = \cdots = 0.$$

- So, to second order we can approximate

$$y(x) \simeq a_0 + a_2 x^2.$$

- We have fixed $y(1) = y_1$, so we can simplify to get

$$y(x) \simeq a_0 + (y_1 - a_0)x^2.$$

Example 13.1 (Ritz and the Catenary)

$$y \simeq a_0 + (y_1 - a_0)x^2$$
$$y' \simeq 2(y_1 - a_0)x$$

- Taking into account $y(1) = 2$ we get $a_0 + a_2 = 2$. We can substitute into the functional

$$J_p[y] = mg \int_{x_0}^{x_1} y\sqrt{1 + y'^2}dx$$

  and integrate to get a function $J_p[a_2]$ with respect to a_2.

- But this function is pretty complicated.

Example 13.1 (Ritz and the Catenary)

From Maple we have the value for $J_p[a_2]$, $(a := a_2)$

```
> f(x) := (2 - a + a·x^2)·sqrt(1 + 4·a^2·x^2) :
> int(f(x), x=-1..1)
```

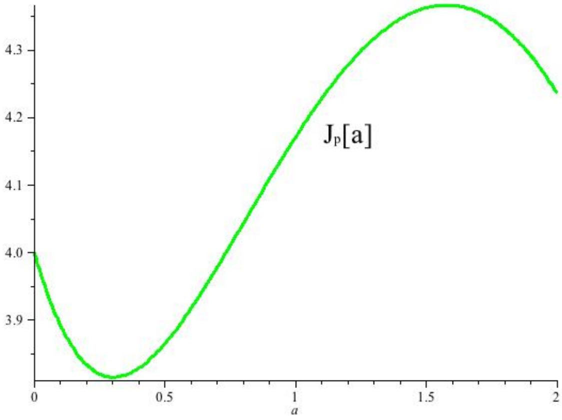
$$\frac{1}{64}\,\frac{1}{a^2}\Big(\big(16\,a^2 \ln\big((-2\,a+\sqrt{1+4\,a^2}\,\mathrm{csgn}(a))\,\mathrm{csgn}(a)\big)+128\,\sqrt{1+4\,a^2}\,a^2\,\mathrm{csgn}(a) - 64\,a^3\sqrt{1+4\,a^2}\,\mathrm{csgn}(a) - 32\,a\ln\big((-2\,a+\sqrt{1+4\,a^2}\,\mathrm{csgn}(a))\,\mathrm{csgn}(a)\big)+\ln\big((-2\,a+\sqrt{1+4\,a^2}\,\mathrm{csgn}(a))\,\mathrm{csgn}(a)\big) - 4\sqrt{1+4\,a^2}\,a\,\mathrm{csgn}(a)+8\,(1+4\,a^2)^{3/2}\,a\,\mathrm{csgn}(a) - 16\,a^2\ln\big((2\,a+\sqrt{1+4\,a^2}\,\mathrm{csgn}(a))\,\mathrm{csgn}(a)\big)+32\,a\ln\big((2\,a+\sqrt{1+4\,a^2}\,\mathrm{csgn}(a))\,\mathrm{csgn}(a)\big) - \ln\big((2\,a+\sqrt{1+4\,a^2}\,\mathrm{csgn}(a))\,\mathrm{csgn}(a)\big)\big)\,\mathrm{csgn}(a)\Big) \tag{1}$$

```
>
```

Example 13.1 (Ritz and the Catenary)

- Its a pain to find the zeros of $\frac{dJ_p}{da}$, but its easy to plot, and find them numerically.

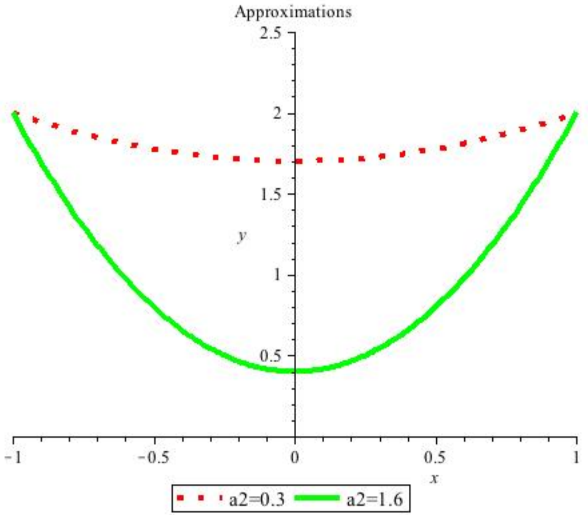
Example 13.1 (Ritz and the Catenary)



