Necessary and Sufficient Conditions for Weak Local Variational Extrema

If $\bar{x} \in \mathbb{R}^n$ is a local minimum of $f : \mathbb{R}^n \to \mathbb{R}$ then the gradient of f vanishes at $\bar{x}$ and its Hessian matrix is positive-semidefinite. Strengthening the latter condition to strict inequality, namely positive-definite Hessian, provides a sufficient condition for a local minimum. In the variational case the necessary conditions are identical: if $\bar{u}$ is a weak local minimum of J then the first variation $\delta J(\bar{u}, \cdot)$ vanishes, implying a variant of the Euler-Lagrange equations and transversality conditions, and the second variation $\delta^2 J(\bar{u}, \cdot)$ is positive-semidefinite, implying a variant of the Legendre condition. However, strengthening the latter condition to strict inequality, namely positive-definite second variation, is not a sufficient condition for weak local minimizers in function spaces. Section 2 lays down the basics of differentiation in function spaces concluding (Section 2.4) with minimum regularity requirements for a variational optimization problem of arbitrary order k in n space dimensions to be well defined (Definition 2.6). Sections 3 and 4.1 review standard first and second order necessary conditions for local extrema in addition to an unsurprising Legendre condition for $k = 2$ (Theorem 4.4). Section 4.2 contains five second order sufficient conditions for arbitrary k-th order variational minimization problems: **1)** *positive-semidefiniteness of second variation in a weak neighborhood of a minimizer (Theorem 4.5),* **2)** *strong positivity of second variation at a minimizer with respect to $\|\cdot\|_{W^{k,\infty}}$ (Theorem 4.6) and* **3)** *with respect to $\|\cdot\|_{W^{k,2}}$ (Theorem 4.7),* **4)** *Strong positive-definiteness of the Hessian of the Lagrangian at a minimizer (Theorem 4.8), and* **5)** *Jacobi's sufficient condition for $k = 1$ (Theorem 4.9).*

Amir Kadivar, January 2018

§ 1 Introduction

We are interested in necessary and sufficient conditions for local extrema of functionals $J : W \to \mathbb{R}$ of the form

$$J(u) \coloneqq \int_\Omega L\left(x, u(x), Du(x), \dots, D^k u(x)\right) dx = \int_\Omega L^{(u)}(x) dx \tag{1}$$

Additionally, boundary conditions could also imposed on u and its derivatives, the simplest case being a prescribed value for u on $\partial\Omega$ (i.e constrained boundary).

The variational integral (1) has three important characteristics:

- $n = 1, 2, \dots$ the number of space dimensions,
- $k = 0, 1, 2, \dots$ the order of the *Lagrangian* L,
- $\Omega \subset \mathbb{R}^n$ the open domain of integration.

The tools we use and the key results depend very little on the number of dimensions n. The following, however, have important consequences on many results and their proofs:

- *The order k of the Lagrangian*: practically we are mostly concerned with the cases $k = 1, 2$. However, results that hold without modification for all k are presented in the general case. The form of the Euler-Lagrange equations and the proof of the Legendre necessary condition varies across different k and are only presented here for $k = 1, 2$.
- *Boundedness of Ω*: as we shall see, although the natural norm for our type of problem is $\|\cdot\|_{W^{k,\infty}(\Omega)}$, continuity in u of the first and second variations $\delta J(u, \varphi), \delta^2 J(u, \varphi)$, which are desired for Fréchet differentiability and Taylor approximation theorems, are related to $\|\varphi\|_{W^{k,1}(\Omega)}$ and $\|\varphi\|_{W^{k,2}(\Omega)}$. In the study of sufficient second order conditions the assumption of boundedness of Ω is crucial in formulating statements involving different norms.
- *Constraints on $\partial\Omega$*: we pay very little attention to boundary conditions and in general only assume that whatever boundary condition is imposed leaves us with a Banach subspace of the function space we start with. The only results here that depend on boundary conditions are the first order necessary conditions in section 3.

In this section we summarize the relevant notions of differentiability in Banach spaces and find the natural Banach space for (1) to be well-defined and analyzable using tools of differential calculus (section 2.4, definition 2.6).

Notation

For clarity we use the notation $L(x, z, p, q, \dots)$ to distinguish arguments of L from derivatives of u and define

$$\begin{aligned} L^{(u)} : \quad \Omega &\to \mathbb{R} \\ x &\mapsto L\left(x, u(x), Du(x), \dots, D^k u(x)\right) \end{aligned}$$

For appropriately differentiable functions $u : \Omega \to \mathbb{R}$ we use the standard multi-index notation to denote differential operators; namely for any $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$ we define $D^\alpha u \coloneqq \partial_{x_1^{\alpha_1}} \ldots \partial_{x_n^{\alpha_n}} u$ and $|\alpha| \coloneqq \alpha_1 + \ldots + \alpha_n$. This means we have two notations for indexing $D^\cdot$: for a multi-index α, D^α is the real valued partial derivative, and for an integer k, D^k refers to the k-th-order tensor containing all k-th order partials. For convenience we also allow D^k to stand for its vectorized, flattened form, e.g. the Hessian $D^2 u$ contains all partials $D^\alpha u$ for $|\alpha| \leq 2$ in a symmetric $n \times n$ matrix and can be flattened as an n^2 vector if needed. We use the notation L_α to denote the partial derivative of L with respect to its real valued variable that corresponds to $D^\alpha u$. Similarly L_p refers to the n-vector of all partials of L with respect to p, L_q to the partial derivatives of L with respect to q in $n \times n$-matrix or n^2-vector form, and so on. Finally, the special case $n = 1$ is often used in examples for simplicity; in these cases the physicist's notation $\dot{u}, \ddot{u}, \ldots$ is used.

Norms and Function Spaces

Definition 1.1. L_p ***norms.*** *For any* $1 \leq p \leq \infty$ *define*

$$\|u\|^p_{L^p(\Omega)} \coloneqq \int_\Omega |u|^p dx, \quad \|u\|_{L^\infty(\Omega)} \coloneqq \operatorname*{ess\,sup}_\Omega |u|$$

Recall that $\|u\|_{L^\infty(\Omega)} < \infty$ iff u is essentially bounded, i.e. bounded except maybe at a set of measure zero, and convergence in $\|\cdot\|_{L^\infty(\Omega)}$ is equivalent to uniform convergence in Ω.

Definition 1.2. $W^{k,p}$ ***norms.*** *For any* $k = 0, 1, \ldots$ *and any* $1 \leq p \leq \infty$

$$\|u\|^p_{W^{k,p}(\Omega)} \coloneqq \int_\Omega \sum_{|\alpha| \leq k} |D^\alpha u|^p dx = \sum_{|\alpha| \leq k} \|D^\alpha u\|^p_{L^p(\Omega)}$$

$$\|u\|_{W^{k,\infty}(\Omega)} \coloneqq \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty} = \sum_{|\alpha| \leq k} \operatorname*{ess\,sup}_\Omega |D^\alpha u|$$

We note that convergence of $u_n \to u$ in $\|\cdot\|_{W^{k,p}(\Omega)}$ is equivalent to

$$\forall |\alpha| \leq k : |D^\alpha (u_n - u)| \to 0 \text{ in } \quad \|\cdot\|_{L^p(\Omega)}$$

Consequently, $u_n \to u$ in $\|\cdot\|_{W^{k,\infty}(\Omega)}$ is equivalent to

$$\forall |\alpha| \leq k : D^\alpha u_n \to D^\alpha u \text{ uniformly over } \Omega$$

Definition 1.3. *We are concerned with the following function spaces for any* $1 \leq p \leq \infty$ *and any*

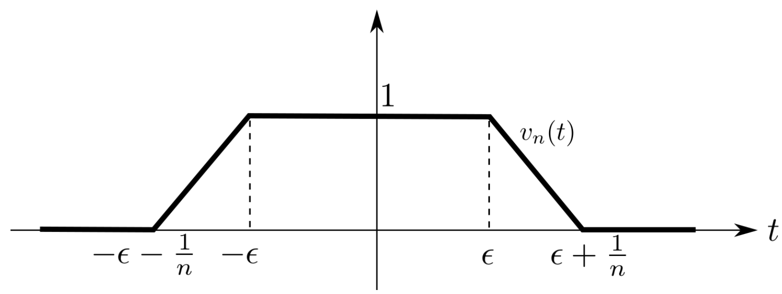

Figure 1: A sequence of continuous functions v_n such that $v_n \to v$ in $\|\cdot\|_{W^{0,1}(\Omega)} = \|\cdot\|_{L^1(-1,1)}$ where $v = \mathbf{1}_{(-\epsilon,\epsilon)}$ is not continuous. This shows that $C^0(\overline{\Omega})$ is not closed in $W^{0,1}(\Omega) = L^1(\Omega)$ even if Ω is bounded.

$k = 0, 1, \ldots$

$$\begin{aligned} C^k(\Omega) &:= \left\{ u : \Omega \to \mathbb{R} \;\middle|\; \forall |\alpha| \le k : D^\alpha u \textit{ defined and continuous} \right\} \\ L^p(\Omega) &:= \left\{ u : \Omega \to \mathbb{R} \;\middle|\; \|u\|_{L^p(\Omega)} < \infty \right\} \\ W^{k,p}(\Omega) &:= \left\{ u : \Omega \to \mathbb{R} \;\middle|\; \|u\|_{W^{k,p}(\Omega)} < \infty \right\} \\ &= \left\{ u : \Omega \to \mathbb{R} \;\middle|\; \forall |\alpha| \le k : \|D^\alpha u\|_{L^p(\Omega)} < \infty \right\} \end{aligned}$$

We accept two standard results (Theorems 1.1 and 1.2) without proof [2].

▶ **Theorem 1.1** *Let $\Omega \subset \mathbb{R}^n$ be an open set. For any $1 \le p \le \infty$ and any $k = 0, 1, \ldots$*

1) $L^p(\Omega)$ with $\|\cdot\|_{L^p(\Omega)}$ is a Banach space,
2) $W^{k,p}(\Omega)$ with $\|\cdot\|_{W^{k,p}(\Omega)}$ is a Banach space,
3) $C^k(\overline{\Omega}) \cap W^{k,\infty}(\Omega)$ with $\|\cdot\|_{W^{k,\infty}(\Omega)}$ is a Banach space.

Remark. In case (3) the intersection is only necessary when Ω is unbounded since otherwise $C^k(\overline{\Omega})$ is a proper subspace of $W^{k,\infty}(\Omega)$. To prove its completeness it suffices to show that it is closed under $\|\cdot\|_{W^{k,\infty}(\Omega)}$. Note that the last claim does not hold for $p < \infty$ (Fig. 1).

▶ **Theorem 1.2** *Let $\Omega \subset \mathbb{R}^n$ be bounded.*

1) For any $1 \le p < q \le \infty$, the L^q norm dominates the L^p norm (section 5.1), that is, there exists $C > 0$ such that for all $u \in L^p(\Omega)$ we have

$$\|u\|_{L^p(\Omega)} \le C \|u\|_{L^q(\Omega)}$$

Consequently, we get $L^q(\Omega) \subset L^p(\Omega)$.

2) ***Poincaré Inequality,*** *For any $1 \le p \le \infty$ there exists a $C > 0$ such that for all $u \in L^p(\Omega)$ we have*

$$\|u\|_{L^p(\Omega)} \le C \|Du\|_{L^p(\Omega)}$$

§ 2 Differentiation in Banach Spaces

In this section we consider an arbitrary functional $f : W \to \mathbb{R}$ on a Banach space W. In section 2.4 we translate our findings to regualrity requirements that we will impose on the variational problem of interest.

