Introduction to the Calculus of Variations
Lecture Notes
Lecture 2

Swarnendu Sil

Spring 2021, IISc

Classical Methods

Model Classical Problem

The classical problems and methods were all concerned with the case $n = 1$. The case $n \geq 2$, the so-called **multiple integrals in the calculus of variations**[1] is much harder and it took time to develop the tools needed to address it. The study of 'multiple integrals' began primarily with the Dirichlet integral and were satisfactorily resolved only with the modern direct methods.

Problem with prescribed Dirichlet condition Let $a, b \in \mathbb{R}$, $a < b$ and $\alpha, \beta \in \mathbb{R}^N$ be given. Let $f \in C^2\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$ be a given Lagrangian density.

We are interested in finding a function $u : [a,b] \to \mathbb{R}^N$ which satisfies the Dirichlet boundary conditions

$$u(a) = \alpha \qquad \text{and} \qquad u(b) = \beta,$$

and minimizes the following Lagrangian

$$I(u) = \int_a^b f(t, u(t), \dot{u}(t)) \, \mathrm{d}t$$

Remark 1. *When $n = 1$, it is more usual to write the independent variable as t rather than x. The reason for this is the fact that often in these type of problems, the independent variable is indeed time. We shall follow the usual custom and write t (or s).*

Now the **question is to chose the class of admissible functions**. So far, we have largely ignored this issue in presenting examples of variational problems,

[1] The name multiple integrals comes from the fact that before the advent of measure theory, it was a common practice to denote integration over a region of $\mathbb{R}^2$ as a 'double integral', written typically with two integral signs and integration over a portion of $\mathbb{R}^3$ as a 'triple integral', written typically with three integral signs e.g.

$$\iint\limits_A f(x,y) \, \mathrm{d}x\mathrm{d}y, \ \iiint\limits_V f(x,y,z) \, \mathrm{d}x\mathrm{d}y\mathrm{d}z \qquad \text{etc.}$$

but it is an immensely important issue in the calculus of variations. Firstly, **existence of minimizers in general depends on our choice of the class of admissible functions**. It is entirely possible that with a particular choice of the class of admissible functions, no minimizer exists whereas the same is not true for some other choice. In fact, **even the values of the infima can be different** for different choices! This is known as the **Lavrentiev gap phenomenon**, discovered by Mikhail Lavrentiev in 1926. Even when a minimizer exists for two different choices, it is possible that **a good choice of the class of admissible function would allow us to prove existence of minimizers, whereas a not so good choice might not!** So we would do best to make our choices wisely. The technical advantage gained by a judicious choice might be substantial.

Coming back to the problem at hand, certainly the space

$$X = \left\{ u \in C^1 \left([a,b] ; \mathbb{R}^N \right) : u(a) = \alpha,\ u(b) = \beta \right\}$$

is a possible choice. So we would be considering the following problem

$$(P) \qquad \inf_{u \in X} \left\{ I(u) = \int_a^b f(t, u(t), \dot{u}(t)) \ \mathrm{d}t \right\} = m.$$

where X is defined as above. We would come back to the problem later to reformulate it using a different choice.

Euler-Lagrange Equations

Now we are interested in deriving the equations that (at least with some assumptions) the critical points of a functional must satisfy. We all know that if $\bar{x} \in \mathbb{R}^n$ is a critical point for a C^1 function $F : \mathbb{R}^n \to \mathbb{R}$, then $\nabla F(\bar{x}) = 0$. The analogue of this statement in the context of functionals says that regular enough critical points of regular enough functionals must satisfy a differential equation, called the Euler-Lagrange equations for the functional.

Fundamental lemma of Calculus of Variations

Before going on to the derivation of the Euler-Lagrange equations, we would need an important lemma, called the fundamental lemma of Calculus of Variations.

Lemma 2 (Fundamental lemma of the calculus of variations)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and $u \in L^1_{loc}\left(\Omega; \mathbb{R}^N\right)$ be such that*

$$\int_\Omega \langle u(x), \psi(x) \rangle \ \mathrm{d}x = 0, \qquad \text{for every } \psi \in C_c^\infty\left(\Omega; \mathbb{R}^N\right) \tag{1}$$

then $u = 0$ a.e. in Ω.

