# NOTES ON THE CALCULUS OF VARIATIONS

A. D. WENDLAND

Notes on chapter 8 of *Partial Differential Equations* by L. C. Evans evans:pde [1].

## Contents



---

## 1. Introduction

1.1. **Basic Ideas.** We suppose that we wish to solve some particular partial differential equation, which we write in the abstract form

$$A[u] = 0. \tag{1.1}$$ {eq:1-1}

In this above equation, $A[\cdot]$ denotes a given, possibly nonlinear partial differential operator and $u$ is the unknown. Recall that there is of course no general theory for solving all such PDE.

The *calculus of variations* identifies an important class of such nonlinear problems that may be solved using relatively simple techniques from nonlinear functional analysis. We call this class of problems the *variational problems*, that is, PDE of the form (eq:1-1 1.1), where the nonlinear differential operator $A[\cdot]$ is the "derivative" of some appropriate *energy functional* $I[\cdot]$. Symbolically, we write

$$A[\cdot] = I'[\cdot]. \tag{1.2}$$ {eq:1-2}

*Date*: March 20, 2023.

Then problem (1.1) becomes

$$I'[u] = 0. \tag{1.3}$$

{eq:1-3}

The idea of the formulation in (1.3) is that we can now recognize solutions of the (possibly nonlinear) PDE (1.1) as being critical points of $I[\cdot]$. In certain circumstances, these critical points may be relatively easy to find: if, for instance, the functional $I[\cdot]$ has a minimum at $u$, then presumably (1.3) holds and thus $u$ is a solution of the original PDE (1.1). *The idea is that on the one hand, it is usually extremely difficult to solve (1.1) directly. On the other, it may be much easier to discover minimizers (or other critical points) of the functional $I[\cdot]$.*

Additionally, many of the laws of physics and other scientific disciplines arise directly as variational principles.

---

1.2. **First Variation, Euler-Lagrange Equation.** Let $\Omega \subseteq \mathbb{R}^n$ be a bounded, open set with $\mathcal{C}^\infty$ boundary $\partial\Omega$.

**Definition 1.1** (Lagrangian)**.** *The* **Lagrangian** *is a* $\mathcal{C}^2(\overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n, \mathbb{R})$ *function,*

$$L : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n.$$

We will write

$$L = L(x, z, p) = L(x_1, \dots, x_n, z, p_1, \dots, p_n)$$

for $x \in \Omega$, $z \in \mathbb{R}$, and $p \in \mathbb{R}^n$. Here "$z$" is the variable for which we substitute $w(x)$ below, and "$p$" is the variable for which we substitute $Dw(x)$. We also set

$$\begin{cases} D_x L := \left(\frac{\partial}{\partial x_1} L, \dots, \frac{\partial}{\partial x_n} L\right) = (L_{x_1}, \dots, L_{x_n}), \\ D_z L := \frac{\partial}{\partial z} L = L_z, \\ D_p L := \left(\frac{\partial}{\partial p_1} L, \dots, \frac{\partial}{\partial p_n} L\right) = (L_{p_1}, \dots, L_{p_n}). \end{cases}$$

We now assume that the energy functional in (1.2) has the explicit form

$$I[w] := \int_\Omega L(x, w(x), Dw(x)) \, dx, \tag{1.4}$$

{eq:1-4}

for all smooth functions $w : \overline{\Omega} \to \mathbb{R}$ satisfying some given boundary condition, say,

$$w|_{\partial\Omega} = g. \tag{1.5}$$

{eq:1-5}

We now additionally assume that some particular smooth function $u : \overline{\Omega} \to \mathbb{R}$, satisfying the boundary condition $u|_{\partial\Omega} = g$, is a minimizer of $I[\cdot]$ among all functions $w$ satisfying the boundary condition (1.5). We will show that this function $u$ is then automatically a solution of a certain nonlinear PDE, called the *Euler–Lagrange Equation*.

To prove this, choose any smooth function $\phi \in \mathcal{C}_c^\infty(\Omega)$ and consider the real–valued function

$$i(\tau) := I[u + \tau\phi], \quad \tau \in \mathbb{R}. \tag{1.6}$$

{eq:1-6}

We call the term $\tau\phi$ the *variation* of the function $u$.

Since $u$ is a minimizer of $I[\cdot]$ and $u + \tau\phi = u = g$ on $\partial\Omega$ (because $\phi$ has compact support on $\Omega$), we observe that $i(\cdot)$ has a minimum at $\tau = 0$. Therefore

$$i'(0) = 0. \tag{1.7}$$

{eq:1-7}

Computing this first derivative explicitly by writing out

$$i(\tau) = \int_\Omega L(x, u + \tau\phi, Du + \tau D\phi)\ dx, \tag{1.8}$$

{eq:1-8}

we have by the chain rule and differentiation under the integral sign

$$\begin{aligned} i'(\tau) &= \frac{d}{d\tau}\left[\int_\Omega L(x, u + \tau\phi, Du + \tau D\phi)\ dx\right] \\ &= \int_\Omega \frac{\partial}{\partial\tau}\left[L(x, u + \tau\phi, Du + \tau D\phi)\right]\ dx \\ &= \int_\Omega L_z(x, u + \tau\phi, Du + \tau D\phi)\phi + \sum_{i=1}^n L_{p_i}(x, u + \tau\phi, Du + \tau D\phi)\phi_{x_i}\ dx. \end{aligned}$$

**Definition 1.2** (First Variation)**.** *The derivative $i'(\tau)$ of $i(\tau)$ as defined in* (1.8)*, that is,*

$$i'(\tau) = \frac{\partial}{\partial\tau} I[u + \tau\phi],$$

*is called the* **first variation** *of the functional $I[\cdot]$.*

We note here that the first variation of $I[\cdot]$ is recognizable as the Gateaux derivative of $I[\cdot]$.

Letting $\tau = 0$, we see from (1.7) and the assumption that $u$ is a minimizer of $I[\cdot]$ that

$$0 = i'(0) = \int_\Omega L_z(x, u, Du)\phi + \sum_{i=1}^n L_{p_i}(x, u, Du)\phi_{x_i}\ dx.$$

Since $\phi$ has compact support in $\Omega$, integration by parts on the second term gives

$$\begin{aligned} 0 = i'(0) &= \int_\Omega L_z(x, u, Du)\phi - \sum_{i=1}^n (L_{p_i}(x, u, Du))_{x_i}\phi\ dx + \int_{\partial\Omega}\sum_{i=1}^n L_{p_i}(x, u, Du)\phi\nu^i\ dS \\ &= \int_\Omega \left[L_z(x, u, Du) - \sum_{i=1}^n (L_{p_i}(x, u, Du))_{x_i}\right]\phi\ dx, \end{aligned}$$

where $\nu = (\nu^1, \ldots, \nu^n)$ as usual denotes the outward pointing unit normal vector field along $\partial\Omega$. Since this equality holds for all text functions $\phi \in \mathcal{C}_c^\infty(\Omega)$, we conclude that $u$ solves the (possibly) nonlinear PDE

$$L_z(x, u, Du) - \sum_{i=1}^n (L_{p_i}(x, u, Du))_{x_i} = 0 \quad \text{on } \Omega. \tag{1.9}$$

{eq:1-9}

**Definition 1.3** (Euler–Lagrange Equation)**.** *For the energy functional $I[\cdot]$ as defined in* (1.4)*, the equation*

$$L_z(x, u, Du) - \sum_{i=1}^n (L_{p_i}(x, u, Du))_{x_i} = 0 \quad \text{on } \Omega$$

*is called the* **Euler–Lagrange equation** *associated with $I[\cdot]$.*

We observe that the Euler–Lagrange equation (1.9) is a quasilinear, second–order PDE in divergence form.

In summary, any smooth minimizer $u$ of $I[\cdot]$ is a solution of the Euler–Lagrange equation. Conversely, we can try to find a solution of the Euler–Lagrange PDE (1.9) by finding minimizers of the energy functional $I[\cdot]$ as defined in (1.4).

**Example 1.1** (Dirichlet's Principle)**.** *Put*

$$L(x,z,p) := \frac{1}{2}|p|^2.$$

*Then* $L_{p_i}(x,z,p) = p_i$, $i = 1,\dots,n$ *and* $L_z(x,z,p) = 0$. *Thus the Euler–Lagrange equation associated with the functional*

$$I[w] := \frac{1}{2}\int_\Omega |Dw|^2 \ dx = \int_\Omega L(x,w,Dw) \ dx$$

*is*

$$\Delta u = 0 \quad \text{on } \Omega.$$

*This is Dirichlet's principle.*

**Example 1.2.** *Sometimes we wish to convert a given PDE into a variational problem, that is, to recover a Lagrangian from a given PDE. Motivated by the previous example, consider the Laplacian*

$$\Delta u = 0 \quad \text{on } \Omega.$$

*Thus we want to define a function* $L$ *such that*

$$0 = \Delta w = L_z(x,w,Dw) - \sum_{i=1}^n (L_{p_i}(x,w,Dw))_{x_i}.$$

*We "guess" that*

$$\sum_{i=1}^n (L_{p_i}(x,w,Dw))_{x_i} = \Delta w.$$

*Taking* $(L_{p_i}(x,w,Dw))_{x_i} := \partial_{x_i}^2 w$ *and integrating with respect to* $x_i$, *we have*

$$L_{p_i}(x,w,Dw) = \partial_{x_i} w + C(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n).$$

*Thus we have*

$$\sum_{i=1}^n L_{p_i}(x,w,Dw) = \operatorname{div} w.$$

*Now taking* $L_{p_i}(x,w,Dw) := \partial_{x_i} w$ *and integrating with respect to* $\partial_{x_i} w$, *it follows*

$$L(x,w,Dw) = \frac{1}{2}(\partial_{x_i} w)^2 + C(\partial_{x_1} w,\dots,\partial_{x_{i-1}} w, \partial_{x_{i+1}} w,\dots,\partial_{x_n} w).$$

*Hence,*

$$L(x,w,Dw) = \frac{1}{2}|Dw|^2,$$

*which is the Lagrangian from the previous example.*

**Example 1.3** (Generalized Dirichlet's Principle)**.** *Take*

$$L(x, z, p) := \frac{1}{2} \sum_{i,j=1}^{n} a^{ij}(x) p_i p_j - z f(x),$$

*where* $a^{ij} = a^{ji}$, $i, j = 1, \ldots, n$. *Then*

$$L_{p_i}(x, z, p) = \frac{1}{2} \sum_{j=1}^{n} \sum_{i=1}^{n} a^{ij} p_j = \sum_{j=1}^{n} a^{ij} p_j,$$

$j = 1, \ldots, n$, *so that*

$$(L_z(x, z, p))_{x_i} = \frac{\partial}{\partial x_i} \left[ \sum_{j=1}^{n} a^{ij} p_j \right] = \sum_{j=1}^{n} (a^{ij} p_j)_{x_i},$$

*and*

$$L_z(x, z, p) = -f(x).$$

*Thus the Euler–Lagrange equation associated with the functional*

$$I[w] := \int_\Omega \frac{1}{2} \sum_{i,j=1}^{n} a^{ij} w_{x_i} w_{x_j} - wf \; dx$$

*is the divergence structure linear equation*

$$-\sum_{i,j}^{n} (a^{ij} u_{x_j})_{x_i} = f \quad \text{on } \Omega.$$

We will see later that the uniform ellipticity condition on the $a^{ij}$, $i, j = 1, \ldots, n$ is a natural further assumption required to prove the existence of a minimizer. Consequently from the nonlinear viewpoint of the calculus of variations, the divergence structure form of a linear second–order elliptic PDE is completely natural.

**Example 1.4** (Nonlinear Poisson Equation)**.** *Assume that we are given a smooth function* $f : \mathbb{R} \to \mathbb{R}$, *and define its antiderivative* $F(z) := \int_0^z f(\xi) \; d\xi$. *Take*

$$L(x, z, p) := \frac{1}{2}|p|^2 - F(z).$$

*Then* $L_{p_i}(x, z, p) = p_i$, $i = 1, \ldots, n$, *so that* $(L_{p_i}(x, z, p))_{x_i} = \partial_{x_i} p_i$, $i = 1, \ldots, n$, *and* $L_z(x, z, p) = -f(z)$. *Thus the Euler–Lagrange equation associated with the functional*

$$I[w] := \int_\Omega \frac{1}{2}|Dw|^2 - F(w) \; dx$$

*is the nonlinear Poisson equation*

$$-\Delta u = f(u) \quad \text{on } \Omega.$$

**Example 1.5** (Minimal Surfaces)**.** *Put*

$$L(x, z, p) := (1 + |p|^2)^{1/2},$$

*so that*

$$I[w] := \int_\Omega (1 + |Dw|^2)^{1/2} \; dx$$

*is the area of the graph of the function* $w : \Omega \to \mathbb{R}$. *We compute*

$$L_{p_i}(x, z, p) = \frac{1}{2}(1 + |p|^2)^{1/2} \cdot 2p_i = \frac{p_i}{(1 + |p|^2)^{1/2}},$$

$i = 1, \ldots, n$, *and* $L_z(x, z, p) = 0$, *so that the associated Euler–Lagrange equation is*

$$\sum_{i=1}^{n} \left( \frac{u_{x_i}}{(1 + |Du|^2)^{1/2}} \right)_{x_i} = 0 \quad \text{on } \Omega. \tag{1.10}$$

{eq:1-10}

*This PDE is called the minimal surface equation. The expression*

$$\operatorname{div} \left( \frac{Du}{(1 + |Du|^2)^{1/2}} \right)$$

*on the LHS of* (eq:1-10 1.10) *is* $n$ *times the mean curvature of* $u$. *Thus a minimal surface has zero mean curvature.*

---

1.3. **Second Variation.** We continue the calculations from §1.2 by computing the *second variation* of the functional $I[\cdot]$ at the function $u$. This we find by observing that since $u$ gives a minimum for $I[\cdot]$, we must have

$$i''(0) \geq 0,$$

where $i$ is defined as above by (eq:1-6 1.6). Recall from that the first variation of $I[\cdot]$ is given by

$$i'(\tau) = \int_\Omega L_z(x, u + \tau\phi, Du + \tau\phi)\phi + \sum_{i=1}^{n} L_{p_i}(x, u + \tau\phi, Du + \tau D\phi)\phi_{x_i} \, dx.$$

Calculating the second derivative explicitly, once again by applying the chain rule and differentiation under the integral sign, we find

$$\begin{aligned}
i''(\tau) &= \frac{d}{d\tau}\left[ \int_\Omega L_z(x, u + \tau\phi, Du + \tau\phi)\phi + \sum_{i=1}^{n} L_{p_i}(x, u + \tau\phi, Du + \tau D\phi)\phi_{x_i} \, dx \right] \\
&= \int_\Omega \frac{\partial}{\partial\tau}\left[L_z(x, u + \tau\phi, Du + \tau D\phi)\phi\right] + \sum_{i=1}^{n} \frac{\partial}{\partial\tau}\left[L_{p_i}(x, u + \tau\phi, Du + \tau D\phi)\phi_{x_i}\right] \, dx \\
&= \int_\Omega L_{zz}(x, u + \tau\phi, Du + \tau D\phi)\phi^2 + \sum_{i=1}^{n} L_{zp_i}(x, u + \tau\phi, Du + \tau D\phi)\phi\phi_{x_i} + \\
&\qquad \sum_{i=1}^{n} L_{p_i z}(x, u + \tau\phi, Du + \tau D\phi)\phi_{x_i}\phi + \sum_{i=1}^{n}\sum_{j=1}^{n} L_{p_i p_j}(x, u + \tau\phi, Du + \tau D\phi)\phi_{x_i}\phi_{x_j} \, dx \\
&= \int_\Omega L_{zz}(x, u + \tau\phi, Du + \tau D\phi)\phi^2 + 2\sum_{i=1}^{n} L_{p_i z}(x, u + \tau\phi, Du + \tau D\phi)\phi_{x_i}\phi + \\
&\qquad \sum_{i,j}^{n} L_{p_i p_j}(x, u + \tau\phi, Du + \tau D\phi)\phi_{x_i}\phi_{x_j} \, dx.
\end{aligned}$$

**Definition 1.4** (Second Variation)**.** *The second derivative $i''(\tau)$ of $i(\tau)$ as defined in (1.8), that is,*

$$i''(\tau) = \frac{\partial^2}{\partial \tau^2} I[u + \tau\phi],$$

*is called the* **second variation** *of the functional $I[\cdot]$.*

Again now letting $\tau = 0$, we obtain the inequality

$$0 \leq i''(0) = \int_\Omega L_{zz}(x,u,Du)\phi^2 + 2\sum_{i=1}^n L_{p_i z}(x,u,Du)\phi_{x_i}\phi + \sum_{i,j=1}^n L_{p_i p_j}(x,u,Du)\phi_{x_i}\phi_{x_j}\, dx. \tag{1.11}$$

{eq:1-11}

This holds for all test functions $\phi \in \mathcal{C}_c^\infty(\Omega)$.

