Supplements 1 to Chapter IV

E. Gekeler
11/10/06

To Variational Calculus

Lemma 1 *(Lagrange, Fundamentallemma of Calculus of Variations)*
Let f, $g \in \mathcal{C}([0,T];\mathbb{R}^n)$ and $h \in \mathcal{C}_0^1([0,T];\mathbb{R}^n)$ then

$$
\begin{aligned}
&(1^\circ) \quad \forall\, h : \int_0^T f(t)^T h(t)\,dt = 0 && \Longrightarrow \quad f \equiv 0\,,\\
&(2^\circ) \quad \forall\, h : \int_0^T f(t)^T \dot h(t)\,dt = 0 && \Longrightarrow \quad f = \text{ constant},\\
&(3^\circ) \quad \forall\, h : \int_0^T [f(t)^T h(t) + g(t)^T \dot h(t)]\,dt = 0 && \Longrightarrow \quad g \in C^1([0\,,\,T];\mathbb{R}^n)\\
& && \quad \textit{and } f = \dot g\,.
\end{aligned}
$$

Proof. (1°) following [Amann]. We suppose that $f \neq 0$. Then there exists an i = 1:n so that $f_i \neq 0$. Since f_i continuous, there exist an $x_0 \in (0,T)$ and an $\varepsilon > 0$ so that $U_\varepsilon := (x_0 - \varepsilon, x_0 + \varepsilon)$ is still contained entirely in $[0,T]$ and that $f_i(t) \neq 0$ for all $t \in U_\varepsilon$. Now we choose a function $h_i \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ such that $\text{supp}(h_i) \subset U_\varepsilon$ and $h_i > 0$. Then

$$h := [0, \dots, 0, h_i, 0, \dots, 0]^T \in \; C_0^1([0,T],\mathbb{R}^n)$$

and

$$\int_0^T f(t)^T h(t)\,dt = \int_0^T f_i(t) h_i(t)\,dt = \int_{x_0-\varepsilon}^{x_0+\varepsilon} f_i(t) h_i(t)\,dt \neq 0$$

since $f_i h_i$ continuous, does not change sign, and does not disappear. This contradiction yields the assertion.

For the existence of a function h_i with the desired properties see [Amann], Bemerkung (2.12).

(2°) Following [Clegg],[Kosmol], p. 106. Let $c := T^{-1} \int_0^T f(t)\,dt \in \mathbb{R}^n$ then

$$\int_0^T (f(t) - c)\,dt = \int_0^T f(t)\,dt - \int_0^T dt = 0\,, \;\; \text{choose } h(t) = \int_0^t (f(t) - c)\,dt$$

then $\dot h(t) = f(t) - c$ and $h(0) = h(T) = 0$. Now, by assumption

$$
\begin{aligned}
&\int_0^T (f(t) - c)^T (f(t) - c)\,dt = \int_0^T (f(t) - c)^T \dot h(t)\,dt\\
&= \int_0^T f(t)^T \dot h(t)\,dt - c^T \int_0^T \dot h(t)\,dt = 0 - c^T (h(b) - h(a)) = 0.
\end{aligned}
$$

Therefore $|f(t) - c| = 0$ for all t hence $f(t) = c$.

Slight generalization possible [Clegg].

(3°) [Kosmol], p. 107, Let

$$a(t) := \int_0^t f(\tau)\,d\tau \in \mathbb{R}^n\,,$$

and $h \in \mathcal{C}^1([0,T],\mathbb{R}^n)$ where $h(a) = h(b) = 0$. Partial integration yields

$$\int_0^T f(t)^T h(t)\,dt = a(T)^T h(T) - a(0)^T h(0) - \int_0^T a(t)^T \dot{h}(t) = 0 - \int_0^T a(t)^T \dot{h}(t)\,dt$$

By assumption

$$\int_0^T f(t)^T h(t)\,dt = -\int_0^T g(t)^T \dot{h}(t)\,dt$$

therefore we obtain

$$\int_0^T g(t)^T \dot{h}(t)\,dt = \int_0^T a(t)^T \dot{h}(t)\,dt \implies \int_0^T (g(t)^T - a(t))^T \dot{h}(t)\,dt = 0$$

But $g(t) - a(t)$ constant by (2°) thus there exists a $c \in \mathbb{R}^n$ so that $g(t) = c + a(t)$. By the main theorem of differential and integral calculus, $a(t)$ is differentiable with derivative $f(t)$, therefore $\dot{g}(t) = f(t)$.