Proof. Since $\psi = \left(0, \ldots, 0, \underbrace{\psi_i}_{i-\text{th place}}, 0, \ldots, 0\right) \in C_c^\infty\left(\Omega; \mathbb{R}^N\right)$ for any $\psi_i \in C_c^\infty\left(\Omega\right)$, we can argue componentwise. So it is enough to prove for $N = 1$. Now pick $K \subset \Omega$ be compact arbitrarily. It is enough to show $u = 0$ a.e. in K. Set

$$v := \begin{cases} \operatorname{sgn} u & \text{if } x \in K, \\ 0 & \text{if } x \in \mathbb{R}^n \setminus K. \end{cases}$$

Here sgn is the sign or signum function defined as

$$\operatorname{sgn} x = \begin{cases} +1 & \text{if } x > 0, \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases}$$

Mollify to find a sequence $\{v_s\} \subset C_c^\infty\left(\Omega\right)$ such that $\left\|v_s\right\|_{L^\infty} \leq \left\|v\right\|_{L^\infty}$ and $v_s \to v$ in L^1. Then, up to a subsequence $v_s \to v$ a.e.

Now, since $v_s \in C_c^\infty\left(\Omega\right)$, for every s, we have

$$\int_\Omega u\left(x\right) v_s\left(x\right) \mathrm{d}x = 0 \qquad \text{for every } s \geq 1.$$

By dominated convergence theorem, we have

$$\int_\Omega u\left(x\right) v\left(x\right) \mathrm{d}x = 0 \quad \Rightarrow \int_K \left|u\right| \mathrm{d}x = 0 \quad \Rightarrow u = 0 \ \text{ a.e. in } K.$$

$\square$

Note that the power of lemma derives from its quite weak hypothesis (of u being only L^1_{loc}). Indeed, the lemma holds with the stronger conclusion $u = 0$ for all $x \in \Omega$ if u is assumed to be continuous, i.e. $u \in C^0\left(\Omega\right)$ and the proof is elementary (Check!). Even slightly stronger hypothesis would have allowed a simpler proof. Just as an illustration, let us supply a different proof for the following L^2 version.

Lemma 3. *Let $\Omega \subset \mathbb{R}^n$ be open and $u \in L^2\left(\Omega; \mathbb{R}^N\right)$ be such that*

$$\int_\Omega \left\langle u\left(x\right), \psi\left(x\right)\right\rangle \ \mathrm{d}x = 0, \quad \text{for every } \psi \in C_c^\infty\left(\Omega; \mathbb{R}^N\right) \tag{2}$$

then $u = 0$ a.e. in Ω.

Proof. Fix $\varepsilon > 0$ arbitrary. Using the fact that $C_c^\infty\left(\Omega\right)$ is dense in $L^2\left(\Omega\right)$, there exists $\psi \in C_c^\infty\left(\Omega\right)$ such that

$$\left\|u - \psi\right\|_{L^2(\Omega)} \leq \varepsilon. \tag{3}$$

Now we have

$$\|u\|_{L^2(\Omega)}^2 = \int_\Omega \langle u, u\rangle \overset{(2)}{=} \int_\Omega \langle u, u-\psi\rangle$$
$$\overset{\text{Hölder}}{\leq} \|u\|_{L^2(\Omega)} \|u-\psi\|_{L^2(\Omega)}$$
$$\overset{(3)}{=} \varepsilon \|u\|_{L^2(\Omega)}.$$

Since $\varepsilon > 0$ is arbitrary, we must have $\|u\|_{L^2(\Omega)} = 0$. This obviously implies $u = 0$ a.e. in Ω. □

The lemma also has several variants. A very commonly used one is as follows.

Lemma 4. *Let $\Omega \subset \mathbb{R}^n$ be open and $u \in L^1_{loc}\left(\Omega; \mathbb{R}^N\right)$ be such that*

$$\int_\Omega \langle u(x), \psi(x)\rangle \, \mathrm{d}x = 0, \qquad \textit{for every } \psi \in C_c^\infty\left(\Omega; \mathbb{R}^N\right) \textit{ with } \int_\Omega \psi = 0, \quad (4)$$

then $u =$ constant a.e. in Ω.

We would not prove it here. Instead, we have asked you to try and supply a proof for the even more general version.