We can extract useful information from the inequality (1.11) as follows. First, note that after a standard approximation argument that the estimate (1.11) is valid for any Lipschitz continuous function $\phi$ vanishing on $\partial\Omega$. This is because for an open, bounded set $\Omega \subseteq \mathbb{R}^n$ with $\partial\Omega \in \mathcal{C}^1$, $\mathrm{Lip}(\Omega, \mathbb{R}) = W^{1,\infty}(\Omega)$. We then fix $\xi \in \mathbb{R}^n$ and define

$$v(x) := \epsilon\rho\left(\frac{\xi \cdot x}{\epsilon}\right)\zeta(x), \quad x \in \Omega, \tag{1.12}$$

{eq:1-12}

where $\zeta \in \mathcal{C}_c^\infty(\Omega)$ and $\rho : \mathbb{R} \to \mathbb{R}$ is the "periodic triangular function" defined by

$$\rho(x) := \begin{cases} x, & 0 \leq x \leq \frac{1}{2}, \\ 1 - x, & \frac{1}{2} \leq x \leq 1, \end{cases} \qquad \rho(x+1) = \rho(x), \quad x \in \mathbb{R}.$$

Note that

$$|\rho'| = 1 \quad \mathcal{L}^1 - \text{a.e.} \tag{1.13}$$

{eq:1-13}

Observe further that by the chain rule,

$$\begin{aligned} v_{x_i}(x) &= \rho'\left(\frac{\xi \cdot x}{\epsilon}\right)\xi_i\zeta(x) + \epsilon\rho\left(\frac{\xi \cdot x}{\epsilon}\right)\zeta_{x_i}(x) \\ &= \rho'\left(\frac{\xi \cdot x}{\epsilon}\right)\xi_i\zeta(x) + \mathcal{O}(\epsilon), \quad \epsilon \to 0, \end{aligned}$$

since $|\rho(x)| \leq 1$ for all $x \in \mathbb{R}$ and $\max_{x\in\Omega}|\zeta_{x_i}(x)| < +\infty$ because $\zeta \in \mathcal{C}_c^\infty(\Omega)$. Similarly, note also that

$$v(x) = \mathcal{O}(\epsilon), \quad \epsilon \to 0.$$

Thus, substituting $v$ (cf. (1.12)) into the estimate (1.11) yields

$$0 \leq \int_\Omega \sum_{i,j=1}^n L_{p_i p_j}(x,u,Du)(\rho')^2\xi_i\xi_j\zeta^2 + \mathcal{O}(\epsilon)\, dx.$$

Recalling (1.13) and taking the limit as $\epsilon \to 0$, we obtain the inequality

$$0 \leq \int_\Omega \sum_{i,j=1}^n L_{p_i p_j}(x,u,Du)\xi_i\xi_j\zeta^2\, dx.$$

Since this estimate holds for all $\zeta \in \mathcal{C}_c^\infty(\Omega)$, it follows

$$\sum_{i,j=1}^n L_{p_i p_j}(x,u,Du)\xi_i\xi_j \geq 0, \quad \xi \in \mathbb{R}^n, \quad x \in \Omega. \tag{1.14}$$

{eq:1-14}

We will see later that this necessary condition suggests the assumption that the Lagrangian $L$ is convex in its third argument, which is required for the existence theory.

---

## 1.4. **Systems.**

1.4.1. *Euler–Langrange Equations.* The previous considerations generalize easily to the case of systems. Recall that $\mathbb{R}^{m\times n}$ denotes the space of real $m \times n$ matrices. Assume that the smooth Lagrangian

$$L : \overline{\Omega} \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$$

is given.

We will write

$$L = L(x, z, P) = L(x_1, \dots, x_n, z^1, \dots, z^m, p_1^1, \dots, p_n^m)$$

for $x \in \Omega$, $z \in \mathbb{R}^m$, and $P \in \mathbb{R}^{m\times n}$, where

$$P := \begin{pmatrix} p_1^1 & \cdots & p_n^1 \\ & \ddots & \\ p_1^m & \cdots & p_n^m \end{pmatrix}.$$

We are employing superscripts to denote rows, as this notational convention simplifies the following formulas.

We associate with $L$ the functional

$$I[\boldsymbol{w}] := \int_\Omega L(x, \boldsymbol{w}(x), D\boldsymbol{w}(x))\ dx, \tag{1.15}$$ {eq:1-15}

defined for smooth functions $\boldsymbol{w} : \overline{\Omega} \to \mathbb{R}^m$, $\boldsymbol{w} = (w^1, \dots, w^m)$, satisfying some given boundary condition

$$\boldsymbol{w}|_{\partial\Omega} = \boldsymbol{g},$$

where $\boldsymbol{g} : \partial\Omega \to \mathbb{R}^m$. Note here that

$$D\boldsymbol{w}(x) = \begin{pmatrix} w_{x_1}^1(x) & \dots & w_{x_n}^1(x) \\ & \ddots & \\ w_{x_1}^m(x) & \dots & w_{x_n}^m(x) \end{pmatrix}$$

is the gradient matrix of $\boldsymbol{w}$ at $x$.

We proceed as we did in the previous sections and show that any smooth minimizer $\boldsymbol{u} = (u^1, \dots, u^m)$ of $I[\cdot]$, given now by (1.15), taking among all functions satisfying the boundary condition $\boldsymbol{w}|_{\partial\Omega} = \boldsymbol{g}$, must solve a certain *system* of nonlinear PDEs. We therefore fix a function $\boldsymbol{\phi} = (\phi^1, \dots, \phi^m) \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^m)$ and write

$$i(\tau) := I[\boldsymbol{u} + \tau\boldsymbol{\phi}].$$

As before,

$$i'(0) = 0.$$

Calculating the first variation explicitly, we find

$$i'(\tau) = \frac{d}{d\tau}\left[\int_\Omega L(x, \boldsymbol{w} + \tau\boldsymbol{\phi}, D\boldsymbol{w} + \tau D\boldsymbol{\phi})\ dx\right] = \int_\Omega \frac{\partial}{\partial\tau}\left[L(x, \boldsymbol{w} + \tau\boldsymbol{\phi}, D\boldsymbol{w} + \tau D\boldsymbol{\phi})\right]\ dx$$

$$= \int_\Omega \sum_{k=1}^m L_{z^k}(x, \boldsymbol{w} + \tau\boldsymbol{\phi}, D\boldsymbol{w} + \tau D\boldsymbol{\phi})\phi^k + \sum_{i=1}^n \sum_{k=1}^m L_{p_i^k}(x, \boldsymbol{w} + \tau\boldsymbol{\phi}, D\boldsymbol{w} + \tau D\boldsymbol{\phi})\phi_{x_i}^k \; dx.$$

Setting $\tau = 0$, we derive the equality

$$0 = i'(0) = \int_\Omega \sum_{k=1}^m L_{z^k}(x, \boldsymbol{u}, D\boldsymbol{u})\phi^k + \sum_{i=1}^n \sum_{k=1}^m L_{p_i^k}(x, \boldsymbol{u}, D\boldsymbol{u})\phi_{x_i}^k \; dx.$$

Integrating by parts, we have

$$0 = i'(0) = \int_\Omega \sum_{k=1}^m L_{z^k}(x, \boldsymbol{u}, D\boldsymbol{u})\phi^k - \sum_{i=1}^n \sum_{k=1}^m (L_{p_i^k}(x, \boldsymbol{u}, D\boldsymbol{u}))_{x_i}\phi^k \; dx$$

holding for all test functions $\boldsymbol{\phi} \in \mathcal{C}_c^\infty(\Omega, \mathbb{R}^m)$. Thus we conclude that $\boldsymbol{u}$ solves the coupled, quasilinear system of PDE

$$L_{z^k}(x, \boldsymbol{u}, D\boldsymbol{u}) - \sum_{i=1}^n \Big( L_{p_i^k}(x, \boldsymbol{u}, D\boldsymbol{u}) \Big)_{x_i} = 0 \quad \text{on } \Omega, \quad k = 1, \ldots, m. \tag{1.16}$$ {eq:1-16}

**Definition 1.5** (Euler–Lagrange Equations)**.** *For the energy functional $I[\cdot]$ as defined in* eq:1-15 *(1.15), the equations*

$$L_{z^k}(x, \boldsymbol{u}, D\boldsymbol{u}) - \sum_{i=1}^n \Big( L_{p_i^k}(x, \boldsymbol{u}, D\boldsymbol{u}) \Big)_{x_i} = 0 \quad \text{on } \Omega, \quad k = 1, \ldots, m,$$

*are called the* **Euler–Lagrange equations** *associated with $I[\cdot]$.*

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

1.4.2. *Null Lagrangians.* It turns out to be interesting to study certain systems of nonlinear PDEs for which *every* smooth function is a solution.

**Definition 1.6** (Null Lagrangian)**.** *The function $L : \overline{\Omega} \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is called a* **null Lagrangian** *if the system of Euler–Lagrange equations*

$$L_{z^k}(x, \boldsymbol{u}, D\boldsymbol{u}) - \sum_{i=1}^n \Big( L_{p_i^k}(x, \boldsymbol{u}, D\boldsymbol{u}) \Big)_{x_i} = 0 \quad \text{on } \Omega, \quad k = 1, \ldots, m, \tag{1.17}$$ {eq:1-17}

*is solved by all smooth functions $u : \Omega \to \mathbb{R}^m$.*

The importance of null Lagrangians is that the corresponding energy functional

$$I[\boldsymbol{w}] = \int_\Omega L(x, \boldsymbol{w}, D\boldsymbol{w}) \; dx$$

depends only on the boundary conditions.

t1.1 **Theorem 1.1** (Null Lagrangians and Boundary Conditions)**.** *Let $L : \overline{\Omega} \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be a null Lagrangian. Assume that $\boldsymbol{u}, \widetilde{\boldsymbol{u}} \in \mathcal{C}^2(\overline{\Omega}, \mathbb{R}^m)$ are two functions such that*

$$\boldsymbol{u} \equiv \widetilde{\boldsymbol{u}} \quad \text{on } \partial\Omega. \tag{1.18}$$ {eq:1-18}

*Then*

$$I[\boldsymbol{u}] = I[\widetilde{\boldsymbol{u}}]. \tag{1.19}$$ {eq:1-19}

*Proof.* Define

$$i(\tau) := I[\tau\boldsymbol{u} + (1-\tau)\widetilde{\boldsymbol{u}}], \quad \tau \in [0,1].$$

Note that $\tau\boldsymbol{u} + (1-\tau)\widetilde{\boldsymbol{u}} \in \mathcal{C}^2(\overline{\Omega}, \mathbb{R}^m)$, and thus satisfies the system of Euler–Lagrange equations

$$L_{z^k}(x, \tau\boldsymbol{u}+(1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u}+(1-\tau)D\widetilde{\boldsymbol{u}}) - \sum_{i=1}^{n}\Big(L_{p_i^k}(x, \tau\boldsymbol{u} + (1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u} + (1-\tau)D\widetilde{\boldsymbol{u}})\Big)_{x_i} = 0$$

on $\Omega$, $k = 1, \ldots, m$. Therefore,

$$\begin{aligned} i'(\tau) = \int_\Omega \sum_{k=1}^{m} & L_{z^k}(x, \tau\boldsymbol{u} + (1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u} + (1-\tau)D\widetilde{\boldsymbol{u}})(u^k - \widetilde{u}^k) + \\ & \sum_{i=1}^{n}\sum_{k=1}^{m} L_{p_i^k}(x, \tau\boldsymbol{u} + (1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u} + (1-\tau)D\widetilde{\boldsymbol{u}})(u^k_{x_i} - \widetilde{u}^k_{x_i})\, dx. \end{aligned}$$

Integrating by parts on the second term, we obtain

$$\begin{aligned} i'(\tau) &= \int_\Omega \sum_{k=1}^{m} L_{z^k}(x, \tau\boldsymbol{u} + (1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u} + (1-\tau)D\widetilde{\boldsymbol{u}})(u^k - \widetilde{u}^k) - \\ &\qquad \sum_{i=1}^{n}\sum_{k=1}^{m} \Big(L_{p_i^k}(x, \tau\boldsymbol{u} + (1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u} + (1-\tau)D\widetilde{\boldsymbol{u}})\Big)_{x_i}(u^k - \widetilde{u}^k)\, dx + \\ &\qquad \int_{\partial\Omega}\sum_{i=1}^{n}\sum_{k=1}^{m} L_{p_i^k}(x, \tau\boldsymbol{u} + (1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u} + (1-\tau)D\widetilde{\boldsymbol{u}})(u^k - \widetilde{u}^k)\nu^i\, dS \\ &= \sum_{k=1}^{m}\int_\Omega \Bigg[L_{z^k}(x, \tau\boldsymbol{u} + (1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u} + (1-\tau)D\widetilde{\boldsymbol{u}}) \;- \\ &\qquad \sum_{k=1}^{m}\Big(L_{p_i^k}(x, \tau\boldsymbol{u} + (1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u} + (1-\tau)D\widetilde{\boldsymbol{u}})\Big)\Bigg](u^k - \widetilde{u}^k)\, dx + \\ &\qquad \int_{\partial\Omega}\sum_{i=1}^{n}\sum_{k=1}^{m} L_{p_i^k}(x, \tau\boldsymbol{u} + (1-\tau\widetilde{\boldsymbol{u}}), \tau D\boldsymbol{u} + (1-\tau)D\widetilde{\boldsymbol{u}})(u^k - \widetilde{u}^k)\nu^i\, dS \\ &= 0, \end{aligned}$$

where the last equality holds by the assumption (1.18) and the fact that $\tau\boldsymbol{u} + (1-\widetilde{\ })\widetilde{\boldsymbol{u}}$ solves the system of Euler–Lagrange equations. Thus identity (1.19) follows by observing

$$I[\boldsymbol{u}] = i(1) = i(0) = I[\widetilde{\boldsymbol{u}}].$$

The proof is complete. □

In the scalar case $m = 1$, the only null Lagrangians are the cases where $L$ is linear in the variable $p$. For the case of systems, that is, when $m > 1$, however, there are certain nontrivial examples.