Theorem 1 *[Amann] Let $\mathcal{X}$ be a normed vector space, $\mathcal{U} \subset \mathcal{X}$ a subspace, $w \in \mathcal{X}$ arbitrary, $\mathcal{V} = w + \mathcal{U}$ an affine subspace, $\mathcal{D} \subset \mathcal{V}$ open. Further, let $J : \mathcal{D} \to \mathbb{R}$, $x^* \in \mathcal{D}$ and let $\delta J(x^*, \cdot) : \mathcal{U} \to \mathbb{R}$ exist. Then*

$$\forall\, x \in \mathcal{D} : J(x^*) \le J(x) \Longrightarrow \forall\, h \in \mathcal{U} : \delta J(x^*; h) = 0.$$

Variational Problem: Let $0 < T < \infty$, $\mathcal{M} \subset \mathbb{R}^m \times \mathbb{R}^m$ open, $a, b \in \mathbb{R}^m$,

$$\begin{aligned} &\mathcal{Z} := \{x \in \mathcal{C}^1([0,T];\mathbb{R}^m),\ x(0) = a,\ x(T) = b,\ (x(t),\dot{x}(t)) \subset \mathcal{M}\ \forall\, t \in [0,T]\},\\ &L \in \mathcal{C}^1([0,T] \times \mathcal{M};\mathbb{R}),\ J : \mathcal{Z} \ni x \mapsto J(x) := \textstyle\int_0^T L(t, x(t), \dot{x}(t))\,dt. \end{aligned}$$

Then the problem

$$\min\{J(x),\ x \in \mathcal{Z}\} \tag{1}$$

is called variational problem with fixed boundary conditions. Obviously this problem has not always a solution and in case a solution exists, it must not be unique. In the sequel, we derive a *necessary* condition for the existence of a solution. This condition is known under the name EULER equations. But, at first we write the problem in a more abstract form.

Theorem 2 (1°) *$\mathcal{X} := \mathcal{C}^1([0,T];\mathbb{R}^m)$ is a* BANACH *space w.r.t. the norm.*

$$\|x\|_{\mathcal{X}} := \max_{0 \le t \le T} |x(t)| + \max_{0 \le t \le T} |\dot{x}(t)|.$$

(2°) *$\mathcal{D}_M := \{x \in \mathcal{X},\ (x(t),\dot{x}(t)) \in \mathcal{M}\ \forall\, t \in [0,T]\}$ is open in $\mathcal{X}$.*
(3°) *$\mathcal{U} := \mathcal{C}_0^1([0,T];\mathbb{R}^m) := \{x \in \mathcal{X}, x(0) = x(T) = 0\}$ is a closed subspace in $\mathcal{X}$ w.r.t. the norm $\|\cdot\|_{\mathcal{X}}$.*
(4°) *Let $w \in \mathcal{X}$ arbitrary such that $w(0) = a$, $w(T) = b$, then $\mathcal{V} = w + \mathcal{U}$ is an affine subspace (linear manifold) in $\mathcal{X}$.*

Proof. [Amann] Lemmata 2.4, 2.5, Aufg. 1.2.4, p. 31.
Let now $\mathcal{D} = \mathcal{D}_M \cap \mathcal{V}$ then $\mathcal{D}$ open in $\mathcal{V}$ and, using the above notations and assumptions, Theorem 1 yields

$$\forall\, x \in \mathcal{D} : J(x^*) \le J(x) \implies \forall\ \in \mathcal{U} : \delta J(x^*; h) = 0\,.$$

Theorem 3 *Adopt the above notations and assumptions, then*

$$\delta J(x;h) = \int_0^T [D_2L(t,x(t),\dot{x}(t))h(t) + D_3L(t,x(t),\dot{x}(t))\dot{h}(t)]dt$$

$\forall\, x \in \mathcal{D}\ \ \forall\, h \in \mathcal{U}$, *and* $\delta J(x;\,\cdot\,)$ *is the* FRECHET-*derivative of* J *at the point* x.

