# Calculus of Variations

Liping Liu

February 22, 2017

## 1 Divergence Theorem for scalar, vector and tensor fields

**Tensor Fields** Let $\Omega \subset \mathbb{R}^n$ be an open domain.

- $\varphi : \Omega \to \mathbb{R}$ is a scalar field;
- $\mathbf{v} : \Omega \to \mathbb{R}^m$ is a vector field;
- $\mathbf{T} : \Omega \to \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^k)$ is a tensor field.

Notation convention: it is often convenient to denote vectors and tensors in index notation, e.g., $v_i$ ($v_i = \mathbf{v} \cdot \mathbf{e}_i$) and $T_{pi}$ ($T_{pi} = \hat{\mathbf{e}}_p \mathbf{T} \mathbf{e}_i$), where the bases $\{\hat{\mathbf{e}}_p, p = 1, \cdots, m\}$ and $\{\mathbf{e}_i, i = 1, \cdots, n\}$ are usually not specified but tacitly understood.

**Differentiation** Let $\varphi$ be a scalar field on $\Omega \subset \mathbb{R}^n$. For any $\mathbf{a} \in \mathbb{R}^n$,

$$\begin{aligned}
&D\varphi(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}, \\
&(D\varphi(\mathbf{x}))(\mathbf{a}) = \lim_{\varepsilon \to 0} \frac{\varphi(\mathbf{x} + \varepsilon \mathbf{a}) - \varphi(\mathbf{x})}{\varepsilon}.
\end{aligned}$$

Definition: $\varphi$ is differentiable on $\Omega$ if $D\varphi(\mathbf{x}) \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ for all $\mathbf{x} \in \Omega$

$$\begin{aligned}
\varphi(\mathbf{x} + \varepsilon \mathbf{a}) &= \varphi(\mathbf{x}) + \varepsilon (D\varphi(\mathbf{x}))(\mathbf{a}) + o(\varepsilon) \\
&= \varphi(\mathbf{x}) + \varepsilon \nabla \varphi(\mathbf{x}) \cdot \mathbf{a} + o(\varepsilon) \qquad \forall\, \mathbf{x} \in \Omega, \mathbf{a} \in \mathbb{R}^n.
\end{aligned}$$

and

$$\nabla \varphi(\mathbf{x}) = \sum_{i=1}^n \varphi_{,i} \mathbf{e}_i, \qquad \varphi_{,i} = \mathbf{e}_i \cdot \nabla \varphi = \lim_{\varepsilon \to 0} \frac{\varphi(\mathbf{x} + \varepsilon \mathbf{e}_i) - \varphi(\mathbf{x})}{\varepsilon} = \frac{\partial \varphi(x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n)}{\partial x_i}.$$

Definition: $\mathbf{v} : \Omega \to \mathbb{R}^m$ is differentiable on $\Omega$ if every component is differentiable

$$\mathbf{v}(\mathbf{x}) = \sum_p v_p(\mathbf{x}) \hat{\mathbf{e}}_p \,.$$

Definition:

$$\begin{aligned}
&\nabla \mathbf{v}(\mathbf{x}) = D\mathbf{v}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m \\
&(\nabla \mathbf{v}(\mathbf{x}))(\mathbf{a}) = \sum_{p=1}^m \hat{\mathbf{e}}_p \nabla v_p(x) \cdot \mathbf{a}
\end{aligned}$$

Thus,

$$\nabla \mathbf{v}(\mathbf{x}) = \sum_{p,i} v_{p,i} \hat{\mathbf{e}}_p \otimes \mathbf{e}_i \ .$$

<u>Divergence:</u> If $m = n$, $\mathrm{div}(\mathbf{v}) = \mathrm{Tr}(\nabla \mathbf{v})$, i.e.,

$$\mathrm{div}(\mathbf{v}) = v_{i,i} \mathbf{e}_i \cdot \mathbf{e}_i = v_{i,i}.$$

Further, if $\mathbf{T} : \Omega \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ and

$$\mathbf{T}(x) = T_{pi} \hat{\mathbf{e}}_p \otimes \mathbf{e}_i.$$

Then

$$\begin{aligned} &\mathrm{div}(\mathbf{T}) : \mathbb{R}^m \to \mathbb{R} \\ &\mathrm{div}(\mathbf{T})(\mathbf{a}) = T_{pi,i} \mathbf{a} \cdot \hat{\mathbf{e}}_p. \end{aligned}$$

One may identify $\mathrm{div}(\mathbf{T})$ with a vector field $\Omega \to \mathbb{R}^m$ (instead of $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R})$). With an abuse of notation, we write

$$\mathrm{div}(\mathbf{T}) = T_{pi,i} \hat{\mathbf{e}}_p.$$

<u>**Field of class $C^0, C^1, C^2, \cdots, C^\infty$**</u>

♦ CLAIM: Assume $\varphi,\ \mathbf{v},\ \mathbf{u},\ \mathbf{T} : \Omega \to \mathbb{R}, \mathbb{R}^n, \mathbb{R}^n, \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$, $\mathbf{w} : \Omega \to \mathbb{R}^m$ are smooth fields on $\Omega$. The following identities hold:

1. $\nabla(\varphi \mathbf{v}) = \mathbf{v} \otimes (\nabla \varphi) + \varphi \nabla \mathbf{v}$;
2. $\mathrm{div}(\varphi \mathbf{v}) = (\nabla \varphi) \cdot \mathbf{v} + \varphi \mathrm{div} \mathbf{v}; \quad \nabla \cdot \mathbf{v} = \mathrm{div} \mathbf{v}$
3. $\nabla(\mathbf{v} \cdot \mathbf{u}) = (\nabla \mathbf{v})^T \mathbf{u} + (\nabla \mathbf{u})^T \mathbf{v}$
4. $\mathrm{div}(\mathbf{v} \otimes \mathbf{u}) = \mathbf{v} \mathrm{div}(\mathbf{u}) + (\nabla \mathbf{v}) \mathbf{u}$
5. $\mathrm{div}(\mathbf{T}^T \mathbf{w}) = \mathbf{T} \cdot \nabla \mathbf{w} + \mathbf{w} \cdot \mathrm{div} \mathbf{T}$
6. $\mathrm{div}(\varphi \mathbf{T}) = \varphi \mathrm{div} \mathbf{T} + \mathbf{T} \nabla \varphi$

