Chapter 3

Lecture 3

3.1 Euler-Lagrange equations in field theory

3.1.1 Preliminaries

Definition 3.1 (Field)
Let us consider two topological spaces $\mathcal{M}$ and $\mathcal{N}$. Let us consider a linear space $\mathcal{F}$ of functions defined by some proposition $\mathbb{P}$, i.e.

$$\mathcal{F} = \{\phi | \phi : \mathcal{M} \longrightarrow \mathcal{N} \quad , \quad \mathbb{P} \text{ is true for } \phi\}.$$

An element $\phi \in \mathcal{F}$ is called a field.

In what follows we are going to consider $\mathcal{M}$ a bounded domain D in the Euclidean space $\mathbb{R}^n$ with Euclidean coordinates (which we will denote with $\{x^\alpha\}_{\alpha=1,\dots,n}$). $\mathcal{N}$ will be $\mathbb{R}$. We will denote as usual with ∂D the boundary of D. $\mathcal{F}$ will be a space of functions

$$\phi^j : D \subset \mathbb{R}^n \longrightarrow \mathbb{R} \quad , \quad j = 1, \dots, N$$

sufficiently regular, for example of class C^2 (this is our $\mathbb{P}$ above).

Definition 3.2 (Functional of fields)
A functional $S[\boldsymbol{\phi}]$ *of the fields $\boldsymbol{\phi} = (\phi^j(x^\mu))_{j=1,\dots,N} \in \mathcal{F}$ is an integral of a function $\mathcal{S}$ of ϕ^j, $\partial_\nu \phi^k$ and, eventually, x^ρ:*

$$S[\boldsymbol{\phi}] = \int_D \mathcal{S}(x^\mu, \phi^j(x^\nu), \partial_\rho \phi^k(x^\sigma)) d^n x.$$

We will now see how some equations from our fields can be defined using a *stationary functional principle* starting from a given functional of the fields, as defined above. We first will give the following

Definition 3.3 (Field fluctuation (or variation))
Let us consider $\boldsymbol{\phi} \in \mathcal{F}$. Let $\delta\boldsymbol{\phi}$ be a differentiable function over D with the following properties:

1. *$\delta\boldsymbol{\phi}$ is finite on D;*

2. *there exists* $N \subset D$*, such that* $N \cup \partial N \subset D$ *and* $\delta\boldsymbol{\phi} \equiv 0$ *on* $D \setminus N$ *(thus in particular* $\delta\boldsymbol{\phi} \equiv 0$ *on* ∂D*).*

3. $\exists \bar{\epsilon} \in \mathbb{R}^+$ *such that* $\forall\, 0 < \epsilon < \bar{\epsilon}$*,* $\boldsymbol{\phi} + \epsilon\delta\boldsymbol{\phi} \in \mathcal{F}$*.*

Then $\delta\boldsymbol{\phi}$ *is called a* fluctuation *of the field* $\boldsymbol{\phi}$*.*

3.1.2 Functional Derivatives

In terms of the fluctuations of the fields we can define the variation of a functional. Intuitively this is the change in the functional due to a fluctuation in the fields, and can be formalized as below.

Definition 3.4 (Finite variation of a functional)
Let us consider a field theory defined, according to the above notation, as a set of fields $\boldsymbol{\phi} \in \mathcal{F}$ *on* D*. Let* $S[\boldsymbol{\phi}]$ *be a functional of the fields* $\boldsymbol{\phi}$*. The expression*

$$\Delta S[\boldsymbol{\phi}] = S[\boldsymbol{\phi} + \delta\boldsymbol{\phi}] - S[\boldsymbol{\phi}]$$

is the variation of the functional S *due to the field fluctuation* $\delta\boldsymbol{\phi}$*. In what follows we are going to use the notation*

$$\Delta_\epsilon S[\boldsymbol{\phi}] = S[\boldsymbol{\phi} + \epsilon\delta\boldsymbol{\phi}] - S[\boldsymbol{\phi}].$$

With a procedure similar to the one we use for the derivative of a function we can now define the functional derivative of a functional.