Beweis [Amann], p. 21, [Heuser] II, § 191.
In Theorem 3, D_2L is the gradient of the mapping $x \mapsto L(t,x,y)$ (in m variables) for fixed (t,y) and D_3L is the gradient of the mapping $y \mapsto L(t,x,y)$ for fixed (t,x).

Theorem 4 *Let moreover the mapping* $[0,T] \ni t \mapsto D_3L(t,x(t)\dot{x}(t)) \in \mathbb{R}^m$ *be continuously differentiable then*

$$\delta J(x;h) = \int_0^T [D_2L(t,x(t)\dot{x}(t)) - \frac{d}{dt}D_3L(t,x(t)\dot{x}(t))]h(t)dt + D_3L(t,x(t),\dot{x}(t))h(t)|_0^T$$
$$= \int_0^T [D_2L(t,x(t),\dot{x}(t)) - \frac{d}{dt}D_3L(t,x(t),\dot{x}(t))]h(t)dt.$$

Proof. Partial integration regarding $h \in \mathcal{U}$.
An application of Lemma 1(3°) to Theorem 3 and 4 shows that the assumption of Theorem 4 can be cancelled.
Summary: Under the above assumptoins

$$\delta J(x^*;v) = 0\ \forall\, v \in C_0^1([0,T];\mathbb{R}^m)$$
$$\Longleftrightarrow [\operatorname{grad}_x L - \frac{d}{dt}\operatorname{grad}_{\dot{x}} L](t,x^*(t),\dot{x}^*(t)) = 0\ \forall\, t \in [0,T].$$

Examples. (1.) EULER equation in some special cases where $m = 1$ (abbreviated notation).

$$\int_0^T L(t,\dot{x})dt = \text{ extr! } \Longrightarrow \frac{d}{dt}L_{\dot{x}} = c \Longrightarrow \dot{x} = f(t,c),$$

if $L_{\dot{x}}$ invertible w.r.t. $\dot{x}$.

$$\int_0^T L(\dot{x})dt = \text{ Extr.! } \Longrightarrow L_{\dot{x}}(\dot{x}) = c \Longrightarrow \dot{x} = konst \Longrightarrow x \text{ straight line.}$$

$$\int_0^T L(x,\dot{x})dt = \text{ Extr.! } \Longrightarrow L_x - L_{\dot{x}x}\dot{x} - L_{\dot{x}\dot{x}}\ddot{x} = 0.$$

Multiplication by $\dot{x}$ yields

$$L_x\dot{x} - L_{\dot{x}x}\dot{x}\dot{x} - L_{\dot{x}\dot{x}}\dot{x}\ddot{x} = 0 \text{ or } \frac{d}{dt}(L - \dot{x}L_{\dot{x}}) = 0$$

and after integration the DUBOIS-REYMOND condition

$$L - \dot{x}L_{\dot{x}} = constant \quad \text{implicit differential equation for } x.$$

(2.) Mass point in central field (dimension $m = 3$). According to NEWTON's law (axiom) for $x(t) \in \mathbb{R}^3$

$$\begin{aligned} m\ddot{x} &= f(x) = -\operatorname{grad} U(x) && U \text{ potential energy,}\\ T &= \frac{m}{2}|\dot{x}|^2 && \text{kinetic energy,}\\ E &= T + U = \frac{m}{2}|\dot{x}|^2 + U(x) && \text{total energy (constant).} \end{aligned}$$

According to Example 1(c)

$$-m\ddot{x} - \operatorname{grad} U(x) = -(m\ddot{x} + \operatorname{grad} U(x)) = 0$$

