Optimality gaps and regularity for one-dimensional variational problems

Florian Noah Grün

Spring 2021

Abstract

In this Bachelor thesis we will provide a light treatment to the topic of optimality gaps and regularity for integral functionals. First we will establish the classical existence result by Tonelli, via the so called direct method. Proceedingly we discuss examples of gaps between the infima over different classes of functions. We start with the classical Lavrentiev phenomenon (gap between AC and *Lipschitz*) and continue with gaps involving C^k. In the last section we will look at different regularity conditions, that is conditions such that gaps between certain infima do not occur and end with some applications to ODE's.

Acknowledgements

First of all I have to mention my supervisor Andrey Ghulchak, who was the lecturer for the introductory course in the Calculus of Variations in Spring 2020 and introduced me to the topic. Without his helpful guidance during the last semester I could not have written this thesis. Our weekly discussions were insightful and I learned a lot through his tips and suggestion of mathematical, educational and editorial nature. Moreover I want to thank the library staff at Matematikcentrum for providing a comfortable study environment and helpful resources.

Populärvetenskaplig sammanfattning

Variationskalkyl handlar om problemet att bestämma det minsta värdet av en funktion som har en funktion som input, nämligen en så kallad funktional. Det vill säga, man minimerar $f(x)$ där x själv är en funktion. Funktionalen brukar vara i form av en integral och x är ett element av ett funktionsrum, till exempel rummet av alla kontinuerliga funktioner. Först och främst kommer frågan om existens; finns det ett minimum eller inte? Om svaret här är positivt, kommer nästa steg; hur kan man hitta den där "bästa" funktionen? Tyvärr är det inte så lätt. Problemet är att lösningen kan existera i ett mycket större rum än det man sökt lösningen och det går inte garantera att en explicit lösning hittas alls. Det är som en klyfta mellan lösningen och sökningsområdet, det så kallade *Lavrentiev gap*. Om man inte kan hitta lösningen eller komma godtyckligt nära, då hjälper existensen inte särskilt mycket. Därför är ansatsen att kräva olika villkor på funktionalen så att lösningen också existerar i ett mindre rum som kan genomsökas lättare. Inom dessa *regularitetsvillkor* finns det fortfarande aktiv forskning med omfattande tillämpningar.

Contents

1 Introduction

1.1 Historical context

The Calculus of Variations grew out of the infinitesimal calculus developed by Newton and Leibniz in the late 17th century. One of the first and most famous problems was the Brachistochrone problem, posed by Johann Bernoulli as a challenge for mathematicians all over Europe in 1696. It asks for the curve of fastest decent in the two dimensional plane from a higher point to a lower point under influence of a unitary gravitational field. Among others, Isaac Newton himself showed that the solution curve is a cycloid [9, p. 34]. The next step was due to Euler and Lagrange, the so called Euler-Lagrange equation, see next section. This lead to a new understanding of physical theories like the least action principle or Fermat's principle of least time in optics. However so far the theory was only concerned with necessary conditions and it took until Weierstrass to assert sufficient conditions and to ask about general existence of optimal solutions. Hilbert then stated as his 20th problem the question whether problems have always solutions and as his 19th whether those solutions are particularly smooth (existence and regularity)(see also [10]). Tonelli (1915) was one of the first to work on the issue and published a proof, known as the "direct method", to show existence under certain assumptions (see chapter 3). What followed were various contributions in the field, also to regularity. Francis Clarke and his supervisor R.T Rockafellar made important contributions to non-smooth analysis and thus less/non smooth variational problems could be approached. Amongst others, common applications are analytical mechanics and minimal surfaces. The latter case is connected to Plateau's problem and soap film problems, finding minimal surfaces prescribed to certain boundary conditions. In 2002 a complete proof of a longstanding conjecture, the double-bubble conjecture was published. Another direction was then to include differential equations, the subjects of optimal control and dynamical programming arouse. Here the main protagonists were among others Lev Pontrayagin and Richard Bellman. Optimal control has again wideranging applications from economics and operational research to robotics and control engineering. For a detailed history of the Calculus of Variations we refer to [9].

1.2 The classical theory

We state now the basic problem (denoted often by (P)) as

$$\text{minimize } J(x) = \int_a^b \Lambda(t, x(t), x'(t))dt, \quad x \in S[a,b], x(a) = A, x(b) = B, \tag{P}$$

where $S[a,b]$, e.g $C^2[a,b]$, denotes a certain class of functions and $x(a) = A, x(b) = B$ are normally referred to as boundary conditions. Note that here we are interested in a **global minimum**, i.e. x_* such that $J(x_*) \le J(x)$ where x can be any function in $S[a,b]$. In this thesis we are only looking at the case where $x(t)$ is scalar valued. As with one-variable optimization, we could also look for local minima, solely those are normally of minor interest.

Definition 1.1. A **weak local minimizer** in a class $S \subseteq AC$ is a function $x \in S$ s.t. there exists $\delta > 0$ and $J(x) \le J(y)$ for all $y \in S$ whenever $|y(t) - x(t)| + |y'(t) - x'(t)| < \delta$ for almost all t.
A **strong local minimizer** in a class S is a function $x \in S$ s.t. there exists $\delta > 0$ and $J(x) \le J(y)$ for all $y \in S$ whenever $|y(t) - x(t)| < \delta, \quad \forall t$.

In that regard, a weak local minimizer x_* is optimal compared to all other functions x that are sufficiently close to x_* not only in their function values but also in their derivatives. A strong local minimizer however respects also functions that have totally different derivative (e.g. strong oscillations around x_*). Any strong minimizer is thus a weak minimizer, any global minimizer a strong one, but the converse does not hold. In this section we are mostly concerned with some of the classical results before the turn of the 20th century. We will state some classical theorems, mostly without proof and point out their weaknesses.

1.2.1 First results

A first necessary condition was found by Euler and later again by Lagrange. The standard proof nowadays follows Lagrange's idea to use a variation a small perturbation to the optimal function, differentiation under the integral sign and integration by parts. It can be found in almost any basic textbook on the subject, see for example [12, p.32].

Theorem 1.2 (Euler-Lagrange). *If $x_* \in C^2[a,b]$ and the Lagrangian $\Lambda(t,x,v)$ has continuous second derivatives then the solution x_* to the minimization problem satisfies the* ***Euler-Lagrange equation***

$$\Lambda_x(t, x_*(t), x_*'(t)) - \frac{d}{dt}\left(\Lambda_v(t, x_*(t), x_*'(t))\right) = 0, \qquad x_*(a) = A, x_*(b) = B. \tag{EL}$$

The approach is easily generalized to higher derivatives and several variables. Note that with several variables we obtain a PDE, with vector valued functions a system of ODEs and with both a system of PDEs. Two variants of the basic problem are flexible endpoints, which lead to the so known *transversality conditions* (see [12, p.197]) and further constraints on the function like the isoperimetric constraint which leads to the famous Queen Dido problem (see [12, p. 107]).

Remark 1.3. Solutions to the EL are called **extremals**, in the same fashion as extremal points where the first derivative is zero in one-variable analysis. In both cases minimizers are extremals but not vice versa and thus when solving the EL, we do not know whether we have found a local or global minimum or maybe some other type of extremal. There are various sufficient conditions that ensure at least local minima. In the following exposition however we are concerned with global minima and if they even exist.

Theorem 1.4. *If $\Lambda(t,x,v)$ is convex (Definition 2.16) and twice differentiable in $(x,v), \forall t \in [a,b]$, then any extremal (solution of EL) is the global minimum.*

Proof. By Proposition 2.18 and the convexity condition we can write (seeing t as a parameter)

$$\Lambda(t, x+\epsilon_1, v+\epsilon_2) \geq \Lambda(t,x,v) + \Lambda_x(t,x,v)\epsilon_1 + \Lambda_v(t,x,v)\epsilon_2.$$

Let now h be a variation s.t. $x(t) = x_*(t) + h(t)$ and $h(a) = h(b) = 0$, then the integral becomes

$$\begin{aligned} J(x) &= \int_a^b \Lambda(t, x_* + h, x_*' + h')dt \\ &\geq \int_a^b \Lambda(t, x_*, x_*')dt + \int_a^b \Lambda_x(t, x_*, x_*')h(t) + \Lambda_v(t, x_*, x_*')h'(t)dt. \end{aligned}$$

Now by the EL, $\Lambda_x(t, x_*, x_*') = \frac{d}{dt}\Lambda_v(t, x_*, x_*')$, so the second integral becomes

$$\begin{aligned}
&\int_a^b \Lambda_x(t, x_*, x_*')h(t) + \Lambda_v(t, x_*, x_*')h'(t)dt \\
&= \int_a^b \frac{d}{dt}\Lambda_v(t, x_*, x_*')h(t) + \Lambda_v(t, x_*, x_*')h'(t)dt \\
&= \int_a^b \frac{d}{dt}\left(\Lambda_v(t, x_*, x_*')h(t)\right)dt = \Lambda_v(t, x_*, x_*')h(t)\Big|_a^b = 0.
\end{aligned}$$

□

First problems arise quickly when considering only smooth functions. Take for example the following scenario:

$$\text{minimize } J(x) = \int_{-1}^{1} (x'(t)^2 - 1)^2 dt \qquad \text{subject to } x(-1) = x(1) = 1.$$

It is clear that the function $x_*(t) = |t|$ is a global minimizer, that satisfies the boundary conditions. However for any smooth function $x(t)$ satisfying the boundary conditions there exists a small interval where $x'(t) \neq \pm 1$, hence $J(x) > 0$ for any smooth function x. But $x_* = |t|$ can be approximated arbitrary well in L^1-norm by smooth functions ($\forall \epsilon > 0, \exists x \in C^\infty$ s.t. $|J(x) - J(x_*)| < \epsilon$), the infimum over smooth functions is 0, but it is not attained and the EL-equation does not give the desired solution.