*Proof:* Tacitly, an orthonormal basis $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\} \subset \mathbb{R}^n$ and an orthonormal basis $\{\hat{\mathbf{e}}_1, \cdots, \hat{\mathbf{e}}_m\} \subset \mathbb{R}^m$ are chosen and fixed. Notation: Einstein summation, i.e., summation over double index is understood. For example, to show 5, we have

$$\mathrm{div}(\mathbf{T}^T \mathbf{w}) = (T_{pi} w_p)_{,i} = T_{pi,i} w_p + T_{pi} w_{p,i} = \mathbf{w} \cdot \mathrm{div} \mathbf{T} + \mathbf{T} \cdot \nabla \mathbf{w}$$

■

**Curl operator:** Let $\mathbf{v} : \Omega \to \mathbb{R}^3$.

$$\text{curl}\mathbf{v} = \nabla \times \mathbf{v} = \begin{vmatrix} e_1 & e_2 & e_3 \\ \partial_1 & \partial_2 & \partial_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \sum_{i=1}^{3} \mathcal{E}_{ijk} v_{k,j} \mathbf{e}_i,$$

where *Levi-Civita* symbol is defined as

$$\mathcal{E}_{ijk} = \begin{cases} 1 & \text{if } (ijk) = (123),\ (231),\ (312), \\ -1 & \text{if } (ijk) = (132),\ (213),\ (321), \\ 0 & \text{otherwise.} \end{cases}$$

We notice that $\mathcal{E}_{ijk}$ is antisymmetric, i.e.,

$$\mathcal{E}_{ijk} = -\mathcal{E}_{ikj}, \quad \mathcal{E}_{ijk} = -\mathcal{E}_{jik}, \quad etc.$$

A useful identity between Kronecker symbol and Levi-Civita symbol is

$$\mathcal{E}_{pij}\mathcal{E}_{pkl} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}.$$

Let $\Omega \subset \mathbb{R}^3$ be a domain in $\mathbb{R}^3$. Assume that $\mathbf{v} : \Omega \to \mathbb{R}^3$, $\varphi : \Omega \to \mathbb{R}$ are smooth fields.

♦ CLAIM: the following identities hold:

1. $\nabla \times \nabla\varphi = 0$.
2. $\text{div}(\nabla \times \mathbf{v}) = 0$.
3. If $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$
4. $\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \Delta\mathbf{v}$.
5. $(\nabla \times \mathbf{v}) \times \mathbf{a} = [\nabla\mathbf{v} - (\nabla\mathbf{v})^T]\mathbf{a}$.
6. $\text{div}(\mathbf{u} \times \nabla \times \mathbf{v}) = (\nabla \times \mathbf{v}) \cdot (\nabla \times \mathbf{u}) - \mathbf{u} \cdot (\nabla \times \nabla \times \mathbf{v})$.

*Proof:*

∎

**Divergence Theorem**

Let $\Omega$ be a smooth simply connected domain in $\mathbb{R}^n$, $\mathbf{v} : \Omega \to \mathbb{R}^m$ is a smooth vector field on $\Omega$. Then we have

$$\int_{\partial\Omega} \mathbf{v} \otimes \mathbf{n}\, da = \int_{\Omega} \nabla\mathbf{v} dv, \tag{1}$$

where $\mathbf{n} : \partial\Omega \to \mathbb{R}^n$ is the outward unit normal on the boundary $\partial\Omega$. If $m = n$, take the trace of Eq. (1), we have

$$\int_{\partial\Omega} \mathbf{v} \cdot \mathbf{n}\, da = \int_{\Omega} \text{div}\mathbf{v} dv. \tag{2}$$

For a smooth tensor field $\mathbf{T} : \Omega \to \mathrm{Lin}(\mathbb{R}^n; \mathbb{R}^m)$, we have

$$\int_{\partial\Omega} \mathbf{T}\mathbf{n}\, da = \int_{\Omega} \mathrm{div}\mathbf{T} dv. \tag{3}$$

♦ PROVIDE a heuristic proof for (1) with $\Omega$ being a rectangle in two dimensions.

**Implications of divergence theorem in physics and mechanics**

1. Gauss theorem.

2. Stokes theorem or Green formula. In $\mathbb{R}^3$, let $\mathcal{C}$ be a directional closed space curve given by $\{\hat{\mathbf{x}}(s) : s \in [0, L]\}$, $s$ is the arc-length parameter, $\mathcal{S}$ be a curved surface with $\mathcal{C}$ being its boundary, and $\varphi, (\mathbf{v})$ be smooth scalar (vector) field defined on the entire space. Then we have the identities:

$$\int_{\mathcal{C}} \varphi d\mathbf{x} := \int_0^L \varphi(\hat{\mathbf{x}}(s)) \frac{d\hat{\mathbf{x}}(s)}{ds} ds = \int_{\mathcal{S}} \mathbf{n} \times \nabla\varphi,$$
$$\int_{\mathcal{C}} \mathbf{v} \cdot d\mathbf{x} := \int_0^L \mathbf{v}(\hat{\mathbf{x}}(s)) \cdot \frac{d\hat{\mathbf{x}}(s)}{ds} ds = \int_{\mathcal{S}} \mathbf{n} \cdot \mathrm{curl}\mathbf{v},$$

where $\mathbf{n}$ is the unit normal on $\mathcal{S}$ that follows the right-hand rule with the directional curve $\mathcal{C}$.