§ 2.1 Directional, Gâteaux, and Fréchet Derivatives

Definition 2.1. *The* ***directional derivative*** *or the* ***first variation*** *of f at $u \in W$ along $\varphi \in W$ is*

$$\delta f(u, \varphi) \coloneqq \frac{d}{ds}\Big|_{s=0} f(u + s\varphi) = \lim_{s\to 0} \frac{1}{s}\Big[f(u + s\varphi) - f(u)\Big] \tag{2}$$

► **Theorem 2.1** ***Integration of first variation along lines [1].*** *Suppose $u, v \in W$ and denote by γ the line segment connecting u and $u + v$, i.e the curve $\gamma : t \mapsto u + tv$ satisfying $\gamma(0) = u$ and $\gamma(1) = u + tv$. If the first variation of f along v is defined everywhere on γ then we have:*

$$f(u + v) - f(u) = \int_0^1 \delta f(\gamma(t), v)dt = \int_0^1 \delta f(u + tv, v)dt$$

Proof. Define $h(t) \coloneqq f(\gamma(t)) = f(u + tv)$. We can verify that

$$\dot{h}(t) = \delta f(\gamma(t), v)$$

The claim now follows by applying the fundamental theorem of calculus to $h : [0, 1] \to \mathbb{R}$. □

Definition 2.2. *If all directional derivatives of f at u exist and the map $\varphi \mapsto \delta f(u, \varphi)$ is linear, we call it the* ***Gâteaux derivative*** *[1, 5] of f at u, denoted by $Df(u)$. It is easy to verify that if the Gâteaux derivative exists it is unique and it agrees with all directional derivatives.*

The first variation limit (2) can also be written as

$$\lim_{s\to 0} \frac{1}{s}\Big[f(u + s\varphi) - f(u) - \delta f(u, s\varphi)\Big] = 0$$

Thus f is Gâteaux differentiable iff there exists a linear map $A : \varphi \mapsto \delta f(u, \varphi)$ satisfying:

$$\lim_{s\to 0} \frac{1}{s}\Big[f(u + s\varphi) - f(u) - sA\varphi\Big] = 0 \tag{3}$$

The first variation is always *homogeneous* in φ, that is, for all $\alpha \in \mathbb{R}$ we have $\delta f(u, \alpha\varphi) = \alpha\delta f(u, \varphi)$. Therefore, Gâteaux differentiability is equivalent to the requirement that the first variation be addi-

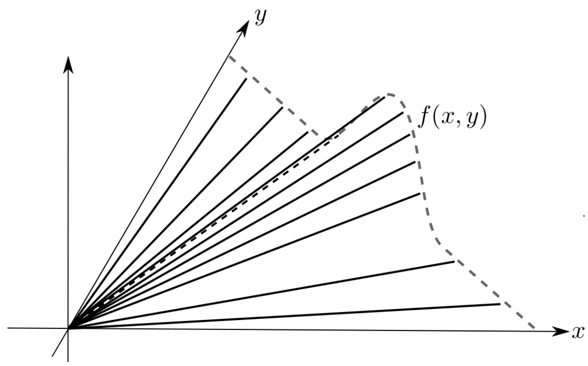

Figure 2: A function $f : \mathbb{R}^2 \to \mathbb{R}$ with all directional derivatives defined but such that the map $\varphi \mapsto \delta f(0, \varphi)$ is not additive and thus not linear; f is defined to in polar coordinates $f(r, \theta) \coloneqq r\eta\left(2\theta - \frac{\pi}{2}\right)$ where η is the standard mollifier (section 5.4).

tive in φ which is not always true (Fig. 2). Linear maps, however, are always Gâteaux differentiable with the Gâteaux derivative coinciding with the functional itself, i.e $Df(u)(\varphi) = \delta f(u, \varphi) = f(\varphi)$.

§ 2.1.1 Bounded Gâteaux Derivative

Linear maps in finite dimensions are always continuous; in infinite dimensions, however, continuity is equivalent to boundedness of the map on the unit ball (i.e finite operator norm, see section 5.2). It is easy to construct Gâteaux differentiable functions whose derivative is not continuous (bounded): any unbounded linear map $f : W \to \mathbb{R}$ (see 5.1) has unbounded Gâteaux derivative.

If the map $\varphi \mapsto \delta f(u, \varphi)$ is linear and bounded (i.e f has bounded Gâteaux derivative) then it is an element of the continuous dual W^* of W, often referred to as the *gradient* of f at u [5]. Furthermore, if $\|\cdot\|_W$ is induced by an inner product, for instance if $W = \mathbb{R}^n$ or $W = W^{k,2}(\Omega)$, then, by the Riesz representation theorem [2], the gradient can be canonically identified with an element of W.

§ 2.1.2 Fréchet Differentiability

Definition 2.3. *A bounded linear map A is the **Fréchet derivative** [1, 5] of f at u, denoted by $df(u)$, if we have*

$$\lim_{\varphi \to 0} \frac{1}{\|\varphi\|}\Big[f(u + \varphi) - f(u) - A\varphi\Big] = 0 \tag{4}$$

It is easy to verify that if the Fréchet derivative exists it is unique and agrees with the Gâteaux derivative and thus with all directional derivatives.

It is often useful to separate the effect of $\|\varphi\| \to 0$ from the changes in direction in the limit as $\varphi \to 0$ in (4). For any φ write $s \coloneqq \|\varphi\|$ and $\hat{\varphi} = \varphi/s$, and rewrite the limit (4) as follows: for any sequence of unit vectors $\hat{\varphi}_n$ and any sequence of reals $s_n \to 0$, we have

$$\frac{1}{s_n}\Big[f(u + s_n\hat{\varphi}_n) - f(u) - s_n A\hat{\varphi}_n\Big] \to 0 \text{ as } n \to \infty \tag{5}$$

The Gâteaux limit (3) and the Fréchet differentiability conditions (4), (5) are identical except for the fact that in (3) the perturbation is forced to approach the origin along the line $s\varphi$ for some fixed φ. Fréchet differentiability essentially demands that the Gâteaux limit (3) be uniform on the unit ball $\|\varphi\| = 1$. Therefore, if the Gâteaux derivative exists and is bounded, it does not follow that the Fréchet derivative exists.

Example 2.1 bounded Gâteaux derivative without Fréchet differentiability.

We construct a prototype using a simple functional $f : W \to \mathbb{R}$ such that $\varphi \mapsto \delta f(u, \varphi)$ is a bounded (continuous) linear map, yet f is not Fréchet differentiable. Let $\Omega \subset \mathbb{R}$ be open, and define:

$$f(u) := \int_\Omega u^2(t)dt$$

Clearly we must have $W \subset L^2(\Omega)$ with a norm to be specified. The first variation is

$$\delta f(u, \varphi) = 2\int_\Omega \varphi u dt$$

which is clearly linear in φ for any u. Therefore, f is Gâteaux differentiable and its Gâteaux derivative at u is a bounded map iff $\int_\Omega \varphi u dt$ is unbounded over unit ball in W. The LHS of the Gâteaux limit (3), or the Fréchet limit (5) is

$$\frac{1}{s}\Big[f(u + s\varphi) - f(u) - \delta f(u, s\varphi)\Big] = s\int_\Omega \varphi^2 dt$$

For Fréchet differentiability we require that the limit (5) holds for all $s_n \to 0$ and any sequence $\hat{\varphi}_n$ on the unit ball. To get a functional f that is not Fréchet differentiable $\int_\Omega \varphi^2 dt$ must be unbounded over unit ball in W. Our requirements for $\Omega, W, \|\cdot\|_W, u$ are

$$\begin{aligned} &W \subset L^2(\Omega) && \text{with norm such that:} \\ &\int_\Omega \varphi u dt && \text{bounded over unit ball in } W \\ &\int_\Omega \varphi^2 dt && \text{unbounded over unit ball in } W \end{aligned}$$

For instance, take $\Omega = (0, 1)$, $W = L^2(\Omega) \subset L^1(\Omega)$ with $\|\cdot\|_W = \|\cdot\|_{L^1(\Omega)}$ and $u \equiv 1$. The sequence $\hat{\varphi}_n = n\mathbf{1}_{(0,\frac{1}{n})}$ on the $\|\cdot\|_{L^1(\Omega)}$ unit ball of W is such that

$$\int_\Omega \hat{\varphi}_n^2 dt \to \infty \text{ as } n \to \infty$$

and thus $f : W \to \mathbb{R}$ defined above is not Fréchet differentiable at $u \equiv 1$ despite having a bounded Gâteaux derivative.

As can be seen in definitions and as demonstrated in the previous example, the existence and value of first variations and Gâteaux derivatives are independent of the norm imposed on W. The norm, however, is part of the information captured in Fréchet differentiability. The example above also shows that a function $f : W \to \mathbb{R}$ may be Fréchet differentiable with respect to some norms on W and not others.

▶ **Theorem 2.2** ***Sufficient condition for Fréchet differentiability [1, 5].*** *If $f : W \to \mathbb{R}$ is Gâteaux differentiable with bounded Gâteaux derivative in some neighborhood U of u, that is $\varphi \mapsto \delta f(v, \varphi)$ is a bounded linear map for all $v \in U$, and the map $v \mapsto \delta f(v, \varphi)$ is continuous at u for all $\varphi \in W$, then f is Fréchet differentiable at u.*

Proof. Suppose f is Gâteaux differentiable everywhere in U, denoted by $Df(\cdot)$. For any $\varphi \in W$ we write Theorem 2.1 for u and φ:

$$f(u + \varphi) - f(u) = \int_0^1 \delta f(u + t\varphi, \varphi) dt$$

Subtracting $\delta f(u, \varphi)$ from both sides gives:

$$f(u + \varphi) - f(u) - \delta f(u, \varphi) = \int_0^1 \Big[\delta f(u + t\varphi, \varphi) - \delta f(u, \varphi)\Big] dt$$

Therefore:

$$\begin{aligned}\Big|f(u + \varphi, f(u) - \delta f(u, \varphi)\Big| &\le \int_0^1 \Big|\delta f(u + t\varphi, \varphi) - \delta f(u, \varphi)\Big| dt \\ &\le \|\varphi\| \int_0^1 \|Df(u + t\varphi) - Df(u)\| \, dt\end{aligned}$$

where the norm under the integral is the operator norm (section 5.1). We now wish to show that if $\delta f(u, \varphi)$ is continuous in u then

$$\int_0^1 \|Df(u + t\varphi) - Df(u)\| \, dt \to 0 \ , \ \text{ as } \ \|\varphi\| \to 0$$

For this we write:

$$\begin{aligned}\int_0^1 \|Df(u + t\varphi) - Df(u)\| \, dt &\le \sup_{0 \le t \le 1} \|Df(u + t\varphi) - Df(u)\| \\ &= \sup_{0 \le t \le 1} \sup_{\|\psi\|=1} \Big|\delta f(u + t\varphi, \psi) - \delta(u, \psi)\Big|\end{aligned}$$

Now since for all $\psi \in W$ the map $v \mapsto \delta f(v, \psi)$ is continuous at u the first variation $\delta f(u + t\varphi, \psi)$

approaches $\delta f(u,\psi)$ as $\|\varphi\| \to 0$ for all ψ and all t. □

§ 2.2 Second Order Derivatives

Definition 2.4. *We define the second directional derivative of f at u along φ and ψ to be the first variation of $\delta f(u,\psi)$ at u along φ, namely:*

$$\delta^2 f(u,\varphi,\psi) \coloneqq \frac{d}{ds}\Big|_{s=0} \delta f(u+s\psi,\varphi) = \frac{\partial^2}{\partial s \partial t}\Big|_{s=t=0} \delta f(u+s\psi+t\varphi)$$

*We reserve the term **second variation** for the case when $\varphi = \psi$ and write:*

$$\delta^2 f(u,\varphi) \coloneqq \frac{d^2}{ds^2}\Big|_{s=0} f(u+s\varphi) \tag{6}$$

*As before, the map $\varphi,\psi \mapsto \delta^2 f(u,\varphi,\psi)$ is always homogeneous and thus is bilinear iff it is additive in φ and ψ in which case we call it the **second Gâteaux derivative** [1, 5] of f at u, denoted $D^2 f(u)$.*

Definition 2.5. *If f is Fréchet differentiable at a neighborhood $U \subset W$ of u with Fréchet derivative $df(\cdot)$ then the **second Fréchet derivative** [1, 5] of f at u is the Fréchet derivative of $df(\cdot)$ at u, i.e a bounded bilinear map $B : W \times W \to \mathbb{R}$ such that:*

$$\lim_{\psi \to 0} \frac{1}{\|\psi\|} \Big[df(u+\psi)(\varphi) - df(u) - B(\varphi,\psi) \Big] = 0$$

As in the first order case, all derivatives if they exist, are unique and imply and coincide with weaker notions of derivative (Fréchet implies Gâteaux and Gâteaux implies directional). Also as before, the Gâteaux derivative need not be continuous (bounded) and even if so, it need not imply second order Fréchet differentiability. However, for higher order derivatives too, sufficient conditions akin to Theorem 2.2 exist with similar proofs.

▶ **Theorem 2.3** ***Sufficient condition for second Fréchet differentiability [1, 5].*** *Suppose that in some neighborhood $U \subset W$ of u the functional $f : W \to \mathbb{R}$ is Fréchet differentiable and its second Gâteaux derivative $\delta^2 f(u,\cdot,\cdot)$ is a bounded bilinear map. Then if the map $v \mapsto \delta^2 f(v,\varphi,\psi)$ is continuous at u for all $\varphi,\psi \in W$ then f is second Fréchet differentiable at u.*

§ 2.3 Taylor Approximation Theorems

The definition of Fréchet differentiability is precisely the requirement for Taylor expansion theorem in Banach spaces. The following weak forms are all we need here (for more general cases see [1, 5]).