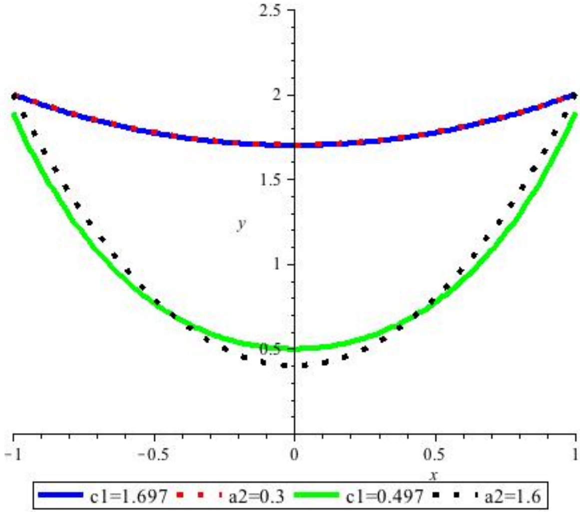
Stationary points

- local max: $a = a_2 \simeq 1.6$
- local min: $a = a_2 \simeq 0.3$

Example 13.1 (Ritz and the Catenary)



Example 13.1 (Ritz and the Catenary)



Ritz and the Catenary

Doesn't just give us an approximation to the extremal curves, its also give us some insight into the nature of these extremals. If

- approximations are near to the actual extrema
- There are no other extrema so close by
- The functional is smooth (it can't have jumps either)

Then the type of extrema we get for the approximation will be the same for the real extrema, i.e.,

- local max: $a_2 \simeq 1.6 \;\Rightarrow$ local max for $c_1 = 0.497$
- local min: $a_2 \simeq 0.3 \;\Rightarrow$ local min for $c_1 = 1.697$

More than one independent variables

$2D$ Case:

We are approximating a surface with series of functions, e.g.

$$z(x,y) \simeq z_n(x,y) = \phi_0(x,y) + \sum_{i=1}^{n} c_i \phi_i(x,y)$$

where

- $\phi_0(x,y)$ satisfies the boundary conditions, e.g.

  $$\phi_0(x,y) = z_0(x,y) \quad \text{for} \quad (x,y) \in \partial\Omega,$$

  the boundary of the region on interest Ω,

- and the $\phi_i(x,y)$ satisfy the homogeneous boundary conditions

  $$\phi_i(x,y) = 0 \quad \text{for} \quad (x,y) \in \partial\Omega.$$

$2D$ Case:

- As before, we approximate the functional by

$$J[z] \simeq J[z_n] = J_n(c_1, \ldots, c_n).$$

- As before we determine the c_j by requiring that the partial derivatives are zero, e.g.

$$\frac{\partial J_n}{\partial c_i} = 0$$

  for all $i = 1, 2, \ldots, n$.

Kantorovich's Method

- Approximate with

$$z(x, y) \simeq z_n(x, y) = \phi_0(x, y) + \sum_{i=1}^{n} c_i(x)\phi_i(x, y).$$

- Again the ϕ_i are suitable chosen, but the c_i are no longer constants, but rather functions of one independent variable.
- This allows a larger class of functions to be used.

Kantorovich's Method

- Note that the integral function

$$J[z_n] = \iint\limits_{\Omega} z_n(x,y)dxdy = \sum_{i=0}^{n} \int c_i(x) \left[\int\limits_{y_0(x)}^{y_1(x)} \phi_i(x,y)dy \right] dx$$

- We integrate the inner integral, and get

$$J[z_n] = \sum_{i=0}^{n} \int c_i(x)\Phi_i(x)dx.$$

- Now we just have a function of x, and so we may apply the Euler-Lagrange machinery.
- The method approx. separates the variables x and y.

Example 13.2

Find the extremals of

$$J[z(x,y)] = \int_{-b}^{b}\int_{-a}^{a} \left(z_x^2 + z_y^2 - 2z \right) dxdy$$

with $z = 0$ on the boundary.