We explain a bit of notation for the following result. If $A$ is an $n \times n$ matrix, we denote by

$$\operatorname{cof} A$$

the *cofactor* matrix, whose $(k,i)^{\text{th}}$ entry is $(\operatorname{cof} A)_i^k = (-1)^{i+k} d(A)_i^k$, where $d(A)_i^k$ is the determinant of the $(n-1)\times(n-1)$ matrix obtained by deleting the $k^{\text{th}}$ row and $i^{\text{th}}$ column of $A$.

l1.1 **Lemma 1.1** (Divergence–Free Rows)**.** *Let $\boldsymbol{u}\mathbb{R}^n \to \mathbb{R}^n$ be a smooth function. Then*

$$\sum_{i=1}^{n} (\operatorname{cof} D\boldsymbol{u})^n_{i,x_i} = 0, \quad k = 1,\dots,n. \tag{1.20}$$ {eq:1-20}

*Proof.*

(i). From linear algebra we recall the identity

$$(\det P)I = P^T(\operatorname{cof} P), \quad P \in \mathbb{R}^{n\times n}, \tag{1.21}$$ {eq:1-21}

that is,

$$(\det P)\delta_{ij} = \sum_{i=1}^{n} p_i^k (\operatorname{cof} P)_j^k, \quad i,j = 1,\dots,n. \tag{1.22}$$ {eq:1-22}

Thus in particular

$$\partial_{p_m^k} \det P = (\operatorname{cof} P)_m^k, \quad k,m = 1,\dots,n. \tag{1.23}$$ {eq:1-23}

(ii). Now set $P = D\boldsymbol{u}$ in (1.22), differentiate with respect to $x_j$, and sum $j = 1,\dots,n$, to find

$$\sum_{j,k,m=1}^{n} \delta_{ij} (\operatorname{cof} D\boldsymbol{u})_m^k u^k_{x_m x_j} = \sum_{k,j=1}^{n} u^k_{x_i x_j} (\operatorname{cof} D\boldsymbol{u})_j^k + u^k_{x_i} (\operatorname{cof} D\boldsymbol{u})^k_{j,x_j},$$

for $i = 1,\dots,n$. This identity simplifies to

$$\sum_{i=1}^{n} u^n_{x_i} \left( \sum_{j=1}^{n} (\operatorname{cof} D\boldsymbol{u})^k_{j,x_j} \right) = 0, \quad i = 1,\dots,n. \tag{1.24}$$ {eq:1-24}

(iii). Now if $\det D\boldsymbol{u}(x_0) \neq 0$, we deduce from (1.24) that

$$\sum_{j=1}^{n} (\operatorname{cof} D\boldsymbol{u})^k_{j,x_j} = 0, \quad k = 1,\dots,n$$

at $x_0$. If instead $\det D\boldsymbol{u}(x_0) = 0$, we choose a number $\epsilon > 0$ so small that $\det(D\boldsymbol{u}(x_0) + \epsilon I) \neq 0$, apply steps (i) and (ii) to $\widetilde{\boldsymbol{u}} := \boldsymbol{u} + \epsilon x$, and send $\epsilon \to 0$. The proof is complete. ☐

t1.2 **Theorem 1.2** (Determinants as Null Lagrangians)**.** *The determinant function*

$$L(P) = \det P, \quad P \in \mathbb{R}^{n\times n},$$

*is a null Lagrangian.*

*Proof.* We must show that for any smooth function $\boldsymbol{u} : \Omega \to \mathbb{R}^n$,

$$\sum_{i=1}^{n} \Big( L_{p_i^k}(D\boldsymbol{u}) \Big)_{x_i} = 0, \quad k = 1,\dots,n.$$

By (1.23), we have $L_{p_i^k} = (\operatorname{cof} P)_i^k$, $i,k = 1,\dots,n$. But then it follows by Lemma (1.1) that

$$\sum_{i=1}^{n} \Big( L_{p_i^k}(D\boldsymbol{u}) \Big)_{x_i} = \sum_{i=1}^{n} (\operatorname{cof} D\boldsymbol{u})^k_{i,x_i} = 0, \quad k = 1,\dots,n,$$

as required. The proof is complete. ☐

1.4.3. *Application.* A nice application is an analytic proof of Brouwer's Fixed Point Theorem.

t1.3 **Theorem 1.3** (Brouwer's Fixed Point Theorem)**.** *Assume that*

$$\boldsymbol{u} : \overline{B(0,1)} \to \overline{B(0,1)}$$

*is continuous, where $B(0,1)$ denotes the open unit ball in $\mathbb{R}^n$. Then $\boldsymbol{u}$ has a fixed point, that is, there exists a point $x \in \overline{B(0,1)}$ such that*

$$\boldsymbol{u}(x) = x.$$

*Proof.*

(i). Write $B := \overline{B(0,1)}$. We first of all show that there does not exist a smooth function

$$\boldsymbol{w} : B \to \partial B \tag{1.25}$$ {eq:1-25}

such that

$$\boldsymbol{w}(x) = x \tag{1.26}$$ {eq:1-26}

for all $x \in \partial B$.

(ii). Suppose by contradiction that such a function $\boldsymbol{w}$ exists. Denote by $\widetilde{\boldsymbol{w}}$ the identity function on $B$, so that $\widetilde{\boldsymbol{w}}(x) = x$ for all $x \in B$. By (1.26), $\boldsymbol{w} = \widetilde{\boldsymbol{w}}$ on $\partial B$. Since the determinant is a null Lagrangian (cf. (1.2)), we have by Theorem (1.1)

$$\int_B \det D\boldsymbol{w}\ dx = \int_B \det D\widetilde{\boldsymbol{w}}\ dx = \mathcal{L}^n(B) \neq 0. \tag{1.27}$$ {eq:1-27}

On the other hand, (1.25) implies that $|\boldsymbol{w}(x)| = 1$ for all $x \in B$, so that $|\boldsymbol{w}|^2 \equiv 1$. Differentiating, we find

$$(D\boldsymbol{w})^T \boldsymbol{w} = \mathbf{0}. \tag{1.28}$$ {eq:1-28}

Since $|\boldsymbol{w}| = 1$, (1.28) implies that 0 is an eigenvalue of $(D\boldsymbol{w})^T$ for each $x \in B$. Therefore $\det D\boldsymbol{w} \equiv 0$ in $B$. This contradicts (1.27), and therefore no smooth function $\boldsymbol{w}$ satisfying (1.25) and (1.26) can exist.

(iii). Next we show that there does not exist any continuous function satisfying (1.25) and (1.26). Suppose again that such a function $\boldsymbol{w}$ does exist. We may then continuously extend $\boldsymbol{w}$ by setting $\boldsymbol{w}(x) := x$ if $x \in \mathbb{R}^n \setminus B$. Observe that $\boldsymbol{w}(x) \neq 0$ for all $x \in \mathbb{R}^n$. Fix $\epsilon > 0$ so small that $\boldsymbol{w}_1 := \eta_\epsilon * \boldsymbol{w}$ satisfies $\boldsymbol{w}_1(x) \neq 0$ for all $x \in \mathbb{R}^n$. Note also that since $\eta_\epsilon$ is radial, we have $\boldsymbol{w}_1(x) = x$ if $x \in \mathbb{R}^n \setminus \overline{B(0,2)}$, for $\epsilon > 0$ sufficiently small. Then

$$\boldsymbol{w}_2 := \frac{2\boldsymbol{w}_1}{|\boldsymbol{w}_1|}$$

would be a smooth mapping satisfying (1.25) and (1.26) with the ball $\overline{B(0,2)}$ replacing $\overline{B(0,1)}$, a contradiction to steps (i) and (ii).

(iv). Finally, suppose that $\boldsymbol{u} : B \to B$ is continuous but has no fixed point. Define the mapping $\boldsymbol{w} : B \to \partial B$ by setting $\boldsymbol{w}(x)$ to be the point on $\partial B$ hit by the ray emanating from $\boldsymbol{u}(x)$ and passing through $x$. Note that this mapping is well–defined since $\boldsymbol{u}(x) \neq x$ for all $x \in B$. Moreover, $\boldsymbol{w}$ is continuous and satisfies (1.25) and (1.26).

But this is a contradiction to step (iii). The proof is complete. □

## 2. Existence of Minimizers

In this section we present some conditions on the Lagrangian $L$ which ensure that the functional $I[\cdot]$ will indeed have a minimizer, at least within an appropriate Sobolev space.

### 2.1. **Coercivity, Lower Semicontinuity.**

We begin with some insights as to when the functional (cf. (1.4))

$$I[w] := \int_\Omega L(x, w(x), Dw(x))\, dx, \tag{2.1}$$

defined for appropriate functions $w : \Omega \to \mathbb{R}$, satisfying a given boundary condition (cf. (1.5))

$$w|_{\partial\Omega} = g, \tag{2.2}$$

should have a minimizer.

---

2.1.1. *Coercivity.* We first of all note that even a smooth function $f$ mapping $\mathbb{R}$ to $\mathbb{R}$ and bounded below need not attain its infimum. For example, consider the functions $f(x) := e^x$ or $f(x) := (1 + x^2)^{-1}$. These examples suggest that we in general will need some hypothesis controlling the functional $I[w]$ for "large" functions $w$. The most effective way to ensure this is to hypothesize that $I[w]$ "grows rapidly as $|w| \to +\infty$."

To be specific, we assume that

$$1 < q < +\infty \tag{2.3}$$

is fixed. We then suppose that there exist constants $\alpha > 0$ and $\beta \geq 0$ such that

$$L(x, z, p) \geq \alpha |p|^q - \beta \tag{2.4}$$

for all $x \in \Omega$, $z \in \mathbb{R}$, and $p \in \mathbb{R}^n$. Inserting $w(x)$ for $z$ and $Dw(x)$ for $p$ in (2.1), we have from (2.4) therefore

$$I[w] \geq \alpha \|Dw\|^q_{L^q(\Omega)} - \gamma, \tag{2.5}$$

where $\gamma := \beta \mathcal{L}^n(\Omega)$. In this instance, we see that $I[w] \to +\infty$ as $\|Dw\|_{L^q(\Omega)} \to +\infty$. This behavior is called *coercivity.*

**Definition 2.1** (Coercivity)**.** *The condition*

$$I[w] \geq \alpha \|Dw\|^q_{L^q(\Omega)} - \gamma$$

*in* (2.5)*, for constants $\alpha > 0$ and $\gamma \geq 0$, is called a* **coercivity condition** *on $I[\cdot]$.*

Recall that we want to find minimizers of the functional $I[\cdot]$. Observe from the inequality (2.5) that so far we only suppose that $Dw \in L^q(\Omega)$. Thus it seems reasonable to define $I[w]$ not only for smooth functions $w$, but also for functions $w$ in the Sobolev space $W^{1,q}(\Omega)$ that satisfy the boundary condition (cf. (2.2)) $w|_{\partial\Omega} = g$ in the trace sense. The wider the class of functions $w$ for which $I[w]$ is defined, the more candidates we have for a minimizer.

We denote

$$\mathcal{A} := \left\{ w \in W^{1,q}(\Omega) : w|_{\partial\Omega} = g \right\}$$

to denote the class of *admissible* functions $w$, where $w|_{\partial\Omega} = g$ is understood in the trace sense. Note that by (2.4) $I[w]$ is defined, but may be equal to $+\infty$, for each $w \in \mathcal{A}$.

2.1.2. *Lower Semicontinuity.* First observe that a continuous function $f : \mathbb{R} \to \mathbb{R}$ satisfying a coercivity condition

$$\lim_{|x| \to +\infty} f(x) = +\infty$$

attains its infimum. However, the functional $I[w]$ in general will not. To understand the issue, set

$$m := \inf_{w \in \mathcal{A}} I[w]. \tag{2.6}$$

{eq:2-6}

By definition of an infimum, there exists a sequence $\{u_k\}_{k=1}^{+\infty} \subseteq \mathcal{A}$ so that

$$\lim_{k \to +\infty} I[u_k] = m. \tag{2.7}$$

{eq:2-7}

**Definition 2.2** (Minimizing Sequence)**.** *A sequence $\{u_k\}_{k=1}^{+\infty} \subseteq \mathcal{A}$ such that*

$$\lim_{k \to +\infty} I[u_k] = \inf_{w \in \mathcal{A}} I[w]$$

*(cf. (2.7)) is called a* **minimizing sequence***.*

We now want to show that some subsequence of the minimizing sequence $\{u_k\}_{k=1}^{+\infty}$ actually converges to a minimizer. For this, however, we need some kind of compactness, and this is certainly an issue since the space $W^{1,q}(\Omega)$ is infinite–dimensional. If we apply the coercivity inequality (2.5), we can conclude only that the minimizing sequence $\{u_k\}_{k=1}^{+\infty}$ lies in a bounded subset of $W^{1,q}(\Omega)$. To see this, we make the following remark.

**Remark.** *Suppose that $I[\cdot]$ satisfies the coercivity condition*

$$I[w] \geq \alpha \|Dw\|_{L^q(\Omega)}^q - \gamma$$

*for constants $\alpha > 0$ and $\gamma \geq 0$, and let $\{u_k\}_{k=1}^{+\infty}$ be a minimizing sequence for $I[\cdot]$. Then $\{u_k\}_{k=1}^{+\infty}$ is bounded in $W^{1,q}(\Omega)$.*

*Proof.* Since $I[u_k] \to \inf_{w \in \mathcal{A}} I[w] =: m$, there exists $k_0 \in \mathbb{N}$ so large that for all $k \geq k_0$,

$$I[u_k] \leq m + 1.$$

Thus for all $k \in \mathbb{N}$,

$$I[u_k] \leq M_1 + m + 1,$$

where $M_1 := \max\{I[u_1], \ldots, I[u_{k_0 - 1}]\}$. Therefore, for all $k \in \mathbb{N}$, we have by the coercivity condition (2.5)

$$\|Du_k\|_{L^q(\Omega)} \leq \left( \frac{M_1 + m + 1 + \gamma}{\alpha} \right)^{1/q} < +\infty.$$

It remains to show that $\{u_k\}_{k=1}^{+\infty}$ is bounded in $L^q(\Omega)$. We have by the Poincaré inequality

$$\|u_k\|_{L^q(\Omega)} + \|(u_k)_\Omega\|_{L^q(\Omega)} \leq C \|Du_k\|_{L^q(\Omega)}$$

for some constant $C > 0$, where $(u_k)_\Omega := ⨍_\Omega u_k \, d\mathcal{L}^n$. Since $u_k \in L^q(\Omega)$ and $\Omega$ is bounded, we have that

$$\|(u_k)_\Omega\|_{L^q(\Omega)} = |(u_k)_\Omega| \mathcal{L}^n(\Omega)^{1/q} < +\infty.$$

Putting $M_2 := \sup_{k \in \mathbb{N}} |(u_k)_\Omega| < +\infty$, it follows

$$\begin{aligned} \|u_k\|_{W^{1,q}(\Omega)} &= \|u_k\|_{L^q(\Omega)} + \|Du_k\|_{L^q(\Omega)} \leq (C+1)\|Du_k\|_{L^q(\Omega)} + M_2 \mathcal{L}^n(\Omega)^{1/q} \\ &\leq (C+1) \left( \frac{M_1 + m + 1 + \gamma}{\alpha} \right)^{1/q} + M_2 \mathcal{L}^n(\Omega)^{1/q} \end{aligned}$$

$$< +\infty,$$

as required. The proof is complete. □

Thus we see that any minimizing sequence $\{u_k\}_{k=1}^{+\infty} \subseteq \mathcal{A}$ is bounded in $W^{1,q}(\Omega)$. But this does *not* imply that there exists any subsequence which converges in $W^{1,q}(\Omega)$.

We therefore turn our attention to the *weak topology* on $W^{1,q}(\Omega)$. Let us first recall a definition from measure theory and an important theorem from functional analysis.