Definition 3.5 *(Functional derivative of a functional)*
Given a functional $S[\boldsymbol{\phi}]$ *of the fields* $\boldsymbol{\phi}$*, the first functional derivative of* $S[\boldsymbol{\phi}]$*,*

$$\frac{\delta S[\boldsymbol{\phi}]}{\delta\boldsymbol{\phi}(x)} = \left(\frac{\delta S[\boldsymbol{\phi}]}{\delta\phi^j(x^\mu)}\right)_{j=1,\dots,N}$$

is defined as

$$\begin{aligned}\int_D \left\langle \frac{\delta S[\boldsymbol{\phi}]}{\delta\boldsymbol{\phi}(x)}, \delta\boldsymbol{\phi}(x)\right\rangle d^n x \quad &\equiv \quad \int_D \sum_j^{1,N} \frac{\delta S[\boldsymbol{\phi}]}{\delta\phi^j(x^\mu)}\delta\phi^j(x^\rho) d^n x \\ &\overset{\text{def.}}{=} \lim_{\epsilon\to 0}\frac{\Delta_\epsilon S[\boldsymbol{\phi}]}{\epsilon} \equiv \left.\frac{dS[\boldsymbol{\phi}+\epsilon\delta\boldsymbol{\phi}]}{d\epsilon}\right]_{\epsilon=0}. \end{aligned} \tag{3.1}$$

In terms of the first functional derivative we can define a stationary point (which is a *field configuration*) for a functional $S[\boldsymbol{\phi}]$. Technically it is called an extremal.

3.1.3 Estremals and field equations

Definition 3.6 (Extremal of a functional)
A (set of) field(s) $\boldsymbol{\phi}_0 \in \mathcal{F}$ *is an* extremal *for the functional* $S[\boldsymbol{\phi}]$ *if for all field fluctuations* $\delta\boldsymbol{\phi}$ *it holds*

$$\frac{\delta S[\boldsymbol{\phi}_0]}{\delta\boldsymbol{\phi}} \equiv 0.$$

Just for later convenience, analogously to what is done for finite degrees of freedom systems, we define the Euler-Lagrange equations for the fields.

Definition 3.7 (Euler-Lagrange equations)
Let $\boldsymbol{\phi}$ be a (set of) field(s) in a theory (defined as above) described by the action functional

$$S[\boldsymbol{\phi}] = \int_D \mathcal{S}(x^\mu, \phi^j(x^\nu), \partial_\rho \phi^k(x^\sigma)) d^n x.$$

The differential equations

$$\frac{\partial \mathcal{S}}{\partial \phi^k} - \sum_{\mu}^{1,n} \partial_\mu \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) = 0 \tag{3.2}$$

are called the Euler-Lagrange equations *for* $\boldsymbol{\phi}$[1]*.*

All the above set of definitions is a premise to the following proposition, which is crucial in setting up a relation between the Euler-Lagrange equations of a system and the variation of the functional associated to the field theory.

Proposition 3.1 (Conditions for an extremal)
A field $\boldsymbol{\phi}_0$ is an extremal of the functional $\mathcal{S}[\boldsymbol{\phi}]$ if and only if it satisfies the system of Euler-Lagrange equations for $\boldsymbol{\phi}$.

Proof:

Let us consider the ϵ-finite variation of the functional $S[\boldsymbol{\phi}]$, that we write with the explicit dependence from the fields and their derivatives:

$$\Delta_\epsilon S[\phi] = \int_D \left[\mathcal{S}(x^\mu; \phi^j + \epsilon \delta \phi^j; \partial_\nu \phi^k + \epsilon \partial_\alpha \delta \phi^k) - \mathcal{S}(x^\mu; \phi^j; \partial_\nu \phi^k) \right] d^n x.$$

We will now develop the first term in square brackets, which is a function of ϵ, as a Taylor series in ϵ, stopping at first order. We thus get

$$\begin{aligned} &\mathcal{S}(x^\mu; \phi^j + \epsilon \delta \phi^j; \partial_\nu \phi^k + \epsilon \partial_\alpha \delta \phi^k) = \\ &\quad = \mathcal{S}(x^\mu; \phi^j; \partial_\alpha \phi^k) + \\ &\qquad + \sum_j^{1,N} \frac{\partial \mathcal{S}}{\partial \phi^j} (\delta \phi^j) \epsilon \\ &\qquad + \sum_k^{1,N} \sum_\mu^{1,n} \frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} (\partial_\mu \delta \phi^k) \epsilon + \mathcal{O}(\epsilon). \end{aligned}$$

Then

$$\Delta_\epsilon S[\phi] = \epsilon \int_D \left[\sum_j^{1,N} \frac{\partial \mathcal{S}}{\partial \phi^j} \delta \phi^j + \sum_k^{1,N} \sum_\mu^{1,n} \frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \partial_\mu \delta \phi^k + \mathcal{O}(\epsilon) \right] d^n x.$$