- The Euler-Lagrange equation reduces to the Poisson equation, e.g.

$$F_z - \frac{d}{dx}F_{z_x} - \frac{d}{dy}F_{z_y} = 0$$

$$-2 - \frac{d}{dx}(2z_x) - \frac{d}{dy}(2z_y) = 0$$

$$z_{xx} + z_{yy} = -1$$

Example 13.2

- Approximate

$$z_1(x,y) = c(x)\left(b^2 - y^2\right)$$

- Note $z_1(x, \pm b) = 0$ (as required) and

$$\left(\frac{\partial z_1}{\partial x}\right)^2 = \left[c'(x)\left(b^2 - y^2\right)\right]^2$$
$$= c'(x)^2\left[b^4 - 2b^2y^2 + y^4\right],$$
$$\left(\frac{\partial z_1}{\partial y}\right)^2 = [c(x)2y]^2$$
$$= 4c(x)^2y^2$$

Example 13.2

Hence, we approximate

$$J[z(x,y)] \simeq J[z_1(x,y)] = \int_{-b}^{b}\int_{-a}^{a} \left(z_x^2 + z_y^2 - 2z\right) dxdy$$

$$= \int_{-a}^{a} \left[\int_{-b}^{b} \left[c'(x)^2 \left(b^2 - y^2\right)^2 + 4c(x)^2 y^2 - 2c(x)\left(b^2 - y^2\right)\right] dy\right] dx$$

$$= \int_{-a}^{a} \left[c'(x)^2 \left(b^4 y - 2b^2 y^3/3 + y^5/5\right) + 4c(x)^2 y^3/3 \right.$$

$$\left. + 2c(x)\left(b^2 y - y^3/3\right)\right]_{-b}^{b} dx$$

$$= \int_{-a}^{a} \left[\frac{16}{15} b^5 c'(x)^2 + \frac{8}{3} b^3 c(x)^2 - \frac{8}{3} b^3 c(x)\right] dx$$

Example 13.2

- So we can write

$$J[z(x,y)] \simeq J[z_1(x,y)] = J[c(x)] = \int_{-a}^{a} F(x,c,c')dx$$

- We can use the simple Euler-Lagrange equation, where

$$F(x,c,c') = \frac{16}{15}b^5 c'(x)^2 + \frac{8}{3}b^3 c(x)^2 - \frac{8}{3}b^3 c(x)$$

$$\frac{\partial F}{\partial c} = \frac{16}{3}b^3 c(x) - \frac{8}{3}b^3$$

$$\frac{\partial F}{\partial c'} = \frac{32}{15}b^5 c'(x) \qquad \frac{d}{dx}\frac{\partial F}{\partial c'} = \frac{32}{15}b^5 c''(x)$$

Example 13.2

- The Euler-Lagrange equation

$$\frac{16}{3}b^3c(x) - \frac{8}{3}b^3 - \frac{32}{15}b^5c''(x) = 0$$
$$c''(x) - \frac{5}{2b^2}c(x) = -\frac{5}{4b^2}$$

- Solutions

$$c(x) = k_1 \cosh\left(\sqrt{\frac{5}{2}}\frac{x}{b}\right) + k_2 \sinh\left(\sqrt{\frac{5}{2}}\frac{x}{b}\right) + \frac{1}{2}$$

Example 13.2

- Note that the function must be zero on the boundary, so $z(\pm a, y) = 0$.
- We look for an even function $c(x)$, and so $k_2 = 0$.
- Also $c(\pm a) = 0$, so

$$c(a) = k_1 \cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right) + \frac{1}{2}$$

$$-\frac{1}{2} = k_1 \cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right)$$

$$k_1 = -\frac{1}{2\cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right)}$$

Example 13.2

- Solution

$$z_1(x,y) = \frac{1}{2}(b^2 - y^2)\left(1 - \frac{cosh\left(\sqrt{\frac{5}{2}}\frac{x}{b}\right)}{cosh\left(\sqrt{\frac{5}{2}}\frac{a}{b}\right)}\right)$$