EULER's equation of the variational problem

$$J(x) = \int_{t_1}^{t_2} L(x, \dot{x})dt = \int_{t_1}^{t_2} [\frac{m}{2}|\dot{x}|^2 - U(x)]dt = \text{ extr.!} \tag{2}$$

where $L = T - U$ is the LAGRANGE function! This result is called HAMILTON's principle of least action (dimension of J = energy · time) and is generally valid. Introducing $y = m\dot{x} = L_{\dot{x}}$ for new (additinal) variable (mass m constant) then we obtain the differential system

$$\begin{aligned} \dot{x} &= \operatorname{grad}_y H &&:= y/m && \text{(definition)},\\ \dot{y} &= -\operatorname{grad}_x H &&:= -\operatorname{grad} U(x) && \text{NEWTON's law.} \end{aligned}$$

It follows immediately that $H = E$ is an invariant of the system:

$$\begin{aligned} &konst = E = \frac{1}{2}y^T\dot{x} + U(x) = \frac{1}{2m}y^Ty + U(x) = H(x,y)\\ &= T + U = 2T - (T - U) = y^T\dot{x} - L. \end{aligned}$$

Because $H = y^T\dot{x} - L$, the HAMILTON function H is the LEGENDRE transformation of the LAGRANGE function L w.r.t. $\dot{x}$.
In general however not $\dot{x}$ is introduced for new variable but $\partial L/\partial\dot{x}$ where in generalized coordinate systems commonly q is written in place of x for the vector of space variables. Then the LEGENDRE transformation H of L is defined by $H(t,p,q) = p\dot{q} - L(t,q,\dot{q})$ where it is assumed that $p := \partial L/\partial\dot{q}$ is resolvable w.r.t. $\dot{q}$ and that the result is inserted in H.

Consider the variational problem:

$$J(y) = \int_\alpha^\beta y(x)^n(1 + y'(x)^2)^{1/2}dx = \text{ Extr.! } \forall\, y \in C^1([\alpha,\beta];\mathbb{R}),\ y(\alpha) = a,\ y(\beta) = b. \tag{3}$$

where x (instead t) is the independent variable and y is the dependent variable. Here the case $n = 1,\ 1/2,\ 0,\ -1/2,\ -1$ are of particular interest.
Case 1. n = 0. Shortest connection between two points in the plane. Because $L_y = 0$, the EULER equation yields immediately

$$\frac{d}{dx}L_{y'}(\ldots) = \frac{d}{dx}[y'(1 + (y')^2)^{-1/2}] = 0 \Longrightarrow y'' = 0.$$

$y = py + q$ is the unique straight line through the points (α, a) and (β, b).
Case 2. $n = 1$. See § 4.1; cf. also [Bryson-Ho], p. 64,65.
Case 3. $n = -1/2$. See § 4.1; cf. also [Clegg], p. 49; [Kosmol], §4.2.
Case 4. $n = 1/2$. Free motion in a homogeneous field, e.g. parabola trajectory.

In the following lemma let $(T^*, x^*) \in \mathbb{R}_+ \times C^1[0, T_0]$ be a solution of (**??**) hence $g(T^*, x^*(T^*)) = 0$. The vector space $C^1([0,\,T_0];\mathbb{R}^n)$ is equipped with the norm $\|x\| := \|x\|_\infty + \|\dot{x}\|_\infty$ being a BANACH space by this way, and we suppose that $\emptyset \neq \mathcal{U}_\varepsilon(x^*) = \{x,\ \|x - x^*\| < \varepsilon\}$ be a neighborhood of x^* in that space.

Lemma 2 *Let*

$$\frac{d}{dt}g(t,x^*(t))\Big|_{t=T^*} = g_t(T^*,x^*(T^*)) + \nabla_x g(T^*,x^*(T^*))\dot{x}^*(T^*) \neq 0\,,$$

then there exists a $\varepsilon > 0$ *and a function* $\varphi \in C^1(\mathcal{U}_\varepsilon;\mathbb{R})$ *such that*

$$g(\varphi(x), x(\varphi(x))) = 0 \text{ und } \varphi(x^*) = T^*\,.$$

Proof. By the Implicit Function Theorem in BANACH spaces the equation $G(t,x) := g(t,x(t)) = 0 \in \mathbb{R}$ is resolvable w.r.t. t in a neighborhood of $(T^*, x^*(T^*))$ without defect of smoothness if