**Definition 2.3** (Weak Convergence in $L^q(\Omega)$)**.** *A sequence $\{f_k\}_{k=1}^{+\infty} \subset L^q(\Omega)$ is said to converge weakly to a function $f \in L^q(\Omega)$, written*

$$f_k \rightharpoonup f \quad \text{in } L^q(\Omega),$$

*if*

$$\lim_{k\to+\infty} \int_\Omega f_k g \; d\mathcal{L}^n = \int_\Omega fg \; d\mathcal{L}^n$$

*for every $g \in L^p(\Omega)$, where $q$ and $p$ are (Hölder) conjugate exponents, $\frac{1}{q} + \frac{1}{p} = 1$.*

**Theorem** (Eberlein–Smulyan Theorem)**.** *A Banach space $X$ is reflexive if and only if every bounded sequence in $X$ has a weakly convergent subsequence.*

Since we are assuming that $1 < q < +\infty$, so that $L^q(\Omega)$ is a reflexive Banach space, we have by the Eberlein–Smulyan Theorem that there exists a subsequence $\{u_{k_j}\}_{j=1}^{+\infty} \subset L^q(\Omega)$ such that $u_{k_j} \rightharpoonup u$ weakly in $L^q(\Omega)$. Similarly there exists a further subsequence $\{u_{k_{j_l}}\}_{l=1}^{+\infty} \subset L^q(\Omega)$ such that $Du_{k_{j_l}} \rightharpoonup u'$ weakly in $L^q(\Omega, \mathbb{R}^n)$. But since we have also $\{u_{k_{j_l}}\}_{l=1}^{+\infty} \subset \{u_{k_j}\}_{j=1}^{+\infty}$, and $u_{k_j} \rightharpoonup u$ in $L^q(\Omega)$, we have that $Du = u'$. Therefore, given a minimizing sequence $\{u_k\}_{k=1}^{+\infty} \subseteq \mathcal{A}$, there exists a subsequence $\{u_{k_j}\}_{j=1}^{+\infty} \subseteq \{u_k\}_{k=1}^{+\infty}$ and a function $u \in W^{1,q}(\Omega)$ so that

$$\begin{cases} u_{k_j} \rightharpoonup u \text{ weakly in } L^q(\Omega), \\ Du_{k_j} \rightharpoonup Du \text{ weakly in } L^q(\Omega, \mathbb{R}^n). \end{cases} \tag{2.8}$$

{eq:2-8}

We denote (2.8) by writing

$$u_{k_j} \rightharpoonup u \text{ weakly in } W^{1,q}(\Omega). \tag{2.9}$$

{eq:2-9}

**Definition 2.4** (Weak Convergence in $W^{1,q}(\Omega)$)**.** *A sequence $\{f_k\}_{k=1}^{+\infty} \subset W^{1,q}(\Omega)$ is said to converge weakly to a function $f \in W^{1,q}(\Omega)$, written*

$$f_{k_j} \rightharpoonup f \textit{ weakly in } W^{1,q}(\Omega),$$

*if*

$$\begin{cases} f_{k_j} \rightharpoonup f \textit{ weakly in } L^q(\Omega), \\ Df_{k_j} \rightharpoonup Df \textit{ weakly in } L^q(\Omega, \mathbb{R}^n). \end{cases}$$

Furthermore, it is true that $u|_{\partial\Omega} = g$ in the trace sense, so that $u \in \mathcal{A}$. *Note that this implies that the trace operator is continuous with respect to weak convergence in $W^{1,q}(\Omega)$.*

Consequently by shifting to the weak topology on $W^{1,q}(\Omega)$ we have actually recovered enough compactness from the coercivity inequality (2.5) to deduce weak convergence (2.5) in $W^{1,q}(\Omega)$ up to a subsequence for an appropriate minimizing sequence $\{u_k\}_{k=1}^{+\infty} \subseteq \mathcal{A}$. But now another difficulty arises, for without any a priori knowledge of the functional $I[\cdot]$, this weak convergence is not enough to pass to the limit in (2.7). That is, in essentially all cases of interest, the functional $I[\cdot]$ *is not continuous with respect to weak convergence in* $W^{1,q}(\Omega)$. In other words, we *cannot* deduce from (2.9) and (2.7) that

$$I[u] = \lim_{j\to+\infty} I[u_{k_j}] = \inf_{w\in\mathcal{A}} I[w], \tag{2.10}$$ {eq:2-10}

and thus claim that $u$ is a minimizer of $I[\cdot]$. The problem is that the functional $I[\cdot]$ *is* continuous with respect to *pointwise convergence,* and $Du_{k_j} \rightharpoonup Du$ does *not* imply $Du_{k_j} \to Du$ $\mathcal{L}^n$–a.e., even up to a subsequence. It is quite possible for instance that the (weak) gradients $Du_{k_j}$, although bounded in $L^q(\Omega)$, may oscillate more and more wildly as $k_j \to +\infty$.

However, we do not really need the full strength of (2.10). In fact, it is enough to know only that

$$I[u] \le \liminf_{j\to+\infty} I[u_{k_j}]. \tag{2.11}$$ {eq:2-11}

Then from (2.6) and (2.7) it would follow

$$I[u] \le \liminf_{j\to+\infty} I[u_{k_j}] = \lim_{k\to+\infty} I[u_k] = \inf_{w\in\mathcal{A}} I[w].$$

But since the RHS is an infimum, and $u \in \mathcal{A}$,

$$\inf_{w\in\mathcal{A}} I[w] \le I[u].$$

Consequently $I[u] = \inf_{w\in\mathcal{A}} I[w]$ and $u$ is indeed a minimizer of $I[\cdot]$.

**Definition 2.5** (Sequential Weak Lower Semicontinuity in $W^{1,q}(\Omega)$)**.** *We say that a functional* $I[\cdot]$ *is* **sequentially weakly lower semicontinuous** *on* $W^{1,q}(\Omega)$ *provided that*

$$I[u] \le \liminf_{k\to+\infty} I[u_k]$$

*whenever*

$$u_k \rightharpoonup u \quad \textit{weakly in } W^{1,q}(\Omega).$$

Our goal is to now identify reasonable conditions on the nonlinear term $L(x, w(x), Dw(x))$ which imply that the functional $I[\cdot]$ is sequentially weakly lower semicontinuous.

---

2.2. **Convexity.** Recall that from our analysis of the second variation (cf. §1.3) that we derived the inequality

$$\sum_{i,j=1}^{n} L_{p_ip_j}(x, u(x), Du(x))\xi_i\xi_j \ge 0, \quad \xi \in \mathbb{R}^n, \quad x \in \Omega$$

holding as a necessary condition whenever $u$ is a smooth minimizer. This inequality suggests that it is reasonable to assume that $L$ is convex in its third argument.

t2.1 **Theorem 2.1.** *Assume that* $L : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ *is smooth, bounded below, and in addition*

$$\textit{the mapping } p \mapsto L(x, z, p) \textit{ is convex}$$

*for each* $x \in \Omega, z \in \mathbb{R}$. *Then* $I[\cdot]$ *is sequentially weakly lower semicontinuous on* $W^{1,q}(\Omega)$.

*Proof.*

(i). Choose any sequence $\{u_k\}_{k=1}^{+\infty}$ such that

$$u_k \rightharpoonup u \quad \text{weakly in } W^{1,q}(\Omega). \tag{2.12}$$ {eq:2-12}

Note that at least one such sequence exists. Set $l := \liminf_{k\to+\infty} I[u_k]$. We must show that

$$I[u] \leq \liminf_{k\to+\infty} I[u_k] = l. \tag{2.13}$$ {eq:2-13}

(ii). Since weakly convergent sequences are bounded, it follows from (2.12) that

$$\sup_{k\in\mathbb{N}} \|u_k\|_{W^{1,q}(\Omega)} < +\infty. \tag{2.14}$$ {eq:2-14}

Passing to a subsequence if necessary, we may as well also suppose that

$$\lim_{k\to+\infty} I[u_k] = l. \tag{2.15}$$ {eq:2-15}

Furthermore, since $W^{1,q}(\Omega) \subset\subset L^q(\Omega)$, we have from (2.14) that $u_k \to u$ strongly in $L^q(\Omega)$ up to a further subsequence. Thus, passing if necessary up to yet another subsequence,

$$u_k \to u \quad \mathcal{L}^n - \text{a.e. on } \Omega. \tag{2.16}$$ {eq:2-16}

(iii). Fix $\epsilon > 0$. Since $u_k \to u$ $\mathcal{L}^n-$a.e. on $\Omega$ (cf. (2.16)), Egorov's Theorem implies that there exists an $\mathcal{L}^n-$measurable set $E_\epsilon$ such that

$$u_k \to u \quad \text{uniformly on } E_\epsilon, \tag{2.17}$$ {eq:2-17}

with

$$\mathcal{L}^n(\Omega \setminus E_\epsilon) \leq \epsilon. \tag{2.18}$$ {eq:2-18}

We may assume that $E_\epsilon \subseteq E_{\epsilon'}$ for $0 < \epsilon' < \epsilon$. Now define

$$F_\epsilon := \left\{ x \in \Omega : |u(x)| + |Du(x)| \leq \frac{1}{\epsilon} \right\}. \tag{2.19}$$ {eq:2-19}

Then

$$\mathcal{L}^n(\Omega \setminus F_\epsilon) \to 0 \quad \text{as } \epsilon \to 0, \tag{2.20}$$ {eq:2-20}

since $u \in W^{1,q}(\Omega)$ implies that $|u(x)|, |Du(x)| < +\infty$ $\mathcal{L}^n-$a.e. on $\Omega$. We finally set

$$G_\epsilon := E_\epsilon \cap F_\epsilon, \tag{2.21}$$ {eq:2-21}

and notice from (2.18) and (2.20) that $\mathcal{L}^n(\Omega \setminus G_\epsilon) \to 0$ as $\epsilon \to 0$, for

$$\mathcal{L}^n(\Omega \setminus G_\epsilon) = \mathcal{L}^n((\Omega \setminus E_\epsilon) \cup (\Omega \setminus F_\epsilon)) \leq \mathcal{L}^n(\Omega \setminus E_\epsilon) + \mathcal{L}^n(\Omega \setminus F_\epsilon).$$

(iv). Now let us observe that since $L$ is bounded below, we may as well assume that

$$L \geq 0, \tag{2.22}$$ {eq:2-22}

for otherwise we could apply the following arguments to $\widetilde{L} := L + \beta \geq 0$ for some appropriate constant $\beta > 0$. Consequently

$$I[u_k] = \int_\Omega L(x, u_k, Du_k)\, dx \geq \int_{G_\epsilon} L(x, u_k, Du_k)\, dx.$$

Since $L$ is smooth and convex in the third argument, notice that

$$L(x, u_k, Du_k) \geq L(x, u_k, Du) + D_pL(x, u_k, Du) \cdot (Du - Du_k).$$

Thus

$$I[u_k] \geq \int_{G_\epsilon} L(x, u_k, Du)\ dx + \int_{G_\epsilon} D_pL(x, u_k, Du) \cdot (Du - Du_k)\ dx. \tag{2.23}$$

{eq:2-23}

By (2.17), (2.19), (2.21), and the continuity of $L$, we make the following two observations:

(i) $L(x, u_k, Du) \to L(x, u, Du)$ uniformly on $G_\epsilon$;
(ii) $L(x, u, Du) \in L^q(\Omega)$.

Thus by Lebesgue's Dominated Convergence Theorem,

$$\lim_{k\to+\infty} \int_{G_\epsilon} L(x, u_k, Du)\ dx = \int_{G_\epsilon} L(x, u, Du)\ dx. \tag{2.24}$$

{eq:2-24}

For the second term on the RHS of (2.23), note again by (2.17), (2.19), (2.21), the smoothness of $L$, and the boundedness of $D_pL$ and $\Omega$ that

(i) $D_pL(x, u_k, Du) \to L(x, u, Du)$ uniformly on $G_\epsilon$;
(ii) $D_pL(x, u, Du) \in L^p(\Omega)$.

Appealing once again to Lebesgue's Dominated Convergence Theorem,

$$\begin{aligned} \lim_{k\to+\infty} \int_{G_\epsilon} D_pL(x, u_k, Du) \cdot (Du - Du_k)\ dx &= \int_{G_\epsilon} D_pL(x, u, Du) \cdot \left\{ \lim_{k\to+\infty} Du - Du_k \right\}\ dx \\ &= 0, \end{aligned} \tag{2.25}$$

{eq:2-25}

where the last inequality follows from the weak convergence $Du_k \rightharpoonup Du$ in $L^q(\Omega)$. Thus by (2.24) and (2.25), we have from (2.23) that

$$l = \lim_{k\to+\infty} I[u_k] \geq \int_{G_\epsilon} L(x, u, Du)\ dx.$$

Note that this inequality holds for any $\epsilon > 0$. Thus taking the limit as $\epsilon \to 0$ and recalling the assumption (2.22), it follows by the Monotone Convergence Theorem that

$$\liminf_{k\to+\infty} I[u_k] \geq \lim_{\epsilon\to 0} \int_{G_\epsilon} L(x, u, Du)\ dx = \int_{\Omega} L(x, u, Du)\ dx = I[u],$$

as required. The proof is complete. □

**Remark.** *It is important to understand how the above proof deals with the weak convergence $Du_k \rightharpoonup Du$. The key is the convexity inequality (2.23), on the RHS of which $Du_k - Du$ appears linearly. Weak convergence is compatible with linear expressions, and so the limit (2.25) holds. Remember that it is not in general true that if $Du_k \rightharpoonup Du$, then $Du_k \to Du$ a.e., even upon passing to a subsequence.*

*The convergence of $u_k$ to $u$ in $L^q(\Omega)$ is much stronger, and lets us indeed conclude that $u_k \to u$ $\mathcal{L}^n$−a.e. Thus we do not need any convexity assumption concerning the mapping $z \mapsto L(x, z, p)$.*

We can now establish that the functional $I[\cdot]$ has a minimizer among the functions in the admissible set $\mathcal{A}$. We first recall another important theorem from functional analysis.

**Theorem** (Mazur's Theorem)**.** *A convex subspace $K$ of a normed linear space $X$ is closed if and only if $K$ is weakly sequentially closed.*

We now state and prove the main existence theorem.

t2.2 **Theorem 2.2** (Existence of Minimizer)**.** *Assume that $L : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n$ is convex in the third argument and satisfies the coercivity inequality (cf. (2.4))*

$$L(x, z, p) \geq \alpha |p|^q - \beta$$

*for constants $\alpha > 00$ and $\beta \geq 0$. Suppose also that the admissible set $\mathcal{A}$ is nonempty.*

*Then there exists at least one function $u \in \mathcal{A}$ such that*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

*Proof.*

(i). Set $m := \inf_{w \in \mathcal{A}} I[w]$. If $m = +\infty$, we are done, and thus we assume that $m$ is finite. Choose a minimizing sequence $\{u_k\}_{k=1}^{+\infty} \subseteq \mathcal{A}$. Then

$$I[u_k] \to m \quad \text{as } k \to +\infty. \tag{2.26}$$ {eq:2-26}

(ii). We may assume that $\beta = 0$ in the coercivity inequality (2.4), for otherwise we could just as well consider $\widetilde{L} := L + \beta$. Thus $L(x, z, p) \geq \alpha |p|^q$, and so

$$I[w] = \int_\Omega L(x, w(x), Dw(x))\, dx \geq \alpha \int_\Omega |Dw(x)|^q\, dx. \tag{2.27}$$ {eq:2-27}

Since $m$ is finite, we conclude from (2.26) and (2.27) that

$$\sup_{k \in \mathbb{N}} \|Du_k\|_{L^q(\Omega)} < +\infty. \tag{2.28}$$ {eq:2-28}

Indeed, there exists $k_0 \in \mathbb{N}$ so large that for all $k \geq k_0$,

$$I[u_k] \leq m + 1,$$

so that

$$\|Du_k\|_{L^q(\Omega)}^q \leq \frac{m+1}{\alpha}$$

whenever $k \geq k_0$. Thus, for all $k \in \mathbb{N}$,

$$\|Du_k\|_{L^q(\Omega)}^q \leq M + \frac{m+1}{\alpha},$$

where $M := \max\{\|Du_1\|_{L^q(\Omega)}^q, \ldots, \|Du_{k_0-1}\|_{L^q(\Omega)^q}\}$.

(iii). Now fix any function $w \in \mathcal{A}$. Since $u_k \in \mathcal{A}$ for all $k \in \mathbb{N}$, $u_k$ and $w$ both equal $g$ on $\partial\Omega$ in the trace sense, and so we have $u_k - w \in W_0^{1,q}(\Omega)$. Therefore an application of Poincaré's inequality gives

$$\begin{aligned}
\|u_k\|_{L^q(\Omega)} &\leq \|u_k - w\|_{L^q(\Omega)} + \|w\|_{L^q(\Omega)} \\
&\leq C_1 \|Du_k - Dw\|_{L^q(\Omega)} + \|w\|_{L^q(\Omega)} \\
&\leq C_1 \|Du_k\|_{L^q(\Omega)} + C_1 \|Dw\|_{L^q(\Omega)} + \|w\|_{L^q(\Omega)} \\
&\leq C_1 \|Du_k\|_{L^q(\Omega)} + (1 + C_1) \|w\|_{W^{1,q}(\Omega)} \\
&\leq C,
\end{aligned}$$

by (2.28). Hence, $\sup_{k \in \mathbb{N}} \|u_k\|_{L^q(\Omega)} < +\infty$. This and (2.28) imply that the minimizing sequence $\{u_k\}_{k=1}^{+\infty}$ is bounded in $W^{1,q}(\Omega)$.