Lemma 5. *Let $\Omega \subset \mathbb{R}^n$ be open and $u, v \in L^1_{loc}\left(\Omega; \mathbb{R}^N\right)$ be such that*

$$\int_\Omega \langle u(x), \psi(x)\rangle \, \mathrm{d}x = 0, \quad (5)$$

holds for all for every $\psi \in C_c^\infty\left(\Omega; \mathbb{R}^N\right)$ satisfying the condition

$$\int_\Omega \langle v(x), \psi(x)\rangle \, \mathrm{d}x = 0. \quad (6)$$

Then there exists a number $\lambda \in \mathbb{R}$ such that we have $u = \lambda v$ a.e. in Ω.

Clearly, both Lemma 2 and Lemma 4 are special cases of Lemma 5, as can be readily seen by taking $v \equiv 0$ and $v \equiv 1$, respectively. Another important corollary is the *Du Bois-Raymond's lemma.*

Lemma 6 (Du Bois-Raymond's lemma)**.** *Let $\Omega \subset \mathbb{R}^n$ be open bounded and connected and $u \in L^1_{loc}\left(\Omega; \mathbb{R}^N\right)$ be such that*

$$\int_\Omega \langle u(x), D_i\psi(x)\rangle \, \mathrm{d}x = 0, \qquad \textit{for } 1 \leq i \leq n, \qquad \textit{for every } \psi \in C_c^\infty\left(\Omega; \mathbb{R}^N\right). \quad (7)$$

Then $u =$ constant a.e. in Ω.

Derivation of the Euler-Lagrange equations

Theorem 7 (Euler-Lagrange equation for Dirichlet boundary value)**.** *Let* $\alpha, \beta \in \mathbb{R}^N$ *be two given vectors and* $f = f(t, u, \xi) \in C^2\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$ *be a given function. Set*

$$X = \left\{u \in C^1\left([a,b]; \mathbb{R}^N\right) : u(a) = \alpha,\ u(b) = \beta\right\}$$

and consider the problem

$$(P) \qquad \inf_{u \in X} \left\{ I(u) = \int_a^b f(t, u(t), \dot{u}(t))\ \mathrm{d}t \right\} = m.$$

If $\bar{u} \in X \cap C^2\left([a,b]; \mathbb{R}^N\right)$ *is a minimizer for* (P)*, then* $\bar{u}$ *satisfies,*

$$(\boldsymbol{EL}) \qquad \frac{d}{dt}\left[f_\xi\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right)\right] = f_u\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right), \qquad \text{for every } t \in (a,b).$$

Writing out explicitly, the above vector equation reads,

$$\frac{d}{dt}\left[f_{\xi_i}\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right)\right] = f_{u_i}\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right), \qquad \text{for every } t \in (a,b).$$

for every $1 \leq i \leq N$*. Thus the* ***Euler-Lagrange*** *equations are a* system *of* N *second order ODEs, which employing matrix notation*[2]*, we write as*

$$\begin{aligned} f_{\xi\xi}\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right)\ddot{\bar{u}}(t) &+ f_{u\xi}\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right)\dot{\bar{u}}(t) \\ &+ f_{t\xi}\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right) - f_u\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right) = 0. \end{aligned}$$

Proof. Pick a $\phi \in C^1_c\left((a,b); \mathbb{R}^N\right)$. Now since $\phi(a) = 0 = \phi(b)$, we have $\bar{u} + h\phi \in X$ for any $h \in \mathbb{R}$. Now we define the function $g : \mathbb{R} \to \mathbb{R}$ by

$$g(h) := I(\bar{u} + h\phi).$$

Now we claim that $g \in C^1(\mathbb{R})$ (Check!) and since $\bar{u}$ is a minimizer, g must

[2] Note that $f_{\xi\xi}$ and $f_{u\xi}$ are $N \times N$ matrices, $\ddot{\bar{u}}$, $\dot{\bar{u}}$, $f_{t\xi}$ and f_u are vectors in $\mathbb{R}^N$, which we view as $N \times 1$ matrices.