Definition 1.5. A function x on $[a, b]$ is **Lipschitz** if there exist $L \in \mathbb{R}$ such that

$$|x(t) - x(s)| \leq L|s - t| \qquad \forall s, t \in [a, b].$$

Theorem 1.6 (integral Euler-Lagrange equation)**.** *If x_* is a weak local minimizer for the basic problem over the class $Lip[a, b]$ then x_* satisfies for some $c \in \mathbb{R}^n$*

$$\Lambda_v(t, x_*(t), x_*'(t)) = c + \int_a^t \Lambda_x(s, x_*(s), x_*'(s))ds, \qquad t \in [a, b] \quad a.e. \tag{IEL}$$

Note that the integral Euler-Lagrange equation holds for vector valued functions as well. The partial derivatives are than replaced by gradients and we have a system of integral equations. For a proof see [6, p. 308,Theorem 15.2]. However the IEL and considering *Lip*-functions instead of C^2 is the furthest we can go with necessary conditions (that is a constructive way of finding a solution), working under the assumption that the minimum exists in the first place. Existence/Nonexistence of the minimum is the central topic of chapter 3.

Another important theorem is the so called *fundamental lemma of the calculus of variations*, credited to duBois-Raymond which can be found in detail in [6, p.184].

Theorem 1.7. *Let $f \in L^1[a, b]$. If*

$$\int_a^b f(t)\phi'(t)dt = 0 \qquad \forall \phi \in Lip[a, b] : \phi(a) = \phi(b) = 0,$$

then $f(t) = C \in \mathbb{R}$ almost everywhere.

1.2.2 Regularity

The previous example leads naturally to the question of regularity of solutions, i.e. is a minimizer in $Lip[a, b]$ maybe more than merely Lip like C^k? Thus we want to find conditions on the functional or the Lagrangian that "lift" the solution from a larger class (here Lip) into a smaller, better class (here C^1). The results here are often refered to as *higher regularity* as the starting point is the fairly "high" Lipschitz class.

Theorem 1.8. *Let $x_* \in Lip[a, b]$ satisfy the integral Euler equation and let the Lagrangian $\Lambda(t, x, v)$ be strictly convex in v for almost every t. Then x_* is actually of type $C^1[a, b]$.*

Proof. We know that Lipschitz functions are differentiable almost everywhere but the derivative may not be defined at points in a set of measure zero and $x'_*(t)$ may be a discontinuous function with holes. We now want to find a function $\bar{v}$ that agrees with x'_* almost everywhere, but is continuous. As the fundamental theorem of Calculus applies to Lipschitz functions we then write

$$x_*(t) = x_*(a) + \int_a^t x'_*(s)ds = x_*(a) + \int_a^t \bar{v}(s)ds \qquad \forall t \in [a, b].$$

Differentiating now x_* gives back $\bar{v}$, a continuous function, $x_* \in C^1[a, b]$.
We define $W = \{t \in (a, b) : x'_*(t) \text{ exists } \}$ and $p(t) = \Lambda_v(t, x_*(t), x'_*(t)),\ t \in W$. Now we fix a point t_0 in $[a, b]$ and let $\{s_i\}$, $\{t_i\}$ be two sequences in W converging to t_0. Those sequences exist because W is dense(even of full measure) in $[a, b]$. Now construct the following two sequences $\{S_i\}, \{T_i\}$ (based on $\{s_i\}$, $\{t_i\}$) by

$$S_i = x'_*(s_i), \qquad\qquad T_i = x'_*(t_i).$$

As $x'_*(t)$ is bounded, the above sequences are bounded and thus by the Bolzano-Weierstrass theorem have convergent subsequences $\{S_{i_k}\}, \{T_{i_k}\}$ that converge to the limits l_s, l_t. We use those limits with the continuity of Λ_v(being the LHS of (IEL),the RHS is continuous because it is an integral) and x_* in t to pass to the limit and obtain

$$\lim_{k\to\infty} \Lambda_v(s_{i_k}, x_*(s_{i_k}), x'_*(s_{i_k})) = \Lambda_v(t_0, x_*(t_0), l_s) = p(t_0),$$
$$\lim_{k\to\infty} \Lambda_v(t_{i_k}, x_*(t_{i_k}), x'_*(t_{i_k})) = \Lambda_v(t_0, x_*(t_0), l_t) = p(t_0).$$

But the strict convexity of $\Lambda(t_0, x_*(t_0), v)$ in v implies, by an elementary result from convexity (2.21) that $\Lambda_v(t_0, x_*(t_0), v)$ is injective in v, i.e. $l_s = l_t$. Now define $\bar{v}(t_0) := \lim_{i\to\infty} x'_*(s_i)$, where s_i is any sequence in W converging to t_0. This holds also for $t_0 \notin W$ as W is dense in $[a, b]$. The Lipschitz condition implies that the derivative is bounded when it exists, i.e. $\lim_{i\to\infty} x'_*(s_i) \neq \pm\infty$. Now $\bar{v}(t_0)$ is a continuous function even at the aforementioned holes as its value at a hole is defined as the limit when approaching from either side. Interesting to note is that x'_* did not have jump discontinuities, only point(removable) discontinuities. On the points in W it obviously agrees with x'_*, as $\forall t_0 \in W$ we can take the constant sequence $\{t_0\}_i$. To conclude $x_*(t)$ can be written as $x_*(a) + \int_a^t \bar{v}(x)ds$ too, so the derivative is continuous, $x_* \in C^1[a, b]$. □

High enough regularity enables us to use the EL equation as a differential equation, instead of working with the integral Euler equation. It is then possible to assert even higher regularity to the solution in some cases. The next result stated as Theorem 15.7 in [6], due to Hilbert and Weierstrass can thus be seen as a Corollary to the previous case

Corollary 1.9. *I $x_* \in Lip[a,b]$ is a solution of the IEL equation, $\Lambda \in C^m (m \geq 2)$ and $\Lambda_{vv}(t, x_*(t), v) > 0$ for all $t \in [a,b], v \in \mathbb{R}^n$, then x_* is of the type C^m .*

Proof. By the positive definiteness in v, convexity in v follows and thus $x_* \in C^1[a,b]$ by the theorem. Now $t \to \Lambda_x(t, x_*(t), x_*'(t))$ is continuous on $[a,b]$ (composition of continuous functions) thus the RHS of the (IEL), the so called *costate* $p(t)$, is a C^1 function. Fix $t_0 \in [a,b]$ and look at the function $v \to \Lambda_v(t_0, x_*(t_0), v)$ which is injective by the assumption $\Lambda_{vv} > 0, \forall t_0 \in [a,b]$. Thereby any equation of the form $\Lambda_v(t_0, x_*(t_0), v) = K$ has a unique solution v. This surely holds for

$$\Lambda_v(t_0, x_*(t_0), v) = K = c + \int_a^{t_0} \Lambda_x(s, x_*(s), x_*'(s))ds =: p(t_0), \tag{$\dagger$}$$

with the unique solution $v = x_*'(t_0)$. Look now at the C^1 function

$$F(t, v) = \Lambda_v(t, x_*(t), v) - p(t),$$

clearly $\frac{\partial}{\partial v} F(t, v) \neq 0 \;\; \forall t \in [a,b], v \in \mathbb{R}$ and $F(t_0, x_*'(t_0)) = 0$ by ($\dagger$). The implicit function theorem gives now that x_*' is of type C^1 in a sufficiently small neighborhood around t_0. But this holds for all $t_0 \in [a,b]$ and thus $v(t) = x_*'(t) \in C^1[a,b]$ and hence $x_* \in C^2[a,b]$. This leads to $t \to c + \int_a^t \Lambda(s, x_*(s), x_*'(s))ds$ is actually in $C^2[a,b]$ as well. If now $m \geq 3$, iterating the argument leads to $x_*' \in C^2[a,b]$, $x_* \in C^3[a,b]$ and $t \to c + \int_a^t \Lambda(s, x_*(s), x_*'(s))ds \in C^3[a,b]$. Continuing until $x_*' \in C^{m-1}[a,b]$ gives $x_* \in C^m[a,b]$.

$\square$

Corollary 1.10. *If we have a Lagrangian that is of class C^m for every $m \in \mathbb{N}$, i.e. in C^∞, for example a polynomial Lagrangian, then under the previous assumptions we have $x_* \in C^\infty[a,b]$.*

2 Background in Functional Analysis

Now we take a step back and look at the minimization problem from a more abstract perspective. The chapter is based on [6, Chapter 2 and 5] is not the main part of the exposition and bears little relation to the actual Calculus of Variations. The reader may skip it and refer to it when needed. Firstly, we generalize the basic problem (P) to (P'):

$$\text{minimize } f(x) \text{ subject to } x \in A. \tag{P'}$$

The set A is now in most cases a subset of a function space but could also be just some subset of $\mathbb{R}^n$, e.g. minimizing a function from unit circle to $\mathbb{R}$. The function f takes in our case values on the extended real line $\mathbb{R} \cup \{\infty\}$ usually denoted by $\mathbb{R}_\infty$.
We start the next section by defining some elementary concepts about normed spaces and continue in the second section with the notion of convexity and lower semicontinuity. Especially the lower semicontinuity helps us later to assert existence. Consequently we look at the notion of weak convergence, later used in asserting that a certain "weak limit", the actual solution to (P'), of a minimizing sequence exists.