(iv). Since $W^{1,q}(\Omega)$ is a reflexive Banach space, the Eberlein–Smulyan Theorem implies that there exist a subsequence $\{u_{k_j}\}_{j=1}^{+\infty} \subseteq \{u_k\}_{k=1}^{+\infty}$ and a function $u \in W^{1,q}(\Omega)$ such that

$$u_{k_j} \rightharpoonup u \quad \text{weakly in } W^{1,q}(\Omega).$$

We verify next that $u \in \mathcal{A}$. To see this, note that for $w \in \mathcal{A}$ as above, $u_k - w \in W_0^{1,q}(\Omega)$ for all $k \in \mathbb{N}$. Since $W_0^{1,q}(\Omega)$ is a closed, convex subspace of $W^{1,q}(\Omega)$, it follows by Mazur's Theorem that $W_0^{1,q}(\Omega)$ is weakly closed. Thus $u - w \in W_0^{1,q}(\Omega)$. Consequently $u|\,|_{\partial\Omega} = g$ in the trace sense, and $u \in \mathcal{A}$.

Finally, recall from Theorem (2.1) that since $L$ is convex in the third argument, $I[\cdot]$ is sequentially weakly lower semicontinuous on $W^{1,q}(\Omega)$. Therefore

$$I[u] \le \liminf_{j\to+\infty} I[u_{k_j}] = \inf_{w\in\mathcal{A}} I[w].$$

But on the other hand, since $u \in \mathcal{A}$, it follows

$$\inf_{w\in\mathcal{A}} I[w] \le I[u],$$

and therefore

$$I[u] = \min_{w\in\mathcal{A}} I[w].$$

The proof is complete. □

We now turn to the problem of uniqueness. In general there can be many minimizers, and so if we want uniqueness of a minimizer, we require further assumptions. For instance, suppose that

$$L(x,z,p) = L(x,p) \text{ does not depend on } z \tag{2.29}$$ {eq:2-29}

and also that there exists $\theta > 0$ such that

$$\sum_{i,j=1}^{n} L_{p_i p_j}(x,p)\xi_i\xi_j \ge \theta|\xi|^2, \quad x \in \Omega, \quad p,\xi \in \mathbb{R}^n. \tag{2.30}$$ {eq:2-30}

If $L$ satisfies (2.30), we say that $L$ is *uniformly convex*.

**Definition 2.6** (Uniformly Convex)**.** *If there exists a constant $\theta > 0$ such that $L : \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ satisfies*

$$\sum_{i,j=1}^{n} L_{p_i p_j}(x,z,p)\xi_i\xi_j \ge \theta|\xi|^2, \quad x \in \Omega, \quad z \in \mathbb{R}, \quad p,\xi \in \mathbb{R}^n,$$

*(cf. (2.30)), we say that the mapping $p \mapsto L(x,z,p)$ is* **uniformly convex** *for each $x \in \Omega$ and $z \in \mathbb{R}$.*

t2.3 **Theorem 2.3** (Uniqueness of Minimizer)**.** *Suppose that*

$$L(x,z,p) = L(x,p) \text{ does not depend on } z$$

*and that the mapping $p \mapsto L(x,p)$ is uniformly convex for each $x \in \Omega$ (cf. (2.29), (2.30)). Then a minimizer $u \in \mathcal{A}$ of $I[\cdot]$ is unique.*

*Proof.*

(i). Assume that $u, \widetilde{u} \in \mathcal{A}$ are both minimizers of $I[\cdot]$ over $\mathcal{A}$. Then $v := \frac{1}{2}(u + \widetilde{u}) \in \mathcal{A}$. We claim that

$$I[v] \le \frac{I[u] + I[\widetilde{u}]}{2} \tag{2.31}$$ {eq:2-31}

with a strict inequality, unless $u = \widetilde{u}$ a.e.

(ii). To see that (2.31) holds, note that by the uniform convexity assumption (cf. (2.30)) that we have

$$L(x,p) \geq L(x,r) + D_pL(x,r)\cdot(p-r) + \frac{\theta}{2}|p-r|^2, \quad x \in \Omega, \quad p,r \in \mathbb{R}^n. \tag{2.32}$$

{eq:2-32}

Setting $r := \frac{Du+D\widetilde{u}}{2}$, $p := Du$, and integrating over $\Omega$, we obtain

$$\begin{aligned} I[u] = \int_\Omega L(x,Du)\,dx &\geq \int_\Omega L\left(x, \frac{Du+D\widetilde{u}}{2}\right)\,dx\, + \\ &\int_\Omega D_pL\left(x, \frac{Du+D\widetilde{u}}{2}\right)\cdot\left(\frac{Du-D\widetilde{u}}{2}\right)\,dx + \frac{\theta}{8}\int_\Omega |Du-D\widetilde{u}|^2\,dx \\ &= I[v] + \int_\Omega D_pL\left(x, \frac{Du+D\widetilde{u}}{2}\right)\cdot\left(\frac{Du-D\widetilde{u}}{2}\right)\,dx + \frac{\theta}{8}\int_\Omega |Du-D\widetilde{u}|^2\,dx. \end{aligned} \tag{2.33}$$

{eq:2-33}

Similarly, setting $r := \frac{Du+D\widetilde{u}}{2}$, $p := D\widetilde{u}$, and integrating over $\Omega$ gives

$$\begin{aligned} I[\widetilde{u}] = \int_\Omega L(x,D\widetilde{u})\,dx &\geq \int_\Omega L\left(x, \frac{Du+D\widetilde{u}}{2}\right)\,dx\, + \\ &\int_\Omega D_pL\left(x, \frac{Du+D\widetilde{u}}{2}\right)\cdot\left(\frac{D\widetilde{u}-Du}{2}\right)\,dx + \frac{\theta}{8}\int_\Omega |D\widetilde{u}-Du|^2\,dx \\ &= I[v] + \int_\Omega D_pL\left(x, \frac{Du+D\widetilde{u}}{2}\right)\cdot\left(\frac{D\widetilde{u}-Du}{2}\right)\,dx + \frac{\theta}{8}\int_\Omega |D\widetilde{u}-Du|^2\,dx. \end{aligned} \tag{2.34}$$

{eq:2-34}

Adding (2.33) and (2.34) and dividing by 2, we deduce

$$\frac{I[u]+I[\widetilde{u}]}{2} \geq I[v] + \frac{\theta}{8}\int_\Omega |Du-D\widetilde{u}|^2\,dx \geq I[v],$$

with a strict inequality unless $Du = D\widetilde{u}$ a.e. This proves the claim in (2.31).

(iii). Since

$$I[u] = I[\widetilde{u}] = \min_{w\in\mathcal{A}} I[w] \leq I[v],$$

we deduce that

$$\frac{I[u]+I[\widetilde{u}]}{2} = I[v],$$

and thus $Du = D\widetilde{u}$ a.e. on $\Omega$. Since $u, \widetilde{u} \in \mathcal{A}$, $u|_{\partial\Omega} = \widetilde{u}|_{\partial\Omega} = g$ in the trace sense, and thus $u = \widetilde{u}$ a.e. on $\Omega$. The proof is complete. □

---

2.3. **Weak Solutions of Euler–Lagrange Equation.** We want to demonstrate that any minimizer $u \in \mathcal{A}$ of the functional $I[\cdot]$ solves the Euler–Lagrange equation in some suitable sense. Notice that this does not follow from the calculations in §8.1 since we do not know that $u$ is smooth, only that $u \in W^{1,q}(\Omega)$. To do this, we will need some growth conditions on $L$ and its derivatives. We suppose that

$$|L(x,z,p)| \leq C(1+|z|^q+|p|^q) \tag{2.35}$$

{eq:2-35}

and also

$$\begin{cases} |D_pL(x,z,p)| \leq C(1+|z|^{q-1}+|p|^{q-1}), \\ |D_zL(x,z,p)| \leq C(1+|z|^{q-1}+|p|^{q-1}) \end{cases} \tag{2.36}$$

{eq:2-36}

for some constant $C > 0$ and all $x \in \Omega$, $z \in \mathbb{R}$, and $p \in \mathbb{R}^n$.

**Motivation for definition of weak solution.** We recall the definition of the boundary value problem for the Euler–Lagrange PDE associated with the functional $I[\cdot]$, which for a smooth minimizer $u$ is as follows:

$$\begin{cases} L_z(x,u,Du) - \sum_{i=1}^{n} (L_{p_i}(x,u,Du))_{x_i} = 0 \quad \text{on } \Omega, \\ u|_{\partial\Omega} = g. \end{cases} \tag{2.37}$$

{eq:2-37}

Multiplying (2.37) by a test function $\phi \in \mathcal{C}_c^\infty(\Omega)$ and integrating by parts, we obtain the equality

$$\begin{aligned} 0 &= \int_\Omega L_z(x,u,Du)\phi - \sum_{i=1}^{n} (L_{p_i}(x,u,Du))_{x_i}\phi \; dx \\ &= \int_\Omega L_z(x,u,Du)\phi + \sum_{i=1}^{n} L_{p_i}(x,u,Du)\phi_{x_i} \; dx. \end{aligned} \tag{2.38}$$

{eq:2-38}

Now assume that $u \in W^{1,q}(\Omega)$. Then by (2.36) we see that

$$|D_p L(x,u,Du)| \le C(|1 + |u|^{q-1} + |Du|^{q-1}) \in L^{q'}(\Omega),$$

where $q' := \frac{q}{q-1}$, $\frac{1}{q} + \frac{1}{q'} = 1$. Indeed, we observe

$$\begin{aligned} \|D_p L(x,u,Du)\|_{L^{q'}(\Omega)}^{q'} &\le C^{q'} \int_\Omega \left(1 + |u|^{q-1} + |Du|^{q-1}\right)^{q'} \; dx \\ &\le 3^{q'} C^{q'} \int_\Omega \max\{1, |u|^q, |Du|^q\} \; dx \\ &< +\infty. \end{aligned}$$

Similarly by (2.36)

$$|D_z L(x,u,Du)| \le C(1 + |u|^{q-1} + |Du|^{q-1}) \in L^{q'}(\Omega). \tag{2.39}$$

{eq:2-39}

Consequently using a standard approximation argument, we see that the equality (2.38) is valid for any $\phi \in W_0^{1,q}(\Omega)$. this motivates the following.

**Definition 2.7** (Weak Solution of Euler–Lagrange Equation)**.** *We say that $u \in \mathcal{A}$ is a weak solution of the boundary–value problem (2.37) for the Euler–Lagrange equation provided that*

$$\int_\Omega L_z(x,u,Du) + \sum_{i=1}^{n} L_{p_i}(x,u,Du)\phi_{x_i} \; dx = 0$$

*for all $\phi \in W_0^{1,q}(\Omega)$.*

t2.4 **Theorem 2.4** (Solution of Euler–Lagrange Equation)**.** *Assume that $L(x,z,p)$ satisfies the growth conditions given by (2.35) and (2.36), and $u \in \mathcal{A}$ satisfies*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

*Then $u$ is a weak solution of the Euler–Lagrange equation (2.37).*

*Proof.* We proceed as we did previously, taking care about differentiating inside the integrals. Fix any $\phi \in W_0^{1,q}(\Omega)$ and set

$$i(\tau) := I[u + t\phi], \quad \tau \in \mathbb{R}.$$

By (2.35), we see that $i(\tau)$ is finite for all $\tau \in \mathbb{R}$.

Now let $\tau \neq 0$ and write the difference quotient

$$\frac{i(\tau) - i(0)}{\tau} = \int_\Omega \frac{L(x, u + \tau\phi, Du + \tau D\phi) - D(x, u, Du)}{\tau}\, dx =: \int_\Omega L^\tau(x)\, dx, \tag{2.40}$$

{eq:2-40}

where

$$L^\tau(x) := \frac{L(x, u(x) + \tau\phi(x), Du(x) + \tau D\phi(x)) - L(x, u(x), Du(x))}{\tau}$$

for $\mathcal{L}^n$–a.e. $x \in \Omega$. Clearly

$$L^\tau(x) \to L_z(x, u, Du)\phi + \sum_{i=1}^n L_{p_i}(x, u, Du)\phi_{x_i} \tag{2.41}$$

{eq:2-41}

a.e. on $\Omega$ as $\tau \to 0$, for

$$\begin{aligned}\lim_{\tau\to 0} L^\tau(x) = \lim_{\tau\to 0} &\left\{ \frac{L(x, u(x) + \tau\phi(x), Du(x) + \tau D\phi(x)) - L(x, u(x), Du(x) + \tau D\phi(x))}{\tau} \right\} + \\ &\lim_{\tau\to 0} \left\{ \frac{L(x, u(x), Du(x) + \tau D\phi(x)) - L(x, u(x), Du(x))}{\tau} \right\}.\end{aligned}$$

Furthermore

$$\begin{aligned} L^\tau(x) &= \frac{1}{\tau}\left( L(x, u(x) + \tau\phi(x), Du(x) + \tau D\phi(x)) - L(x, u(x), Du(x)) \right) \\ &= \frac{1}{\tau} \int_0^\tau \frac{d}{ds} L(x, u + s\phi, Du + sD\phi)\, ds \\ &= \frac{1}{\tau} \int_0^\tau L_z(x, u + s\phi, Du + sD\phi)\phi + \sum_{i=1}^n L_{p_i}(x, u + s\phi, Du + sD\phi)\phi_{x_i}\, ds. \end{aligned}$$

Next we observe that, by Young's inequality and the growth conditions given by (2.36), for $q' := \frac{q}{q-1}$ we have for any $M \geq |\tau|$

$$\begin{aligned} |L_z(x, u + s\phi, Du + sD\phi)\phi| &\leq \frac{|L_z(x, u + s\phi, Du + sD\phi)|^{q'}}{q'} + \frac{|\phi|^q}{q} \\ &\leq C \left(1 + |u + s\phi|^{q-1} + |Du + sD\phi|^{q-1}\right)^{q'} + \frac{|\phi|^q}{q} \\ &\leq C \left(|\phi|^q + \max\{1, |u + M\phi|^q, |Du + MD\phi|^q\}\right) \\ &\leq C(1 + |u|^q + |\phi|^q + |Du|^q + |D\phi|^q), \end{aligned}$$

and similarly

$$\sum_{i=1}^n |L_{p_i}(x, u + s\phi, Du + sD\phi)\phi_{x_i}| \leq C(1 + |u|^q + |\phi|^q + |Du|^q + |D\phi|^q).$$

Hence

$$|L^\tau(x)| \leq \frac{C}{\tau} \int_0^\tau (1 + |u|^q + |\phi|^q + |Du|^q + |D\phi|^q)\, ds$$

$$= C(1 + |u|^q + |\phi|^q + |Du|^q + |D\phi|^q),$$

so evidently $L^\tau \in L^1(\Omega)$. Consequently we may apply Lebesgue's Dominated Convergence Theorem to conclude from (2.40) and (2.41) that $i'(0)$ exists, and

$$\begin{aligned} i'(0) &= \lim_{\tau\to 0} \frac{i(\tau) - i(0)}{\tau} = \lim_{\tau\to 0} \int_\Omega L^\tau(x)\, dx = \int_\Omega \lim_{\tau\to 0} L^\tau(x)\, dx \\ &= \int_\Omega L_z(x, u, Du)\phi + \sum_{i=1}^n L_{p_i}(x, u, Du)\phi_{x_i}\, dx. \end{aligned}$$

But then since $i(\cdot)$ has a minimum for $\tau = 0$, we have that $i'(0) = 0$, and thus

$$\int_\Omega L_z(x, u, Du)\phi + \sum_{i=1}^n L_{p_i}(x, u, Du)\phi_{x_i}\, dx = 0,$$

so that $u$ is a weak solution of the Euler–Lagrange equation, as required. The proof is complete. □

**Remark.** *In general, the Euler–Lagrange equation (2.37) will have other solutions which do not correspond to minima of $I[\cdot]$; see §2.5, for instance. However, in the special case that the mapping $(z, p) \mapsto L(x, z, p)$ is convex for any $x \in \Omega$, then each weak solution is in fact a minimizer.*

To see this, suppose that $u \in \mathcal{A}$ solves the Euler–Lagrange PDE

$$\begin{cases} L_z(x, u, Du) - \sum\limits_{i=1}^n (L_{p_i}(x, u, Du))_{x_i} = 0 \quad \text{on } \Omega, \\ u|_{\partial\Omega} = g \end{cases} \tag{2.42}$$

{eq:2-42}

in the weak sense and select any $w \in \mathcal{A}$. Using the convexity of the mapping $(z, p) \mapsto L(x, z, p)$, we have

$$L(x, w, r) \geq L(x, z, p) + D_z L(x, z, p) \cdot (w - z) + D_p L(x, z, p) \cdot (r - p).$$

Let $p = Du(x)$, $r = Dw(x)$, $z = u(x)$, $w = w(x)$, and integrate over $\Omega$ to find

$$I[w] \geq I[u] + \int_\Omega D_z L(x, u, Du) \cdot (w - u) + D_p L(x, u, Du) \cdot (Dw - Du)\, dx.$$

In view of (2.42) the second term on the RHS is zero, and therefore $I[u] \leq I[w]$ for each $w \in \mathcal{A}$.