- If we want a more exact approximation, we could try

$$z_2(x,y) = (b^2 - y^2)c_1(x) + (b^2 - y^2)^2 c_2(x).$$

Remarks

- Obviously, quality of solution depends on
  - family of functions chosen
  - number of terms used, n
- Could test convergence by increasing n and seeing the difference in
$$|J[y_{n+1}] - J[y_n]|,$$
but this is not guaranteed to be a good indication.
- A better way to assess convergence is to have a lower bound
$$\text{lower bound } \leqslant J[y] \leqslant \text{ upper bound}$$

Introduction in Optimal Control Problems

Formulation of control problems

We break a control problems into two parts

1. The system state: $\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^t$

   The system state describes the system (e.g. position and velocity of the car in car parking example)

2. The control: $\mathbf{u}(t) = (u_1(t), \ldots, u_m(t))^t$

   We apply the control to the system (e.g. force applied to the car).

The evolution of the system is governed by the set of DEs

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$$

In a control problem we want to get the system to a particular state $\mathbf{x}(t)$ at time t, given initial state $\mathbf{x}(t_0)$.

Optimal control problems

- In an optimal control problem we have still have the system equations
$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$$
and we might wish to get to state $\mathbf{x}(t)$ given initial state $\mathbf{x}(t_0)$, but now we wish to do so while minimizing a functional
$$J[\mathbf{x}, \mathbf{u}] = \int_{t_0}^{t_1} F(t, \mathbf{x}, \mathbf{u}) dt.$$
- That is, we wish to choose a function $\mathbf{u}(t)$ which minimizes the functional $J[\mathbf{x}, \mathbf{u}]$, while satisfying the end-point conditions $\mathbf{x}(t_0) = \mathbf{x}_0$ and $\mathbf{x}(t_1) = \mathbf{x}_1$, and the non-holonomic constaints
$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u}).$$

Optimal control problems

Optimization functional

$$J[\mathbf{x}, \mathbf{u}] = \int_{t_0}^{t_1} F(t, \mathbf{x}, \mathbf{u}) dt$$

Remarks

Note that

- $F(t, \mathbf{x}, \mathbf{u})$ has no dependence on $\dot{\mathbf{u}}$: this is typically because costs depend on the control, not how we change the control, but there might be counter-examples.
- $F(t, \mathbf{x}, \mathbf{u})$ has no dependence on $\dot{\mathbf{x}}$: this is common in control problems, but not universal (we have seen at least one counter example).

Terminal costs

- Sometimes in optimal control we don't fix the end-point $\mathbf{x}(t_1)$, but rather we assign a cost $\phi(t_1, \mathbf{x}(t_1))$ to particular end-points.
- So now we wish to choose a control $\mathbf{u}(t)$ which minimizes the functional

$$J[\mathbf{x}, \mathbf{u}] = \phi(t_1, \mathbf{x}(t_1)) + \int_{t_0}^{t_1} F(t, \mathbf{x}, \mathbf{u}) dt$$

  while satisfying the single end-point condition $x(t_0) = \mathbf{x}_0$, and the non-holonomic constraint $\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}, \mathbf{u})$.
- $\phi(t_1, \mathbf{x}(t_1))$ is called the terminal cost.

System Terminology

- **linear:** the state equations are a set of linear DEs.
- **autonomous:** time doesn't appear explicitly in the state equations (e.g. in $g(\mathbf{x}, \mathbf{u})$, or $F(\mathbf{x}, \mathbf{u})$).
  - also called time-invariant.
- **terminal cost:** the term $\phi(t_1, \mathbf{x}(t_1))$ is called the terminal cost.
- **controllable:** a solution to the control problem exists.
- **stable:** a stable equilibrium solution to the system DEs exists.
  - often we are interested in problems that are unstable, or we wouldn't really need a control.

Control Terminology

- control (driver or automatic)
  - **planned** (open loop)
  - **feedback** (closed loop) control depends on current state
- type of control
  - movement from A to B
  - continuous operations (maintain equilibrium)
- type of cost functional J
  - minimum time
  - minimum fuel
  - quadratic costs
- admissible controls
  - unbounded / bounded / bang-bang

Cost functional examples

- **minimum time:** choose the fastest possible control

$$J[x,u] = \int_{t_0}^{t_1} dt.$$

- **minimum fuel:** fuel is expended by the controller, and we wish to minimize this

$$J[x,u] = \int_{t_0}^{t_1} |u(t)| dt$$

- **quadratic costs:**

$$J[x,u] = \int_{t_0}^{t_1} \left(x^2(t) + \alpha u^2(t) \right) dt$$

Boundary conditions

- End time t_1: can be fixed or free
- End position $\mathbf{x}(t_1)$: can be fixed or free

In the cases with free boundary conditions, we introduce natural, or transversal boundary conditions.