2.1 Preliminaries

Definition 2.1. A vector space X equipped with a norm $\|\cdot\|$ is called a normed space.

Definition 2.2. We call a function $f\colon X \to \mathbb{R}$ from the vector space to the real line a **linear functional** if for $a, b \in \mathbb{R}$

$$f(ax+by) = af(x) + bf(y) \qquad \forall x, y \in X.$$

Definition 2.3. The **dual space** X^* of X is the vector space of all continuous linear functionals of X and is equipped with the dual norm $\|\cdot\|_{X^*}$ defined by

$$\|f\|_{X^*} = \sup_{x \in X, \|x\| \leq 1} f(x).$$

Definition 2.4. A normed space $(X, \|\cdot\|)$ is said to be a **Banach** space if it is complete, i.e. all Cauchy sequences $\{x_n\}$ converge in the norm $\|\cdot\|$ to a limit $x \in X$. That is

$$\forall\{x_n\} : \lim_{i,j\to\infty} \|x_i - x_j\| = 0 \implies \exists x \in X : \lim_{i\to\infty} \|x_i - x\| = 0.$$

Note that the completeness always depends on the norm, so by changing the norm the space may become incomplete. Most space we encounter will, together with their norm, be Banach spaces. Prominent Banach spaces are the space of sequences converging sequences $\ell^p, 1 \leq p \leq \infty$, the Lebesgue spaces L^p (which are important later on) and the space of n-times differentiable functions C^n. Unfortunately only completeness of a space is not enough in our setting, something stronger is needed, reflexivity.

Having the notion of the dual space we can construct the bidual, the dual of the dual space, $(X^*)^*$. Recalling that the dual X^* consists of all continuous linear functionals from X to the real line, we now take all continuous linear functionals that take now an element from X^* to the real line; that is the bidual X^{**}.

Definition 2.5. We define a map $J\colon X \to X^{**}$, often called the *canonical injection*, given by

$$J(x) = z \in X^{**} \text{ such that } z(y) = y(x), \forall y \in X^*.$$

In other words, the real number we get by evaluating z at any element of the dual X^* should be the same as the real number we would get when evaluating the element from the dual y at the preimage $J^{-1}(z)$.

Proposition 2.6. *$J: X \to X^{**}$ is an injective map.*

$$z_1 = z_2 \Leftrightarrow z_1(y) = z_2(y), \forall y \Leftrightarrow y(x_1) = y(x_2), \forall y \Leftrightarrow y(x_1) - y(x_2) = y(x_1 - x_2) = 0, \forall y \Leftrightarrow x_1 = x_2.$$

One can then show that J is norm preserving (the norm on $J(X)$ and X are the same) and thus $J : X \to J(X)$ is an isometry of X and $J(X) \subseteq X^{**}$, its image under J, actually have the same structure, the bidual space contains an "isomorphic copy" of the original space.

Definition 2.7. X is **reflexive** if $X \cong X^{**}$, i.e. $J(X) = X^{**}$.

Reflexivity is a really nice property, when taking successive duals there are only two spaces to consider ($X^{***} = X^*$,...). Were X not reflexive, successive dual space would "grow". Reflexivity gives rise to some powerful theorems from functional analysis.

Theorem 2.8. *The dual space of $\mathbb{R}^n$ is $\mathbb{R}^n$ itself. Furthermore that leads to the fact that R^n is also reflexive.*

Theorem 2.9. *The dual space of $L^p[a,b]$ can be characterized by $L^q[a,b]$ where $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$.*

Proof. Sketch: The idea is to find an isometry between $L^q[a,b]$ and $L^p[a,b]^*$. We define the (functional) map $f_y : L^p[a,b] \to \mathbb{R}$ by the following relation

$$f_y(x) = \int_a^b x(t)y(t)dt, \qquad \forall x \in L^p[a,b].$$

Now we consider the map $I: L^q[a,b] \to L^p[a,b]^*$ given by $I(y) = f_y$. One can then show that this map is injective, norm preserving using Hoelders inequality, and surjective. Thus $L^q[a,b]$ and $L^p[a,b]^*$ are isometric. For a complete proof see [17, Theorem 6.16]. □

We now can take the bidual of $L^p[a,b]$ which is $L^p[a,b]$ again and thus $L^p[a,b]$ is a reflexive space. Proofs of Tonelli's theorem and its extensions rely heavily on the reflexivity of $L^p[a,b]$.

Proposition 2.10. (L^p space embedding) *If $1 \leq p < q \leq \infty$ and $\|f\|_r := \int_a^b |f(t)|^r dt$ then $\|f\|_q \leq C\|f\|_p, C \in \mathbb{R}$. Thus we have the inclusion $L^q[a,b] \subseteq L^p[a,b]$.*

Theorem 2.11. (Lusin's theorem [17, p.56 Theorem 2.23]) *Let $h: [a,b] \to \mathbb{R}$ be measurable and $|h(t)| \leq M$ almost everywhere. Then for every $\epsilon > 0$, there exists a $f \in C[a,b]$ s.t. $\mu(\{t \in [a,b] : h(t) \neq f(t)\}) < \epsilon$ (μ denotes here the one-dimensional Lebesgue measure on $[a,b]$) and $\sup_{t\in[a,b]} |f(t)| \leq \sup_{t\in[a,b]} |h(t)|$. In other words, there exists a compact set $K \subseteq [a,b]$ of measure $b - a - \epsilon$ where $h(t)$ is continuous.*

Theorem 2.12. (Egorov's theorem [18, Theorem 4.17]) *Suppose that $\{f_k\}$ is a sequence of measurable functions that converges(pointwise) almost everywhere in a set E of finite measure to a limit function f. Then given $\epsilon > 0$, there is a closed subset F of E such that $\mu(E \backslash F) < \epsilon$ (μ is the Lebesgue measure) and f_k converges uniformly to f on F.*

2.2 Convexity and lower semicontinuity

Definition 2.13. The **epigraph** of a function $f\colon X \to \mathbb{R}_\infty$ is defined by the set $\mathrm{epi} f = \{(x,r) \in X \times \mathbb{R} : f(x) \leq r\}$.

Definition 2.14. Let $(X, \|\cdot\|)$ be a normed space. A function $f\colon X \to \mathbb{R}_\infty$ is said to be **lower semicontinuous**(lsc) at x_0 if

$$\forall \epsilon > 0, \exists \delta : \text{ if } \|x_0 - x\| < \delta \implies f(x) \geq f(x_0) - \epsilon,$$

or expressed differently $\liminf_{x \to x_0} f(x) \geq f(x_0)$.

It is possible now to define upper semicontinuity in an analogous way and we see that upper and lower semicontinuity together imply continuity; any continuous function is automatically lower semicontinuous. A quick example for a lsc but not continuous function could be the piecewise defined function

$$f(x) = \begin{cases} 1 & \text{if}\, x > 0, \\ 0 & \text{if}\, x \leq 0. \end{cases}$$

Definition 2.15. A set A is called **convex** if for $x, y \in A$ and $t \in [0,1]$ $(1-t)x + ty \in A$.

Definition 2.16. A function $f\colon X \to \mathbb{R}$ is called **convex** if for all $x, y \in X$ it holds that $f((1-t)x + ty) \leq (1-t)f(x) + tf(y), \quad \forall t \in [0,1]$.

Definition 2.17. A function $f\colon X \to \mathbb{R}$ is called **strictly convex** if for all $x, y \in X$ with $x \neq y$ it holds that $f((1-t)x + ty) < (1-t)f(x) + tf(y), \quad t \in (0,1)$.

Proposition 2.18. *A differentiable convex function $f\colon X \to \mathbb{R}$ fulfills the* ***gradient inequality***

$$f(x) - f(y) \geq \langle \nabla f(x), y - x \rangle, \qquad \forall x, y \in X. \tag{GI}$$

This inequality can be generalized from the derivative to the so called subgradient when f is not differentiable at x, see [6, p.59].

Another way to characterize lower semicontinuity and convexity of functions is by means of the epigraph.

Proposition 2.19. *A function $f\colon X \to \mathbb{R}_\infty$ is lsc iff* $\mathrm{epi} f$ *is closed in $X \times \mathbb{R}$.*

Proposition 2.20. *A function $f\colon X \to \mathbb{R}_\infty$ is convex iff* $\mathrm{epi} f$ *is convex in $X \times \mathbb{R}$.*

Proposition 2.21. *The gradient of a strictly convex differentiable function is injective.*

Proof. Strictly convex functions can be characterized by $f(x) - f(y) > \langle \nabla f(y), (x-y) \rangle$ for all $x \neq y$, i.e the graph of the function lies above all tangent planes. Suppose the gradient is not injective. Then for some x_0, y_0 we have $\nabla f(x_0) = K = \nabla f(y_0)$ but $x_0 \neq y_0$. So

$$\begin{aligned} L = f(x_0) - f(y_0) > \langle \nabla f(y_0), (x_0 - y_0) \rangle = (K \cdot (x_0 - y_0)) = K', \\ -L = f(y_0) - f(x_0) > \langle \nabla f(x_0), (y_0 - x_0) \rangle = (K \cdot (y_0 - x_0)) = -K'. \end{aligned}$$

But that entails a contradiction and thus $\nabla f(x)$ is injective. $\square$

Closures of convex sets are convex again (easy to show) and we also have the following important theorem from functional analysis concerning convex sets that we will use in a later on. It can be found in detail in [6, Theorem 2.37]

Theorem 2.22. (strict Hahn-Banach separation) *Let K_1, K_2 be two disjoint non-emtpy convex subsets of a normed vector space X and assume furthermore that K_1 compact and K_2 closed, then*

$$\exists y \in X^*, \exists \lambda \in \mathbb{R}: \qquad y(x_1) < \lambda < y(x_2) \qquad \forall x_1 \in K_1, x_2 \in K_2.$$

The intuition here is quite clear, the convexity prevents the sets to be "surrounded" by each other, in the two-dimensional case it is easy to see that two convex sets can be separated by a line. Furthermore for the finite dimensional case there exists the following results.