have a local minima at 0. Thus we must have $g'(0)=0$. So we compute

$$
\begin{aligned}
0 &= \frac{d}{dh}\left[I\left(\bar{u}+h\phi\right)\right]\Big|_{h=0} \\
&= \frac{d}{dh}\left[\int_a^b f\left(t,\bar{u}\left(t\right)+h\phi\left(t\right),\dot{\bar{u}}\left(t\right)+h\dot{\phi}\left(t\right)\right)\,\mathrm{d}t\right]\Big|_{h=0} \\
&= \int_a^b \left(\frac{d}{dh}\left[f\left(t,\bar{u}\left(t\right)+h\phi\left(t\right),\dot{\bar{u}}\left(t\right)+h\dot{\phi}\left(t\right)\right)\right]\Big|_{h=0}\right)\,\mathrm{d}t \qquad \text{(Justify!)} \\
&= \int_a^b \left(\left[\begin{aligned}&\left\langle f_\xi\left(t,\bar{u}\left(t\right)+h\phi\left(t\right),\dot{\bar{u}}\left(t\right)+h\dot{\phi}\left(t\right)\right),\dot{\phi}\left(t\right)\right\rangle \\ &\quad+\left\langle f_u\left(t,\bar{u}\left(t\right)+h\phi\left(t\right),\dot{\bar{u}}\left(t\right)+h\dot{\phi}\left(t\right)\right),\phi\left(t\right)\right\rangle\end{aligned}\right]\Bigg|_{h=0}\right)\,\mathrm{d}t \\
&= \int_a^b \left[\left\langle f_\xi\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\dot{\phi}\left(t\right)\right\rangle+\left\langle f_u\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\phi\left(t\right)\right\rangle\right]\,\mathrm{d}t \\
&= \int_a^b \left\langle\left[-\frac{d}{dt}f_\xi\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)+f_u\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)\right],\phi\left(t\right)\right\rangle\,\mathrm{d}t,
\end{aligned}
$$

where we have used integration by parts in the last line and the fact that $\phi\left(a\right)=0=\phi\left(b\right)$. We conclude by applying the fundamental lemma. □

Remark 8. 1. *The Euler-Lagrange equation is sometimes called the* ***first variation formula****. The name comes from the fact that given a functional $I(u)$, its* ***first variation*** *at $\bar{u}$ is (its* Gateaux derivative *at $\bar{u}$,)*

$$\delta I\left(\bar{u},\phi\right):=\frac{d}{dh}\left[I\left(\bar{u}+h\phi\right)\right]\Big|_{h=0}.$$

The EL equation is just follows from "first variation $=0$."

This is also the reason for the name Calculus of variations. All we used to do is to compute the first variation and the second variation!

2. *Although the Euler-Lagrange equations themselves are an important and useful tool, probably even more important point here is the method of proof. As we shall see in a moment, different problems might lead to different equations, but the general method is always the same. So we summarize the method below in words.*

   - *We construct a* ***one-parameter family of perturbations in the admissible class of functions****,* **as general as we can**. *This often requires clever ideas. The* art *of deducing Euler-Lagrange equations lies mostly in this crucial step.*
   - *Using the abovementioned perturbation, we reduce the problem to the first order (derivative) condition for a local minima of a* real valued real function.

- *Use integration by parts intelligently on the identity obtained from first order condition to reduce it to a form where the fundamental lemma of calculus of variations can do the rest for us.*

Let us now derive the Euler-Lagrange equation for the **natural boundary conditions.** The reasons for the name will be apparent soon.

Theorem 9 (Euler-Lagrange equation for natural boundary conditions)**.** *Let $f = f(t, u, \xi) \in C^2\left([a, b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$ be a given function. Set*

$$X = \left\{u \in C^1\left([a, b]; \mathbb{R}^N\right)\right\}$$

and consider the problem

$$(P) \qquad \inf_{u \in X} \left\{ I(u) = \int_a^b f(t, u(t), \dot{u}(t)) \ \mathrm{d}t \right\} = m.$$

If $\bar{u} \in X \cap C^2\left([a, b]; \mathbb{R}^N\right)$ is a minimizer for (P), then $\bar{u}$ satisfies,

$$(\boldsymbol{EL}) \begin{cases} \dfrac{d}{dt}\left[f_\xi\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right)\right] = f_u\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right), & \textit{for every } t \in (a, b), \\ \qquad f_\xi\left(a, \bar{u}(a), \dot{\bar{u}}(a)\right) = 0 = f_\xi\left(b, \bar{u}(b), \dot{\bar{u}}(b)\right). \end{cases}$$

Remark 10. *Although the equations $f_\xi\left(a, \bar{u}(a), \dot{\bar{u}}(a)\right) = 0 = f_\xi\left(b, \bar{u}(b), \dot{\bar{u}}(b)\right)$ are clearly boundary conditions, they are indeed part of the Euler-Lagrange equations in this case. Even though we have not encoded any boundary condition at all in our admissible class X, still these conditions popped up, 'on their own' or 'naturally', so to speak. This is the reason these boundary conditions are called natural boundary conditions.*

Proof. The proof is easy as long as we remembered the general scheme. Here we can pick any $\phi \in C^1\left([a, b]; \mathbb{R}^N\right)$ and there is no need to require compact support. Indeed, allowing only perturbations by C_c^1 functions would not be the most general type of perturbation in X that we can have and thus would fail to extract the full information. You can easily check that using $\phi \in C^1\left([a, b]; \mathbb{R}^N\right)$ gives only the differential equation part of the EL equations as before, but fails to extract the boundary condition part.