Example 13.3 Dynamic production

- A producer in purely competitive market
  - A large numbers of independent producers
  - Standardized product, e.g. potatoes
  - Firms are „price takers", i.e. they have no significant control over product price
  - Free entry and exit
  - Free flow of information
- wants to find optimal production path $x(t)$, $0 \leqslant t \leqslant T$.
- production target $x(T) = x_T$
- profit at time t is $\pi(x, \dot{x}, t)$
- maximize profit functional $J[x] = \int\limits_0^T \pi(x, \dot{x}, t)dt$.

Example 13.3 Dynamic production-2

Profit calculation

- quadratic production costs $C_1 = a_1 x^2 + b_1 x + c_1$
  - labor
  - raw materials
- production increase costs $C_2 = a_2 (\dot{x})^2 + b_2 \dot{x} + c_2$
  - new bildings
  - recruiting and training costs
- revenue $r = px$ where p is the constant price per unit
  - $p = const$ due to purely competitive market
- profit at time t is

$$\pi(x, \dot{x}, t) = px - C_1(x) - C_2(\dot{x}).$$

Example 13.3 Dynamic production-3

Problem formulation: maximize total profit

$$J[x] = \int_0^T (px - C_1(x) - C_2(\dot{x}))\, dt$$

subject to $x(0) = 0$ and $x(T) = x_T$.

- notice that the control, and rate of change of state are the same (i.e., $u = \dot{x}$) but we write it as above for simplicity
- autonomous problem
- the control is planned, and has quadratic costs
- admissible controls are unbounded

Example 13.3 Dynamic production-4

Euler-Lagrange equations

$$\begin{aligned}
\frac{\partial \pi}{\partial x} - \frac{d}{dt}\frac{\partial \pi}{\partial \dot{x}} &= 0 \\
p - \frac{\partial C_1}{\partial x} + \frac{d}{dt}\frac{\partial C_2}{\partial \dot{x}} &= 0 \\
p - 2a_1 x - b_1 + \frac{d}{dt}\left[2a_2\dot{x} + b_2\right] &= 0 \\
2a_2\ddot{x} - 2a_1 x + p - b_1 &= 0 \\
\ddot{x} - \frac{a_1}{a_2}x &= \frac{b_1 - p}{2a_2}
\end{aligned}$$

for $a_2 \neq 0$.

Example 13.3 Dynamic production-5

Solution (for $a_1, a_2 \neq 0$)

$$x(t) = Ae^{\sqrt{\frac{a_1}{a_2}}t} + Be^{-\sqrt{\frac{a_1}{a_2}}t} + \frac{b_1 - p}{2a_2}$$

where A and B are determined by the fixed end points $x(0) = x_0$ and $x(T) = x_T$.

This gives the optimal production schedule

- no dependence on c_1 or c_2 (these are constant costs and so shouldn't effect production strategy)
- no dependence on b_2 because this is a linear cost in increasing production, and so occurs regardless of how we increase over time (to get to the final production target $x(T) = x_T$).

Example 13.3 Dynamic production-6

What happens if we make the end point $x(T)$ free, i.e. we don't have a production target at time T?

Then we get a natural boundary condition

$$\left.\frac{\partial \pi}{\partial \dot{x}}\right|_{t=T} = \left.\frac{\partial C_2}{\partial \dot{x}}\right|_{t=T} = \left. 2a_2\dot{x} + b_2\right|_{t=T} = 0$$

So, rearranging, we get

$$\dot{x}(T) = -\frac{b_2}{2a_2}$$

- constants A and B are determined by end-point conditions $x(0) = 0$ and $\dot{x}(T) = -\frac{b_2}{2a_2}$.

- Production costs

$$C_1 = x^2 + 5x$$

- Production increase costs

$$C_2 = 2\dot{x}^2 + 5\dot{x}$$

- $p = 10$
- $T = 1$
- $x_0 = 0, \quad x_T = 1$