We can now perform an integration by parts on the second term:

$$\begin{aligned} &\int_D \sum_k^{1,N} \sum_\mu^{1,n} \frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \partial_\mu \delta \phi^k d^n x = \\ &\quad = \sum_k^{1,N} \sum_\mu^{1,n} \int_D \frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \partial_\mu \delta \phi^k d^n x \end{aligned}$$

[1]Please, note the difference between S and $\mathcal{S}$

$$= \sum_{k}^{1,N} \sum_{\mu}^{1,n} \int_D \partial_\mu \left(\left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) \delta\phi^k \right) d^n x +$$
$$- \sum_{k}^{1,N} \sum_{\mu}^{1,n} \int_D \partial_\mu \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) \delta\phi^k d^n x$$

and rewrite $\Delta_\epsilon S[\phi]$ as

$$\begin{aligned} \Delta_\epsilon S[\phi] &= \epsilon \int_D \left[\sum_{j}^{1,N} \left(\frac{\partial \mathcal{S}}{\partial \phi^j} - \sum_{\mu}^{1,n} \partial_\mu \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^j)} \right) \right) \delta\phi^j \right] d^n x + \int_D \mathcal{O}(\epsilon) dx^n \\ & \quad + \epsilon \sum_{k}^{1,N} \sum_{\mu}^{1,n} \int_D \partial_\mu \left(\left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) \delta\phi^k \right) d^n x. \end{aligned}$$

We have to deal now with the last contribution: in particular we see that the integrand is a derivative with respect to x^μ integrated on the volume of D. We can separate the x^μ integration from the others and we obtain

$$\begin{aligned} & \int_D \partial_\mu \left(\left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) \delta\phi^k \right) d^n x = \\ & \quad = \int_D \partial_\mu \left(\left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) \delta\phi^k \right) d^n x \\ & \quad = \int dx^1 \dots dx^{\mu-1} \dots dx^{\mu+1} \dots dx^n \int_{\bar{x}^\mu_{(0)}}^{\bar{x}^\mu_{(1)}} \partial_\mu \left(\left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) \delta\phi^k \right) dx^\mu \\ & \quad = \int dx^{n-1} \left[\left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) \delta\phi^k \Big]_{\bar{x}^\mu_{(1)}} - \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) \delta\phi^k \Big]_{\bar{x}^\mu_{(0)}} \right]. \end{aligned}$$

We note that the integration limits $\bar{x}^\mu_{(0/1)}$ are functions of the remaining coordinates. Moreover they are on the boundary of ∂D of D. So the two terms inside the square brackets have to be calculated on the boundary of D, where the fluctuations of the fields vanish by definition. Thus the integrand function above, a function of $x^1, \dots, x^{\mu-1}, x^{\mu+1}, \dots, x^n$, vanishes identically. The integral then also vanishes, so that the Taylor expansion from which we started is simply

$$\Delta_\epsilon S[\boldsymbol{\phi}] = \epsilon \int_D \left[\sum_{j}^{1,N} \left(\frac{\partial \mathcal{S}}{\partial \phi^j} - \sum_{\mu}^{1,n} \partial_\mu \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^j)} \right) \right) \delta\phi^j \right] d^n x + \int_D \mathcal{O}(\epsilon) dx^n.$$

From the above expression is easier to calculate $\lim_{\epsilon \to 0} \left(\Delta_\epsilon S[\boldsymbol{\phi}] / \epsilon \right)$:

$$\begin{aligned} \lim_{\epsilon \to 0} \frac{\Delta_\epsilon S[\boldsymbol{\phi}]}{\epsilon} &= \int_D \left[\sum_{j}^{1,N} \left(\frac{\partial \mathcal{S}}{\partial \phi^j} - \sum_{\mu}^{1,n} \partial_\mu \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^j)} \right) \right) \delta\phi^j \right] d^n x \\ &= \int_D \left[\left\langle \left(\frac{\partial \mathcal{S}}{\partial \boldsymbol{\phi}} - \sum_{\mu}^{1,n} \partial_\mu \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \boldsymbol{\phi})} \right) \right), \delta\boldsymbol{\phi} \right\rangle \right] d^n x. \end{aligned}$$

If we compare this expression with the definition of the first functional variation of $S[\boldsymbol{\phi}]$,

$$\int \left\langle \frac{\delta S[\boldsymbol{\phi}]}{\delta \boldsymbol{\phi}}, \delta\boldsymbol{\phi} \right\rangle d^n x,$$

we obtain

$$\frac{\delta S[\phi]}{\delta \phi} = \frac{\partial \mathcal{S}}{\partial \phi} - \sum_{\mu}^{1,n} \partial_\mu \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi)} \right),$$

which is a shorthand for

$$\frac{\delta S[\phi]}{\delta \phi} = \left(\frac{\partial \mathcal{S}}{\partial \phi^k} - \sum_{\mu}^{1,n} \partial_\mu \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) \right)_{k=1,\dots,N} . \tag{3.3}$$