$$\frac{\partial}{\partial t}G(t,x^*)\Big|_{t=T^*} \neq 0\,.$$

□

This function φ is now substituted into the objective function

$$\widetilde{J}(\varphi(x),x) = p(\varphi(x),x(\varphi(x))) + \int_0^{\varphi(x)} q(t,x,\dot{x})\,dt\,,$$

and ensuing the first variation is set equal to zero again. Regarding the EULER equations we then obtain for all test functions v where $v(0) = 0$ by partial integration and by an application of LEIBNIZ' rule

$$\begin{aligned}
0 &= \partial\widetilde{J}(x^*;v) = \frac{d}{d\varepsilon}\widetilde{J}(x^*+\varepsilon v)\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\Big[p\Big(\varphi(x^*+\varepsilon v), x^*\big(\varphi(x^*+\varepsilon v)\big) + \varepsilon v\big(\varphi(x^*+\varepsilon v)\big)\Big) \\
&\quad + \int_0^{\varphi(x^*+\varepsilon v)} q(t, x^*(t)+\varepsilon v(t), \dot{x}^*(t)+\varepsilon\dot{v}(t))\,dt\Big] \\
&= D_1 p(T^*,x^*(T^*))\partial\varphi(x^*;v) + \nabla_2 p\big(T^*,x^*(T^*)\big)\big(v(T^*) + \dot{x}^*(T^*)\partial\varphi(x^*;v)\big) \\
&\quad + q(T^*,x^*(T^*),\dot{x}^*(T^*))\partial\varphi(x^*;v) + \nabla_3 q(T^*,x^*(T^*),\dot{x}^*(T^*))v(T^*) \\
&\quad + \int_0^{T^*=\varphi(x^*)} \big[\nabla_2 q(t,x^*,\dot{x}^*) - \frac{d}{dt}\nabla_3 q(t,x^*,\dot{x}^*)\big]v\,dt\,.
\end{aligned}$$

Suppose first that

$$(\nabla_2 p + \nabla_3 q)v + (\nabla_2 p\dot{x} + p_t + q)\partial\varphi(x^*;v) = 0 \tag{4}$$

where the arguments T^* and x^* are dropped for simplicity, then the Fundamentallemma 1 yields the EULER equations in interval $[0,\, T^*]$ again. But the constraint $g(T,x(T)) = 0 \in \mathbb{R}$ does also hold for sufficiently small $|\varepsilon|$ in the following variated form

$$B(\varepsilon) := g\Big(\varphi(x^*+\varepsilon v), x^*\big(\varphi(x^*+\varepsilon v)\big) + \varepsilon v\big(\varphi(x^*+\varepsilon v)\big)\Big) = 0\,,$$

by the above regularity assumption, hence

$$\begin{aligned}
0 &= \frac{d}{d\varepsilon}B(\varepsilon)\Big|_{\varepsilon=0} \\
&= \frac{\partial g}{\partial t}(T^*,x^*(T^*))\partial\varphi(x^*;v) + \nabla_x g(T^*,x^*(T^*))[v(T^*) + \dot{x}^*(T^*)\partial\varphi(x^*;v)]\,.
\end{aligned}$$

Substitution of the resolution w.r.t. $\partial\varphi(x^*; v)$ into (4) yields — again in abbreviated form

$$\nabla_x p v + \nabla_{\dot{x}} q v + \left[\nabla_x p \dot{x} + p_t + q\right] \frac{\nabla_x g v}{\nabla_x g \dot{x} + g_t} = 0 \tag{5}$$

for arbitrary $v(T^*)$. By this way we obtain the necessary *condition of transversality* for the terminal time T^* at the point $(t, x(t)) = (T^*, x^*(T^*))$, namely

$$\boxed{(\nabla_x g \dot{x}^* + g_t)\,[\nabla_x p + \nabla_{\dot{x}} q] + (\nabla_x p \dot{x}^* + p_t + q)\,\nabla_x g = 0 \in \mathbb{R}_n}\,, \tag{6}$$

in addition to the EULER equation. By Lemma 2 it does make sense to suppose that the denominator in (5) is nonzero. Note also that *no additional properties* at all are required for the variation $\partial\varphi(x; v)$ of φ in the above computation.