Theorem 2.23. [6, Theorem 2.46] *Let D be a convex subset of $\mathbb{R}^n$ and $\alpha \in \partial D$, a point on the boundary of D. Then $\alpha \in \partial \bar{D}$, the boundary of the closure of D, and there exists $0 \neq y \in \mathbb{R}^n$ s.t.*

$$y \cdot (x - \alpha) \leq 0 \qquad \forall x \in \bar{D}.$$

Theorem 2.24. [6, Theorem 2.47] *If X is finite dimensional and K_1, K_2 are disjoint convex subsets, then there exists $0 \neq y \in X^*$ such that*

$$y(x_1) \leq y(x_2) \qquad \forall x_1 \in K_1, x_2 \in K_2.$$

Corollary 2.25. *Let p be a point in $\mathbb{R}^n$ that does not lie in the interior of a non-empty convex set $K \subseteq \mathbb{R}^n$. Then there exists $0 \neq y \in \mathbb{R}^n$ s.t.*

$$y \cdot p \leq y \cdot x \qquad \forall x \in K \qquad\qquad (\cdot \text{ is the usual dot product on } \mathbb{R}^n).$$

Proof. If K is open, let $K_1 = \{p\}$ and $K_2 = K$ and apply 2.24. If K is closed, either $p \notin K$ or $p \in \partial K$. In the first case apply 2.24, in the second case apply 2.23 to get $y \cdot x \leq y \cdot p$. Now choose instead $-y$, the inequality flips and the statement follows. □

2.3 Weak Topologies

We will now turn our focus to different modes of convergence in X, depending on the chosen topology. One should always keep Weierstrass theorem in mind, i.e. a continuous function on a compact interval attains its minimum. The concept here is to introduce topologies coarse enough so that we achieve compactness but still have continuity (i.e. lsc) of the functional. One extreme case is the trivial topology, clearly every set is compact(in the sense that every open cover has a finite subcover) but the inverse image of an open set (condition for continuity) is most likely not a set in the trivial topology. In the discrete topology however we have continuity guaranteed, but usually no compactness (by taking e.g the discrete cover).

So far we have only encountered the norm topology on X, the topology generated by the open balls $B_r(x_0) = \{x \in X : \|x_0 - x\| < r\}$. This topology is also called the **norm/strong topology** T_s.

Definition 2.26. The topology generated by all the sets of the form

$$U(x_0, y, r) = \{x \in X : |y(x - x_0)| < r\}, \qquad x_0 \in X, y \in X^*, r > 0,$$

is called the **weak topology** $\mathcal{T}_w$ on X. The collection of U's form a subbasis for the weak topology.

The weak topology is constructed such that it is the coarsest topology on X that ensures all $y \in X^*$ are continuous. If one set of the form $U(x_0, y_i, r)$ were missing, y_i would not be continuous anymore.

Definition 2.27. We say a sequence x_n converges **weakly** to an element x_* (sometimes denoted by $x_n \xrightarrow{w} x_*$) if it converges wrt. the weak topology. Compare to pointwise convergence $x_n \to x_*$ and uniform convergence $x_n \xrightarrow{u} x_*$.

In the finite case the weak and strong topology are the same but in the infinite case the following holds.

Theorem 2.28. *Let X be an infinite dimensional vector space. Then the weak topology is coarser than the strong topology.*

Proof. It is enough to show that there is an element in the strong topology that is not in the weak topology. We claim that element is the unit ball $B = \{x \in X : \|x\| < 1\}$. Take a weak open set W containing the origin. Each open set contains a basis element, a finite intersection of the subbasis elements from the definition. Hence $\cap_{i=1}^n U(0, y_i, r) \subseteq W$.
Take now the function $y(x) = (y_1(x), y_2(x), ..., y_n(x))$ which maps each element from X to an element in $\mathbb{R}^n$. y can not be an injective function, otherwise the infinite dimensional space X would be isomorphic to the finite dimensional $\mathbb{R}^n$. We thus have some non-zero point x_0 such that $y(x_0) = 0$ and by linearity $\forall a \in \mathbb{R}, ax_0 \in ker(f) \subseteq W$. Thus any open set in the weak topology around the origin contains a whole subspace (i.e the line $ax_0, a \in \mathbb{R}$), but the unit ball in the strong topology is bounded and can not contain whole subspaces (which are unbounded in the norm topology). So the unit ball B is an open set in the strong topology but not contained in the weak one.

□

The next step we need is to connect the convexity and lower semicontinuity to the weak topology.

Theorem 2.29. *If K is a convex subset of a normed space X, then it is strongly closed if and only if it is weakly closed.*

Proof. We want to show equaility between the weak and the strong closure of K. As each closed set in the weak closure is also closed in the strong topology it follows $\bar{K}^s \subseteq \bar{K}^w$. To prove the converse, assume there is a point x_0 in $\bar{K}^w$ but not in $\bar{K}^s$. Now we use the Hahn-Banach separation theorem from the previous section with $K_1 = \{x_0\}$ (one-point sets are always compact) and $K_2 = \bar{K}^s$ (closure wrt. norm of convex set is convex again). Thus for some $y \in X^*$

$$y(x_0) < \lambda < y(x) \quad \forall x \in \bar{K}^s.$$

It follows that $x_0 \notin y^{-1}[\lambda, \infty) =: S$. S is the inverse image of a closed set, hence closed in $\mathcal{T}_w$ and $K \subseteq \bar{K}^w \subseteq S$. So $x_0 \notin S \implies x_0 \notin \bar{K}^w$, a contradiction, x_0 does not exist and we have $\bar{K}^s = \bar{K}^w$. □

Corollary 2.30. *Let $f : X \to \mathbb{R}_\infty$ be convex. Then f is (strongly) lsc iff f is weakly lsc.*

Proof. The epigraph of f is a convex and closed set in the product topology $\mathcal{T}_s \times \mathcal{T}_{standard}$ on the space $X \times \mathbb{R}$. By the theorem the epigraph is then also convex and closed in $\mathcal{T}_w \times \mathcal{T}_{standard}$, the weak topology on $X \times \mathbb{R}$ □

2.4 A first existence theorem

A solution to the problem (P') is an element $\bar{x} \in A \subseteq X$ such that $f(\bar{x}) = \inf_A f$. The following theorem can be seen as the first cornerstone to minimization on function spaces and ties together a lot of the theorems and definitions from chapter 2. The technique of the proof using a minimizing sequence and then weak compactness and weak lower semicontinuity is known as the **direct method**.

Theorem 2.31. *If*

1. *A is a closed convex subset of a reflexive Banach space X,*
2. *$f : X \to \mathbb{R}_\infty$ is convex and lsc,*
3. *$f(a) < \infty$ for some point $a \in A$,*
4. *there exists some $M > \inf_A f$ such that $\{x \in A : f(x) \leq M\} \neq \varnothing$ is bounded,*

then $\inf_A f$ is finite and we have existence of a solution, characterized as the limit of a weakly converging sequence.

We note that the fourth condition can be achieved in different ways, the most obvious way being A bounded itself. But also f such that $\lim_{\|x\|\to\infty} f(x) = \infty$ is sufficient, for a fixed M, x is not allowed to grow indefinitly. Combinations of the two types are also possible, e.g $f = e^x$ and $A = [1, \infty)$.

The proof here uses a heavy result from functional analysis, the Eberlein-Smulian theorem, which can be found in detail in [19, p.141].

Theorem 2.32. Eberlein-Smulian *A Banach space X is reflexive iff every strongly bounded sequence (wrt. $\|\cdot\|_X$) in X contains a subsequence which converges weakly to an element of X.*

It is similar to the characterization theorem for metric spaces, but the weak topology is never metrizable for infinite dimensional spaces. This can be seen from the proof that the weak topoloy is strictly coarser than the strong topology. Each open set in the weak topology contained a whole subspace, which is unbounded in the strong topology.
There are other ways to approach the existence theorem, in [6] it is proven, without referring to 2.32, in a more elementary way.