# 2 Calculus of Variations

## 2.1 Single variable

**Definition 2.1** *Let $u, v \in C^2(a, b)$ and consider a functional*

$$E[u] = \int_a^b f(u(x))dx.$$

*Then the first variation of this functional with respect to the perturbation $v$ is the first derivative*

$$\frac{d}{d\varepsilon}\varphi(\varepsilon)\Big|_{\varepsilon=0}, \qquad \varphi(\varepsilon) := E[u + \varepsilon v].$$

**Definition 2.2** *A function $v \in C_0^\infty(a, b)$ if $v$ is smooth, and $v$ and all its derivatives vanish at $x = a$ and $x = b$.*

**Theorem 1 (Localization theorem)** *Assume $f \in C(a, b)$. If*

$$\int_a^b f v dx = 0$$

*for any $v \in C_0^\infty(a, b)$, then*

$$f(x) = 0 \qquad \forall\, x \in (a, b).$$

**Problem 1** *Consider the functionals:*

$$E[u] = \int_a^b cu^n dx, \quad \int_a^b c(u')^2 dx, \quad \int_a^b cu'u dx, \quad \int_a^b c(u'')^2 dx, \quad \int_a^b [c_1(u')^2 + c_2 u] dx,$$

*where $u : (a, b) \to \mathbb{R}$ and $c, c_1, c_2$ are continuous functions. Suppose that the first variation of the above functionals is equal to zero with respect to any perturbation $v \in C_0^\infty(a, b)$. Find the differential equations necessarily satisfied by $u$.*

*Solution:*

1. If $E[u] = \int_a^b cu^n dx$, for a perturbation $v \in C_0^\infty(a, b)$ the first variation is given by

$$\begin{aligned} \frac{d}{d\varepsilon}\varphi(\varepsilon)\bigg|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\int_a^b c(u + \varepsilon v)^n dx\bigg|_{\varepsilon=0} = 0 \\ \Rightarrow \int_a^b ncu^{n-1}v &= 0. \end{aligned}$$

Since $v \in C_0^\infty(a, b)$ is arbitrary, by the localization theorem (i.e., Theorem 1) we infer

$$nc(x)u(x)^{n-1} = 0 \qquad \forall x \in (a, b).$$

2. If $E[u] = \int_a^b c(u')^2 dx$, for a perturbation $v \in C_0^\infty(a, b)$ the first variation is given by

$$\frac{d}{d\varepsilon}\int_a^b c((u + \varepsilon v)')^2 dx\bigg|_{\varepsilon=0} = 0 \qquad \Rightarrow \qquad \int_a^b 2cu'v' dx = 0.$$

Integrating by part, we find that

$$0 = \int_a^b 2cu'v' dx = cu'v|_a^b - \int_a^b 2(cu')'v dx = -\int_a^b 2(cu')'v dx.$$

Since $v \in C_0^\infty(a, b)$ is arbitrary, by Theorem 1 we conclude that

$$(cu')' = 0, \qquad \forall x \in (a, b).$$

3. If $E[u] = \int_a^b cu'u dx$, for a perturbation $v \in C_0^\infty(a, b)$ the first variation is given by

$$\frac{d}{d\varepsilon}\int_a^b c(u + \varepsilon v)'(u + \varepsilon v) dx\bigg|_{\varepsilon=0} = 0 \qquad \Rightarrow \qquad \int_a^b c(uv' + u'v) dx = 0.$$

Integrating by parts we obtain

$$cuv|_a^b + \int_a^b (cu' - (cu)')v dx = \int_a^b (cu' - (cu)')v dx = 0,$$

and henceforth, by Theorem 1, arrive at

$$cu' - (cu)' = 0, \qquad \forall x \in (a, b)$$

4. If $E[u] = \int_a^b c(u'')^2 dx$, for a perturbation $v \in C_0^\infty(a,b)$ the first variation is given by
$$\frac{d}{d\varepsilon}\int_a^b c((u+\varepsilon v)'')^2 dx\Big|_{\varepsilon=0} = \int_a^b 2cu''v''dx = 0.$$
Integrating by parts we obtain
$$2cu''v'|_a^b - \int_a^b 2(cu'')'dv = -2(cu'')'v|_a^b + \int_a^b 2(cu'')''vdx = 0.$$
Therefore, by Theorem 1 we obtain
$$(cu'')'' = 0 \qquad \forall x \in (a,b).$$

5. If $E[u] = \int_a^b [c_1(u')^2 + c_2 u]dx$, for a perturbation $v \in C_0^\infty(a,b)$ the first variation is given by
$$\frac{d}{d\varepsilon}\int_a^b c_1((u+\varepsilon v)')^2 + c_2(u+\varepsilon v)dx\Big|_{\varepsilon=0}\int_a^b (2c_1u'v' + c_2 v)dx = 0.$$
Integrating by parts we obtain
$$2c_1u'v|_a^b + \int_a^b (-2(c_1u')' + c_2)vdx = 0,$$
which, by Theorem 1, implies
$$-2(c_1u')' + c_2 = 0 \qquad \forall x \in (a,b).$$

■

**Theorem 2 (Necessary conditions for a minimizing function of a functional)** *Let $E[u]$ be a functional given by*
$$E[u] = \int_a^b [\frac{1}{2}c_1|u'|^2 + c_2 u]dx,$$
*where $c_1, c_2$ are given continuous functions on the interval $(a,b)$. If $u_0 \in C^2(a,b)$ is a minimizer in the sense that for all perturbation $v \in C_0^\infty(a,b)$ and all $\varepsilon$ small enough, we have*
$$E[u_0 + \varepsilon v] \geq E[u_0].$$
*Then the minimizer $u_0$ satisfies*
$$-(c_1u_0')' + c_2 = 0 \qquad \text{on } (a,b).$$

*Proof:* Since $u_0$ is a minimizer, the first variation of functional $E[u_0]$ with respect to any perturbation $v \in C_0(a,b)$ should vanish:
$$\tfrac{d}{d\varepsilon}(\textstyle\int_a^b [\frac{1}{2}c_1((u+\varepsilon)')^2 + c_2(u+\varepsilon v)]\Big|_{\varepsilon=0} = 0 \Rightarrow \displaystyle\int_a^b (c_1u_0'v' + c_2 v)dx = 0.$$
Integrating by parts, we have
$$\int_a^b [-(c_1u_0')' + c_2]vdx = 0 \quad \forall x \in (a,b).$$
By the localization theorem (c.f. Theorem 1) we conclude that
$$-(c_1u_0')' + c_2 = 0 \quad \forall x \in (a,b).$$