Now we proceed as before to obtain

$$
\begin{aligned}
0 &= \frac{d}{dh}\left[I\left(\bar{u}+h\phi\right)\right]\Big|_{h=0} \\
&= \int_a^b \left[\left\langle f_\xi\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\dot{\phi}\left(t\right)\right\rangle + \left\langle f_u\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\phi\left(t\right)\right\rangle\right]\ \mathrm{d}t \\
&= \int_a^b \left\langle\left[-\frac{d}{dt}f_\xi\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)+f_u\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)\right],\phi\left(t\right)\right\rangle\ \mathrm{d}t \\
&\qquad\qquad + \left\langle f_\xi\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\phi\left(t\right)\right\rangle\Big|_a^b \\
&= \int_a^b \left\langle\left[-\frac{d}{dt}f_\xi\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)+f_u\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)\right],\phi\left(t\right)\right\rangle\ \mathrm{d}t \\
&\qquad + \left\langle f_\xi\left(b,\bar{u}\left(b\right),\dot{\bar{u}}\left(b\right)\right),\phi\left(b\right)\right\rangle - \left\langle f_\xi\left(a,\bar{u}\left(a\right),\dot{\bar{u}}\left(a\right)\right),\phi\left(a\right)\right\rangle.
\end{aligned}
$$

Now the term inside the integral yields the equation as before by the fundamnetal lemma. The boundary conditions follow from the last two terms by noticing that $\phi(a)$ and $\phi(b)$ are arbitrary. □

Example: Eigenvalue problem To get a better sense of the two types of boundary conditions, we now compute the Euler-Lagrange equation for both the Dirichlet and the natural boundary value problem for a simple functional. Let $n, N = 1$, $\lambda \in \mathbb{R}$ and consider the Lagrangian density

$$f\left(t,u,\xi\right) = \frac{1}{2}\xi^2 - \frac{\lambda^2}{2}u^2.$$

For the homogeneous Dirichlet problem, we define

$$X_{\mathrm{Dir},0} := \left\{u \in C^1\left(\left[0,T\right];\mathbb{R}^N\right) : u\left(0\right) = 0 = u\left(T\right)\right\}.$$

The variational problem is

$$\inf\left\{I\left(u\right) = \int_0^T f\left(t,u\left(t\right),\dot{u}\left(t\right)\right)\ \mathrm{d}t : u \in X_{\mathrm{Dir},0}\right\} = m.$$

C^2 minimizers satisfy the following homogeneous Dirichlet boundary value problem

$$\begin{cases}\ddot{u} = -\lambda^2 u & \text{in } [0,T] \\ u(0) = 0 = u(T). \end{cases} \tag{8}$$

On the other hand, C^2 minimizers of

$$\inf\left\{I\left(u\right) = \int_0^T f\left(t,u\left(t\right),\dot{u}\left(t\right)\right)\ \mathrm{d}t : u \in C^1\left(\left[0,T\right]\right)\right\} = m,$$

satisfy the following homogeneous Neumann problem

$$\begin{cases} \ddot{u} = -\lambda^2 u & \text{in } [0,T] \\ \dot{u}(0) = 0 = \dot{u}(T). \end{cases} \tag{9}$$

Solutions of (8) and (9) are called Dirichlet and Neumann eigenfunctions for the Laplacian (albeit in one dimension the Laplacian is just the one and only second derivative). Can you find what they are in this simple setting?

Constrained minimization and Lagrange multiplier

Often, one has to run into minimization problems where there are additional subsidiary constraints which are not boundary conditions, but some type of pointwise constraint which is required to satisfied throughout the domain or integral constraints on the whole domain.

The technique to deal with such problems are usually called Lagrange multiplier technique. We shall prove the following theorem about integral constraints.