Proof. (of existence theorem 2.4) The third condition gives $f(a) = B \geq \inf_A f$ for some $a \in A$, hence the infimum over A is finite or $-\infty$ and there exists a minimizing sequence x_n, such that $f(x_n) \to \inf_A f$.
The terms of the sequence become either arbitrarily close to $f(a)$ (in that case $f(a) = \inf_A f$) or become even smaller than $f(a)$ (if the infimum is something else). In both cases $\exists N \in \mathbb{N}$ such that $\forall n > N : f(x_n) \leq \inf_A f + \epsilon = M$, the sequence $(f_n)_{n>N} = (f(x_n))_{n>N}$ is bounded in $\mathbb{R}$.
By the fourth condition the sequence (x_n) is then also bounded in X. Now we can employ the Eberlein-Smulian theorem (X is reflexive) to get a subsequence x_{n_k} that converges weakly to an element $\bar{x} \in X$. By theorem 2.3, A being convex and closed implies A is weakly closed, i.e. contains all its weak limit points, thus $\bar{x} \in A$.
Now f is convex and strongly lsc by the second condition and using 2.30 f is also weakly lsc. That in turn means

$$f(x_0) \leq \liminf_{x\to x_0} f(x) \quad \forall x_0 \in X \quad \text{(where } x \text{ converges weakly to } x_0\text{),}$$

in particular

$$f(\bar{x}) \leq \liminf_{x \to \bar{x}} f(x) = \liminf_{k \to \infty} f(x_{n_k}) = \lim_{n \to \infty} f(x_n) = \inf_A f.$$

So $f(\bar{x}) > -\infty$ from the second condition, hence the infimum has to be finite. Moreover as $\bar{x} \in A$ we actually have equality, $f(\bar{x}) = \inf_A f$, and the existence of a solution $\bar{x}$ to (P'). $\square$

3 Existence - The Direct Method

We now return to the actual Calculus of Variations, away from the abstract functional analysis setting we had in the last chapter. The aim of this chapter(based on [6, Chapter 16.1]) is to introduce basic existence theorems for solving the basic problem (P). To see the need for an existence theorem, consider first a counterexample, that is a minimization problem where no solution exists.

Example 3.1.

$$\text{minimize } J(x) = \int_0^1 x(t)x'(t)^2 dt, \qquad x(0) = 0, x(1) = 0, x \in C^2[0,1].$$

The corresponding EL equation is

$$\Lambda_x - \frac{d}{dt}\Lambda_v = -2x(t)x''(t) - x'(t)^2 = 0, \tag{3.1}$$

which has only the trivial solution $x(t) = 0$. This however is not a minimizer, because there is a sequence of functions $x_n(t)$ such that $\inf_x J(x) \leq \lim_{n\to\infty} J(x_n) = -\infty$ by defining

$$x_n(t) = nt(t-1) \qquad \in C^\infty[0,1], \tag{3.2}$$

$$J(x_n) = \int_0^1 nt(t-1)n^2(2t-1)^2 dt = -\frac{n^3}{30}.$$

So when now passing to the limit the infimum is $-\infty$, which can not be attained by any function in $C^2[0,1]$. A solution to the minimization problem does not exist, the extremal from the EL (or IEL) equation is not a minimzer.

Asserting existence is the first and crucial step in finding a solution. There is absolutely no point in optimizing when it is not certain if a given system can be optimized in the first place.

3.1 Tonelli's Theorem

To get to the first theorem concerning the existence of solutions to the basic problem (P) we need to extend our function classes even further, so far we have only had the classes $C^2[a,b]$ and $Lip[a,b]$.

Definition 3.2. A function $x : [a,b] \to \mathbb{R}$ is called **absolutely continuous** if there exists a Lebesgue-integrable function y s.t. $y = x'$ almost everywhere and

$$x(t) = x(a) + \int_a^t y(s)ds = x(a) + \int_a^t x'(s)ds,$$

where the integral is defined in the Lebesgue sense, holds for all $t \in [a,b]$.

Remark 3.3. This is not the usual definition of absolutely continuous functions, however for the one-variable case it is sufficient.

The class of all absolutely continuous functions is abbreviated by AC, absolute continuity is stronger than continuity, in fact we have the following chain of strict inequalities.

$$C^2[a,b] \subsetneq C^1[a,b] \subsetneq Lip[a,b] \subsetneq AC[a,b] \subsetneq C[a,b].$$

In each case we can find examples that justify the strict inequalities.

- The absolute value function $x(t) = |t|$ is not differentiable on any compact interval containing 0, but fulfills the Lipschitz condition.
- The function $x(t) = \sqrt{t}$ is not in $Lip[0,1]$ (derivative unbounded as $t \to 0$), but absolutely continuous on $[0,1]$ as

$$x(0) + \int_0^t x'(s)ds = 0 + \int_0^t \frac{1}{2\sqrt{s}}ds = \lim_{\epsilon\to 0}\int_\epsilon^t \frac{1}{2\sqrt{s}}ds = \sqrt{t} - \lim_{\epsilon\to 0}\sqrt{\epsilon} = \sqrt{t}.$$

- The Cantor function [1] is continuous, not identically zero, but the derivative is zero almost everywhere, hence it can not be recovered by integrating the derivative. It is special in the sense that it increases without having positive derivative outside a set of measure 0.

Tonelli's result now deals with the setting (P) where $S[a,b] = AC[a,b]$, so we minimize the functional J over all admissible arcs in $AC[a,b]$. Working in the class AC makes it possible to define functions as integrals of derivatives which is crucial in the proof of the existence theorem.

Remark 3.4. It may be tempting to use the class of continuous functions $C[a,b]$ directly. However $C[a,b]$ contains a lot of "weird" function like the above mentioned Cantor function, the Weierstrass function and the Minkoswski question mark function.

Theorem 3.5. (Tonelli 1915) *If $\Lambda(t,x,v)$ is continuous in (t,x,v) and convex in v, $\Lambda_v(t,x,v)$ continuous in (t,x,v) and*

$$\Lambda(t,x,v) \geq \alpha|v|^r + \beta \qquad \forall (t,x,v) \in [a,b]\times\mathbb{R}\times\mathbb{R},$$

where $\alpha > 0$ and $r > 1$, then there exists a solution x_* to problem (P) *in the class $AC[a,b]$.*

The last condition is known as *coercivity*, it assures boundedness from below for the Lagrangian. One also says a function $g(t)$ is coercive of degree r when $g(t)$ grows at least as fast as $|t|^r$. The structure of the proof is similar to the proof of 2.31, but needs careful justification for the semicontinuity. The continuity of Λ_v is assumed to simplify the fourth step in the proof, in [6, p321, Theorem 16.2] the theorem is proved without it.

Proof. Assume without loss of generality that $\beta = 0$. Otherwise consider $\bar{\Lambda}(t,x,v) = \Lambda(t,x,v) - \beta$ which fulfills also the conditions on continuity and convexity and denotes a linear shift in the minimization task, which does not affect the minimizer.
Step 1: $\Lambda \geq 0$ makes the integral non-negative, the infimum is not $-\infty$. Furthermore take any continuous test-function $x(t)$ (e.g. the straight arc between the endpoints), Λ is continuous so by continuity of function composition, the integral admits a finite value again, the infimum is bounded above and thus finite, denote its value by $\inf(P) = L$.
Step 2: Now take a minimizing sequence $\{x_i\}$, i.e. $\lim_{i\to\infty} J(x_i) = L$ and for large enough i, e.g. $i > i_1$ we have

$$L + 1 \geq \int_a^b \Lambda(t, x_i, x_i')dt \geq \int_a^b \alpha|x_i'(t)|^r dt \implies x_i \in L^r[a,b].$$

So the sequence $\{x_i'\}_{i_1}^\infty$ is (equi)bounded in $L^r[a,b]$. Following 2.9, $L^r[a,b]$ is reflexive and thus by the Eberlein-Smulian theorem 2.32, we have a subsequence converging weakly to an element v_* in $L^r[a,b]$.
Step 3: Subsequently define $x_* \in AC^r[a,b]$ (usual AC with additionally $v_* \in L^r$, $AC^r[a,b] \subseteq AC[a,b]$) by

$$x_*(t) = A + \int_a^t v_*(s)ds, \qquad t\in[a,b].$$

The weak convergence of x_i' to v_* in $L^r[a,b]$ can be characterized by

$$\int_a^b x_i'(t)y(t)dt \to \int_a^b v_*(t)y(t)dt, \quad \forall y \in L^{r*}, \frac{1}{r} + \frac{1}{r*} = 1, \tag{3.3}$$

where now $\to$ is ordinary convergence on $\mathbb{R}$. See 2.9. Well, certainly for every $t \in [a,b]$ the characteristic function $\chi_{[a,t]}$ is an element of L^{r*} so we can write

$$\int_a^t x_i'(s)ds = \int_a^b x_i'(s)\chi_{[a,t])}(s)ds \to \int_a^b v_*(s)\chi_{[a,t]}(s) = \int_a^t v_*(s)ds.$$

This however implies pointwise convergence of $x_i(t) = A + \int_a^t x_i'(s)ds$ to the above defined $x_*(t)$. Moreover for $t = b$, $B = x_i(b) \to x_*(b) \implies x_*(b) = B$, the end conditions are fulfilled and $x_* \in AC$, so it may be a possible solution.