---

## 2.4. **Systems.**

2.4.1. *Convexity.* We adopt all the notation from §1.4 And consider the existence of minimizers of the functional

$$I[\boldsymbol{w}] := \int_\Omega L(x, \boldsymbol{w}(x), D\boldsymbol{w}(x))\, dx,$$

defined for appropriate functions $\boldsymbol{w} : \Omega \to \mathbb{R}^m$, where $L : \overline{\Omega} \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$ is given.

The theory from §2.2 extends easily to the case of systems. We assume the coercivity inequality

$$L(x, z, P) \geq \alpha|P|^q - \beta, \quad P \in \mathbb{R}^{m\times n}, \quad z \in \mathbb{R}^m, \quad x \in \Omega \tag{2.43}$$

{eq:2-43}

for some constants $\alpha > 0$, $\beta \geq 0$. Define also the admissible set

$$\mathcal{A} := \{\boldsymbol{w} \in W^{1,q}(\Omega, \mathbb{R}^m) : \boldsymbol{w}|_{\partial\Omega} = \boldsymbol{g}\},$$

where $\boldsymbol{w}|_{\partial\Omega} = \boldsymbol{g}$ is understood in the trace sense, the function $\boldsymbol{g} : \partial\Omega \to \mathbb{R}^m$ being given.

We present most theorems here without proof, as the proofs follow almost exactly to the analogous theorems in the previous section.

t2.5 **Theorem 2.5** (Existence of Minimizer)**.** *Assume that $L$ satisfies the coercivity inequality (cf. (2.43))*

$$L(x, z, P) \geq \alpha |P|^q - \beta, \quad P \in \mathbb{R}^{m \times n}, \quad z \in \mathbb{R}^m, \quad x \in \Omega$$

*for some constants $\alpha > 0$, $\beta \geq 0$, and that $L$ is convex in the third argument. Suppose also that the admissible set $\mathcal{A}$ is nonempty.*

*Then there exists $\boldsymbol{u} \in \mathcal{A}$ satisfying*

$$I[\boldsymbol{u}] = \min_{\boldsymbol{w} \in \mathcal{A}} I[\boldsymbol{w}].$$

We also have uniqueness.

t2.6 **Theorem 2.6** (Existence of Minimizer)**.** *Assume that $L(x, z, P) = L(x, P)$ does not depend on $z$ and that the mapping $P \mapsto L(x, P)$ is uniformly convex. Then a minimizer $\boldsymbol{u} \in \mathcal{A}$ of $I[\cdot]$ is unique.*

Let us now also suppose the growth conditions

$$\begin{cases} |L(x, z, P)| \leq C(1 + |z|^q + |P|^q), \\ |D_z L(x, z, P)| \leq C(1 + |z|^{q-1} + |P|^{q-1}), \\ |D_p L(x, z, P)| \leq C(1 + |z|^{q-1} + |P|^{q-1}) \end{cases} \tag{2.44}$$ {eq:2-44}

for some constant $C > 0$ and holding for all $x \in \Omega$, $z \in \mathbb{R}^m$, and $P \in \mathbb{R}^{m \times n}$.

We consider the system of Euler–Lagrange equations

$$\begin{cases} L_{z^k}(x, \boldsymbol{u}, D\boldsymbol{u}) - \sum_{i=1}^{n} (L_{p_i^k}(x, \boldsymbol{u}, D\boldsymbol{u}))_{x_i} = 0 \quad \text{on } \Omega, \\ u^k|_{\partial\Omega} = g^k, \end{cases} \tag{2.45}$$ {eq:2-45}

for $k = 1, \ldots, m$. We call $\boldsymbol{u} \in \mathcal{A}$ a weak solution of the system of Euler–Lagrange equations (2.45) provided that

$$\sum_{k=1}^{m} \int_\Omega L_{z^k}(x, \boldsymbol{u}, D\boldsymbol{u}) \phi^k + \sum_{i=1}^{n} L_{p_i^k}(x, \boldsymbol{u}, D\boldsymbol{u}) \phi^k_{x_i} \, dx = 0$$

for all $\boldsymbol{\phi} \in W_0^{1,q}(\Omega, \mathbb{R}^m)$, $\boldsymbol{\phi} = (\phi^1, \ldots, \phi^m)$.

t2.7 **Theorem 2.7** (Solution of Euler–Lagrange System)**.** *Assume that $L(x, z, P)$ satisfies the growth conditions given by (2.44) and that $\boldsymbol{u} \in \mathcal{A}$ satisfies*

$$I[\boldsymbol{u}] = \min_{\boldsymbol{w} \in \mathcal{A}} I[\boldsymbol{w}].$$

*Then $\boldsymbol{u}$ is a weak solution of the system of Euler–Lagrange equations (2.45).*

2.4.2. *Polyconvexity.* There are some mathematically and physically interesting systems that are not covered by Theorem (2.5) but can nonetheless still be studied using the calculus of variations. These include certain problems where the Lagrangian $L$ is *not* convex in the third argument, but $I[\cdot]$ is still sequentially weakly lower semicontinuous.

l2.1 **Lemma 2.1** (Weak Continuity of Determinants)**.** *Assume that* $n < q < +\infty$ *and*

$$\boldsymbol{u}_k \rightharpoonup \boldsymbol{u} \quad \textit{weakly in } W^{1,q}(\Omega, \mathbb{R}^n).$$

*Then*

$$\det D\boldsymbol{u}_k \rightharpoonup \det D\boldsymbol{u} \quad \textit{weakly in } L^{\frac{q}{n}}(\Omega).$$

*Proof.*

(i). We first recall the matrix identity $(\det P)I = P(\operatorname{cof} P)^T$. Consequently

$$\det P = \sum_{j=1}^{n} p^i_j (\operatorname{cof} P)^i_j, \quad i = 1, \ldots, n.$$

(ii). Now let $\boldsymbol{w} \in \mathcal{C}^\infty(\Omega, \mathbb{R}^n)$, $\boldsymbol{w} = (w^1, \ldots, w^n)$. Then by (i)

$$\det D\boldsymbol{w} = \sum_{j=1}^{n} w^i_{x_j} (\operatorname{cof} D\boldsymbol{w})^i_j, \quad i = 1, \ldots, n. \tag{2.46}$$ {eq:2-46}

Now recall by Lemma (1.1) that $\sum_{j=1}^n (\operatorname{cof} D\boldsymbol{w})^i_{j,x_j} = 0$. Thus by (2.46)

$$\det D\boldsymbol{w} = \sum_{j=1}^{n} \left( w^i (\operatorname{cof} D\boldsymbol{w})^i_j \right)_{x_j}, \quad i = 1, \ldots, n.$$

Consequently the determinant of the gradient matrix of $\boldsymbol{w}$ may be expressed as a divergence. Thus if $\phi \in \mathcal{C}^\infty_c(\Omega)$, we have upon integrating by parts

$$\begin{aligned} \int_\Omega (\det D\boldsymbol{w}) \phi \, dx &= \int_\Omega \sum_{j=1}^{n} \left( w^i (\operatorname{cof} D\boldsymbol{w})^i_j \right)_{x_j} \phi \, dx \\ &= -\sum_{j=1}^{n} \int_\Omega w^i(x) (\operatorname{cof} D\boldsymbol{w})^i_j \phi_{x_j}(x) \, dx, \quad i = 1, \ldots, n. \end{aligned} \tag{2.47}$$ {eq:2-47}

(iii). We have established the identity (2.47) for a smooth function $w \in \mathcal{C}^\infty(\Omega)$. A standard approximation argument thus gives

$$\int_\Omega (\det D\boldsymbol{u}_k) \phi \, dx = -\sum_{j=1}^{n} \int_\Omega u^i_k(x) (\operatorname{cof} D\boldsymbol{u}_k)^i_j \phi_{x_j}(x) \, dx, \quad i = 1, \ldots, n, \tag{2.48}$$ {eq:2-48}

for our sequence $\{\boldsymbol{u}_k\}_{k=1}^{+\infty} \subset W^{1,q}(\Omega, \mathbb{R}^n)$. Now recall from the assumption that $\boldsymbol{u}_k \rightharpoonup \boldsymbol{u}$ in $W^{1,q}(\Omega, \mathbb{R}^n)$, and thus the sequence $\{\boldsymbol{u}_k\}_{k=1}^{+\infty}$ is bounded in $W^{1,q}(\Omega, \mathbb{R}^n)$. Since also $n < q < +\infty$, we have by Morrey's inequality that $\{\boldsymbol{u}_k\}_{k=1}^{+\infty}$ is bounded in $\mathcal{C}^{0,1-n/q}(\Omega, \mathbb{R}^n)$. But since $\mathcal{C}^{0,\alpha}(\Omega, \mathbb{R}^n) \subset\subset \mathcal{C}(\Omega, \mathbb{R}^n)$ by the Arzela–Ascoli Theorem for any $0 < \alpha \le 1$, it follows

$$\boldsymbol{u}_k \to \boldsymbol{u} \quad \text{uniformly in } \Omega,$$

up to a subsequence. Returning to the identity (2.48), we see that we could conclude that

$$\lim_{k \to +\infty} \int_\Omega (\det D\boldsymbol{u}_k) \phi \, dx = -\sum_{j=1}^{n} \int_\Omega u^i(x) (\operatorname{cof} D\boldsymbol{u})^i_j \phi_{x_j}(x) \, dx = \int_\Omega (\det D\boldsymbol{u}) \phi \, dx \tag{2.49}$$ {eq:2-49}

if we knew that

$$\lim_{k\to+\infty}\int_\Omega (\operatorname{cof} D\boldsymbol{u}_k)^i_j\psi\ dx = \int_\Omega (\operatorname{cof} D\boldsymbol{u})^i_j\psi\ dx, \tag{2.50}$$ {eq:2-50}

for each $i,j=1,\dots,n$ and each $\psi\in\mathcal{C}^\infty_c(\Omega)$. Note that $(\operatorname{cof} D\boldsymbol{u}_k)^i_j$ is the determinant of an $(n-1)\times(n-1)$ matrix, which may be analyzed as above by being written as a sum of determinants of appropriate $(n-2)\times(n-2)$ matrices, times uniformly convergent factors. We continue in this fashion and eventually must only show that the entries of the matrices $D\boldsymbol{u}_k$ converge weakly to the corresponding entries of $D\boldsymbol{u}$. But this is obvious, so that (2.50) holds, and thus (2.49) holds also.

(iv). Finally, since $\{\boldsymbol{u}_k\}_{k=1}^\infty$ is bounded in $W^{1,q}(\Omega,\mathbb{R}^n)$ and $|\det D\boldsymbol{u}_k|\le C|D\boldsymbol{u}_k|^n$, we see that $\{\det D\boldsymbol{u}_k\}_{k=1}^{+\infty}$ is bounded in $L^{\frac{q}{n}}(\Omega)$. Hence any subsequence of $\{\det D\boldsymbol{u}_k\}_{k=1}^{+\infty}$ has a weakly convergent subsequence in $L^{\frac{q}{n}}(\Omega)$, which by (2.49) can only converge to $\det D\boldsymbol{u}$. The proof is complete. $\square$

We next use this lemma to establish a sequential weak lower semicontinuity assertion analogous of Theorem (2.1), except that we will not assume that the Lagrangian $L$ is necessarily convex in $P$. Instead let us suppose that $m=n$ and $L$ has the form

$$L(x,z,P)=F(x,z,\det P,P),\quad x\in\Omega,\quad z\in\mathbb{R}^n,\quad P\in\mathbb{R}^{n\times n}, \tag{2.51}$$ {eq:2-51}

where $F:\overline{\Omega}\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^{n\times n}\to\mathbb{R}$ is smooth. We additionally assume that for each fixed $x\in\mathbb{R}^n$ and $z\in\mathbb{R}^n$,

$$\text{the joint mapping } (r,P)\mapsto F(x,z,r,P) \text{ is convex.} \tag{2.52}$$ {eq:2-52}

This is called *polyconvexity.*

**Definition 2.8** (Polyconvex). *Suppose that $m=n$ and the Lagrangian $L:\overline{\Omega}\times\mathbb{R}^n\times\mathbb{R}^{n\times n}\to\mathbb{R}$ has the form*

$$L(x,z,P)=F(x,z,\det P,P),\quad x\in\Omega,\quad z\in\mathbb{R}^n,\quad P\in\mathbb{R}^{n\times n},$$

*where $F:\overline{\Omega}\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^{n\times n}\to\mathbb{R}$ is smooth. We say that $L$ is* **polyconvex** *provided that for each fixed $x\in\mathbb{R}^n$ and $z\in\mathbb{R}^n$,*

$$\textit{the joint mapping } (r,P)\mapsto F(x,z,r,P) \textit{ is convex.}$$

t2.8 **Theorem 2.8** (Weak Lower Semicontinuity of Polyconvex Functionals). *Suppose that $n<q<+\infty$. Assume also that $L$ is bounded below and polyconvex. Then $I[\cdot]$ is sequentially weakly lower semicontinuous on $W^{1,q}(\Omega,\mathbb{R}^n)$.*

*Proof.* Choose any sequence $\{\boldsymbol{u}_k\}_{k=1}^{+\infty}\subset W^{1,q}(\Omega,\mathbb{R}^n)$ such that

$$\boldsymbol{u}_k\rightharpoonup\boldsymbol{u}\quad\text{weakly in } W^{1,q}(\Omega,\mathbb{R}^n). \tag{2.53}$$ {eq:2-53}

According to Lemma (2.1),

$$\det D\boldsymbol{u}_k\rightharpoonup\det D\boldsymbol{u}\quad\text{weakly in } L^{\frac{q}{n}}(\Omega). \tag{2.54}$$ {eq:2-54}

Arguing exactly as in the proof of Theorem (2.1), and adopting the same notation, we have by the polyconvexity assumption

$$\begin{aligned} I[\boldsymbol{u}_k] &= \int_\Omega L(x,\boldsymbol{u}_k,D\boldsymbol{u}_k)\ dx \ge \int_{G_\epsilon} L(x,\boldsymbol{u}_k,D\boldsymbol{u}_k)\ dx \\ &= \int_{G_\epsilon} F(x,\boldsymbol{u}_k,\det D\boldsymbol{u}_k,D\boldsymbol{u}_k)\ dx \end{aligned}$$

$$\geq \int_{G_\epsilon} F(x, \boldsymbol{u}_k, \det D\boldsymbol{u}, D\boldsymbol{u})\, dx + \int_{G_\epsilon} F_r(x, \boldsymbol{u}_k, \det D\boldsymbol{u}, D\boldsymbol{u}) \cdot (\det D\boldsymbol{u}_k - \det D\boldsymbol{u}) + F_p(x, \boldsymbol{u}_k, \det D\boldsymbol{u}, D\boldsymbol{u}) \cdot (D\boldsymbol{u}_k - D\boldsymbol{u})\, dx.$$

Thus by reasoning as in the proof of Theorem (2.1), we deduce from (2.53) and (2.54) that the limit of the second integral on the RHS is zero as $k \to +\infty$. The proof is complete. □

As before, we immediately get existence.

t2.9 **Theorem 2.9** (Existence of Minimizers for Polyconvex Functionals)**.** *Assume that $n < q < +\infty$ and that $L$ satisfies the coercivity inequality (2.43) and is polyconvex. Suppose also that the admissible set $\mathcal{A}$ is nonempty.*