■

## 2.2 Multiple variables

**Theorem 3 (Necessary conditions for a stationary point of a functional)** *Let $E[u]$ be a functional given by*

$$E[\mathbf{u}] = \int_\Omega [W(\nabla \mathbf{u}) + f(\mathbf{u})] d\mathbf{x},$$

*where $W : \mathbb{R}^{n\times m} \to \mathbb{R}$ and $f : \mathbb{R}^m \to \mathbb{R}$ are given continuous differentiable functions. If $\mathbf{u} \in C^2(\Omega; \mathbb{R}^m)$ is a stationary point in the sense that for all perturbation $\mathbf{v} \in C_0^\infty(\Omega; \mathbb{R}^m)$, we have*

$$\frac{d}{d\varepsilon} E[\mathbf{u} + \varepsilon \mathbf{v}]\Big|_{\varepsilon=0} = 0. \tag{4}$$

*Then the stationary state $\mathbf{u}$ satisfies*

$$-\mathrm{div}[D_{\mathbf{F}} W(\nabla \mathbf{u})] + D_{\mathbf{u}} f(\mathbf{u}) = 0 \qquad \text{in } \Omega, \tag{5}$$

*where*

$$D_{\mathbf{F}} W(\nabla \mathbf{u}) := \frac{\partial W(\mathbf{F})}{\partial \mathbf{F}}\Big|_{\mathbf{F}=\nabla \mathbf{u}}, \qquad D_{\mathbf{u}} f(\mathbf{u}) := \frac{\partial f(\mathbf{u})}{\partial \mathbf{u}}.$$

*Proof:* From the definition, we have that for any $\mathbf{v} \in C_0^\infty(\Omega; \mathbb{R}^m)$,

$$\begin{aligned}
\frac{d}{d\varepsilon} E[\mathbf{u} + \varepsilon \mathbf{v}]\Big|_{\varepsilon=0} &= \frac{d}{d\varepsilon} \int_\Omega [W(\nabla \mathbf{u} + \varepsilon \mathbf{v}) + f(\mathbf{u} + \varepsilon \mathbf{v})] d\mathbf{x}\Big|_{\varepsilon=0} \\
&= \int_\Omega [D_{\mathbf{F}} W(\nabla \mathbf{u}) \cdot \nabla \mathbf{v}) + D_{\mathbf{u}} f(\mathbf{u}) \cdot \mathbf{v}] d\mathbf{x},
\end{aligned}$$

where, in index form,

$$\sigma_{pi} := [D_{\mathbf{F}} W(\nabla \mathbf{u})]_{pi} = \frac{\partial W(\mathbf{F})}{\partial F_{pi}}\Big|_{\mathbf{F}=\nabla \mathbf{u}}, \quad t_p := [D_{\mathbf{u}} f(\mathbf{u})]_p = \frac{\partial f(\mathbf{u})}{\partial u_p},$$

and

$$\begin{aligned}
\frac{d}{d\varepsilon} E[\mathbf{u} + \varepsilon \mathbf{v}]\Big|_{\varepsilon=0} &= \int_\Omega \sigma_{pi} v_{p,i} + t_p v_p d\mathbf{x} \\
&= \int_\Omega (-\sigma_{pi,i} + t_p) v_p d\mathbf{x},
\end{aligned}$$

where the second equality follows from the divergence theorem. Since the test function $\mathbf{v} \in C_0^\infty(\Omega; \mathbb{R}^m)$ is arbitrary, by (4) and the localization theorem we obtain

$$-\sigma_{pi,i} + t_p = 0 \qquad \text{in } \Omega,$$

which is precisely the index form of (5).

■

# 3 Applications in mechanics, physics and geometry

## 3.1 Hamilton's principle and equation of motion

### 3.1.1 Discrete system

For a particle in a conservative force field, i.e., the force on the particle at $\mathbf{x}$ is given by

$$\mathbf{f} = -\nabla V(\mathbf{x}).$$

Then the Lagrangian of the particle is given by

$$L[\mathbf{x}] = \frac{1}{2}m|\dot{\mathbf{x}}|^2 - V(\mathbf{x}),$$

and the action functional is

$$S[\mathbf{x}] = \int_{t_0}^{t_1} \frac{1}{2}m|\dot{\mathbf{x}}|^2 - V(\mathbf{x})dt.$$

We also specify the initial condition:

$$\dot{\mathbf{x}}(0) = \mathbf{x}_0, \qquad \ddot{\mathbf{x}}(0) = \mathbf{v}_0.$$

The actual motion of the particle shall be a *stationary state* of the action functional, i.e.,

$$\frac{d}{d\varepsilon}S[\mathbf{x}+\varepsilon\mathbf{x}_1]\Big|_{\varepsilon=0} = 0.$$

By divergence theorem, we find

$$0 = \int_{t_0}^{t_1} m\dot{\mathbf{x}}\cdot\dot{\mathbf{x}}_1 - \mathbf{x}_1\cdot\nabla V(\mathbf{x})dt = \int_{t_0}^{t_1} \mathbf{x}_1\cdot(-m\ddot{\mathbf{x}} - \nabla V(\mathbf{x})dt \qquad \forall\, \mathbf{x}_1(t)$$

the localization theorem we conclude that

$$m\ddot{\mathbf{x}} + \nabla V(\mathbf{x}) = 0 \quad \text{i.e.} \quad m\ddot{\mathbf{x}} = \mathbf{f} = -\nabla V(\mathbf{x}).$$

For general mechanical system, it may have many degrees of freedom, which are described by *generalized coordinates* $\mathbf{q} = (q_1, \cdots, q_k)$. Denote by $\dot{\mathbf{q}} = (\dot{q}_1, \cdots, \dot{q}_k)$ the generalized velocity. To derive the equation of motion by the Hamilton principle, we shall first identify the Lagrangian of the system( i.e., the kinetic energy minus the potential of the system) in terms of $\mathbf{q}, \dot{\mathbf{q}}$:

$$L = L[\mathbf{q}, \dot{\mathbf{q}}, t], \qquad S[\mathbf{q}] = \int_{t_0}^{t_1} L[\mathbf{q}, \dot{\mathbf{q}}, t]dt.$$