Step 4: We proceed now by showing the lower semicontinuity of the functional $J(x)$. Each x_i is continuous and thus measurable, Egorov's theorem 2.12 applies, there is a compact set $K_1 \subseteq [a,b]$ where $x_i \xrightarrow{u} x$ and $\mu([a,b]\backslash K_1) < \epsilon_1$. Moreover by Lusin's theorem there is a compact set $K_2 \subseteq [a,b]$ where $x_*' := v_*$ is continuous (x_* is trivially continuous on K as it is continuous on $[a,b]$) and $\mu([a,b]\backslash K_2) < \epsilon_2$. Construct now the compact set $K = K_1 \cap K_2$ big enough so that by the absolute continuity of the Lebesgue integral $|\int_{I\backslash K}| < \epsilon$, it follows

$$\int_K \Lambda(t, x_*(t), x_*'(t))dt \geq \int_a^b \Lambda(t, x_*(t), x_*'(t))dt - \epsilon$$

As $\Lambda(t,x,v) \geq 0$, using the convexity and differentiablity of Λ in v (**??**)p. 11]gradientinequality gives

$$\begin{aligned}
J(x_i) &= \int_a^b \Lambda(t, x_i(t), x_i'(t))dt \\
&\geq \int_K \Lambda(t, x_i(t), x_i'(t))dt \\
&\geq \int_K \Lambda(t, x_i, x_*')dt + \int_K \Lambda_v(t, x_i, x_*')(x_i' - x_*')dt \\
&= \int_K \Lambda(t, x_i, x_*')dt + \int_K \Lambda_v(t, x_*, x_*')(x_i' - x_*')dt \\
&+ \int_K \left(\Lambda_v(t, x_i, x_*') - \Lambda_v(t, x_*, x_*')\right)(x_i' - x_*')dt = I_1 + I_2 + I_3.
\end{aligned}$$

Now $\Lambda_v(t, x_*(t), x_*'(t))$ is continuous on $[a,b]$ and thus element of L^{r^*}, so by the weak convergence (3.3)

$$I_2 = \int_K \Lambda_v(t, x_*, x_*')(x_i' - x_*')dt \to 0 \text{ as } i \to \infty.$$

For I_3, note that $x_*, x_i \in AC[a,b]$, so $x_*', x_i' \in L^1[a,b]$ and x_i is bounded in $L^r[a,b] \subseteq L^1[a,b]$. Furthermore $|\Lambda_v(t, x_i, x_*') - \Lambda_v(t, x_*, x_*')|$ converges to 0 uniformly on K because Λ_v was continuous and $x_i \xrightarrow{u} x_*$ on K. Thereby

$$\begin{aligned}
|I_3| &= \left| \int_K (\Lambda_v(t, x_i, x_*) - \Lambda_v(t, x_*, x_*))(x_i' - x_*')dt \right| \\
&\leq \max_{t \in K}\{|\Lambda_v(t, x_i, x_*) - \Lambda_v(t, x_*, x_*)|\} \int_K |(x_i' - x_*')| \to 0 \text{ as } i \to \infty.
\end{aligned}$$

It follows now

$$
\begin{aligned}
L = \liminf_{i\to\infty} J(x_i) &\geq \liminf_{i\to\infty} \int_K \Lambda(t, x_i, x_i')dt \\
&\geq \liminf_{i\to\infty} \int_K \Lambda(t, x_i, x_*')dt + \liminf_{i\to\infty}(I_2 + I_3) \\
&= \liminf_{i\to\infty} \int_K \Lambda(t, x_i, x_*')dt \\
&= \lim_{i\to\infty} \int_K \Lambda(t, x_i, x_*')dt \\
&= \int_K \Lambda(t, x_*, x_*')dt \\
&\geq \int_I \Lambda(t, x_*, x_*')dt - \epsilon = J(x_*) - \epsilon.
\end{aligned}
$$

But since ϵ was arbitrary, $L = \liminf_{k\to\infty} J(x_i) \geq J(x_*)$. So the infimum L is actually attained by $x_* \in AC[a,b]$ that fulfills the boundary conditions and so problem (P) has a (global) solution. $\square$

As usual the assumptions we make (convexity in v, coercivity of degree $r > 1$) may be stronger than actually needed. We now show two examples where those conditions are weakened and problem (P) does not have a solution. We consider first the case where the convexity condition is not satisfied.

Example 3.6.

$$
\text{minimize } J(x) = \int_0^1 |x(t)| + \left(x'(t)^2 - 1\right)^2 dt, \qquad x \in AC[0,1], x(0) = 0, x(1) = 0.
$$

We can see that $\Lambda(t, x, v) = |x| + (v^2 - 1)^2$ is continuous and fulfills the coercivity condition with e.g. $r = 4, \alpha = 0.1, \beta = -2$. But $(v^2 - 1)^2$ is not convex, the epigraph is not a convex set. Choose a sequence of sawtooth functions given by

$$
x_n(t) = \sum_{k=0}^{2n-1} \left((-1)^k t - (-1)^k \frac{k}{2n} + \frac{1 + (-1)^{k+1}}{4n} \right) \chi_{[\frac{k}{2n}, \frac{k+1}{2n})}.
$$

The expression in its closed form seems complicated but for each n, $x_n(t)$ is a nonnegative piecewise linear function that looks like a sawtooth function above the x-axis with n jags, each of height $1/n$. Increasing n leads to more but lower jags.
The slope stays constant with ± 1, so evaluating the functional gives

$$
J(x_n) = \int_0^1 \left(|x_n(t)| + (1-1)^2\right) dt = \int_0^1 x_n(t)dt = 2n \cdot \int_0^{\frac{1}{2n}} t = \frac{1}{4n}.
$$

Thus letting $n \to \infty$, the infimum is 0. But for this infimum to be attained we would need a (continuous) function $x : J(x) = 0$ such that $x(t) = 0$ and $x'(t) = \pm 1$ almost everywhere, which does not exist. Hence the problem (3.6) has no solution.

In the next counterexample we want to demonstrate that the degree of coercivity in Tonelli's theorem needs to be necessarily strictly greater than 1, coercivity of $r = 1$ is not enough. It is given as Exercise 16.5 in [6].

Example 3.7. The counterexample consists of minimizing the functional

$$J(x) = \int_0^1 x^2 + g(x')dt \qquad x \in AC[0,1], x(0) = 0, x(1) = 1,$$

where $g(v) = v(1 + \min\{v, 0\})$.

Step 1: We start by writing g in an easier form that is equivalent to the previous form.

$$g(v) = \begin{cases} v & v \geq 0, \\ v + v^2 & v < 0. \end{cases} \tag{3.4}$$

$g(v)$ is continuous differentiable as the only uncertain point is $v = 0$. But the derivative from the right and from the left coincide, both are 1. Furthermore we can see that the second derivative is non-negative everywhere (2 for $x < 0$ and 0 for $x \geq 0$), hence the function is convex in v.
Now we show the inequality $g(v) \geq \max\{v, |v| - 1\}$. It suffices to compare when $v < 0$. On the nonnegative part they are identical.

$$\begin{aligned} \text{Want to show (for } v < 0) : v + v^2 &\geq -v - 1 \\ (1 + v)^2 = 1 + 2v + v^2 &\geq 0. \end{aligned}$$

Step 2 By the inequality from step 1 we can write

$$\begin{aligned} J(x) = \int_0^1 x^2 + g(x')dt &\geq \int_0^1 x^2 + \max\{x', |x'| - 1\}dt \\ &\geq \int_0^1 x^2 + x'dt = \int_0^1 x^2 dt + \int_0^1 x'dt = \int_0^1 x^2 dt + 1 > 1. \end{aligned}$$

The last inequality follows from the continuity of x and thereby x^2 and that $x(1) = 1$. We can always find some small interval $(1 - \delta, 1]$ such that $x^2 \geq \epsilon > 0$ and get some finite area under the integral. That implies that for any admissible arc x, $J(x) > 1$.

Step 3 Let $x_i(t) = t^i, i = 1, ...$ be a sequence of polynomials. Inserting those into J gives

$$J(t^i) = \int_0^1 t^{2i} + g(it^{i-1})dt = \int_0^1 t^{2i} dt + \int_0^1 it^{i-1} dt = 1 + \frac{1}{2i + 1},$$

where we have used the fact that $g(it^{i-1}) = it^{i-1}(1 + \min\{0, it^{i-1}\}) = it^{i-1}$. But when taking the limit here we get $\lim_{i\to\infty} J(x_i) = \lim_{i\to\infty}(1 + \frac{1}{2i+1}) \to 1$. So the infimum is 1.

Step 4 Well now there is a problem. We have shown in step 2 that if there were a minimum, the functional would admit a value greater than 1. But in step 3 we have show that there exist a sequence that approaches the infimum (without ever reaching it). So for any hypothetical minimum that can be admitted, we can find some element of the sequence, where the functional gives a smaller value.

The problem lies in the fact that not all conditions in Tonelli's theorem are fulfilled. We have the continuity and convexity of the Lagrangian from step 1. Assume we had coercivity for some $r > 1$. Then there exists $\alpha > 0, \beta \in \mathbb{R}$

$$\Lambda(t,x,v) = x^2 + g(v) \geq \alpha |v|^r + \beta, \qquad \forall (x,v) \in \mathbb{R} \times \mathbb{R}.$$

But now for $x = 0, v > 0$ the following should hold (using (3.4))

$$\Lambda(t,x,v) = x^2 + g(v) = v \geq \alpha |v|^r + \beta \qquad \forall v > 0.$$

That is not true, for large v, the RHS grows faster, no α or β can compensate for that. For $r = 1$ the statement actually holds clearly for positive v. For negative v we have

$$\Lambda(t,x,v) = x^2 + g(v) \geq g(v) = v + v^2 \geq \alpha |v| + \beta,$$

where the last inequality holds with e.g. $\alpha = 1, \beta = -1$.

To conclude, the coercivity of degree $r > 1$ is not only a trivial necessity in the given proof of Tonelli's theorem, but a crucial condition for existence. As a matter of fact, it is less the coercivity but the superlinearity (which is not fulfilled) of the Lagrangian in v that leads to existence. There are some version that deal with those cases, see [4, Theorem 3.7].

3.2 An extension to Tonelli's Theorem

Tonelli's result can be generalized in several directions. Here we look at a case where the Lagrangian is not necessarily bounded below. The general approach, the *direct method* is however the same.