*Then there exists $\boldsymbol{u} \in \mathcal{A}$ such that*

$$I[\boldsymbol{u}] = \min_{\boldsymbol{w} \in \mathcal{A}} I[\boldsymbol{w}].$$

---

2.5. **Local Minimizers.** We want to determine under what circumstances a critical point of the energy functional $I[\cdot]$ is in fact a minimizer or a local minimizer. Assume that $u$ is a smooth solution of the Euler–Lagrange PDE

$$\begin{cases} L_z(x,u,Du) - \sum_{i=1}^n (L_{p_i}(x,u,Du))_{x_i} = 0 \quad \text{on } \Omega, \\ u|_{\partial\Omega} = g, \end{cases} \tag{2.55}$$ {eq:2-55}

and is therefore a critical point of the functional

$$I[w] := \int_\Omega L(x, w, Dw)\, dx$$

among all functions $w$ satisfying the boundary condition $w|_{\partial\Omega} = g$. We also assume that the mapping

$$p \mapsto L(x,z,p) \text{ is convex.}$$

We will show that if the graph of $x \mapsto u(x)$ lies within a region $R$ generated by a one–parameter family of graphs $x \mapsto u(x,\lambda)$ corresponding to other critical points, then in fact $u$ is a minimizer of $I[\cdot]$, as compared with admissible variations $w$ taking values within $R$. *Note that the functions $u(x,\lambda)$ need not be minimizers of the functional $I[\cdot]$. They do, however, solve the Euler–Lagrange equation if we do not consider the boundary constraint.* More precisely, suppose that $I \subseteq \mathbb{R}$ is an open interval containing 0 and that $\{u(\cdot,\lambda) : \lambda \in I\}$ is a smooth one–parameter family of solutions of the Euler–Lagrange PDE

$$L_z(x, u(x,\lambda), Du(x,\lambda)) - \sum_{i=1}^n (L_{p_i}(x, u(x,\lambda), Du(x,\lambda))_{x_i} = 0 \tag{2.56}$$ {eq:2-56}

within $\Omega$, such that

$$u(x) = u(x,0) \quad x \in \Omega. \tag{2.57}$$ {eq:2-57}

We as follows construct an admissible function $w$ taking values in the region $R$, the union of the graphs of the functions $u(\cdot,\lambda)$ for $\lambda \in I$. Take $\theta : \overline{\Omega} \to I$ to be a smooth function satisfying

$$\theta|_{\partial\Omega} = 0. \tag{2.58}$$ {eq:2-58}

Define then

$$w(x) := u(x, \theta(x)), \tag{2.59}$$ {eq:2-59}

and note that $w|_{\partial\Omega} = u|_{\partial\Omega} = g$. *Note that $w$ is not necessarily a solution of the Euler–Lagrange PDE (2.55).*

t2.10 **Theorem 2.10** (Local Minimizers)**.** *Suppose that $u$ is a smooth solution of the Euler–Lagrange PDE (2.55). Then $u$ is a local minimizer of $I[\cdot]$ within the region $R$, in the sense that*

$$I[u] \leq I[w] \tag{2.60}$$ {eq:2-60}

*for any function $w$ constructed as follows.*

*Take $\theta : \overline{\Omega} \to I$ to be a smooth function satisfying*

$$\theta|_{\partial\Omega} = 0.$$

*Then define*

$$w(x) := u(x, \theta(x)).$$

What the theorem says is that if $u$ is a solution of the Euler–Lagrange PDE and is embedded within a family of other solutions, then $u$ is a minimizer of $I[\cdot]$ among functions $w$ having the form (2.59). In fact, if, say, $u_\lambda > 0$ for $\lambda$ small, we can write any $w$ that is sufficiently close to $u$ pointwise in this form. Note that $Dw$ need not be close to $Du$.

In other words, if we view $u$ and $w$ as elements in an appropriate function space $X$, then what we have is that $I[u] \leq I[w]$ whenever $\|u - w\|_X$ is small.

*Proof.*

(i). We first observe that, by the chain rule,

$$w_{x_i}(x) = u_{x_i}(x, \theta(x)) + u_\lambda(x, \theta(x))\theta_{x_i}(x), \quad i = 1, \ldots, n,$$

where $u_\lambda$ denotes the derivative of $u$ in the second argument. Thus

$$Dw = Du + u_\lambda D\theta,$$

and we see that the convexity of $L$ in its third argument implies that

$$\begin{aligned} I[w] &= \int_\Omega L(x, w, Dw)\ dx = \int_\Omega L(x, w, Du + u_\lambda D\theta)\ dx \\ &\geq \int_\Omega L(x, w, Du) + D_pL(x, w, Du) \cdot (u_\lambda D\theta)\ dx \\ &= \int_\Omega L(x, w, Du) + u_\lambda D_pL(x, w, Du) \cdot D\theta\ dx, \end{aligned} \tag{2.61}$$ {eq:2-61}

where $u$ is evaluated at $(x, \theta(x))$ and $D = D_x$.

(ii). We introduce the vector field $\boldsymbol{b} := (b^1, \ldots, b^n)$, defined by

$$b^i := \int_0^{\theta(x)} u_\lambda(x, \lambda) L_{p_i}(x, u(x, \lambda), Du(x, \lambda))\ d\lambda, \quad i = 1, \ldots, n. \tag{2.62}$$ {eq:2-62}

By the Leibniz integral rule and the product rule, we calculate

$$\begin{aligned} \partial_{x_i} b^i = {} & u_\lambda(x, \theta(x)) L_{p_i}(x, u(x, \theta(x)), Du(x, \theta(x)))\theta_{x_i}(x) + \\ & \int_0^{\theta(x)} \partial_{x_i}\left(u_\lambda(x, \lambda) L_{p_i}(x, u(x, \lambda), Du(x, \lambda))\right)\ d\lambda \\ = {} & u_\lambda(x, \theta(x)) L_{p_i}(x, u(x, \theta(x)), Du(x, \theta(x)))\theta_{x_i}(x) + \end{aligned}$$

$$\int_0^{\theta(x)} u_{\lambda x_i}(x,\lambda) L_{p_i}(x, u(x,\lambda), Du(x,\lambda))\ d\lambda \ + \\ \int_0^{\theta(x)} u_\lambda(x,\lambda)\left(L_{p_i}(x, u(x,\lambda), Du(x,\lambda))\right)_{x_i}\ d\lambda$$

for each $i = 1, \dots, n$. Hence,

$$\begin{aligned}
\operatorname{div} \boldsymbol{b} &= \sum_{i=1}^n u_\lambda(x, \theta(x)) L_{p_i}(x, u(x,\theta(x)), Du(x,\theta(x)))\theta_{x_i}(x) \ + \\
&\quad \sum_{i=1}^n \int_0^{\theta(x)} u_{\lambda x_i}(x,\lambda) L_{p_i}(x, u(x,\lambda), Du(x,\lambda))\ d\lambda \ + \\
&\quad \sum_{i=1}^n \int_0^{\theta(x)} u_\lambda(x,\lambda)\left(L_{p_i}(x, u(x,\lambda), Du(x,\lambda))\right)_{x_i}\ d\lambda \\
&= u_\lambda(x,\theta(x)) D_p L(x, w(x), Du(x,\theta(x))) \cdot D\theta(x) \ + \\
&\quad \sum_{i=1}^n \int_0^{\theta(x)} u_{\lambda x_i}(x,\lambda) L_{p_i}(x, u(x,\lambda), Du(x,\lambda))\ d\lambda \ + \\
&\quad \int_0^{\theta(x)} \sum_{i=1}^n u_\lambda(x,\lambda)\left(L_{p_i}(x, u(x,\lambda), Du(x,\lambda))\right)_{x_i}\ d\lambda.
\end{aligned}$$

But since $u(x,\lambda)$ solves the Euler–Lagrange equation (2.55) in the classical sense, we have

$$\begin{aligned}
\operatorname{div} \boldsymbol{b} &= u_\lambda(x,\theta(x)) D_p L(x, w(x), Du(x,\theta(x))) \cdot D\theta(x) \ + \\
&\quad \int_0^{\theta(x)} \sum_{i=1}^n u_{\lambda x_i}(x,\lambda) L_{p_i}(x, u(x,\lambda), Du(x,\lambda))\ d\lambda \ + \\
&\quad \int_0^{\theta(x)} u_\lambda(x,\lambda) L_z(x, u(x,\lambda), Du(x,\lambda))\ d\lambda.
\end{aligned}$$

Observe next that by the chain rule,

$$\begin{aligned}
\partial_\lambda L(x, u(x,\lambda), Du(x,\lambda)) &= L_z(x, u(x,\lambda), Du(x,\lambda)) u_\lambda(x,\lambda) \ + \\
&\quad \sum_{i=1}^n L_{p_i}(x, u(x,\lambda), Du(x,\lambda)) u_{\lambda x_i}(x,\lambda).
\end{aligned}$$

Hence, the previous calculation and the Fundamental Theorem of Calculus imply that

$$\begin{aligned}
\operatorname{div} \boldsymbol{b} &= u_\lambda(x,\theta(x)) D_p L(x, w(x), Du(x,\theta(x))) \cdot D\theta(x) \ + \\
&\quad \int_0^{\theta(x)} \left(L(x, u(x,\lambda), Du(x,\lambda))\right)_\lambda\ d\lambda \\
&= u_\lambda(x,\theta(x)) D_p L(x, w(x), Du(x,\theta(x))) \cdot D\theta(x) + L(x, u(x,\theta(x)), Du(x,\theta(x))) \ - \\
&\quad L(x, u(x,0), Du(x,0)) \\
&= u_\lambda(x,\theta(x)) D_p L(x, w(x), Du(x,\theta(x))) \cdot D\theta(x) + L(x, w(x), Du(x,\theta(x))) \ - \\
&\quad L(x, u(x), Du(x)),
\end{aligned}$$

in view of (2.57).

(iii). We insert this calculation into (2.61) to obtain

$$I[w] \geq \int_{\Omega} L(x, u(x), Du(x)) + \operatorname{div} \boldsymbol{b} \; dx = I[u] + \int_{\Omega} \operatorname{div} \boldsymbol{b} \; dx.$$

But by the Divergence Theorem,

$$\int_{\Omega} \operatorname{div} \boldsymbol{b} \; dx = \int_{\partial\Omega} \boldsymbol{b} \cdot \nu \; dS = 0,$$

since $\theta|_{\partial\Omega} = 0$ and consequently $\boldsymbol{b} = 0$ on $\partial\Omega$, in view of the definition of $\boldsymbol{b}$ in (2.62). Hence,

$$I[w] \geq I[u],$$

as required. The proof is complete. □

## 3. Regularity

In this section we discuss the smoothness of minimizers of our energy functional $I[\cdot]$. This is in general a difficult topic, and so we need to make a few strong simplifying assumptions. We suppose that our energy functional $I[\cdot]$ has the form

$$I[w] := \int_\Omega L(Dw) - wf \; dx, \tag{3.1}$$

for $f \in L^2(\Omega)$. We also take $q = 2$ and assume also the growth condition

$$|D_pL(p)| \leq C(|p| + 1), \quad p \in \mathbb{R}^n. \tag{3.2}$$

If we temporarily rewrite our Lagrangian $L$ as

$$F(x, w, Dw) := L(Dw) - wf,$$

we see that the corresponding Euler–Lagrange equation of (3.1) is

$$F_z(x, w, Dw) - \sum_{i=1}^n (F_{p_i}(x, w, Dw))_{x_i} = -f - \sum_{i=1}^n (L_{p_i}(Dw))_{x_i} = 0 \quad \text{on } \Omega.$$

Thus any minimizer $u \in \mathcal{A}$ is a weak solution of the Euler–Lagrange PDE

$$-\sum_{i=1}^n (L_{p_i}(Du))_{x_i} = f \quad \text{on } \Omega. \tag{3.3}$$

That is,

$$\int_\Omega \sum_{i=1}^n L_{p_i}(Du)\phi_{x_i} \; dx = \int_\Omega f\phi \; dx \tag{3.4}$$

for all $\phi \in H_0^1(\Omega)$. Note that the LHS of (3.4) is well-defined, since $u \in W^{1,2}(\Omega)$ and the growth condition (cf. (3.2)) implies

$$|D_pL(Du)| \leq C(|Du| + 1) \in L^2(\Omega),$$

for

$$\|D_pL(Du)\|_2^2 \leq C^2 \int_\Omega ||Du| + 1|^2 \; dx \leq C^2 \int_\Omega 1 + 2|Du| + |Du|^2 \; dx < +\infty,$$

since $\Omega \subseteq \mathbb{R}^n$ is bounded. Therefore an application of Hölder's inequality shows that the LHS of (3.4) is finite.

---

3.1. **Second Derivative Estimates.** We now want to show that if $u \in H^1(\Omega)$ is a weak solution of the Euler–Lagrange PDE (3.3), then in fact $u \in H^2_{loc}(\Omega)$. To establish this we need to strengthen our growth conditions on $L$. Let us first of all suppose that

$$|D^2L(p)| \leq C, \quad p \in \mathbb{R}^n. \tag{3.5}$$

Note that here we are assuming that we can in fact take *two* derivatives of $L$.

Additionally let us assume that $L$ is uniformly convex, and so there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^n L_{p_ip_j}(p)\xi_i\xi_j \geq \theta|\xi|^2, \quad p, \xi \in \mathbb{R}^n. \tag{3.6}$$

View the uniform convexity assumption (cf. (3.6)) as a nonlinear analogue of the assumption of uniform ellipticity required for the regularity theory of elliptic PDE. The idea is to try to mimic some of the calculations from the regularity theorems for solutions to linear elliptic PDE.

t3.1 **Theorem 3.1** (Second Derivatives for Minimizers).

*(i). Let $u \in H^1(\Omega)$ be a weak solution of the Euler–Lagrange PDE (3.3), and suppose that $L$ satisfies the growth condition*

$$|D^2L(p)| \leq C, \quad p \in \mathbb{R}^n$$

*and is uniformly convex (cf. (3.5), (3.6)). Then*

$$u \in H^2_{loc}(\Omega).$$

*(ii). If in addition $u \in H^1_0(\Omega)$ and $\partial\Omega \in \mathcal{C}^2$, then*

$$u \in H^2(\Omega)$$

*with the estimate*

$$\|u\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}.$$

Note that (ii) actually provides us with an a priori bound on the $H^2(\Omega)$ norm of a minimizer $u$.