The actual motion of the system shall be a *stationary state* of the action functional, i.e.,

$$\frac{d}{d\varepsilon}S[\mathbf{q}+\varepsilon\mathbf{q}_1]\Big|_{\varepsilon=0} = 0,$$

which, by the divergence theorem and localization theorem, implies the *Lagrange equation*:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} - \frac{\partial L}{\partial \mathbf{q}} = 0.$$

**Example 4** *Three mass points connected by two rigid bars on a rigid surface.*

### 3.1.2 Hamilton's principles for elastic continuum bodies

**Dynamics of an elastic bar**

Consider an elastic bar of length $L$, area $A$ and Young's modulus $E$. Also the two ends are fixed. The deformation state of the bar is described by the displacement $u(x,t)$. Clearly, the Lagrange of bar is given by

$$L = \int_0^L \frac{1}{2}\rho\dot{u}^2 - \frac{1}{2}EAu_x^2 dx.$$

Therefore, the action functional of the system is given by

$$S[u] = \int_{t_0}^{t_1}\int_0^L \frac{1}{2}\rho\dot{u}^2 - \frac{1}{2}EAu_x^2 dxdt.$$

The actual equation of motion of the bar shall be a *stationary state* of the action functional, i.e.,

$$\frac{d}{d\varepsilon}S[u+\varepsilon u_1]\Big|_{\varepsilon=0} = 0,$$

which, by the divergence theorem and localization theorem, implies the *Lagrange equation*:

$$(EAu_x)_x = \rho\ddot{u}.$$

Similarly we can derive the equation of motion for an elastic string and elastic beam as

$$(Tw_x)_x = \rho w_{tt}, \qquad (EIw_{xx})_{xx} + \rho w_{tt} = 0.$$

**Dynamics of a 3D elastic body**

**Dynamics of inviscid fluids: Euler equation**

Consider a ideal fluid in domain $\Omega_0 \subset \mathbb{R}^3$. The state is described by motion $\mathbf{x}(\cdot,t) : \Omega_0 \to \Omega_t$. The Lagrangain is given by

$$L[\mathbf{x}] = \int_\Omega [\frac{1}{2}\rho|\mathbf{v}|^2 + \rho\mathbf{g}\cdot\mathbf{x}] = \int_\Omega \Big[\frac{1}{2}\rho|\frac{d}{dt}\mathbf{x}(\mathbf{X},t)|^2 + \rho\mathbf{g}\cdot\mathbf{x}(\mathbf{X},t)\Big]d\mathbf{X}$$

Because of incompressibility, we have

$$\mathrm{div}\mathbf{v} = 0 \qquad \text{in } \Omega,$$

which, together with the conservation of mass

$$\rho_t + \mathrm{div}(\rho\mathbf{v}) = 0,$$

implies

$$\frac{d\rho(\mathbf{x},t)}{dt} = 0.$$

**Method of Lagrange's multiplier**

## 3.2 Principle of minimum free energy

In thermodynamics, the Second Law asserts that the entropy of a closed (no energy flux across the boundary) isolated (no particle flux across the boundary) system is monotonically increasing and attains its maximum in equilibrium. For a system in contact with a heat bath at a constant temperature, the Second Law implies the principle of minimum free energy, i.e., the free energy of an isolated system at a constant temperature shall be monotonically decreasing and reach its minimum in equilibrium.

For most of elastic problem the system is tacitly assumed to be at constant temperature and hence no heat flux occurs during the process. An exception includes plasticity where dissipation becomes significant. The equilibrium equations for the system follow directly from calculus of variation once the total free energy of the system is identified (or postulated). For example, in 3D linearized elasticity we describe the thermodynamic state of the elastic body $\Omega$ by displacement $\mathbf{u} : \Omega \to \mathbb{R}^3$. The boundary of the body $\partial\Omega$ is subdivided into $\Gamma_D$ and $\Gamma_N$; the Dirichlet boundary condition $\mathbf{u} = \mathbf{u}_0$ on $\Gamma_D$ is imposed whereas a Neumann-type boundary condition applied traction= $\mathbf{t}_0$ on $\Gamma_N$. In this setup, the free energy of the system can be identified as

$$F[\mathbf{u}] = \int_\Omega [W(\nabla\mathbf{u}) - \mathbf{b} \cdot \mathbf{u}]dv - \int_{\Gamma_N} \mathbf{t}_0 \cdot \mathbf{u} ds,$$

where the internal energy density function $W : \mathbb{R}^{3\times3} \to \mathbb{R}$ is given by

$$W(\nabla\mathbf{u}) = \frac{1}{2}\nabla\mathbf{u} \cdot \mathbf{C}\nabla\mathbf{u}.$$

For a equilibrium state we have

$$\frac{d}{d\varepsilon}F[\mathbf{u} + \varepsilon\mathbf{u}_1]\Big|_{\varepsilon=0} = 0 \qquad \forall \text{ admissible } \mathbf{u}_1.$$

The above first-variation condition implies the standard boundary value problem in linearized elasticity:

$$\begin{cases} -\operatorname{div}(\mathbf{C}\nabla\mathbf{u}) = \mathbf{b} & \text{in } \Omega, \\ (\mathbf{C}\nabla\mathbf{u})\mathbf{n} = \mathbf{t}_0 & \text{on } \Gamma_N, \\ \mathbf{u} = \mathbf{u}_0 & \text{on } \Gamma_D. \end{cases}$$

## 3.3 Minimal surface

Consider a space curve $\mathcal{C}$ that can be parameterized as

$$z = f_0(x, y), \qquad (x, y) \in \mathcal{C}_P,$$

where $\mathcal{C}_p$ is the planar curve on the $xy$-plane. Assume that $\mathcal{C}_P$ is a simple closed curve and denote by $\Omega_P$ the enclosed area. We are interested in surfaces that have minimum surface area among all surfaces with $\mathcal{C}$ as the boundary.