▶ **Theorem 2.4** ***First order Tarylor approximation.*** *Let $f : W \to \mathbb{R}$ be Fréchet differentiable at*

$u \in W$. For any φ define the constant $c_\varphi \in \mathbb{R}$ such that:

$$f(u+\varphi) = f(u) + df(u)\varphi + c_\varphi \|\varphi\|$$

Then $c_\varphi \to 0$ as $\varphi \to 0$.

▶ **Theorem 2.5** ***Second order Tarylor approximation.*** *Let $f : W \to \mathbb{R}$ be second order Fréchet differentiable at $u \in W$. For any φ define the constant $c_\varphi \in \mathbb{R}$ such that:*

$$f(u+\varphi) = f(u) + df(u)\varphi + \frac{1}{2}d^2 f(u)\varphi + c_\varphi \|\varphi\|^2$$

Then $c_\varphi \to 0$ as $\varphi \to 0$.

It is not surprising that Gâteaux differentiability is not enough for the above Taylor approximation theorems since Gâteaux derivatives are agnostic to the norm imposed on W.

§ 2.4 Regularity Requirements for Variational Optimization Problems

We now turn to our variation problem of finding local extrema of $J : W \to \mathbb{R}$ given by

$$J(u) = \int_\Omega L\big(x, u(x), Du(x), \ldots, D^k u(x)\big)dx = \int_\Omega L^{(u)}(x)dx$$

Minimal conditons on the domain Ω, the Lagrangian L, the function space W and its topology for the problem to be appropriately well defined are discussed.

$L^{(u)}$ well defined

For the integrand $L^{(u)}$ to be defined everywhere on Ω we require that $u \in C^k(\Omega)$. If boundary conditions are also imposed on u and its derivatives at $\partial\Omega$, we strengthen this to $u \in C^k(\overline{\Omega})$. Note that since the pointwise derivatives of u are passed as arguments to L we cannot expand W to contain functions with distributional derivatives alone. Since we wish to be able to say something about changes in L in response to perturbations in u, the natural norm to impose on W is one of the $\|\cdot\|_{W^{k,p}(\Omega)}$ norms where k is the order of the Lagrangian L. Therefore, we conclude by setting $W = C^k(\overline{\Omega}) \cap W^{k,p}(\Omega)$ with p and a norm to be specified.

J well defined, integrability of $L^{(u)}$

A k-th order Lagrangian L is a functional of the form

$$L : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n\times n} \times \ldots \mathbb{R}^{\overbrace{n \times \ldots \times n}^{k}} \to \mathbb{R}$$

Whether or not J is defined at $u : W \to \mathbb{R}$ depends on whether $L^{(u)} : \Omega \to \mathbb{R}$ is integrable.

Suppose we require that $L^{(u)}$ be integrable for any u. Without demanding continuity no integrability requirement on L in its original domain can guarantee integrability of $L^{(u)}$. This is because the subset of the domain of L over which J integrates is an n-dimensional submanifold and thus has measure zero. Even if we demand continuity and integrability from L, that is, say $L \in C^0 \cap L^1$, we cannot necessarily guarantee integrability for $L^{(u)}$

Example 2.2 Continuity and integrability of L does not guarantee integrability of $L^{(u)}$.

Let $n = 1, k = 0$. To see why integrability alone is not enough consider $L(x, z) = x\mathbf{1}_{\{0\}}(z)$ which satisfies any integrability condition since it is almost everywhere zero in its domain $\mathbb{R}^2$. Yet for the zero function $u \equiv 0$ over $\boldsymbol{\Omega} = \mathbb{R}$ we get a non-integrable $L^{(u)}$.

We now consider a continuous and integrable Lagrangian. Let $n = 1, k = 0, \boldsymbol{\Omega} = (0, 1)$, and let $h : \mathbb{R} \to \mathbb{R}$ be any continuous and locally bounded function such that $h(z) = 0$ only at $z = 0$ and $1/h$ is integrable in $\mathbb{R}$. Define

$$L(x, z) = x^{h(z)-1}$$

The zero function $u \equiv 0$ leads to a non-integrable $L^{(u)} : x \mapsto 1/x$. However, L is continuous everywhere in $\boldsymbol{\Omega} \times \mathbb{R}$ and integrable too:

$$\int_{\mathbb{R}\times\Omega} |L(x,z)| dxdz = \int_{\mathbb{R}} \int_{\Omega} L(x,z) dxdz = \int_{\mathbb{R}} \int_0^1 x^{h(z)-1} dxdz = \int_{\mathbb{R}} \frac{1}{h(z)} dz < \infty$$

We have, however, a sufficient (albeit too strong) condition for integrability of $L^{(u)}$:

▶ **Lemma 2.6** *If $\Omega \subset \mathbb{R}^n$ is bounded and $L \in C^0 \cap L^1$ then $L^{(u)}$ is integrable and thus J is defined for any $u \in W$.*

Continuity of J

None of the limits defining various derivatives discussed above exist if $J : W \to \mathbb{R}$ is not continuous. In order for J to be continuous at every $u \in W$ we must be able to control changes in L based on changes in u, the latter being represented in terms of the norm on W. Among the $W^{k,p}$ norms only $p = \infty$ allows us to make this connection.

▶ **Lemma 2.7** *If L is continuous and W is equipped with $\|\cdot\|_{W^{k,\infty}(\Omega)}$ then J is continuous.*

Proof.

$$\begin{aligned} |J(u) - J(v)| &= \left| \int_{\Omega} [L^{(u)} - L^{(v)}] dx \right| \leq \int_{\Omega} |L^{(u)} - L^{(v)}| dx \\ &= \int_{\Omega} \Big| L\big(x, u(x), Du(x), \dots\big) - L\big(x, v(x), Dv(x), \dots\big) \Big| dx \end{aligned}$$

Since L is continuous the RHS approaches zero as $\|u-v\|_{W^{k,\infty}(\Omega)} \to 0$ and thus J is continuous. $\square$

Example 2.3 Continuity of $J: W \to \mathbb{R}$ not guaranteed in any $\|\cdot\|_{W^{k,p}}$ for $p < \infty$.

Let $n = 1, k = 0$ and $\Omega = (1, \infty)$. For any $p < \infty$ consider the Lagrangian $L(x,z) = z^{2p}/n^{\frac{1}{4p}}$ and the the sequence of functions $u_n = x^{-\frac{1}{2p}-\frac{1}{n}}$.

$$\int_\Omega |u_n|^p dx = \frac{1}{n^{\frac{1}{4}}} \int_\Omega \frac{1}{x^{\frac{p}{n}+\frac{1}{2}}} dx = -\frac{1}{n^{\frac{1}{4}}} \frac{1}{\frac{p}{n}-\frac{1}{2}} x^{\frac{1}{2}-\frac{p}{n}} \Big|_1^\infty \to 0$$

and

$$J(u_n) = \int_\Omega u_n^{2p} dx = \frac{1}{n^{\frac{1}{2}}} \int_\Omega \frac{1}{x^{\frac{p}{n}+1}} dx = -\frac{\sqrt{n}}{p} x^{-\frac{p}{n}} \Big|_1^\infty \to \infty$$

Therefore we have $u_n \to 0$ in $\|\cdot\|_{W^{k,p}}$ while $J(u_n) \to \infty$ and $J(0) = 0$.

Additionally, when manipulating first or second variation integrals (2), (6) we wish to be able to bring the derivative inside the integral to relate properties of variations to those of L, u and φ. For this reason we require $L \in C^1$ when studying the first variation and $L \in C^2$ for the second variation.

Choice of Ω bounded and $p = \infty$

For multiple reasons we restrict our attention to the case where Ω is bounded and $p = \infty$ for the norm on $W = C^k(\overline{\Omega}) \cap W^{k,p}(\Omega)$.

- If Ω is bounded then $L \in C^0 \cap L^1$ is enough for J to be well defined on W (lemma 2.6).
- Only with $p = \infty$ we get continuity of $J: W \to \mathbb{R}$ with respect to $\|\cdot\|_{W^{k,p}(\Omega)}$ (lemma 2.7).
- Only with $p = \infty$ we get $C^k(\overline{\Omega})$ to be a closed subspace and hence a Banach space (Theorem 1.1). Furthermore, if Ω is bounded we can simply write $W = C^k(\overline{\Omega})$ since $C^k(\overline{\Omega}) \subset W^{k,\infty}(\Omega)$.
- As we shall see, to get continuity of $\delta J(u, \varphi)$ in u which is required for Fréchet differentiability (Theorems 2.2, 2.3) we need $\varphi \in W^{k,1}(\Omega)$ (Theorem 3.1). Similarly, for continuity of $\delta^2 J(u, \varphi)$ in u we need $\varphi \in W^{k,2}(\Omega)$ (Theorem 4.1) Both of these conditions are automatically satisfied when $p = \infty$ and Ω is bounded and are messy otherwise.

This norm generates the smallest neighborhoods amongst all $\|\cdot\|_{W^{k,p}}$ norms and thus extrema with respect to $\|\cdot\|_{W^{k,\infty}}$ are the "weakest" extrema in the sense that all $\|\cdot\|_{W^{k,p}}$ extrema are $\|\cdot\|_{W^{k,\infty}}$ extrema but not vice versa (see section 5.1 on domination relations amongst norms). The norm $\|\cdot\|_{W^{k,\infty}}$ is known as the "weak norm on $C^k(\overline{\Omega})$" in contrast to the "strong norm" $\|\cdot\|_{W^{0,\infty}}$ which only controls changes in u itself and not its derivatives.

Definition 2.6. ***Regularity conditions for Variational Problem (1).***

For the variational problem (1) we consider the Banach space $W := C^k(\overline{\Omega}) \cap W^{k,\infty}(\Omega)$ with the norm $\|\cdot\|_{W^{k,\infty}}$. We further assume that L is bounded and continuous. Boundedness of Ω will be noted when it is required.

Perturbations

By a mollification argument (section 5.4) we restrict our attention to $\varphi \in C^\infty(\overline{\Omega}) \subset W$ which is known to be dense in $W^{k,p}(\Omega)$. In the constrained boundary case, a similar argument allows us to restrict our attention to $\varphi \in C_c^\infty(\Omega) \subset W$ [2].

§ 3 First Variation

Suppose the conditions in definition 2.6 for variational problem (1) are satisfied and that additionally, the Lagrangian $L \in C^1 \cap W^{1,\infty}$, i.e. continuously differentiable with bounded derivatives.

The definition of first variation (2) reads:

$$\delta J(u,\varphi) = \frac{d}{ds}\Big|_{s=0} \int_\Omega L^{(u+s\varphi)} dx = \int_\Omega \frac{d}{ds}\Big|_{s=0} L^{(u+s\varphi)} dx$$

Since L is C^1 we get

$$\delta J(u,\varphi) = \int_\Omega \begin{bmatrix} \varphi & D\varphi & \dots \end{bmatrix} \begin{bmatrix} L_z^{(u)} \\ L_p^{(u)} \\ \vdots \end{bmatrix} dx = \int_\Omega \sum_{|\alpha|\le k} D^\alpha \varphi L_\alpha^{(u)} dx \tag{7}$$

Clearly $\varphi \mapsto \delta J(u,\varphi)$ is linear. In addition, since L and its derivatives are bounded, the first variation is also bounded and thus J is Gâteaux differentiable everywhere. To get Fréchet differentiability and thus Taylor approximation, we need $u \mapsto \delta J(u,\varphi)$ to be continuous for all φ (Theorem 2.2).