Theorem 3.8. [6, Exercise 16.10] *Let $\Lambda(t,x,v)$ be continuous in (t,x,v), convex in v, $\Lambda_v(t,x,v)$ continuous in (t,x,v) and*

$$\Lambda(t,x,v) \geq \alpha_1 |v|^r - \gamma_1 |x|^s + \beta_1 \qquad \forall (t,x,v) \in [a,b] \times \mathbb{R}^n \times \mathbb{R}^n,$$

where $r > 1$, $r > s > 0$, $\gamma_1 \geq 0$, $\alpha_1 > 0$. Then the basic problem (P) *has a solution in $AC[a,b]$.*

Proof. The proof is similar to Theorem 3.5, but we need to argue more carefully. Some arguments here are repeated, sometimes it is referred to the proof of Theorem 3.5.

Step 1′:

By taking the clearly admissible arc consisting of a straight line we see that the infimum is finite or negative infinity so a minimizing sequence $\{x_i\}$ exists, such that $\lim_{i\to\infty} \int_a^b \Lambda(t,x_i,v_i) = \inf(P)$. Thus for some constant M_1

$$\int_a^b \Lambda(t,x_i,x_i')dt \leq M_1. \tag{†}$$

Step 2′:

We want to make use of the following two inequalities for the cases $s > 1$ and $s \leq 1$. As x_i is in the space of absolutely continuous functions we write

$$x_i(t) = x_i(a) + \int_a^t x_i'(\tau)d\tau = A + \int_a^t x_i'(\tau)d\tau.$$

Case 1: $s \geq 1$

$$\begin{aligned}|x_i(t) - A| &\leq |\int_a^t x_i'(\tau)d\tau| \leq \int_a^t |x_i'(\tau)|d\tau \leq \int_a^b |x_i'(\tau)|d\tau \\ &\leq \left(\int_a^b |x_i'(\tau)|^s d\tau\right)^{\frac{1}{s}} \cdot \left(\int_a^b 1^{s^*} d\tau\right)^{\frac{1}{u}} = \left(\int_a^b |x_i'(\tau)|^s d\tau\right)^{\frac{1}{s}} \cdot C_1.\end{aligned} \tag{3.5}$$

where in the last inequality Hoelders inequality was used and u denoted the Hoelder conjugate ($\frac{1}{s} + \frac{1}{u} = 1$). Taking both sides to the s'th power gives

$$|x_i(t) - A|^s \leq C_2 \cdot \int_a^b |x_i'(\tau)|^s d\tau. \tag{3.6}$$

Now integrate both sides wrt. t. The RHS is constant wrt. t, integration only multiplies with $b - a$, but switching variables from τ to t leads to

$$\int_a^b |x_i(t) - A|^s dt \leq C_3 \cdot \int_a^b |x_i'(t)|^s dt. \tag{3.7}$$

Rewrite $\int_a^b |x_i(t)|^s dt$ in the following way:

$$\begin{aligned}\int_a^b |x_i(t)|^s dt &= \int_a^b |x_i(t) - A + A|^s dt \leq \int_a^b (|x_i(t) - A| + |A|)^s dt \\ &\leq \int_a^b 2^{s-1}|x_i(t) - A|^s + 2^{s-1}|A|^s dt \\ &\leq 2^{s-1}C_3 \cdot \int_a^b |x_i'(t)|^s dt + \int_a^b 2^{s-1}|A|^s = C_4 \int_a^b |x_i'(t)|^s dt + C_5,\end{aligned} \tag{3.8}$$

by making use of the known result $|x + y|^s \leq 2^{s-1}(|x|^s + |y|^s), s \geq 1$ in the second inequality and equation (3.7) in the third. The result is of great importance, so we state it again

$$\int_a^b |x_i(t)|^s dt \leq C \int_a^b |x_i'(t)|^s dt + D. \tag{3.9}$$

Case 2: $s < 1$

Here we can use the fact $|x|^s \leq |x| + 1$ and thus rewrite

$$\int_a^b |x_i(t)|^s dt \leq \int_a^b |x_i(t)| + 1 dt = \int_a^b |x_i(t)| dt + D_1. \tag{3.10}$$

Now employ equation (3.9) with $s = 1$ on the RHS to get

$$\int_a^b |x_i(t)|^s dt \leq C \int_a^b |x_i'(t)| dt + D. \tag{3.11}$$

In both cases there is a way to estimate the integral of x_i with an integral of x_i'. We can now continue to work with equation (†).

$$\begin{aligned} M_1 &\geq \int_a^b \Lambda(t, x_i, x_i')dt \geq \int_a^b \alpha_1 |x_i'(t)|^r - \gamma_1 |x_i(t)|^s + \beta_1 dt \\ &\geq \int_a^b \alpha_1 |x_i'(t)|^r dt - \gamma_1 \ C \int_a^b |x_i'(t)|^{\tilde{s}} dt + D_1 + \int_a^b \beta_1 \\ &= \alpha_1 \int_a^b |x_i'(t)|^r - \gamma_2 \int_a^b |x_i'(t)|^{\tilde{s}} dt + \beta_2 \\ &\implies M_2 \geq \int_a^b |x_i'(t)|^r - \gamma_3 |x_i'(t)|^{\tilde{s}} dt, \end{aligned} \tag{††}$$

where $r > \tilde{s} = \max\{s, 1\}$, a shorthand of notation to deal with both cases.
If we could now show that $|x_i'(t)|^r - \gamma_3 |x_i'(t)|^{\tilde{s}} \geq \frac{1}{2}|x_i'(t)|^r + E$ for some (probably really small) constant E, we are essentially done. For $y \geq 0$, $r > \tilde{s} \geq 1$

$$\begin{aligned} y^r - \gamma y^{\tilde{s}} &\geq \frac{1}{2} y^r + D, \\ \frac{1}{2} y^r - \gamma y^{\tilde{s}} &\geq D. \end{aligned}$$

But choosing $D = \min\{\frac{1}{2}y^r - \gamma y^{\tilde{s}}\}$ suffices. The minimum exist as $f(0) = 0$, $f(y)$ continuous and $\lim_{y\to\infty} f(y) = +\infty$ for $f(y) = \frac{1}{2}y^r - \gamma y^{\tilde{s}}$. Thus from equation (††)

$$M_2 \geq \int_a^b |x_i'(t)|^r - \gamma_3 |x_i'(t)|^{\tilde{s}} dt \geq \frac{1}{2} \int_a^b |x_i'(t)|^r dt + (b-a)D. \tag{† † †}$$

But moving the constants to the other sides, that implies actually that $M_3 \geq \int_a^b |x_i'(t)|^r dt$ and x_i' is actually in $L^r[a,b]^n$ and we can invoke the Eberlein-Smulian theorem on the reflexive space L^r as in the original proof.

Step 3′:
Define now the admissible x_* in the same way as in Theorem 3.5

$$x_*(t) = A + \int_a^t v_*(s)ds, \qquad \forall t \in [a,b].$$

The weak convergence of x_i' implies that $x_i(t) = A + \int_a^t x'(s)ds$ converges pointwise to x_* and $x_i(a) = x_*(a) = A$, $x_i(b) = x_*(b) = B$.

Step 4′:
As in Theorem 3.5 we construct the compact set K where $x_i \xrightarrow{u} x_*$ and x_*' continuous. Also

$$|x_i(t)| = \left| A + \int_a^t x_i'(s)ds \right| \leq |A| + \|x_i'\|_{L^1[a,b]} = M_0,$$

which implies that the sequence $\{x_i\}$ is equibounded in $[a,b]$. Besides the infimum is finite as

$$\Lambda(t, x_i(t), x_i'(t)) \geq -\gamma_1 |x_i(t)|^s + \beta_1 \geq -\gamma_1 M_0^s + \beta_1 = M.$$

So noting that the minimizing sequence $\{x_i\}$ of the functional $J(x) = \int \Lambda(t, x(t), x'(t))dt$ is also a minimizing sequence of $\bar{J}(x) = \int \Lambda(t, x(t), x'(t)) - M dt$, the integrand in $\bar{J}$ is then non-negative and

$$\bar{J}(x_i) \geq \int_K \Lambda(t, x_i(t), x'_*(t)) - M dt + I_2 + I_3,$$

where $I_2, I_3 \to 0$. It then follows (see argument in 3.5)

$$\liminf_{i \to \infty} \bar{J}(x_i) \geq \bar{J}(x_*) \implies \liminf_{i \to \infty} J(x_i) \geq J(x_*).$$

Hence the finite infimum is attained by $x_* \in AC[a, b]$ and the basic problem (P) has a solution.

$\square$

4 Lavrentiev phenomenon

In this section we will treat examples of the so called Lavrentiev phenomenon, i.e. functionals where the infimum over larger classes is strictly less than the infimum over smaller classes. Intuitively the Lavrentiev phenomenon is strange, any absolutly continuous function can be approximated uniformly even by polynomials, nevertheless the value of the functional does not become arbitrary close.

4.1 Gap between AC and Lip

The standard setting (originally considered by Mikhail Lavrentiev in 1926) deals with AC and Lip functions. This behaviour, that one may have different minimizers in AC and Lip is usually undesired and a lot of research is being done to assure that the Lavrentiev gap does not occur. Especially for numerical solver that use certain finite element methods, the unboundedness of the derivative in AC can lead to "losing" a solution, so to speak. For most finite element methods the solution is approximated by piecewise linear arcs which are Lipschitz functions, thus the minimum in AC can not be achieved. Here we are only concerned with one-dimensional problems, the Lavrentiev gap however does also occur in higher dimensions. In [14] a multivariable example is given where the Lagrangian is even autonomous.