For simplicity, we restrict ourselves to surfaces that admit Monge parametrization $z = w(x, y)$. The area functional $A[w]$ is then given by

$$A[w] = \int_{\Omega_P} (1 + |\nabla w|^2)^{1/2}) dxdy.$$

The Euler-Lagrange equation is given by

$$\begin{cases} \operatorname{div}\Big[\dfrac{\nabla w}{\sqrt{1+|\nabla w|^2}}\Big] = 0 & \text{in } \ \Omega_P, \\ w = f_0(x,y) & \text{on } \ \mathcal{C}_P. \end{cases}$$

We remark that the first of the above equation is equivalent to

$$H = \text{mean curvature} = 0 \qquad \text{on } \ \Omega_P.$$

### 3.4 Geodesics on a curved surface

Consider a surface element given by

$$x^p = x^p(u^1, u^2), \quad (p = 1, 2, 3) \qquad \forall\, (u^1, u^2) \in U \subset \mathbb{R}^2.$$

Then

$$ds^2 = dx^p \delta_{pq} dx^q = g_{ij} du^i du^j, \tag{6}$$

where the metric tensor

$$g_{ij} = x^p_{,i} \delta_{pq} x^q_{,j}.$$

Consider two fixed points $A$ and $B$ on the surface $\mathcal{S}$, we are interested in the curve that have $A, B$ as two end points and have minimum curve length.

From (6), the length of curves on $\mathcal{S}$ is given by

$$L[u^1(t), u^2(t)] = \int_0^L (g_{ij} \frac{du^i}{dt} \frac{du^j}{dt})^{1/2} dt.$$

Without loss of generality, we may choose an arc-length parametrization such that

$$\gamma(t) = g_{ij} \frac{du^i}{dt} \frac{du^j}{dt} = 1.$$

If not, we change the parametrization to $t' = \int_0^t \gamma^{1/2}(s) ds$. Then the curve with minimum curve length shall satisfy

$$\begin{aligned}
0 &= \frac{d}{d\varepsilon} L[u^1(t) + \varepsilon v^1(t), u^2(t) + \varepsilon v^2(t)]\Big|_{\varepsilon=0} \\
&= \frac{1}{2} \int_0^L [g_{ij,k} \frac{du^j}{dt} \frac{du^i}{dt} v^k + 2 g_{ij} \frac{du^j}{dt} \frac{dv^i}{dt}] dt \\
&= \frac{1}{2} \int_0^L \{g_{ij,m} \frac{du^i}{dt} \frac{du^j}{dt} - 2[g_{mj} \frac{d^2 u^j}{dt^2} + g_{mj,k} \frac{du^j}{dt} \frac{du^k}{dt}]\} v^m dt
\end{aligned}$$

where we have used

$$\frac{d}{dt}[g_{ij} \frac{du^j}{dt} v^i] = v^i \frac{d}{dt}[g_{ij} \frac{du^j}{dt}] + g_{ij} \frac{du^j}{dt} \frac{dv^i}{dt}$$

and

$$\frac{d}{dt}[g_{ij} \frac{du^j}{dt}] = g_{ij} \frac{d^2 u^j}{dt^2} + g_{ij,k} \frac{du^j}{dt} \frac{du^k}{dt} = 0.$$

Therefore, the equation for geodesics is given by

$$g_{ij,m} \frac{du^i}{dt} \frac{du^j}{dt} - 2[g_{mj} \frac{d^2 u^j}{dt^2} + g_{mj,k} \frac{du^j}{dt} \frac{du^k}{dt}] = 0,$$

which is equivalent to

$$\frac{d^2 u^i}{dt^2} + g^{im} \frac{1}{2}(g_{mj,k} + g_{mk,j} - g_{jk,m}) \frac{du^j}{dt} \frac{du^k}{dt} = \frac{d^2 u^i}{dt^2} + \Gamma^i_{jk} \frac{du^j}{dt} \frac{du^k}{dt} = 0.$$

### 3.5 Ginzburg-Landau model

In the Ginzburg-Landau model of phase transition, the thermodynamic state of the system is described by an order parameter $\phi : \Omega \to \mathbb{R}$ and the free energy is given by

$$F[\phi] = \int_\Omega \Big[\frac{\kappa}{2}|\nabla\phi|^2 + a\phi^2(1-\phi^2)\Big]dv.$$

The equilibrium state shall be such that

$$\frac{d}{d\varepsilon}F[\phi+\varepsilon\varphi]\Big|_{\varepsilon=0} = 0 \qquad \forall\,\varphi.$$

The above variational principle implies the following Euler-Lagrange equation:

$$-\kappa\Delta\phi - 4a\phi^3 + 2a\phi = 0 \qquad \text{in } \Omega.$$

### 3.6 Variational principle for electrodynamics

The classic Maxwell equations in vacuum are given by (in SI unit)

$$\begin{aligned} &\nabla\cdot\mathbf{D} = \rho, \qquad &&\nabla\times\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t} = 0, \qquad &&\mathbf{D} = \epsilon_0\mathbf{E}, \\ &\nabla\cdot\mathbf{B} = 0, \qquad &&\nabla\times\mathbf{H} = \mathbf{J} + \frac{\partial\mathbf{D}}{\partial t}, \qquad &&\mathbf{B} = \mu_0\mathbf{H}. \end{aligned} \tag{7}$$