▶ **Theorem 3.1** *If $\|\varphi\|_{W^{k,1}(\Omega)} < \infty$ then the map $u \mapsto \delta J(u,\varphi)$ is continuous in u.*

Proof. Based on (7) we have

$$|\delta J(u,\varphi) - \delta J(v,\varphi)| \le \int_\Omega \sum_{|\alpha|\le k} |D^\alpha \varphi| \left| L_\alpha^{(u)} - L_\alpha^{(v)} \right| dx$$

Now since $L \in C^1$, for any $\epsilon > 0$ we can find a $\delta > 0$ such that $\|u - v\|_{W^{k,\infty}(\Omega)} < \delta$ implies

$$\forall |\alpha| \le k, \quad \forall x \in \Omega : |L_\alpha^{(u)} - L_\alpha^{(v)}| < \epsilon$$

For such a δ and $\|u - v\|_{W^{k,\infty}(\Omega)} < \delta$ we get

$$|\delta J(u,\varphi) - \delta J(v,\varphi)| \le \epsilon \int_\Omega \sum_{|\alpha|\le k} |D^\alpha \varphi| dx = \epsilon \|\varphi\|_{W^{k,1}(\Omega)}$$

□

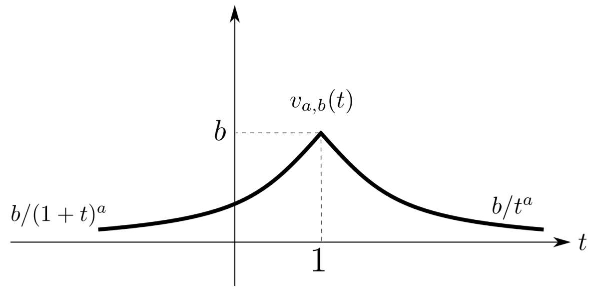

Figure 3: Family of functions $v_{a,b}(t)$ for Example 3.1.

Corollary 3.1.1. The requirement $\|\varphi\|_{W^{k,1}(\Omega)} < \infty$ is always satisfied when Ω is bounded since $L^\infty(\Omega) \subset L^1(\Omega)$ (Theorem 1.2). Therefore, when Ω is bounded J is everywhere Fréchet differentiable.

Example 3.1 $\|\varphi\|_{W^{k,1}(\Omega)} < \infty$ **necessary for continuity of** $\delta J(\cdot, \varphi)$ **when** Ω **unbounded.**

Let $n = 1, k = 0$, $\Omega = \mathbb{R}$ and consider the variational problem:

$$J(u) = \int_{\mathbb{R}} u^2 dt$$

The first variation is

$$\delta J(u, \varphi) = 2 \int_{\mathbb{R}} \varphi u dt$$

Consider the family of functions $v_{a,b}(t)$ depicted in Fig. 3. We wish to find a perturbation $\varphi : W^{k,\infty}(\Omega) \setminus W^{k,1}(\Omega)$ and a sequence of functions u_m such that $u_m \to u$ in $\|\cdot\|_{W^{k,\infty}}$ but $\delta J(u_m, \varphi) \not\to \delta J(u, \varphi)$. Pick $u \equiv 0$ and let

$$\varphi = v_{\alpha,1}(t), \quad u_m = v_{a_m,b_m}(t)$$

Note that $\delta J(u, \varphi) = 0$. We thus seek α and sequences b_m, a_m such that $J(u_m)$ is well defined, all u_m are in $W^{k,\infty}(\Omega)$, $\|u_m\|_{L^\infty} \to 0$, and $\|\dot{u_m}\|_{L^\infty} \to 0$, yet

$$\delta J(u_m, \varphi) = 2 \int_{\mathbb{R}} \varphi u dt = 4 \int_1^\infty \frac{b_m}{t^{\alpha + a_m}} \not\to 0$$

For $J(u_m)$ to be well defined we need $a_m > \frac{1}{2}$. For $\|u_m\|_{L^\infty} \to 0$ we need $b_m \to 0$ and for $\|\dot{u_m}\|_{L^\infty} \to 0$ we need $a_m b_m \to 0$. For the first variation to not converge to 0 as $b_m \to 0$ while maintaining $u_m \in W^{k,\infty}(\Omega)$ we require $\alpha + a_m \uparrow \frac{1}{2}$ and for $\varphi : W^{k,\infty}(\Omega) \setminus W^{k,1}(\Omega)$ we require $0 < \alpha \leq 1$. All these conditions can be satisfied with:

$$\alpha \in \left(0, \frac{1}{2}\right), \quad b_m \downarrow 0, \quad a_m \uparrow 1 - \alpha$$

▶ **Theorem 3.2** ***Essential first order necessary condition.*** *Suppose the conditions in definition 2.6 for variational problem (1) are satisfied as well as $L \in C^1 \cap W^{1,\infty}$. Suppose $\bar{u} \in W$ is a local extremum of J and that J is Fréchet differentiable at $\bar{u}$ (e.g. if Ω bounded or if $u\Big|_{\partial\Omega}$ is constrained such that we limit ourselves to $\varphi \in C_c^\infty(\Omega)$). Then the derivative of J at $\bar{u}$ is identically zero:*

$$\forall \varphi \in C^\infty(\Omega) : \delta J(\bar{u}, \varphi) = 0$$

Proof. Suppose the first variation does not vanish for some unit $\hat{\varphi}$, say $\delta J(\bar{u}, \hat{\varphi}) = C \neq 0$. By Taylor approximation (Theorem 2.4) we have for any $\alpha \in \mathbb{R}$:

$$J(u + \alpha\hat{\varphi}) = J(u) + C\alpha + c_\alpha \alpha$$

with $c_\alpha \to 0$ as $\alpha \to 0$. Since $C \neq 0$ we can always get $\bar{u} + \alpha\hat{\varphi}$ arbitrarily close to $\bar{u}$ such that $J(\bar{u} + \alpha\hat{\varphi}) < J(\bar{u})$ or $J(\bar{u} + \alpha\hat{\varphi}) > J(\bar{u})$ as required to contradict local extremality of $\bar{u}$. □

§ 3.1 Euler-Lagrange Equations

By restricting our attention to perturbations $\varphi \in C_c^\infty(\Omega) \subset C^\infty(\Omega)$ and applying an appropriate variant of the fundamental lemma of calculus of variations (section 5.3) we obtain the Euler-Lagrange equations. In the constrained boundary case that is all we can extract from the essential first order necessary condition 3.2. Here we spell out the Euler-Lagrange equations for $k = 1, 2$.

▶ **Theorem 3.3** *Suppose the hypothesis of Theorem 3.2 is satisfied. If $k = 1$ we have*

$$L_z^{(\bar{u})} - \operatorname{div}(L_p^{(\bar{u})}) = 0 \quad in \quad \Omega$$

and if $k = 2$ we have

$$L_z^{(\bar{u})} - \operatorname{div}(L_p^{(\bar{u})}) + \sum_{i,j=1}^{n} \partial_{x_i x_j} L_{q_i q_j}^{(\bar{u})} = 0 \quad in \quad \Omega$$

Proof. Apply Theorem 5.6 when $k = 1$ and Theorem 5.7 when $k = 2$ to (7). □

§ 3.2 Transversality Conditions

When no boundary constraints are imposed on the variational problem (1) we can extract more information about a local extrema by also considering $\varphi \in C^\infty(\Omega)$ that do not vanish on the boundary Ω. Here we spell out the transversality conditions for $k = 1, 2$.

▶ **Theorem 3.4** *Let Ω be bounded with C^1 boundary and suppose the hypothesis of Theorem 3.2 is* *satisfied. If $k = 1$ we have*

$$L_p^{(\bar{u})} \cdot \hat{\mathbf{n}} \quad on \quad \partial\Omega$$

and if $k = 2$ we have

$$\operatorname{div}\left(L_q^{(\bar{u})}\hat{\mathbf{n}}\right) = \hat{\mathbf{n}} \cdot \left[L_p^{(\bar{u})} - \left(\operatorname{div}(L_{q_{ij}}^{(\bar{u})})_i\right)_j\right] \quad on \quad \partial\Omega$$

where $\hat{\mathbf{n}}$ is the outward unit normal, and L_q is the $n \times n$ matrix of partials of L with respect to q_{ij} corresponding to $\partial_{x_i}\partial_{x_j}u$.

Proof. When $k = 1$ by Theorem 3.3 and equation (12) in the proof of Theorem 5.6 we get:

$$\forall \varphi \in C^\infty(\Omega) : \int_{\partial\Omega} \varphi L_p^{(\bar{u})} \cdot \hat{\mathbf{n}} dS = 0$$

Another application of the fundamental lemma (Theorem 5.5) gives the desired result.

When $k = 2$ by Theorem 3.3 and equation (13) in the proof of Theorem 5.7 we get

$$\int_{\partial\Omega} D\varphi \cdot L_q^{(\bar{u})}\hat{\mathbf{n}} dx + \int_{\partial\Omega} \varphi\hat{\mathbf{n}} \cdot \left[L_p^{(\bar{u})} - \left(\operatorname{div}(L_{q_{ij}}^{(\bar{u})})_i\right)_j\right] dx$$

By an application of the fundamental lemma (Theorem 5.6) we get the desired result. □

§ 4 Second Variation

Suppose the conditions in definition 2.6 for variational problem (1) are satisfied and that additionally, the Lagrangian $L \in C^2 \cap W^{2,\infty}$, i.e twice continuously differentiable with bounded derivatives upto second order. As before, the definition of second directional derivative with the assumption that L is C^2 gives:

$$\begin{aligned}\delta^2 J(u, \varphi, \psi) &= \int_\Omega \begin{bmatrix}\varphi & D\varphi^\top & \dots\end{bmatrix} \begin{bmatrix} L_{zz} & L_{zp}^\top & \dots \\ L_{zp} & L_{pp} & \dots \\ \vdots & \ddots & \dots \end{bmatrix} \begin{bmatrix}\psi \\ D\psi \\ \vdots\end{bmatrix} dx \\ &= \int_\Omega \sum_{|\alpha|,|\beta| \le k} D^\alpha \varphi L_{\alpha,\beta}^{(u)} D^\beta \psi dx\end{aligned} \tag{8}$$

Clearly $\varphi, \psi \mapsto \delta^2 J(u, \varphi, \psi)$ is bilinear. In addition, since L and its second derivatives are bounded, the second variation is also bounded and thus J is second order Gâteaux differentiable everywhere. To get Fréchet differentiability and thus Taylor approximation, we need $u \mapsto \delta J^2(u, \varphi, \psi)$ to be continuous for all φ, ψ (Theorem 2.3).

▶ **Theorem 4.1** *If $\|\varphi\|_{W^{k,2}(\Omega)}, \|\psi\|_{W^{k,2}(\Omega)} < \infty$ then the map $u \mapsto \delta^2 J(u, \varphi, \psi)$ is continuous.*

Proof. Based on (8) we have

$$|\delta^2 J(u,\varphi,\psi) - \delta^2 J(v,\varphi,\psi)| \leq \int_\Omega \sum_{|\alpha|,|\beta|\leq k} |D^\alpha \varphi| \left| L^{(u)}_{\alpha,\beta} - L^{(v)}_{\alpha,\beta} \right| |D^\beta \psi| dx$$

Now since $L \in C^2$, for any $\epsilon > 0$ we can find a $\delta > 0$ such that $\|u - v\|_{W^{k,\infty}(\Omega)} < \delta$ implies

$$\forall |\alpha| \leq k, \quad \forall x \in \boldsymbol{\Omega} : |L^{(u)}_{\alpha,\beta} - L^{(v)}_{\alpha,\beta}| < \epsilon$$

For such a δ and $\|u - v\|_{W^{k,\infty}(\Omega)} < \delta$ we get

$$\begin{aligned}
|\delta J^2(u,\varphi,\psi) - \delta^2 J(v,\varphi,\psi)| &\leq \epsilon \int_\Omega \sum_{|\alpha|,|\beta|\leq k} |D^\alpha \varphi||D^\beta \psi| dx \\
&\leq \epsilon \int_\Omega \sum_{|\alpha|\leq k} |D^\alpha \varphi|^2 dx + \int_\Omega \sum_{|\alpha|\leq k} |D^\alpha \varphi|^2 dx \\
&= \epsilon \left[\|\varphi\|^2_{W^{k,2}(\Omega)} + \|\psi\|^2_{W^{k,2}(\Omega)} \right]
\end{aligned}$$

where in the last step we have used the Cauchy-Schwarz inequality. □

Corollary 4.1.1. The requirement $\|\varphi\|_{W^{k,2}(\Omega)}, \|\psi\|_{W^{k,2}(\Omega)} < \infty$ is always satisfied when $\boldsymbol{\Omega}$ is bounded since $L^\infty(\boldsymbol{\Omega}) \subset L^2(\boldsymbol{\Omega})$ (Theorem 1.2). Therefore, when $\boldsymbol{\Omega}$ is bounded J is automatically second order Fréchet differentiable everywhere.