Theorem 11 (Lagrange Multiplier)**.** *Let* $f = f(t, u, \xi) \in C^2\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$, $g = g(t, u, \xi) \in C^2\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$, $\alpha, \beta \in \mathbb{R}^N$ *be given and*

$$X = \left\{ \begin{array}{l} u \in C^1\left([a,b]; \mathbb{R}^N\right) : u(a) = \alpha,\ u(b) = \beta, \\ \qquad\qquad\qquad\qquad\qquad \displaystyle\int_a^b g(t, u(t), \dot{u}(t))\ \mathrm{d}t = 0 \end{array} \right\}.$$

Now consider the problem

$$(P) \qquad \inf_{u \in X} \left\{ I(u) = \int_a^b f(t, u(t), \dot{u}(t))\ \mathrm{d}t \right\} = m.$$

If $\bar{u} \in X \cap C^2\left([a,b]; \mathbb{R}^N\right)$ *is a minimizer for* (P), *then there exists a* $\lambda \in \mathbb{R}$, *called the* ***Lagrange multiplier****, such that* $\bar{u}$ *satisfies,*

$$\left(\frac{d}{dt}\left[f_\xi\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right)\right] - f_u\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right) \right) \\ = \lambda \left(\frac{d}{dt}\left[g_\xi\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right)\right] - g_u\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right) \right)$$

for every $t \in (a,b)$, *provided*

$$\frac{d}{dt}\left[g_\xi\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right)\right] \not\equiv g_u\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right)$$

in (a,b).

Remark 12. *1. Sometimes the equation is written as (parameter is $-\lambda$)*

$$\frac{d}{dt}\left[f_{\xi}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)+\lambda g_{\xi}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)\right]$$
$$=f_{u}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)+\lambda g_{u}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right).$$

*2. The theorem basically says: The **constrained minimization** of*

$$\int_a^b f\left(t,u\left(t\right),\dot{u}\left(t\right)\right)\ \mathrm{d}t$$

with the constraint

$$\int_a^b g\left(t,u\left(t\right),\dot{u}\left(t\right)\right)\ \mathrm{d}t=0,$$

*is equivalent to the **unconstrained minimization** of*

$$\int_a^b f\left(t,u\left(t\right),\dot{u}\left(t\right)\right)\ \mathrm{d}t+\lambda\int_a^b g\left(t,u\left(t\right),\dot{u}\left(t\right)\right)\ \mathrm{d}t$$

for some $\lambda\in\mathbb{R}$.

Proof. We begin by claiming that there exists a $\psi\in C_c^{\infty}\left((a,b);\mathbb{R}^N\right)$ such that

$$\int_a^b\left[\left\langle g_{\xi}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\dot{\psi}\left(t\right)\right\rangle+\left\langle g_{u}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\psi\left(t\right)\right\rangle\right]\ \mathrm{d}t\neq 0. \tag{10}$$

Suppose this is not the case. Then integrating by parts, for all $\psi\in C_c^{\infty}\left((a,b);\mathbb{R}^N\right)$, we have,

$$\begin{aligned}0&=\int_a^b\left[\left\langle g_{\xi}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\dot{\psi}\left(t\right)\right\rangle+\left\langle g_{u}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\psi\left(t\right)\right\rangle\right]\ \mathrm{d}t\\&=\int_a^b\left\langle\frac{d}{dt}\left[g_{\xi}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)\right]-g_{u}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\psi\left(t\right)\right\rangle\ \mathrm{d}t.\end{aligned}$$

But this implies, by the fundamnetal lemma, Since

$$\frac{d}{dt}\left[g_{\xi}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)\right]-g_{u}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)\equiv 0$$

in (a,b), which contradicts our hypothesis. Once we have found a $\psi\in C_c^{\infty}\left((a,b);\mathbb{R}^N\right)$ satisfying (10), we can of course normalize if necessary to obtain $\psi\in C_c^{\infty}\left((a,b);\mathbb{R}^N\right)$ such that

$$\int_a^b\left[\left\langle g_{\xi}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\dot{\psi}\left(t\right)\right\rangle+\left\langle g_{u}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right),\psi\left(t\right)\right\rangle\right]\ \mathrm{d}t=1.$$

So from now, we shall assume that we have chosen such a ψ.