Since $\nabla\cdot\mathbf{B} = 0$, we can have $\nabla\times\mathbf{A} = \mathbf{B}$, and hence

$$\nabla\times(\mathbf{E} + \frac{\partial}{\partial t}\mathbf{A}) = 0.$$

Therefore, we have $\mathbf{E} + \frac{\partial}{\partial t}\mathbf{A} = -\nabla\varphi$. In other words, there exist $(\varphi, \mathbf{A}) : \mathbb{R}^3 \to \mathbb{R}\times\mathbb{R}^3$ such that

$$\mathbf{E} = -\nabla\varphi - \frac{\partial}{\partial t}\mathbf{A}, \qquad \mathbf{B} = \nabla\times\mathbf{A},$$

and two of the Maxwell equations are automatically satisfied. Then in vacuum the rest of Maxwell equations can be obtained as the Euler-Lagrange equations of the variational principle:

$$\frac{d}{d\varepsilon}S[\mathbf{A}+\varepsilon\mathbf{A}_1, \varphi+\varepsilon\varphi_1] = 0 \qquad \forall\,\mathbf{A}_1, \varphi_1, \tag{8}$$

where

$$\begin{aligned} S[\mathbf{A},\varphi] &= -\int_{t_0}^{t_1}\int_{\mathbb{R}^3}\Big[\frac{\epsilon_0}{2}|\mathbf{E}|^2 - \frac{1}{2\mu_0}|\mathbf{B}|^2 - \rho\varphi + \mathbf{J}\cdot\mathbf{A}\Big]d\mathbf{x}dt \\ &= -\int_{t_0}^{t_1}\int_{\mathbb{R}^3}\Big[\frac{\epsilon_0}{2}|\nabla\varphi + \mathbf{A}_t|^2 - \frac{1}{2\mu_0}|\nabla\times\mathbf{A}|^2 - \rho\varphi + \mathbf{J}\cdot\mathbf{A}\Big]d\mathbf{x}dt \end{aligned} \tag{9}$$

By standard first-variation calculus, we find the Euler-Lagrange equations of (8) are given by

$$\nabla\cdot[-\epsilon_0(\nabla\varphi + \mathbf{A}_t)] - \rho = 0, \qquad -\frac{1}{\mu_0}\nabla\times\nabla\times\mathbf{A} + \frac{\partial}{\partial t}[-\epsilon_0(\nabla\varphi + \mathbf{A}_t) + \mathbf{J} = 0,$$

which is precisely equivalent to

$$\nabla\cdot\mathbf{D} = \rho, \qquad \nabla\times\mathbf{H} = \mathbf{J} + \frac{\partial}{\partial t}\mathbf{D}.$$

**Question**

Can we interpret

$$K.E. = \int_{\mathbb{R}^3} \Big[\frac{1}{2\mu_0}|\nabla\times\mathbf{A}|^2 - \mathbf{J}\cdot\mathbf{A}\Big]d\mathbf{x},$$
$$P.E. = \int_{\mathbb{R}^3} \Big[\frac{\epsilon_0}{2}|\nabla\varphi + \mathbf{A}_t|^2 - \rho\varphi\Big]d\mathbf{x}$$

as the kinetic energy and potential energy, respectively? What would be their implication for materials /collections of charges?

Free charges moving in space:

- Internal energy $U \equiv 0$;
- Kinetic energy: $\int_{\mathbb{R}^3} \frac{1}{2}\rho|\mathbf{v}|^2$
- Potential energy:

**4-space formulation of electrodynamics.** We introduce Minkovski space $\mathbb{R}^{1,3}$ equipped with inner probduct

$$\langle\mathbf{a},\mathbf{b}\rangle = \eta_{\nu\mu}a^\nu b^\mu,$$

where the metric tensor $\eta_{\nu\mu} =$ diag$(+1,-1,-1,-1)$. The coordinates of a point in $\mathbb{R}^{1,3}$ is identified as

$$x^\nu = (ct, x^1, x^2, x^3).$$

Let

$$J^\nu = (c\rho, J^1, J^2, J^3), \qquad A^\nu = (\varphi, cA^1, cA^2, cA^3).$$

Then

$$F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha = \eta^{\alpha\nu}\partial_\nu A^\beta - \eta^{\beta\nu}\partial_\nu A^\alpha = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -cB_z & cB_y \\ E_y & cB_z & 0 & -cB_x \\ E_z & -cB_y & cB_x & 0 \end{bmatrix},$$

and $F_{\alpha\beta} = \eta_{\alpha\nu}\eta_{\beta\mu}F^{\nu\mu}$, i.e.,

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha = \partial_\alpha\eta_{\beta\nu}A^\nu - \partial_\beta\eta_{\alpha\nu}A^\nu = \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{bmatrix}.$$

By (9), the action functional can be rewritten as

$$S[A^\nu] = \int_{\mathbb{R}^{1,3}} L(A^\nu)d^4x$$

where the Lagrange density function is indentified as

$$L[A^\nu] = \frac{1}{c}[\frac{\epsilon_0}{4}F^{\alpha\beta}F_{\alpha\beta} + \frac{1}{c}J_\nu A^\nu] = \frac{1}{c}[\frac{\epsilon_0}{2}C^{\alpha\mu\beta\nu}\partial_\alpha A_\mu\partial_\beta A_\nu + \frac{1}{c}J^\nu A_\nu],$$

where

$$C^{\alpha\mu\beta\nu} = \eta^{\alpha\beta}\eta^{\mu\nu} - \eta^{\alpha\nu}\eta^{\mu\beta}.$$

# Excercise

1. Assume $\varphi,\ \mathbf{v},\ \mathbf{u},\ \mathbf{T} : \Omega \to \mathbb{R}, \mathbb{R}^n, \mathbb{R}^n, \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m),\ \mathbf{w} : \Omega \to \mathbb{R}^m$ are smooth fields on $\Omega$. Show the following identities by indicial notation.

   (a) $\nabla(\varphi\mathbf{v}) = \mathbf{v} \otimes (\nabla\varphi) + \varphi\nabla\mathbf{v}$;
   
   (b) $\mathrm{div}(\varphi\mathbf{v}) = (\nabla\varphi) \cdot \mathbf{v} + \varphi\mathrm{div}\mathbf{v}; \quad \nabla \cdot \mathbf{v} = \mathrm{div}\mathbf{v}$
   
   (c) $\nabla(\mathbf{v} \cdot \mathbf{u}) = (\nabla\mathbf{v})^T\mathbf{u} + (\nabla\mathbf{u})^T\mathbf{v}$
   
   (d) $\mathrm{div}(\mathbf{v} \otimes \mathbf{u}) = \mathbf{v}\mathrm{div}(\mathbf{u}) + (\nabla\mathbf{v})\mathbf{u}$
   
   (e) $\mathrm{div}(\mathbf{T}^T\mathbf{w}) = \mathbf{T} \cdot \nabla\mathbf{w} + \mathbf{w} \cdot \mathrm{div}\mathbf{T}$
   
   (f) $\mathrm{div}(\varphi\mathbf{T}) = \varphi\mathrm{div}\mathbf{T} + \mathbf{T}\nabla\varphi$
   
   (g) $\nabla \times \nabla\varphi = 0$.
   