§ 4.1 Necessary Conditions

▶ **Theorem 4.2** ***Essential second order necessary condition.*** *Suppose the conditions in definition 2.6 for variational problem (1) are satisfied as well as $L \in C^2 \cap W^{2,\infty}$. Suppose $\bar{u} \in W$ is a local minimum (maximum) of J and that J is twice Fréchet differentiable at $\bar{u}$ (e.g. if $\boldsymbol{\Omega}$ bounded). Then the second variation is positive (negative) semidefinite at $\bar{u}$, i.e.*

$$\forall \varphi \in C^\infty(\boldsymbol{\Omega}) : \delta^2 J(\bar{u},\varphi) \geq 0 \quad (\leq 0)$$

Proof. Without loss of generality we will only consider the case where $\bar{u}$ is a local minimum. Suppose the second variation $\delta^2 J(\bar{u},\hat{\varphi}) = -2C < 0$ for some unit $\hat{\varphi}$. By Taylor approximation (Theorem 2.5), and noting that $\delta J(\bar{u},\cdot)$ is identically zero by Theorem 3.2, we have for any $\alpha \in \mathbb{R}$

$$J(u + \alpha\hat{\varphi}) = J(u) - C\alpha^2 + c_\alpha \alpha^2$$

with $c_\alpha \to 0$ as $\alpha \to 0$. Since $C \neq 0$ we can always get $\bar{u} + \alpha\hat{\varphi}$ arbitrarily close to $\bar{u}$ such that $J(\bar{u} + \alpha\hat{\varphi}) < J(\bar{u})$. □

▶ **Theorem 4.3** ***Legendre Condition for*** $k = 1$**.** *If the hypothesis of Theorem 4.2 is satisfied then*

$$L_{pp}^{(\bar{u})} \succeq 0 \quad (\preceq 0) \quad \textit{in} \;\; \mathbf{\Omega}$$

where $\succeq$ *denotes positive-semidefiniteness of the* $n \times n$ *matrix.*

Proof. We write the second variation integral (8) for $k = 1$.

$$\begin{aligned} \delta^2 J(u,\varphi) &= \int_\Omega \begin{bmatrix} \varphi & D\varphi^\top \end{bmatrix} \begin{bmatrix} L_{zz}^{(\bar{u})} & {L_{zp}^{(\bar{u})}}^\top \\ L_{zp}^{(\bar{u})} & L_{pp}^{(\bar{u})} \end{bmatrix} \begin{bmatrix} \varphi \\ D\varphi \end{bmatrix} dx \\ (*) &= \int_\Omega \left[\varphi^2 L_{zz}^{(\bar{u})} + 2\varphi L_{zp}^{(\bar{u})} \cdot D\varphi + D\varphi^\top L_{pp}^{(\bar{u})} D\varphi \right] dx \end{aligned}$$

The key idea now is that the last term in $(*)$ is dominant in the sense that we can find custom perturbations (to be specified) φ_N such that

$$\left| D\varphi_N^\top L_{pp}^{(\bar{u})} D\varphi_N \right| \to \infty \quad \text{as} \quad N \to \infty$$

while $\left|\varphi_N^2 L_{zz}^{(\bar{u})}\right|$ and $\left|\varphi_N L_{zp}^{(\bar{u})} \cdot D\varphi_N\right|$ remain bounded. From this we will argue that if $L_{pp}^{(\bar{u})}(x) \succeq 0$ is violated at any x we can create perturbations that drive the dominant term to $-\infty$ hence making the second variation negative contradicting the standard necessary condition.

Suppose $L_{pp}^{(\bar{u})}(x_\circ) \not\succeq 0$ for some $x_\circ$; fix a negative eigenvalue $-\lambda$ with a unit eigenvector $\phi \in \mathbb{R}^n$. This means:

$$\phi^\top L_{pp}^{(\bar{u})}(x_\circ)\phi = -\lambda < 0$$

which by continuity of $L_{pp}^{(\bar{u})}$ implies that over some open ball $B := B(x_\circ, r)$ we have

$$\phi^\top L_{pp}^{(\bar{u})}(x)\phi < -\frac{\lambda}{2} < 0$$

We now consider the perturbations

$$\varphi_N(x) := f(N)\sin(N\phi \cdot x)\mathbf{1}_B(x)$$

with $f(N) > 0$ to be determined. For any vector a we have $D\sin(a \cdot x) = \cos(a \cdot x)a$. Thus

$$\begin{aligned} \delta^2 J(\bar{u},\varphi_N) &= \int_B f(N)^2 \sin^2(\phi \cdot x) L_{zz}^{(\bar{u})}\Big|_{x,\bar{u}} dx \\ &\quad + \int_B N f(N)^2 \sin(2\phi \cdot x)\phi \cdot L_{zp}^{(\bar{u})}\Big|_{x,\bar{u}} dx \\ &\quad + \int_B N^2 f(N)^2 \cos^2(\phi \cdot x)\phi^\top L_{pp}^{(\bar{u})}\Big|_{x,\bar{u}} \phi \\ &\le C_1 \left[f(N)^2 + N f(N)^2 \right] - C_2 N^2 f(N)^2 \end{aligned}$$

where $C_1 = \text{vol}(B) \max \left\{ \sup |L_{zz}^{(\bar{u})}|, \sup |L_{zp}^{(\bar{u})}| \right\}$ and $C_2 = \text{vol}(B)\lambda/2$. By choosing $f(N) = N^{-3/4}$ we get

$$\delta^2 J(\bar{u}, \varphi_N) \leq C_1 \left[\frac{1}{N\sqrt{N}} + \frac{1}{\sqrt{N}} \right] - C_2\sqrt{N}$$

The first term on the right hand side vanishes as $N \to \infty$ while the second term goes to $-\infty$, that is, for large enough N we get $\delta^2 J(\bar{u}, \varphi_N) < 0$. □

▶ **Theorem 4.4** ***Legendre Condition for*** $k = 2$**.** *If the hypothesis of Theorem 4.2 is satisfied then*

$$L_{qq}^{(\bar{u})} \succeq 0 \quad (\preceq 0) \quad \textit{in} \;\; \Omega$$

where $\succeq$ *denotes positive-semidefiniteness of the* $n^2 \times n^2$ *matrix.*

Proof. We proceed along the same lines as the proof of Theorem 4.3.

$$\begin{aligned}
\delta^2 J(u, \varphi) &= \int_\Omega \begin{bmatrix} \varphi & D\varphi^\top & D^2\varphi^\top \end{bmatrix} \begin{bmatrix} L_{zz}^{(\bar{u})} & L_{zp}^{(\bar{u})\top} & L_{zq}^{(\bar{u})\top} \\ L_{zp}^{(\bar{u})} & L_{pp}^{(\bar{u})} & L_{pq}^{(\bar{u})\top} \\ L_{zq}^{(\bar{u})} & L_{pq}^{(\bar{u})} & L_{qq}^{(\bar{u})} \end{bmatrix} \begin{bmatrix} \varphi \\ D\varphi \\ D^2\varphi \end{bmatrix} dx \\
&= \int_\Omega \Big[\varphi^2 L_{zz}^{(\bar{u})} + 2\varphi L_{zp}^{(\bar{u})} \cdot D\varphi + D\varphi^\top L_{pp}^{(\bar{u})} D\varphi \\
&\qquad + D\varphi^\top L_{pp}^{(\bar{u})} D\varphi + 2D^2\varphi^\top L_{pq}^{(\bar{u})} D\varphi \\
&\qquad + D^2\varphi^\top L_{qq}^{(\bar{u})} D^2\varphi \Big] dx \qquad (*)
\end{aligned}$$

Again, the idea is that the last term in $(*)$ is dominant in the sense that we can find custom perturbations (to be specified) φ_N such that

$$\left| D^2\varphi_N^\top L_{qq}^{(\bar{u})} D^2\varphi_N \right| \to \infty \quad \text{as} \quad N \to \infty$$

while all other terms remain bounded. Suppose $L_{qq}^{(\bar{u})}(x_\circ) \not\succeq 0$ for some $x_\circ$; fix a negative eigenvalue $-\lambda$ with a unit eigenvector $\phi \in \mathbb{R}^{n \times n}$ which by continuity of $L_{qq}^{(\bar{u})}$ implies that over some open ball $B := B(x_\circ, r)$ we have

$$\phi^\top L_{qq}^{(\bar{u})}(x)\phi < -\frac{\lambda}{2} < 0$$

Now define $\psi(x) = x^T \phi x \mathbf{1}_B(x)$ and consider perturbations of the form

$$\varphi_N(x) = \frac{1}{\sqrt{N}} \psi\Big(N(x - x_\circ)\Big)$$

$$\int_B |\varphi_N| dx = \mathscr{O}\left(\frac{1}{N\sqrt{N}}\right) \to 0$$

$$\int_B |D\varphi_N| dx = \mathscr{O}\left(\frac{1}{\sqrt{N}}\right) \to 0$$
$$\int_B |D^2\varphi_N| dx = \mathscr{O}\left(\sqrt{N}\right) \to \infty$$

Noting that $D^2\varphi_N \parallel \phi$ everywhere and thus

$$\int_B D^2\varphi_N^\top L_{qq}^{(\bar{u})} D^2\varphi_N dx = \int_B |D^2\varphi_N|^2 \phi^\top L_{qq}^{(\bar{u})} \phi dx < -\lambda \mathscr{O}(N) \to -\infty$$

completes the proof (TODO: double check calculations). □

§ 4.2 Sufficient Conditions

This section is the main point of diversion between finite dimensional and infinite dimension Banach spaces. When $W = \mathbb{R}^n$ a sufficient condition for $\bar{x}$ to be a local minimum (maximum) of $f : W \to \mathbb{R}$ is that the gradient of f (i.e. first variation) vanishes at $\bar{x}$ and that the Hessian (i.e. second variation) is positive-definite at $\bar{x}$. This condition is no longer sufficient in the infinite dimensional case. We will first consider multiple examples showcasing what can go wrong in the infinite dimensional case before introducing sufficient conditions.

Example 4.1 Scheefer's example 1 [4] (positive-definite second variation not sufficient).

Let $n = k = 1$ and $\Omega = (-1, 1)$. Define

$$J(u) = \int_{-1}^{1} [t^2\dot{u}^2 + t\dot{u}^3] dt$$

with constrained boundary $u(\pm 1) = 0$. We show that at $\bar{u} \equiv 0$ the first variation vanishes and the second variation is positive-definite; yet $\bar{u} \equiv 0$ is not a local minimum of J. The Lagrangian is $L(t, z, p) = t^2p^2 + tp^3$ and its derivatives are:

$$L_z = 0, \quad L_p = 2t^2p + 3tp^2, \quad L_{pp} = 2t^2 + 6tp$$

evaluated at the candidate $\bar{u} \equiv 0$ we have:

$$L_z^{(\bar{u})} = 0, \quad L_p^{(\bar{u})} = 0, \quad L_{pp}^{(\bar{u})} = 2t^2$$

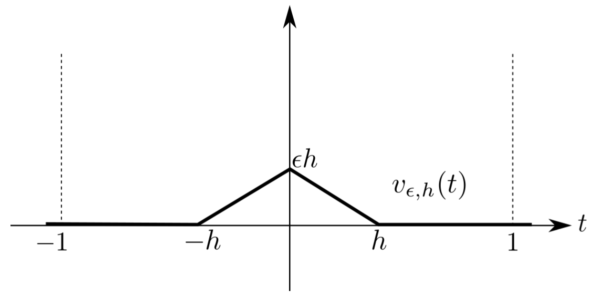

Figure 4: Perturbations $v_{\epsilon,h}$ for Example 4.1.

We also have

$$\begin{aligned}\int_{-1}^{1}|v_{\epsilon,h}|^2dt &= \int_{-h}^{h}\epsilon^2t^2dt = \frac{2}{3}h^3\epsilon^2\\ \int_{-1}^{1}|\dot{v}_{\epsilon,h}|^2dt &= \int_{-h}^{h}\epsilon^2dt = 2h\epsilon^2\\ \int_{-1}^{1}t^2\dot{v}_{\epsilon,h}^2dt &= \int_{-h}^{h}t^2\epsilon^2dt = \frac{2}{3}h^3\epsilon^2\\ \int_{-1}^{1}t\dot{v}_{\epsilon,h}^3ut &= -2\int_{0}^{h}t\epsilon^3dt = -h^3\epsilon^3\end{aligned} \tag{9}$$

The first variation clearly vanishes at $\bar{u}$ and the second variation is

$$\delta^2J(\bar{u},\varphi) = \int_{-1}^{1}L_{pp}^{(\bar{u})}\dot{\varphi}^2dt = \int_{-1}^{1}2t^2\dot{\varphi}^2dt > 0$$

where the last inequality holds for any φ not identically constant on $\boldsymbol{\Omega}$. Now consider the sequence of functions $v_{\epsilon,h}$ shown in Fig. 4. Clearly as $\epsilon, h \to 0$ we have $\epsilon, h \to \bar{u}$ in $\|\cdot\|_{W^{1,\infty}}$. However, we will show that if $\epsilon, h \to 0$ in an appropriately controlled fashion we will can maintain $J(v_{\epsilon,h}) < 0$ as $\epsilon, h \to 0$ thus showing that $\bar{u}$ is not a local minimum of J.