We have thus calculated the first functional variation of the functional and, using this result, we can easily prove the if and only if condition:

$\Rightarrow$) Let ϕ_0 be an extremal for S. Then by definition $\delta S[\phi_0]/(\delta\phi) = 0$, which from (3.3) implies

$$\frac{\partial \mathcal{S}}{\partial \phi^k} - \sum_{\mu}^{1,n} \partial_\mu \left(\frac{\partial \mathcal{S}}{\partial(\partial_\mu \phi^k)} \right) = 0 \quad , \quad k = 1, \dots, N.$$

$\Leftarrow$) If for a given ϕ_0 the Euler-Lagrange equations are satisfied, then $\delta S[\phi_0]/\ (\delta\phi) = 0$.

Note that in this proposition we are speaking of an extremal for a functional, without specifying if that gives a maximum or a minimum.

□

3.2 Stationary action principle

We are now going to connect the above statements with the heuristic picture we gave in lecture 2. Let us consider indeed our discrete system, described for example by the Lagrangian L_θ of equation (2.4). We know from the course of analytical mechanics that the Euler-Lagrange equations associated to L_θ can be derived from a stationary action principle from the action

$$S_\theta = \int dt L_\theta.$$

Using the heuristic limit procedure as we did before we then know that the action functional S_θ will be transformed in an action functional with Lagrangian L which is the integral of the Lagrangian density $\mathcal{L}$, so that, as quickly anticipated in the previous lecture close to the end of section 2.2, the action functional in the continuous case is

$$\mathcal{S}[\Theta] = \int dt \int dl \mathcal{L}(\Theta, \partial\Theta).$$

Thus as the discrete theory was described by equations that could be obtained from the stationarity of the action S_θ, now the continuous theory will be described by *field equations* that can be obtained from the stationarity of the functional $\mathcal{S}[\Theta]$: this means that the solutions of these equations will be extremals for $\mathcal{S}[\Theta]$. Thus the equations for our field are the Euler-Lagrange equations for the field theory described by $\mathcal{L}$. We can compute them easily. The x^μ coordinates are (t, l) and we easily compute

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \Theta(l;t)} &= 0 \\
\frac{\partial \mathcal{L}}{\partial(\partial_t \Theta(l;t))} &= \mu \partial_t \Theta(l;t) \\
\frac{\partial \mathcal{L}}{\partial(\partial_l \Theta(l;t))} &= \kappa \partial_l \Theta(l;t)
\end{aligned}$$

so that

$$\begin{aligned}
\partial_t \left(\frac{\partial \mathcal{L}}{\partial(\partial_t \Theta(l;t))} \right) &= \mu \partial_t^2 \Theta(l;t) \\
\partial_l \left(\frac{\partial \mathcal{L}}{\partial(\partial_l \Theta(l;t))} \right) &= \kappa \partial_l^2 \Theta(l;t)
\end{aligned}$$

and the Euler-Lagrange equations (3.2) then become

$$\mu \partial_t^2 \Theta(l;t) - \kappa \partial_l^2 \Theta(l;t) = 0,$$

which is exactly equation (2.12) that we derived with the heuristic limit procedure in the previous lecture.

If in the future we are going to develop a theory for some fields **and** we will assume it admits a description in terms of an action functional **and** we will give an explicit action functional for the fields, **then** we will be able to determine the field equations by computing the Euler-Lagrange equations of the system, **or**, equivalently, by imposing the stationarity of the action functional under a fluctuations of the fields.

3.3 Synopsis

In this lecture we have studied how the equations of motion of a field theory (Euler-Lagrange equations) can be obtained by using a variational principle. The proofs given assume that the domain of definition of the field theory is a bounded subset of $\mathbb{R}^n$, but choosing properly the space of fields $\mathcal{F}$ suitable generalizations can be given to situations where D is unbounded and fields behave in a sufficiently regular way asymptotically (i.e. in such a way that all defined quantities exist finite). All the formalism has been developed for subsets of $\mathbb{R}^n$, although the definition of field given at the beginning was in far more general settings. It is thus important to stress that the above formulation can be extended from the case of the Euclidean space to more general situations and we will use, without further proofs, some generalized setups in what follows. In particular we will apply all the above to the case in which fields are defined on a Lorentzian manifold.