*Proof.*

(i). Fix any open set $V \subset\subset \Omega$ and then choose an open set $W$ so that $V \subset\subset W \subset\subset \Omega$. Choose by Urysohn's Lemma a smooth cutoff function $\zeta$ satisfying

$$\zeta \equiv 1 \text{ on } V, \quad \zeta \equiv 0 \text{ on } \mathbb{R}^n \setminus W, \quad 0 \leq \zeta \leq 1.$$

Let $|h| > 0$ be small, choose $k \in \{1, \ldots, n\}$, and define the difference quotient

$$D_k^h u(x) := \frac{u(x + he_k) - u(x)}{h}, \quad x \in W.$$

Let us recall that

$$\int_\Omega u D_k^{-h} v \; dx = -\int_\Omega v D_k^h u \; dx$$

for all $u, v \in H^2(\Omega)$, for

$$\begin{aligned}
\int_\Omega u D_k^{-h} v \; dx &= \int_\Omega u(x) \left[\frac{v(x - he_k) - v(x)}{-h}\right] \; dx = \int_\Omega u(x) \left[\frac{v(x) - v(x - he_k)}{h}\right] \; dx \\
&= \int_\Omega \frac{u(x)v(x)}{h} \; dx - \int_\Omega \frac{u(x)v(x - he_k)}{h} \; dx \\
&= \int_\Omega \frac{u(x)v(x)}{h} \; dx - \int_\Omega \frac{u(x + he_k)v(x)}{h} \; dx \\
&= -\int_\Omega \left[\frac{u(x + he_k) - u(x)}{h}\right] v(x) \; dx \\
&= -\int_\Omega v D_k^h u \; dx.
\end{aligned}$$

This is the *integration by parts formula for difference quotients.* Now substitute

$$v := -D_k^{-h}(\zeta^2 D_k^h u)$$

into (3.4) to obtain

$$\begin{aligned}\sum_{i=1}^{n}\int_{\Omega} L_{p_i}(Du)\left(-D_k^{-h}(\zeta^2 D_k^h u)\right)_{x_i}\ dx &= \sum_{i=1}^{n}\int_{\Omega} D_k^h\left(L_{p_i}(Du)\right)(\zeta^2 D_k^h u)_{x_i}\ dx \\ &= -\int_{\Omega} f D_k^{-h}(\zeta^2 D_k^h u)\ dx. \qquad (3.7)\end{aligned}$$

{eq:3-7}

Now observe that

$$\begin{aligned}D_k^h L_{p_i}(Du(x)) &= \frac{1}{h}\int_0^1 \frac{d}{ds} L_{p_i}(sDu(x+he_k)+(1-s)Du(x))\ ds \\ &= \frac{1}{h}\int_0^1 L_{p_i p}(sDu(x+he_k)+(1-s)Du(x))\cdot(Du(x+he_k)-Du(x))\ ds \\ &= \frac{1}{h}\int_0^1 \sum_{j=1}^{n} L_{p_i p_j}(sDu(x+he_k)+(1-s)Du(x))(u_{x_j}(x+he_k)-u_{x_j}(x))\ ds \\ &= \int_0^1 \sum_{j=1}^{n} L_{p_i p_j}(sDu(x+he_k)+(1-s)Du(x)) D_k^h u_{x_j}(x)\ ds \\ &=: \sum_{j=1}^{n} a^{ij,h}(x) D_k^h u_{x_j}(x), \qquad (3.8)\end{aligned}$$

{eq:3-8}

where

$$a^{ij,h}(x) := \int_0^1 L_{p_i p_j}(sDu(x+he_k)+(1-s)Du(x))\ ds, \quad i,j=1,\dots,n. \qquad (3.9)$$

{eq:3-9}

Inserting (3.8) into (3.7) gives the identity

$$\begin{aligned}A_1 + A_2 &:= \sum_{i,j=1}^{n}\int_{\Omega} \zeta^2 a^{ij,h} D_k^h u_{x_j} D_k^h u_{x_i}\ dx + \sum_{i,j=1}^{n}\int_{\Omega} 2\zeta\zeta_{x_i} a^{ij,h} D_k^h u_{x_j} D_k^h u\ dx \\ &= -\int_{\Omega} f D_k^{-h}(\zeta^2 D_k^h u)\ dx =: B. \qquad (3.10)\end{aligned}$$

{eq:3-10}

Now the uniform convexity assumption (cf. (3.6)) implies

$$\begin{aligned}A_1 &= \sum_{i,j=1}^{n}\int_{\Omega} \zeta^2 a^{ij,h} D_k^h u_{x_j} D_k^h u_{x_i}\ dx \\ &= \int_{\Omega}\zeta^2\int_0^1\left(\sum_{i,j=1}^{n} L_{p_i p_j}(sDu(x+he_k)+(1-s)Du(x)) D_k^h u_{x_j} D_k^h u_{x_i}\right)\ ds\ dx \\ &\geq \int_{\Omega}\zeta^2\int_0^1 \theta |D_k^h Du(x)|^2\ ds\ dx \\ &= \theta\int_{\Omega}\zeta^2 |D_k^h Du(x)|^2\ dx. \qquad (3.11)\end{aligned}$$

{eq:3-11}

Furthermore we have by the growth assumption (cf. (3.5)) and Cauchy's inequality with $\epsilon$ that

$$
\begin{aligned}
|A_2| &\le \sum_{i,j=1}^{n} 2\int_\Omega \zeta\zeta_{x_i}|a^{ij,h}||D_k^h u_{x_j}||D_k^h u|\ dx\\
&\le \sum_{i,j=1}^{n} 2\int_W \zeta \int_0^1 |L_{p_ip_j}(sDu(x+he_k)+(1-s)Du(x))|\ ds\ |D_k^h Du_{x_j}||D_k^h u|\ dx\\
&\le C\int_W \zeta|D_k^h Du||D_k^h u|\ dx\\
&\le \epsilon\int_W \zeta^2|D_k^h Du|^2\ dx + \frac{C}{\epsilon}\int_W |D_k^h u|^2\ dx. \qquad (3.12)
\end{aligned}
$$

{eq:3-12}

We now estimate $|B|$. First recall that $\|D^h w\|_{L^2(\Omega)} \le C\|Dw\|_{L^2(\Omega)}$ for any $w \in H^1(\Omega)$. Therefore by Hölder's inequality,

$$|B| \le \|fD_k^{-h}(\zeta^2 D_k^h u)\|_{L^1(\Omega)} \le \|f\|_{L^2(\Omega)}\|D_k^{-h}(\zeta^2 D_k^h u)\|_{L^2(\Omega)} \le C\|f\|_{L^2(\Omega)}\|D(\zeta^2 D_k^h u)\|_{L^2(\Omega)}.$$

Next observe that, since $0 \le \zeta \le 1$,

$$
\begin{aligned}
\|D(\zeta^2 D_k^h u)\|_{L^2(\Omega)}^2 &= \int_\Omega |D(\zeta^2 D_k^h u)|^2\ dx\\
&= \int_\Omega |D(\zeta^2)D_k^h u + \zeta^2 D_k^h Du|^2\ dx\\
&\le C\int_\Omega |D(\zeta^2)D_k^h u|^2 + \zeta^4|D_k^h Du|^2\ dx\\
&\le C\int_\Omega |D_k^h u|^2 + \zeta^2|D_k^h Du|^2\ dx\\
&\le C\left(\|D_k^h u\|_{L^2(\Omega)}^2 + \|\zeta D_k^h Du\|_{L^2(\Omega)}^2\right)\\
&\le C\left(\|Du\|_{L^2(\Omega)}^2 + \|\zeta D_k^h Du\|_{L^2(\Omega)}^2\right).
\end{aligned}
$$

Taking square roots, we find (upon possibly increasing $C$)

$$\|D(\zeta^2 D_k^h u)\|_{L^2(\Omega)} \le C\left(\|Du\|_{L^2(\Omega)} + \|\zeta D_k^h Du\|_{L^2(\Omega)}\right),$$

and so

$$|B| \le C\|f\|_{L^2(\Omega)}\left(\|Du\|_{L^2(\Omega)} + \|\zeta D_k^h Du\|_{L^2(\Omega)}\right).$$

Now applying Cauchy's inequality with $\epsilon$ and choosing $C > \epsilon^2$ gives

$$
\begin{aligned}
|B| &\le \epsilon\|f\|_{L^2(\Omega)}^2 + \frac{C}{\epsilon}\|Du\|_{L^2(\Omega)}^2 + \epsilon\|\zeta D_k^h Du\|_{L^2(\Omega)}^2 + \frac{C}{\epsilon}\|f\|_{L^2(\Omega)}^2\\
&\le \epsilon\|\zeta D_k^h Du\|_{L^2(\Omega)}^2 + \frac{C}{\epsilon}\left(\|f\|_{L^2(\Omega)}^2 + \|Du\|_{L^2(\Omega)}^2\right)\\
&= \epsilon\int_\Omega \zeta^2|D_k^h Du|^2\ dx + \frac{C}{\epsilon}\int_\Omega |f|^2 + |Du|^2\ dx.
\end{aligned}
$$

Next we notice from the above bounds on $A_1, A_2$, and $B$ (cf. (3.11), (3.12)) that

$$\theta\int_\Omega \zeta^2|D_k^h Du|^2\ dx \le A_1 = B - A_2 \le |B| + |A_2|$$

$$\leq 2\epsilon \int_\Omega \zeta^2 |D_k^h Du|^2 \, dx + \frac{C}{\epsilon} \int_\Omega |f|^2 \, dx + \frac{2C}{\epsilon} \int_\Omega |Du|^2 \, dx.$$

Therefore selecting $\epsilon = \frac{\theta}{4}$ shows that

$$\int_\Omega \zeta^2 |D_k^h Du|^2 \, dx \leq C \int_\Omega |f|^2 + |Du|^2 \, dx.$$

(ii). Since $\zeta \equiv 1$ on $V$, we find

$$\int_V |D_k^h Du|^2 \, dx \leq C \int_\Omega |f|^2 + |Du|^2 \, dx$$

for each $k = 1, \ldots, n$ and all sufficiently small $|h| > 0$. Consequently $\|D^h Du\|_{L^2(\Omega)} < +\infty$, and we recall that this implies $Du \in H^1(V)$, and so $u \in H^2(V)$. This is for all $V \subset\subset \Omega$, and therefore $u \in H^2_{loc}(\Omega)$.

(iii). Assuming now that $u \in H^1_0(\Omega)$ is a weak solution of (3.3) and $\partial\Omega \in \mathcal{C}^2$, we may mimic the standard proof for regularity of weak solutions of elliptic PDE up to the boundary to deduce that $u \in H^2(\Omega)$, with the estimate

$$\|u\|_{H^2(\Omega)} \leq C \left( \|f\|_{L^2(\Omega)} + \|u\|_{H^1(\Omega)} \right).$$

Now by since $L$ is uniformly convex (cf. (3.6)), $L$ is strongly convex, and we have

$$(DL(Dw) - DL(0)) \cdot Dw \geq \theta |Dw|^2$$

for all $w \in H^1(\Omega)$. Putting now $\phi = u$ in the Euler–Lagrange equation (3.4), we calculate

$$\begin{aligned}
\int_\Omega fu \, dx &= \int_\Omega \sum_{i=1}^n L_{p_i}(Du) u_{x_i} \, dx \\
&= \int_\Omega DL(Du) \cdot Du \, dx \\
&\geq \int_\Omega \theta |Du|^2 + DL(0) \cdot Du \, dx \\
&= \theta \|Du\|^2_{L^2(\Omega)} + C\|Du\|_{L^1(\Omega)} \\
&\geq C\|Du\|^2_{L^2(\Omega)}.
\end{aligned}$$

On the other hand, we have by Hölder's inequality and Young's inequality

$$\int_\Omega fu \, dx \leq \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \leq \frac{\|f\|^2_{L_2(\Omega)}}{2} + \frac{\|u\|^2_{L^2(\Omega)}}{2}.$$

Therefore

$$\|Du\|^2_{L^2(\Omega)} - \|u\|^2_{L^2(\Omega)} \leq C\|f\|^2_{L^2(\Omega)}.$$

Choosing now $\widetilde{C} > 0$ such that $\widetilde{C} > 2 \left( \frac{\|u\|_{L^2(\Omega)}}{\|f\|_{L^2(\Omega)}} \right)^2$ shows that

$$\|u\|^2_{L^2(\Omega)} + \|Du\|^2_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)} + 2\|u\|^2_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)} + \widetilde{C}\|f\|^2_{L^2(\Omega)}.$$

Hence the estimate

$$\|u\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}$$

follows. The proof is complete. □

**Remark.** *Theorem* (3.1) *tells us that any minimizer* $u \in H_0^1(\Omega)$ *of* (3.3) *is actually in* $H^2(\Omega)$*, provided that* $\partial\Omega$ *is sufficiently smooth. We recall here the Sobolev embedding, which can be found in* [1, §5.6.3, Theorem 6(ii)]*:*

*Let* $\Omega$ *be a bounded open subset of* $\mathbb{R}^n$*, with* $\partial\Omega \in \mathcal{C}^1$*. Assume also* $u \in W^{k,p}(\Omega)$*. If*

$$k > \frac{n}{p},$$

*then* $u \in \mathcal{C}^{k-\lfloor \frac{n}{p} \rfloor -1,\gamma}(\overline{\Omega})$*, where*

$$\gamma := \begin{cases} \lfloor \frac{n}{p} \rfloor + 1 - \frac{n}{p}, & \frac{n}{p} \notin \mathbb{Z}, \\ \text{Any positive number in } (0,1), & \frac{n}{p} \in \mathbb{Z}. \end{cases}$$

*We have in addition the estimate*

$$\|u\|_{\mathcal{C}^{k-\lfloor \frac{n}{p} \rfloor -1,\gamma}(\overline{\Omega})} \le C\|u\|_{W^{k,p}(\Omega)},$$

*where the constant* $C > 0$ *depends only on* $k, p, n, \gamma$*, and* $\Omega$*.*

*Returning to the case of our minimizer* $u \in H_0^1(\Omega)$*, we have* $k = p = 2$*. Thus if* $n < 4$*, then actually* $u \in \mathcal{C}^{1-\lfloor \frac{n}{p} \rfloor,\gamma}(\overline{\Omega})$*, and*

$$\|u\|_{\mathcal{C}^{1-\lfloor \frac{n}{p} \rfloor,\gamma}(\overline{\Omega})} \le C\|f\|_{L^2(\Omega)}.$$

*Note in particular that* $u$ *can have at most one classical derivative, and this is only in the one–dimensional case* $n = 1$*.*

---

3.2. **Remarks on Higher Regularity.** We would like to show that if $L$ is infinitely differentiable, then so is $u$. By analogy with the regularity theory for second–order linear elliptic PDE, it seems natural to try to extend the $H^2_{loc}(\Omega)$ estimate from Theorem (3.1) to obtain further estimates in the higher Sobolev spaces $H^k_{loc}(\Omega)$, for $k = 3, 4, \ldots$.

This method will *not* work for the nonlinear Euler–Lagrange PDE (3.3) however. The reason is as follows. For linear equations we could, roughly speaking, differentiate the equation as many times as needed and still obtain a linear PDE of the same general form as the one we began with. But if we differentiate a nonlinear PDE many times, the resulting increasingly complicated expressions quickly become impossible to handle. Much deeper ideas are called for. We attempt here to give an outline.

To begin, choose a test function $w \in \mathcal{C}_c^\infty(\Omega)$, choose $k \in \{1, \ldots, n\}$, and set $\phi := -w_{x_k}$ in the identity (3.4), where for simplicity we now take $f \equiv 0$. Equation (3.4) thus becomes

$$-\int_\Omega \sum_{i=1}^n L_{p_i}(Du) w_{x_i x_k} \, dx = 0.$$

Knowing that $u \in H^2_{loc}(\Omega)$, integration by parts gives

$$\int_\Omega \sum_{i,j=1}^n L_{p_i p_k}(Du) u_{x_k x_j} w_{x_i} \, dx = 0. \tag{3.13}$$ {eq:3-13}

Next write

$$\widetilde{u} := u_{x_k} \tag{3.14}$$ {eq:3-14}

and

$$a^{ij} := L_{p_i p_k}(Du), \quad i, j = 1, \ldots, n. \tag{3.15}$$ {eq:3-15}

Fix also any $V \subset\subset \Omega$. Then (3.13) becomes

$$\int_\Omega \sum_{i,j=1}^n a^{ij}(x)\widetilde{u}_{x_j} w_{x_i}\, dx = 0. \tag{3.16}$$

After a standard approximation argument, (3.16) holds for any $w \in H_0^1(V)$. This is to say that $\widetilde{u} \in H^1(V)$ is a weak solution of the linear second–order elliptic PDE

$$-\sum_{i,j=1}^n (a^{ij}\widetilde{u}_{x_j})_{x_i} = 0 \quad \text{on } V. \tag{3.17}$$

But we *cannot* apply standard regularity theory from linear elliptic PDE to conclude that $\widetilde{u}$ is smooth, the reason being that we can deduce from (3.5) and (3.15) only that

$$a^{ij} \in L^\infty(V), \quad i, j = 1, \ldots, n.$$

On the other hand, a deep theorem due to DeGiorgi, Moser, and Nash, asserts that any weak solution of (3.17) is in fact locally Hölder continuous for some exponent $\gamma > 0$. Thus if $W \subset\subset V$, we have $\widetilde{u} \in \mathcal{C}^{0,\gamma}(W)$, and recalling the definition of $\widetilde{u}$ (cf. (3.14)), we have

$$u \in \mathcal{C}^{1,\gamma}_{loc}(\Omega).$$

Return now to the definition (3.15). If $L \in \mathcal{C}^{2,\gamma}_{loc}(\Omega)$, we now know $a^{ij} \in \mathcal{C}^{0,\gamma}_{loc}(\Omega)$, $i, j = 1, \ldots, n$. Then (3.3) along with an older theorem of Schauder asserts that in fact

$$u \in \mathcal{C}^{2,\gamma}_{loc}(\Omega).$$

But then returning to (3.15), we see that $a^{ij} \in \mathcal{C}^{1,\gamma}_{loc}(\Omega)$, and reapplying Schauder's estimate implies

$$u \in \mathcal{C}^{3,\gamma}_{loc}(\Omega).$$

We can continue this bootstrapping argument eventually to deduce that $u \in \mathcal{C}^{k,\gamma}_{loc}(\Omega)$ for $k = 1, \ldots,$ and thus $u \in \mathcal{C}^\infty(\Omega)$.

# 4. Constraints

## References

1. Lawrence C. Evans, *Partial differential equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR 2597943