By (9) we have

$$J(v_{\epsilon,h}) = \frac{2}{3}h^3\epsilon^2 - h^2\epsilon^3 = \frac{2}{3}h^2\epsilon^2\left(h - \frac{3}{2}\epsilon\right)$$

Therefore if we set $h = \frac{3}{4}\epsilon$ and let $\epsilon \to 0$ we get $J(v_{\epsilon,h}) < J(\bar{u}) = 0$.

Example 4.2 Strengthened Legendre $L_{pp} \succ 0$ not sufficient.

One might hope that strengthening the Legendre condition, say for $k = 1$, to $L_{pp} \succ 0$ (almost) everywhere would guarantee a local minimum. Scheefer's example 4.1 dismisses this possibility as well since $L_{pp}^{(\bar{u})} = 2t^2$ is strictly positive almost everywhere in $\boldsymbol{\Omega}$. As we shall see, as long as $L_{pp}^{(\bar{u})}$ gets arbitrarily close to 0 somewhere in $\boldsymbol{\Omega}$ perturbations can "hide mass" at that very point.

Example 4.3 Scheefer's example (2).

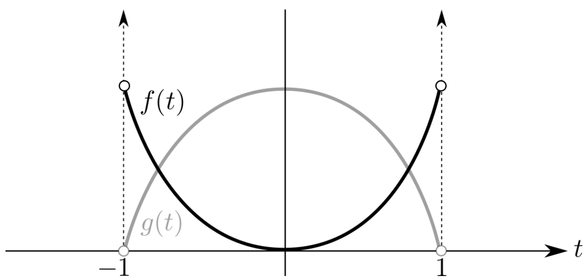

Figure 5: L_{pp} for modified Scheefer Example 4.3, f is $L_{pp}^{(\bar{u})}$ for Scheefer's Example 4.1 and g is that of modified example 4.3.

Suppose we modify Scheefer's example 4.1 as follows:

$$J(u) = \int_{-1}^{1} [(1 - t^2)\dot{u}^2 + t\dot{u}^3]dt$$

Now $L_{pp}^{(\bar{u})} = 2(1 - t^2)$ is positive and bounded away from zero on Ω (Fig. 5). We notice that our perturbations $v_{\epsilon,h}$ can no longer "hide their mass" at the origin:

$$J(v_{\epsilon,h}) = 2\epsilon^2 \left(1 - \frac{1}{3}h^2(h - 3\epsilon)\right)$$

To ensure $J(v_{\epsilon,h}) < 0$ we need to maintain $\frac{1}{3}h^2(h - 3\epsilon) > 1$ as $\epsilon, h \to 0$ which is impossible.

Example 4.4 Scheefer's example (3): L_{pp} bounded away from zero.

If we set

$$J(u) = \int_{-1}^{1} \left[\left(\frac{t^2 + a}{2}\right)\dot{u}^2 + t\dot{u}^3\right] dt$$

for the same perturbations $v_{\epsilon,h}$ as in Example 4.1 we have

$$J(v_{\epsilon,h}) = \epsilon^2 h \left(\frac{1}{3}h^2 - \epsilon + a\right)$$

which can be maintained as negative as $\epsilon, h \to 0$ only if $a = 0$.

Similarly, if we move the point where $v_{\epsilon,h}$ hide their mass by considering perturbations $v_{\epsilon,h}(t - a)$ we get:

$$J(v_{\epsilon,h}) = \frac{2}{3}\epsilon^2 h \left(h^2 - \frac{3}{2}\epsilon h + \frac{1}{3}a^2\right)$$

which can only be made negative as $\epsilon, h \to 0$ if $a = 0$.

Example 4.5 Scheefer's example (4): unconstrained.

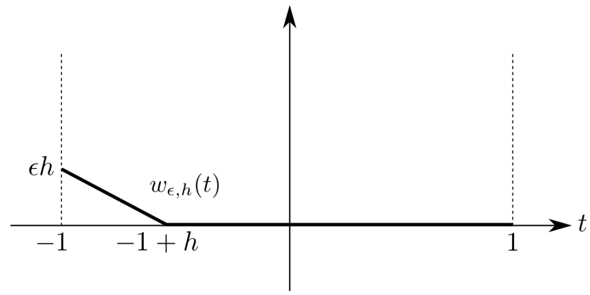

Figure 6: Perturbations $w_{\epsilon,h}$ for Example 4.5 without boundary constraint.

Consider the modified Scheefer's example from Example 4.3:

$$J(u) = \int_{-1}^{1} [(1-t^2)\dot{u}^2 + t\dot{u}^3]dt$$

As before $L_{pp}^{(\bar{u})} = 2(1-t^2)$ is positive and bounded away from zero on Ω (Fig. 5). If we remove the boundary constraint from the variational problem, we can design perturbations that hide their mass at ± 1 where L_{pp} approaches zero. For the perturbations in Fig. 6 we have:

$$J(w_{\epsilon,h}) = -\frac{\epsilon^2 h}{3}\left[h^2 - 3\left(1+\frac{\epsilon}{2}\right)h + 3\epsilon\right]$$

In order to maintain $J(w_{\epsilon,h}) < 0$ we need to have $h^2 - 3\left(1+\frac{\epsilon}{2}\right)h + 3\epsilon > 0$ which can be satisfied by maintaining

$$0 < h < \frac{3}{2}\left[1+\frac{\epsilon}{2} - \sqrt{1-\frac{\epsilon}{3}+\epsilon^2}\right]$$

as $\epsilon, h \to 0$.

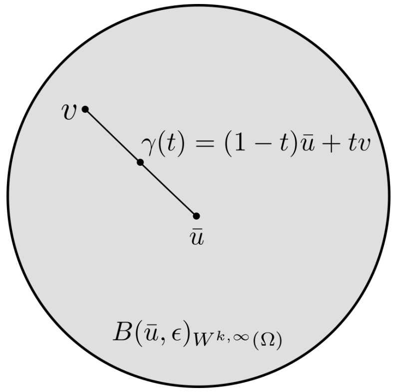

Figure 7: A weak neighborhood of candidate minimum $\bar{u}$.

▶ **Theorem 4.5** ***Positive-semidefinite second variation in a neighborhood is sufficient [4].*** *Suppose the conditions in definition 2.6 for variational problem (1) are satisfied as well as $L \in C^2 \cap W^{2,\infty}$. Suppose $\bar{u} \in W$ is a critical point of J (i.e. first variation vanishes) and in some $\|\cdot\|_{W^{k,\infty}}$ neighborhood U of $\bar{u}$ we have*

$$\forall v \in U,\ \forall \varphi \in C^\infty(\boldsymbol{\Omega}) :\ \delta^2 J(v, \varphi) \geq 0$$

Then $\bar{u}$ is a local minimum of J.

Proof. First note that for any real function $h \in C^2[0,1]$ we have

$$h(1) - h(0) = h'(0) + \int_0^1 (1-t)h''(t)dt$$

because:

$$\begin{aligned}\int_0^1 (1-t)h''(t)dt &= \int_0^1 h''(t)dt - \int_0^1 th''(t)dt \\ &= h'(1) - h'(0) - th'(t)\Big|_0^1 + \int_0^1 h'(t)dt \\ &= h'(1) - h'(0) - h'(1) + h(1) - h(0)\end{aligned}$$

We will now show that for every $v \in U$ we have $J(v) \geq J(\bar{u})$. Let γ be the line segment connecting $\bar{u}$ to v (Fig. 7), i.e. $\gamma : [0,1] \to U$ given by

$$\gamma(t) = \bar{u} + t(v - \bar{u}) = (1-t)\bar{u} + tv$$

Define $h(t) = J(\gamma(t))$ which gives

$$h'(t) = \delta J(\gamma(t), v - \bar{u}), \quad h''(t) = \delta^2 J(\gamma(t), v - \bar{u})$$

Applying the above calculus fact to h completes the proof. □

▶ **Theorem 4.6** ***Strongly positive second variation in*** $\|\cdot\|_{W^{k,\infty}(\Omega)}$ ***is sufficient [3].*** *Suppose the conditions in definition 2.6 for variational problem (1) are satisfied as well as* $L \in C^2 \cap W^{2,\infty}$. *Suppose* $\bar{u} \in W$ *is a critical point of* J *and that* J *is second Fréchet differentiable (e.g. if* Ω *bounded or if* $u\Big|_{\partial\Omega}$ *is constrained such that we limit ourselves to* $\varphi \in C_c^\infty(\Omega)$*). If the second variation is strongly positive in* $\|\cdot\|_{W^{k,\infty}(\Omega)}$*, that is, for some* $K > 0$ *we have*

$$\forall \varphi \in C^\infty(\Omega): \ \delta^2 J(v, \varphi) \geq K \|\varphi\|^2_{W^{k,\infty}(\Omega)}$$

Then $\bar{u}$ *is a local minimum of* J.

Proof. Since J is Fréchet differentiable we can write the Taylor approximation theorem 2.5

$$J(\bar{u} + \varphi) - J(\bar{u}) = \frac{1}{2}\delta^2 J(\bar{u}, \varphi) + \alpha \|\varphi\|^2_{W^{k,\infty}(\Omega)}$$

Now by strong positivity we have:

$$J(\bar{u} + \varphi) - J(\bar{u}) \geq \left(\frac{K}{2} + c_\varphi\right) \|\varphi\|^2_{W^{k,\infty}(\Omega)}$$

and since $c_\varphi \to 0$ as $\varphi \to 0$ while $K > 0$ we can find a small neighborhood of $\bar{u}$, obtained by demanding $|c_\varphi| < K/2$, such that

$$\forall v \in B(\bar{u}, \epsilon) : J(v) - J(\bar{u}) \geq 0$$

□

▶ **Theorem 4.7** ***Strongly positive second variation in*** $\|\cdot\|_{W^{k,2}(\Omega)}$ ***is sufficient [4].*** *Suppose the conditions in definition 2.6 for variational problem (1) are satisfied as well as* $L \in C^2 \cap W^{2,\infty}$. *Suppose* $\bar{u} \in W$ *is a critical point of* J *and we have somehow guaranteed that all perturbations* φ *have bounded* $\|\cdot\|_{W^{k,2}(\Omega)}$ *norm (e.g. if* Ω *bounded or if* $u\Big|_{\partial\Omega}$ *is constrained such that we limit ourselves to* $\varphi \in C_c^\infty(\Omega)$*). If the second variation is strongly positive in* $\|\cdot\|_{W^{k,2}(\Omega)}$*, that is, for some* $K > 0$ *we have*

$$\forall \varphi \in C^\infty(\Omega): \ \delta^2 J(v, \varphi) \geq K \|\varphi\|^2_{W^{k,2}(\Omega)}$$

Then $\bar{u}$ *is a local minimum of* J.

Proof. Since we are not assuming Fréchet differentiability we cannot use Taylor approximation theorems and will rely on the first sufficient condition, Theorem 4.5. First, note that strong positivity can be rewritten as:

$$\forall \varphi \in C^\infty(\Omega) \text{ s.t. } \|\varphi\|_{W^{k,2}(\Omega)} = 1, \ \delta^2 J(v, \varphi) \geq K$$

Therefore if for any φ we find a $\|\cdot\|_{W^{k,2}(\Omega)}$ neighborhood V_φ of $\bar{u}$ such that

$$\forall v \in V_\varphi : \ |\delta^2 J(\bar{u}, \varphi) - \delta^2 J(v, \varphi)| < \frac{K}{2}$$

and thus

$$\forall v \in V_\varphi : \ \delta^2 J(v, \varphi) \geq \frac{K}{2} > 0$$

The only detail that remains before applying Theorem 4.5 is to ensure that the set $\cap_\varphi V_\varphi$ is still an open set. We recall from the proof of Theorem 4.1 that since $L \in C^2$, for any $\epsilon > 0$ we can find a $\delta > 0$ such that $\|\bar{u} - v\|_{W^{k,\infty}(\Omega)} < \delta$ implies

$$|\delta J^2(\bar{u}, \varphi) - \delta^2 J(v, \varphi)| \leq 2\epsilon \|\varphi\|^2_{W^{k,2}(\Omega)}$$

Therefore to get $|\delta^2 J(\bar{u}, \varphi) - \delta^2 J(v, \varphi)| < \frac{K}{2}$ we need to pick $\epsilon > 0$ such that

$$\epsilon \|\varphi\|^2_{W^{k,2}(\Omega)} < \frac{K}{2}$$

which can be done without sending $\epsilon \to 0$ over different φ with $\|\varphi\|_{W^{k,2}(\Omega)} = 1$. □

▶ **Theorem 4.8** ***Strong positive-definiteness of the Hessian of the Lagrangian is sufficient.***

Suppose the conditions in definition 2.6 for variational problem (1) are satisfied as well as $L \in C^2 \cap W^{2,\infty}$. Suppose $\bar{u} \in W$ is a critical point of J and that the matrix

$$A^{(\bar{u})} := \begin{bmatrix} L^{(\bar{u})}_{zz} & {L^{(\bar{u})}_{zp}}^\top & \dots \\ L^{(\bar{u})}_{zp} & L^{(\bar{u})}_{pp} & \dots \\ \vdots & \ddots & \dots \end{bmatrix}$$

is positive definite and has its smallest eigenvalue bounded away from zero everywhere in Ω. Then $\bar{u}$ is a local minimum of J.