   (h) $\mathrm{div}(\nabla \times \mathbf{v}) = 0$.
   
   (i) If $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$
   
   (j) $\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \Delta\mathbf{v}$.
   
   (k) $(\nabla \times \mathbf{v}) \times \mathbf{a} = [\nabla\mathbf{v} - (\nabla\mathbf{v})^T]\mathbf{a}$.

2. Compute the first-variation and assoicated Euler-Lagrange equations for the following functionals:

   (a) Bending of an elastic rod:
   
   $$F[w] = \int_0^L [\frac{1}{2}EI(\frac{d^2w}{dx^2})^2 + qw]dx,$$
   
   where $w : (0, L) \to \mathbb{R}$ satisfies $w(0) = w(L) = w'(0) = w'(L) = 0$ (clamped boundary conditions).
   
   (b) Bending of a plate:
   
   $$F[w] = \int_\Omega [\frac{\kappa_b}{2}|\nabla\nabla w|^2 + \frac{\kappa_g}{2}\det(\nabla\nabla w)]dx,$$
   
   where $w : \mathbb{R}^2 \supset \Omega \to \mathbb{R}$ satisfies $w = \mathbf{n} \cdot \nabla w = 0$ on $\partial\Omega$ (clamped boundary conditions).
   
   (c) 3D elasticity:
   
   $$F[\mathbf{u}] = \int_\Omega [W(\nabla\mathbf{u}) - \mathbf{b} \cdot \mathbf{u}]dv - \int_{\Gamma_N} \mathbf{t}_0 \cdot \mathbf{u}ds,$$
   
   where the internal energy density function $W : \mathbb{R}^{3\times3} \to \mathbb{R}$ is given by
   
   $$W(\nabla\mathbf{u}) = \frac{1}{2}\nabla\mathbf{u} \cdot \mathbf{C}\nabla\mathbf{u} = \frac{1}{2}C_{piqj}u_{p,i}u_{q,j},$$
   
   and displacement $\mathbf{u} : \mathbb{R}^3 \supset \Omega \to \mathbb{R}^3$ satisfies $\mathbf{u} = \mathbf{u}_0$ on $\partial\Omega$ ($\mathbf{u}_0 : \partial\Omega \to \mathbb{R}^3$ is given).
   
   (d) G-L model for superconductivity:
   
   $$F[\phi, \mathbf{A}] = \int_\Omega \Big[\frac{k}{2}|(\nabla - i\mathbf{A})\phi|^2 + \frac{\alpha}{2}\phi^2 + \frac{\beta}{4}\phi^4\Big]dv + \frac{1}{2\mu_0}\int_{\mathbb{R}^3} |\nabla \times \mathbf{A}|^2 dv,$$
   
   where $k, \alpha, \beta, \mu_0$ are real constants, $\phi : \Omega \to \mathbb{C}$ is wave function satisfying $\phi = 0$ on $\partial\Omega$, and $\mathbf{A} : \mathbb{R}^3 \to \mathbb{R}^3$ is the vectorial potential for magnetic field and satisfies $|\mathbf{A}|, |\nabla\mathbf{A}| \to 0$ as $|\mathbf{x}| \to +\infty$.

(e) Minimum surface:

$$A[w] = \int_{\Omega_P} (1 + |\nabla w|^2)^{1/2} dxdy,$$

where $w : \mathbb{R}^2 \supset \Omega_P \to \mathbb{R}$ satisfies $w = f_0(x, y)$ on $\partial\Omega_P$ (i.e., $z = f_0(x, y)$ is the fixed boundary curve of the surface).

(f) Weighted minimum surface:

$$A[w] = \int_{\Omega_P} e^{-\frac{a(x^2+y^2+w^2)}{2}} (1 + |\nabla w|^2)^{1/2} dxdy,$$

where $w : \mathbb{R}^2 \supset \Omega_P \to \mathbb{R}$ satisfies $w = f_0(x, y)$ on $\partial\Omega_P$ (i.e., $z = f_0(x, y)$ is the fixed boundary curve of the surface).

(g) Geodesics:

$$L[u^1(t), u^2(t)] = \int_0^L g_{ij} \frac{du^i}{dt} \frac{du^j}{dt} dt,$$

where $g_{ij}$ is the given metric tensor, and $(u^1, u^2) : (0, L) \to \mathbb{R}^2$ satisfies $(u^1, u^2)|_0 = (u^1_A, u^2_A)$ and $(u^1, u^2)|_L = (u^1_B, u^2_B)$.

(h) Electrodynamics:

$$\begin{aligned} S[\mathbf{A}, \varphi] &= -\int_{t_0}^{t_1} \int_{\mathbb{R}^3} \left[ \frac{\epsilon_0}{2} |\mathbf{E}|^2 - \frac{1}{2\mu_0} |\mathbf{B}|^2 - \rho\varphi + \mathbf{J} \cdot \mathbf{A} \right] d\mathbf{x} dt \\ &= -\int_{t_0}^{t_1} \int_{\mathbb{R}^3} \left[ \frac{\epsilon_0}{2} |\nabla\varphi + \mathbf{A}_t|^2 - \frac{1}{2\mu_0} |\nabla \times \mathbf{A}|^2 - \rho\varphi + \mathbf{J} \cdot \mathbf{A} \right] d\mathbf{x} dt, \end{aligned}$$

where $(\mathbf{A}, \varphi) : \mathbb{R}^3 \times (t_0, t_1) \to (\mathbb{R}^3, \mathbb{R})$ are smooth, and (all derivatives) vanish as $|\mathbf{x}| \to +\infty$.