Now we pick $\phi \in C_c^\infty\left((a,b);\mathbb{R}^N\right)$ arbitrary, ψ as above and for $\varepsilon, h \in \mathbb{R}$, we set

$$F(\varepsilon,h) := \int_a^b f\left(t, \bar{u}(t) + \varepsilon\phi(t) + h\psi(t), \dot{\bar{u}}(t) + \varepsilon\dot{\phi}(t) + h\dot{\psi}(t)\right)\ \mathrm{d}t$$

$$G(\varepsilon,h) := \int_a^b g\left(t, \bar{u}(t) + \varepsilon\phi(t) + h\psi(t), \dot{\bar{u}}(t) + \varepsilon\dot{\phi}(t) + h\dot{\psi}(t)\right)\ \mathrm{d}t$$

We can easily check that $G \in C^1(\mathbb{R}\times\mathbb{R})$. Also, clearly, $G(0,0) = 0$ and we have

$$G_h(0,0) = \int_a^b \left[\left\langle g_\xi\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right), \dot{\psi}(t)\right\rangle + \left\langle g_u\left(t, \bar{u}(t), \dot{\bar{u}}(t)\right), \psi(t)\right\rangle\right]\ \mathrm{d}t = 1.$$

Hence the implicit function theorem implies the existence of a $\varepsilon_0 > 0$ and a function $\bar{h} \in C^1([-\varepsilon_0, \varepsilon_0])$ such that

$$\bar{h}(0) = 0 \qquad \text{and } G\left(\varepsilon, \bar{h}(\varepsilon)\right) = 0 \quad \text{for every } \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

Note that the last equation implies

$$\bar{u} + \varepsilon\phi + \bar{h}(\varepsilon)\psi \in X \qquad \text{for every } \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

It also implies (by differentiating),

$$G_\varepsilon\left(\varepsilon, \bar{h}(\varepsilon)\right) + G_h\left(\varepsilon, \bar{h}(\varepsilon)\right)\bar{h}'(\varepsilon) = 0 \qquad \text{for every } \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

So we deduce

$$\bar{h}'(0) = -G_\varepsilon(0,0). \tag{11}$$

Now once again we use the technique we have already seen. Since $\bar{u}$ is a minimizer, the real valued function on $[-\varepsilon_0, \varepsilon_0]$, given by

$$\varepsilon \mapsto F\left(\varepsilon, \bar{h}(\varepsilon)\right)$$

must have a local minima at $\varepsilon = 0$. So we have

$$0 = \frac{d}{d\varepsilon}\left[F\left(\varepsilon, \bar{h}(\varepsilon)\right)\right]\Big|_{\varepsilon=0} = F_\varepsilon(0,0) + F_h(0,0)\,\bar{h}'(0).$$

So setting $\lambda = F_h(0,0)$ and using (11), we get

$$F_\varepsilon(0,0) - \lambda G_\varepsilon(0,0).$$

This concludes the proof. $\square$

Example: Newton's laws for a point particle

Now we are going to calculate the Euler-Lagrange equation for the motion of a point particle. Let $m > 0$ be the mass and $x(t) \in \mathbb{R}^3$ be the position of a point particle at time $t \in [0,T]$. Let $U : \mathbb{R}^3 \to \mathbb{R}$ written as

$$U(t,x) = U(x)$$

be a given potential energy function. The variational problem is

$$m = \inf \left\{ I(x) := \int_0^T f\left(t, x(t), \dot{x}(t)\right) \, \mathrm{d}t : x(0) = x_0, x(T) = x_1 \right\},$$

where the form of the Lagrangian density is

$$f(t, x, \xi) = \frac{1}{2} m \xi^2 - U\left(x\right).$$

The EL equation reads

$$\begin{aligned} 0 &= \frac{d}{dt}\left[f_\xi\left(t, x(t), \dot{x}(t)\right)\right] - f_x\left(t, x(t), \dot{x}(t)\right) \\ &= \frac{d}{dt}\left[m\dot{x}(t)\right] + \nabla U\left(x(t)\right) = m\ddot{x}(t) + \nabla U\left(x(t)\right). \end{aligned}$$

i.e. the Newton's law of motion,

$$m\ddot{x}(t) = -\nabla U\left(x(t)\right) := F\left(x(t)\right),$$

where $\ddot{x}(t)$ is obviously the 'acceleration' of the particle at time t and the negative gradient of the potential energy $F\left(x(t)\right) = -\nabla U\left(x(t)\right)$ is called 'force'.