Proof. By strong positive-definiteness of $A^{(\bar{u})}$ we have for some $K > 0$

$$\forall x \in \Omega, \ \forall \phi \in \mathbb{R}^m, \ \phi^\top A^{(\bar{u})}(x) \phi \geq K|\phi|^2$$

where m is the dimension of the A matrix. Now by (8) we have

$$\delta^2 J(u, \varphi) = \int_\Omega \begin{bmatrix} \varphi & D\varphi^\top & \dots \end{bmatrix} A^{(\bar{u})} \begin{bmatrix} \varphi \\ D\varphi \\ \vdots \end{bmatrix} dx \geq \int_\Omega K|\varphi|^2 + K|D\varphi|^2 + K|D^2\varphi|^2 + \dots$$

which now reduces our claim to strong positivity in $\|\cdot\|_{W^{k,2}(\Omega)}$ (Theorem 4.7). □

▶ **Theorem 4.9** ***Jacobi's sufficient condition for constrained problem with*** $k = 1$***[3, 4].***

Let $k = 1$ and Ω bounded. Suppose the conditions in definition 2.6 for variational problem (1) with constrained boundary are satisfied as well as $L \in C^2 \cap W^{2,\infty}$. Suppose $\bar{u} \in W$ is a critical point of J and that the strengthened Legendre condition is satisfied at $\bar{u}$, that is

$$L_{pp}^{(\bar{u})} \succ 0 \quad \textit{on } \Omega$$

with the smallest eigenvalue of $L_{pp}^{(\bar{u})}$ bounded away from zero. Further, suppose the Jacobi PDE

$$\left[L_{zz}^{(\bar{u})} - \operatorname{div}\left(L_{zp}^{(\bar{u})}\right)\right] v - \operatorname{div}\left(L_{pp}^{(\bar{u})} Dv\right) = 0$$

has a nonvanishing solution $v : \Omega \to \mathbb{R}$. Then $\bar{u}$ is a local minimum of J.

Proof. We will show by a completing-the-square argument that the second variation is strongly positive in $\|\cdot\|_{W^{k,2}(\Omega)}$ and then apply Theorem 4.7.

Define $Q = L_{zz}^{(\bar{u})} - \operatorname{div}\left(L_{zp}^{(\bar{u})}\right)$ and $P = L_{pp}^{(\bar{u})}$. We first claim that under the hypothesis the system of PDEs in $w : \Omega \to \mathbb{R}^n$

$$Q + \operatorname{div}(w) = w^\top P^{-1} w \tag{10}$$

has a solution. This is a generalization of the Riccati ODE in $n = 1$. Base on this observation, if we seek solutions w of the form

$$w = -\frac{1}{v} P D v$$

then v must be a nonvanishing solution to Jacobi's PDE

$$Qv - \operatorname{div}(PDv) = 0$$

Therefore, by hypothesis, (10) has a solution $w : \Omega \to \mathbb{R}^n$.

Using a similar integration by parts as in Theorem 5.6 we have for any u, φ

$$\delta^2 J(u,\varphi) = \int_\Omega \left[\left(L_{zz}^{(u)} - \operatorname{div}\left(L_{zp}\right)^{(u)}\right)\varphi^2 + D\varphi^\top L_{pp}^{(u)} D\varphi\right] dx = \int_\Omega \left[Q\varphi^2 + D\varphi^\top P D\varphi\right] dx$$

For any $w : \Omega \to \mathbb{R}^n$ we have

$$\forall \varphi \in C_c^\infty(\Omega) : \int_\Omega \operatorname{div}(\varphi^2 w) dx = 0$$

by the divergence theorem. Expanding this gives

$$\forall \varphi \in C_c^\infty(\Omega) : \int_\Omega \left[\varphi^2 \operatorname{div}(w) + 2\varphi D\varphi^\top w\right] dx = 0$$

We now pick w such that it solves (10) and use it to complete the square in the second variation:

$$\begin{aligned}\delta^2 J(\bar{u},\varphi) &= \int_\Omega \left[Q\varphi^2 + D\varphi^\top P D\varphi\right] dx \\ &= \int_\Omega \left[D\varphi^\top P D\varphi + [Q + \operatorname{div}(w)]\varphi^2 + 2\varphi D\varphi^\top\right] dx \\ &= \int_\Omega \left[D\varphi + \varphi P^{-1} w\right]^\top P \left[D\varphi + \varphi P^{-1} w\right] dx \geq 0\end{aligned}$$

where in the last step we have used the fact that w solves (10) and that P is positive-definite. Let $K > 0$ be a lower bound on the smallest eigenvalue of P on $\boldsymbol{\Omega}$. By the strengthened Legendre condition we have: $P - \frac{K}{2} I \succ 0$ everywhere in $\boldsymbol{\Omega}$. Therefore by the same argument as above, replacing P with $P - \frac{K}{2} I$ we get:

$$\int_\Omega Q\varphi^2 + D\varphi^\top \left(P - \frac{K}{2}\right) D\varphi \geq 0$$

which implies

$$\delta^2 J(\bar{u},\varphi) = \int_\Omega Q\varphi^2 + D\varphi^\top P D\varphi \geq \frac{K}{2} \int_\Omega |D\varphi|^2 dx$$

Now boundedness of $\boldsymbol{\Omega}$ and the Poincaré inequality (Theorem 1.2) imply that the second variation at $\bar{u}$ is strongly positive in $\|\cdot\|_{W^{k,2}(\Omega)}$ and our claim follows from Theorem 4.7.

□

Remark. For any fixed $u \in W$, the Jacobi PDE

$$Qv - \operatorname{div}(PDv) = 0$$

where $Q = L_{zz}^{(u)} - \operatorname{div}\left(L_{zp}^{(u)}\right)$ and $P = L_{pp}^{(u)}$ is precisely the Euler-Lagrange equation for the second variation map at u, namely $\tilde{J} : v \mapsto \delta^2 J(u,v)$ [3, 4].

§ 5 Appendix

§ 5.1 Inequivalent Norms

Fix any vector space V. We say that a norm $\|\cdot\|_\downarrow$ is *dominated* by another norm $\|\cdot\|_\uparrow$ if for some $C > 0$ we have:

$$\forall v : \|v\|_\downarrow \leq C \|v\|_\uparrow$$

Two norms are *equivalent* if they both dominate each other, i.e for some $C, C' > 0$:

$$\forall v : C' \|v\|_\uparrow \leq \|v\|_\downarrow \leq C \|v\|_\uparrow$$

Equivalence of norms is clearly an equivalence relation.

▶ **Theorem 5.1** *If* $\dim(V) < \infty$ *then all norms are equivalent.*

Proof. Let $\|\cdot\|$ be any norm on V. Suppose $\dim(V) = n$ and pick any basis $e_1, \ldots, e_n$ such that for all i, $\|e_i\| = 1$. It suffices to show that $\|\cdot\|$ is equivalent to the sup norm with respect to the chosen basis. We have, by properties of a norm, $\|v\| = \|\sum_i v_i e_i\| \leq \sum |v_i| \leq n \max_i |v_i|$. For the opposite direction, for any i we have $\|v\| \geq |v_i|$ and thus $\|v\| \geq \max_i |v_i|$.

□

Corollary 5.1.1. For this reason we never worry about the norm used on $\mathbb{R}^n$, e.g. we ambiguously write $|Du|$ for a C^1 function on $\mathbb{R}^n$ and keep in mind that all norms are equivalent to the standard Euclidean norm.

We now list important consequences of domination relations between norms. We denote by $B(x,r)_\downarrow, B(x,r)_\uparrow$ open balls with respect to each norm.

▶ **Theorem 5.2** *Let* V *be a vector space with two norms such that* $\|\cdot\|_\downarrow \leq C \|\cdot\|_\uparrow$ *for some* $C > 0$. *Then:*

1) *any open ball in* $\|\cdot\|_\uparrow$ *is contained in some open ball in* $\|\cdot\|_\downarrow$.
   Specifically $\forall x, r : B(x,r)_\uparrow \subset B(x, Cr)_\downarrow$.
2) *if* $\bar{v}$ *is a local minimum (maximum) of any function* $f : V \to \mathbb{R}$ *in* $\|\cdot\|_\downarrow$ *then it is also a local minimum (maximum) in* $\|\cdot\|_\uparrow$.
3) *convergence in* $\|\cdot\|_\uparrow$ *implies convergence in* $\|\cdot\|_\downarrow$.

Furthermore, given another normed vector space X *with fixed norm:*

4) *continuity at* $x \in X$ *of* $f : X \to V$ *in* $\|\cdot\|_\uparrow$ *implies continuity in* $\|\cdot\|_\downarrow$
5) *continuity at* $v \in V$ *of* $f : V \to X$ *at in* $\|\cdot\|_\downarrow$ *implies continuity in* $\|\cdot\|_\uparrow$.

Proof. definition chasing. □

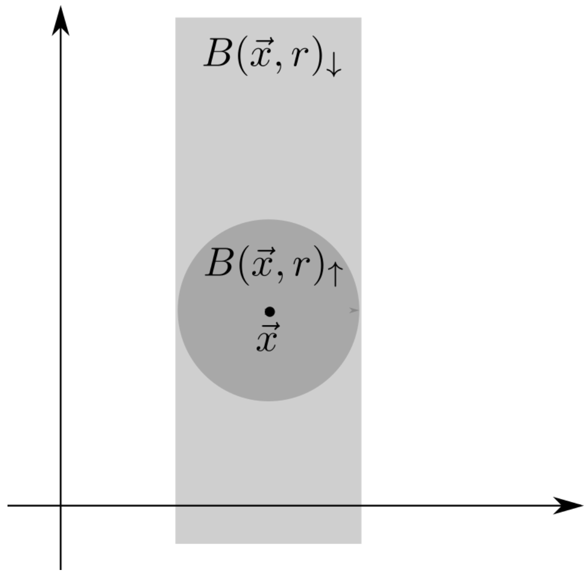

Figure 8: Two inequivalent seminorms on $\mathbb{R}^2$. Open balls in $\|\cdot\|_\uparrow : x \mapsto \sqrt{|x_1|^2 + |x_2|^2}$ can always be contained in open balls in $\|\cdot\|_\downarrow : x \mapsto |x_1|$ but not vice versa.

Remark. In all the above cases the opposite implications do not hold, e.g. open balls in $\|\cdot\|_\downarrow$ are not necessarily contained in some open ball in $\|\cdot\|_\uparrow$. As an analogy to gain intuition we can use seminorms in Euclidean space, which unlike norms, are not all equivalent. Recall that a seminorm satisfies all the properties of a norm except that $\|v\| = 0$ does not necessarily imply $v = 0$ (Fig. 8).

§ 5.2 Continuity of (Multi)linear Functionals

Given a linear functional $f : V \to \mathbb{R}$ the operator norm is defined to be:

$$\|f\| \coloneqq \sup_v \frac{|f(v)|}{\|v\|}$$

which by linearity can also be read as

$$\|f\| \coloneqq \sup_{\|v\|=1} |f(v)|$$

We say that f is bounded if $\|f\| < \infty$.

▶ **Theorem 5.3** *A linear map $f : V \to \mathbb{R}$ is continuous if and only if it is bounded.*

Proof. Continuity of f means $v_n \to v$ always implies $f(v_n) \to f(v)$. By linearity it suffices to only check this at the origin, i.e. continuity is equivalent to whether $v_n \to 0$ always implies $f(v_n) \to 0$.

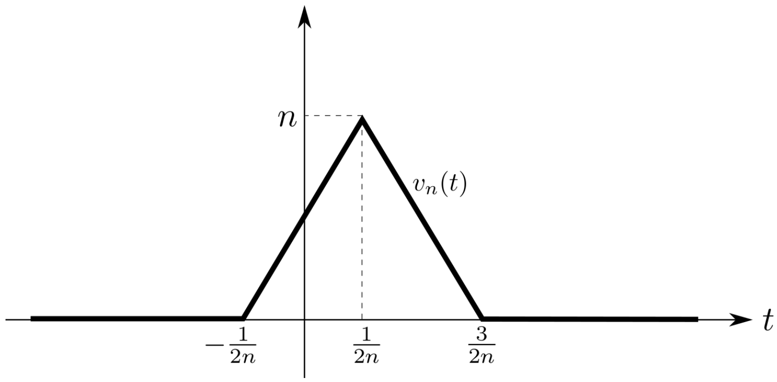

Figure 9: A sequence $v_n \in L^1(0,1)$ with $\|v_n\|_{L^1} = 1$ and unbounded $\dot{v}_n(0)$.

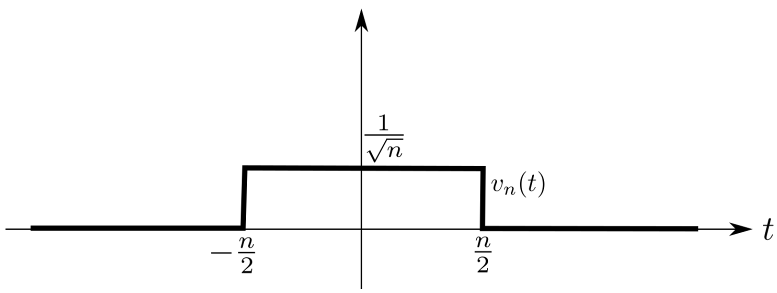

Figure 10: A sequence $v_n \in L^2 \cap L^1$ with $\|v_n\|_{L^2} = 1$ and unbounded $\int v_n dt$.

We now consider any such sequence $v_n = \|v_n\| \, \hat{v}_n$ and write

$$f(v_n) = \|v_n\| \, f(\hat{v}_n)$$

Now $v_n \to 0$ always implies $f(v_n) \to 0$ if and only if f remains bounded on the unit ball.

□

Example 5.1 Unbounded (discontinuous) linear maps.

For $V = L^1(0,1)$ the linear functional $f : v \mapsto \dot{v}(0)$ is unbounded and hence not continuous. Consider the sequence v_n on the unit ball shown in Fig. 9 for which $f(v_n) = n^2 \to \infty$.

As another example, consider $V = L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ with L^2 norm. The linear functional

$$f : v \mapsto \int_{-\infty}^{\infty} v dt$$

is unbounded and hence not continuous. Consider the sequence v_n on the unit ball shown in Fig. 10 for which $f(v_n) = \sqrt{n} \to \infty$.

We can show a similar result for bilinear (or any multilinear) form $f : V \times V \to \mathbb{R}$.

▶ **Theorem 5.4** *Any bilinear map $f : V \times V \to \mathbb{R}$ is continuous in both its arguments if and only it is bounded on the unit ball in V, namely:*

$$\sup_{\|u\|=\|v\|=1} |f(u,v)| < \infty$$

Example 5.2 Unbounded (discontinuous) bilinear maps.

For any unbounded (discontinuous) linear functional $f : V \to \mathbb{R}$ we can produce an unbounded (discontinuous) bilinear form $F : V \times V \to \mathbb{R}$ given by $F(u,v) = f(u)f(v)$.

§ 5.3 Variants of the Fundamental Lemma

▶ **Theorem 5.5** ***Fundamental Lemma of Calculus of Variations.*** *If $\Omega \subset \mathbb{R}^n$ is open, $f : \Omega \to \mathbb{R}$ is continuous, and for all $\varphi \in C_c^\infty(\Omega)$ we have*

$$\int_\Omega \varphi(x) f(x) dx = 0 \tag{11}$$

then f is identically zero over Ω.

Proof. Suppose otherwise. If $f(x_\circ) > 0$ for any $x_\circ \in \Omega$ by continuity of f we can find an open ball $U \subset \Omega$ around $x_\circ$ such that $f(x) > 0$ everywhere in U. Now the function $\varphi = \mathbf{1}_U$ after mollification belongs to $C_c^\infty(\Omega)$ yet

$$\int_\Omega f\varphi dx = \int_U f dx > 0$$

□

▶ **Theorem 5.6** ***Fundamental Lemma Variant 1.*** *Suppose $\Omega \subset \mathbb{R}^n$ is open, $f : \Omega \to \mathbb{R}$ is continuous, $\mathbf{g} : \Omega \to \mathbb{R}^n$ is continuously differentiable, and for all $\varphi \in C_c^\infty(\Omega)$ we have*

$$\int_\Omega \Big[\varphi f + D\varphi \cdot \mathbf{g}\Big] dx = 0$$

Then we have $f = \operatorname{div}(\mathbf{g})$ everywhere in Ω.

Proof. Using the identity $\operatorname{div}(\varphi\mathbf{g}) = \varphi \operatorname{div}(\mathbf{g}) + D\varphi \cdot \mathbf{g}$ we write $D\varphi \cdot \mathbf{g} = \operatorname{div}(\varphi\mathbf{g}) - \varphi \operatorname{div}(\mathbf{g})$ into the integral in hypothesis we get

$$\begin{aligned} 0 = \int_\Omega \Big[\varphi f + D\varphi \cdot \mathbf{g}\Big] dx &= \int_\Omega \varphi \Big[f - \operatorname{div}(\mathbf{g})\Big] dx + \int_\Omega \operatorname{div}(\varphi\mathbf{g}) dx \\ &= \int_\Omega \varphi \Big[f - \operatorname{div}(\mathbf{g})\Big] dx + \int_{\partial\Omega} \varphi\mathbf{g} \cdot \hat{\mathbf{n}} dS \end{aligned} \tag{12}$$

where dS denotes surface integration, $\hat{\mathbf{n}}$ is the unit outward normal, and in the last step we have used the divergence theorem. Now since φ has vanishes on $\boldsymbol{\Omega}$ our claim reduces to the standard version of the fundamental lemma. □

▶ **Theorem 5.7** ***Fundamental Lemma Variant 2.*** *Suppose $\Omega \subset \mathbb{R}^n$ is open, $f : \Omega \to \mathbb{R}$ is continuous, $\mathbf{g} : \Omega \to \mathbb{R}^n$ is continuously differentiable, $\mathbf{F} : \Omega \to \mathbb{R}^{n\times n}$ is twice continuously differentiable with components $\mathbf{F} = (F_{ij})_{i,j}$. If for all $\varphi \in C_c^\infty(\Omega)$ we have*

$$\int_\Omega \Big[\varphi f + D\varphi \cdot \mathbf{g} + D^2\varphi \cdot \mathbf{F}\Big] dx = 0$$

with the dot product of $n \times n$ matrices being in the entrywise sense of $\mathbb{R}^{n^2}$ then everywhere in $\boldsymbol{\Omega}$ we have:

$$f - \operatorname{div}(\mathbf{g}) + \sum_{i,j} \partial_{x_i x_j} F_{ij} = 0$$

Proof. As before for any $\varphi \in C^\infty(\Omega)$ we have:

$$\int_\Omega \Big[\varphi f + D\varphi \cdot \mathbf{g}\Big] dx = \int_\Omega \varphi\Big[f - \operatorname{div}(\mathbf{g})\Big] dx + \int_{\partial\Omega} \varphi \mathbf{g} \cdot \hat{\mathbf{n}} dS$$

For the new term we use similar integration by parts ideas that lead to the identity $D\varphi \cdot \mathbf{g} = \operatorname{div}(\varphi\mathbf{g}) - \varphi \operatorname{div}(\mathbf{g})$ in the previous case:

$$D^2\varphi \cdot \mathbf{F} = \sum_{i,j} \partial_{x_i x_j}\varphi F_{ij} = \sum_{i,j} \partial_{x_i}\left(\partial_{x_j}\varphi F_{ij}\right) - \sum_{i,j} \partial_{x_j}\varphi \partial_{x_i} F_{ij}$$

The first term can be rewritten as

$$\sum_i \partial_{x_i} \sum_j \partial_{x_j}\varphi F_{ij} = \operatorname{div}\left(D\varphi \cdot (F_{ij})_j\right)_i = \operatorname{div}(\mathbf{F} D\varphi)$$

and the second term as

$$\begin{aligned}\sum_{i,j} \partial_{x_j}\varphi \partial_{x_i} F_{ij} &= \sum_{i,j} \partial_{x_j}\left(\varphi \partial_{x_i} F_{ij}\right) - \sum_{i,j} \varphi \partial_{x_j x_i} F_{ij} \\ &= \operatorname{div}\left(\varphi \operatorname{div}(F_{ij})_i\right)_j - \sum_{i,j} \varphi\left(\partial_{x_i x_j} F_{ij}\right)\end{aligned}$$

Combining the two we get

$$D^2\varphi \cdot \mathbf{F} = \operatorname{div}(\mathbf{F} D\varphi) - \operatorname{div}\left(\varphi \operatorname{div}(F_{ij})_i\right)_j + \sum_{i,j} \varphi\left(\partial_{x_i x_j} F_{ij}\right)$$

Putting all the terms together and applying the divergence theorem we have

$$\begin{aligned}\int_\Omega \Big[\varphi f + D\varphi\cdot\mathbf{g} + D^2\varphi\cdot\mathbf{F}\Big]dx = \int_\Omega \varphi\Big[f - \operatorname{div}(\mathbf{g}) + \sum_{i,j}\partial_{x_ix_j}F_{ij}\Big]dx \\ + \int_{\partial\Omega} D\varphi\cdot\mathbf{F}\hat{\mathbf{n}}dx + \int_{\partial\Omega}\varphi\hat{\mathbf{n}}\cdot\Big[\mathbf{g} - \big(\operatorname{div}(F_{ij})_i\big)_j\Big]\,dx\end{aligned} \tag{13}$$

Since φ is compactly supported in Ω our claim reduces to the standard version of the fundamental lemma. □

§ 5.4 Mollification

We often invoke a standard mollification argument either to restrict attention to smooth test functions or to produce simple (counter)examples without worrying about smoothness requirements of the problem. The *standard mollifier* in $\mathbb{R}^n$ is:

$$\eta(x) := C\exp\left\{\frac{1}{|x|^2-1}\right\}\mathbf{1}_{B(0,1)}$$

with constant $C > 0$ chosen such that:

$$\int_{\mathbb{R}^n}\eta dx = 1$$

Let $\Omega \in \mathbb{R}^n$ be an open set. A mollification argument relies on a sequence of mollifiers indexed by $0 < \epsilon < 1$ with support concentrating on the origin as $\epsilon \to 0$:

$$\eta_\epsilon(x) := \frac{1}{\epsilon^n}\eta\left(\frac{x}{\epsilon}\right)$$

all of which are compactly supported in $B(0,1)$ and integrate to 1. For any appropriately integrable function u we define the family of mollifications $u_\epsilon : \Omega_\epsilon \to \mathbb{R}$ as

$$u_\epsilon := \eta_\epsilon * u$$

over $\Omega_\epsilon := \{x \in \Omega | \operatorname{dist}(x,\partial\Omega) < \epsilon\}$. All our mollification arguments are based on the following theorem which we accept without proof.

▶ **Theorem 5.8** *[2] If $u \in L^1_{\text{loc}}(\Omega)$, that is $u|_V \in L^1(V)$ for all $V \subset\subset \Omega$, then we have:*

1) $\forall\epsilon : u_\epsilon \in C^\infty(\Omega_\epsilon)$,
2) $u_\epsilon \to u$ a.e. as $\epsilon \to 0$,
3) If u is additionally continuous, $u_\epsilon \to u$ uniformly on compact subsets of Ω,
4) If u is additionally in $L^p_{\text{loc}}(\Omega)$ for some $p \geq 1$ then $u_\epsilon \to u$ in $L^p_{\text{loc}}(\Omega)$.

References

[1] Melvyn S Berger. *Nonlinearity and Functional Analysis: Lectures on Nonlinear Problems in Mathematical Analysis*. Vol. 74. Academic press, 1977.

[2] Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010.

[3] IM Gelfand and SV Fomin. *Calculus of Variations. 1963*. Prentice-Hall, Inc.

[4] Mariano Giaquinta and Stefan Hildebrandt. *Calculus of Variations I*. Springer-Verlag, 2004.

[5] MZ Nashed. "Some Remarks on Variations and Differentials". In: *The American Mathematical Monthly* 73.4 (1966), pp. 63–76.