4.1.1 Dacorogna/Manià example

Building on [7, Theorem 3.4.6] the problem in question is

$$J(x) = \int_0^1 (t - x(t)^3)^2 x'(t)^{2m} dt, \qquad x(0) = 0, x(1) = 1, m \geq 2.5. \tag{4.1}$$

Throughout this example we interpret $x'(t)^{2m}$ as $(x'(t)^2)^m$ to ensure that it is also defined for negative values. The integrant and thereby also the functional are non-negative, at $x_*(t) = t^{1/3}$, $J(x_*) = 0$, so the infimum (which is attained) in AC is 0. As it turns out however the infimum over the class of Lipschitz functions on $[0, 1]$ (to which x_* does not belong) is strictly greater than 0.

Lemma 4.1. *Define*

$$\mathcal{S}(\alpha, \beta) = \left\{ x \in Lip(\alpha, \beta) : \frac{1}{4} t^{1/3} \leq x(t) \leq \frac{1}{2} t^{1/3}, \quad x(\alpha) = \frac{\alpha^{1/3}}{4}, x(\beta) = \frac{\beta^{1/3}}{2} \right\},$$

where $0 < \alpha < \beta < 1$. That is the set of all Lipschitz functions on the subinterval $[\alpha, \beta]$ that lie "between" $\frac{1}{4}t^{1/3}$ and $\frac{1}{2}t^{1/3}$, starting on $\frac{1}{4}t^{1/3}$ and ending on $\frac{1}{2}t^{1/3}$. Then defining

$$J_1(x) = \int_\alpha^\beta (t - x(t)^3)^2 x'(t)^{2m} dt, \qquad x \in \mathcal{S}, m \geq 2.5,$$

gives $\inf_{x \in \mathcal{S}} J_1(x) \geq \frac{7^2 3^5}{2^{4m+6} 5^5} \approx 0.06 \cdot \frac{1}{16^m} > 0.$

Proof. By the upper bound of $x(t)$ in $\mathcal{S}(\alpha, \beta)$

$$1 - \frac{x^3}{t} \geq 1 - \frac{1}{t} \left(\frac{t^{1/3}}{2} \right)^3 = \frac{7}{2^3}, \qquad \text{which can be used in the functional}$$

$$J_1(x) = \int_\alpha^\beta (t - x(t)^3)^2 x'(t)^{2m} dt = \int_\alpha^\beta t^2 (1 - \frac{x(t)^3}{t})^2 x'(t)^{2m} dt \geq \frac{7^2}{2^6} \int_\alpha^\beta t^2 x'(t)^{2m} dt.$$

Consequently, we perform a variable substitution to eliminate the dependence on t. Let $p = 2m-3$, $q = 2m-1$ and

$$s = t^{p/q} \Leftrightarrow t = s^{q/p} \text{ and } x(t) = \bar{x}(s) = \bar{x}(s(t)) = \bar{x}(t^{p/q})$$

$$x'(t) = \frac{d}{dt}\bar{x}(s(t)) = \bar{x}'(s)s'(t) = \bar{x}'(s)\frac{p}{q}t^{-2/q} = \frac{p}{q}\bar{x}'(s)s^{-2/p}$$

$$dt = \frac{q}{p}s^{2/p}ds,$$

$$\begin{aligned} J_1(x) &\geq \frac{7^2}{2^6}\int_\alpha^\beta t^2 x'(t)^{2m}dt = \frac{7^2}{2^6}\int_{\alpha^{p/q}}^{\beta^{p/q}} s^{2q/p}\left(\frac{p}{q}\bar{x}'(s)s^{-2/p}\right)^{2m}\frac{q}{p}s^{2/p}ds \\ &= \frac{7^2p^q}{2^6q^q}\int_{\alpha^{p/q}}^{\beta^{p/q}} s^{1/p[2q-4m+2]}\bar{x}'(s)^{2m}ds = \frac{7^2p^q}{2^6q^q}\int_{\alpha^{p/q}}^{\beta^{p/q}}\bar{x}'(s)^{2m}ds. \end{aligned}$$

In the next step we apply Jensen's inequality (proof can be found in [13, p.198]) for a continuous and convex function ϕ (here $\phi(x) = x^{2m} = (x^2)^m$). Note here that m needs to be greater or equal to 0.5 to be convex, but we will obtain a larger lower bound for m later on.

$$\text{Jensen's ineq: } \frac{1}{b-a}\int_a^b \phi(x(t))dt \geq \phi\left(\frac{1}{b-a}\int_a^b x(t)dt\right) \implies \int_a^b x'(t)^{2m}dt \geq \frac{b-a}{(b-a)^{2m}}\left(\int_a^b x'(t)dt\right)^{2m}.$$

$$\begin{aligned} J_1(x) &\geq \frac{7^2p^q}{2^6q^q}\int_{\alpha^{p/q}}^{\beta^{p/q}}\bar{x}'(s)^{2m}ds \geq \frac{7^2p^q}{2^6q^q}\frac{1}{(\beta^{p/q}-\alpha^{p/q})^q}\left(\int_{\alpha^{p/q}}^{\beta^{p/q}}\bar{x}'(s)ds\right)^{2m} \\ &= \frac{7^2p^q}{2^6q^q}\frac{(\bar{x}(\beta^{p/q})-\bar{x}(\alpha^{p/q}))^{2m}}{(\beta^{p/q}-\alpha^{p/q})^q} = \frac{7^2p^q}{2^6q^q}\frac{(1/2\cdot\beta^{1/3}-1/4\cdot\alpha^{1/3})^{2m}}{(\beta^{p/q}-\alpha^{p/q})^q} \\ &= \frac{7^2p^q}{2^62^{2m}q^q}\frac{\beta^{2m/3}\left(1-\frac{1}{2}\left(\frac{\alpha}{\beta}\right)^{1/3}\right)^{2m}}{\beta^p\left(1-\left(\frac{\alpha}{\beta}\right)^{3/5}\right)^q} \\ &\geq \frac{7^2p^q}{2^{2(m+3)}q^q}\frac{\beta^{(2m/3-p)}}{2^{2m}} \geq \frac{7^2p^q}{2^{4m+6}q^q}. \end{aligned}$$

However for the last inequality to hold we need $\beta^{2m/3-p} \geq 1$, otherwise the lower bound depends on β, which may be arbitrarily small. Since $0 < \beta < 1$ that implies $0 \geq 2m/3 - p = 2m/3 - 2m + 3$ or in other words $m \geq 2.25$. □

Now we prove that for every $x \in Lip[0,1]$, there exists $0 < \alpha < \beta < 1$ such that $x \in S(\alpha, \beta)$. Well, in the beginning, for small enough t, x has bounded derivative, lies also below both $x_1(t) := \frac{t^{1/3}}{4}$ and $x_2(t) := \frac{t^{1/3}}{2}$. At $t = 1$ however $x(t)$ lies above $x_1(t)$ and $x_2(t)$.

Let $A = \left\{t \in (0,1) : x(t) = \frac{t^{1/3}}{4}\right\}$ and $B = \left\{t \in (0,1) : x(t) = \frac{t^{1/3}}{2}\right\}$. By the intermediate value theorem, both sets are non-empty. Define now $\beta = \min B$, the first time $x(t)$ crosses $x_2(t)$ and $\alpha = \max\{t : t \in A \;\&\; t < \beta\}$, the value where $x(t)$ crosses $x_1(t)$ before increasing to x_2.

To conclude, for $x \in Lip[0,1]$

$$J(x) = \int_0^1 (t - x(t)^3)^2 x'(t)^{2m}dt \geq \int_\alpha^\beta (t-x(t)^3)^2x'(t)^{2m}dt = J_1(x) \geq \frac{7^2p^q}{2^{4m+6}q^q} > 0. \tag{4.2}$$

Thus the infimum over Lipschitz functions is strictly larger than the infimum over absolutely continuous functions as long as $m \geq 2.25$, in paritcular the Lavrentiev phenomenon occurs even for polynomial Lagragiangs by choosing $3 \leq m \in \mathbb{Z}$.

The solution in $AC[0,1]$, $x_*(t) = t^{1/3}$ is however still an analytical solution of (IEL) and (EL), both sides evaluate 0. $x_*'(t)$ may be unbounded when $t \to 0$ but as $\lim_{x\to\infty} x \cdot 0 = 0$ the following still holds.

$$\Lambda_x(t, x(t), x'(t))\Big|_{x(t)=t^{1/3}} = -6x(t)^2(s - x(t)^3)^2 x'(t)^{2m}\Big|_{x(t)=t^{1/3}} = 0,$$
$$\Lambda_v(t, x(t), x'(t))\Big|_{x(t)=t^{1/3}} = 2mx'(t)^{2m-1}(t - x(t)^3)^2\Big|_{x(t)=t^{1/3}} = 0,$$

which implies that both, EL and IEL hold (for $x(t) = t^{1/3}$)

$$0 = \Lambda_x(t, x, x') = \frac{d}{dt}\Lambda_v(t, x, x') = \frac{d}{dt}0 = 0, \tag{EL}$$

$$0 = \Lambda_v(t, x, x') = c + \int_0^t \Lambda_x(s, x, x')ds = c + \int_0^t 0ds = 0 \qquad \text{(choose } c = 0). \tag{IEL}$$

A more severe issue is the fact that even though the Lagrangian is continuous (if $m \in \mathbb{N}$ even a polynomial) and convex in v, the existence of a solution does not come from Tonelli's theorem 3.5, but from our "integral is non-negative so if the integrant is zero the minimum is attained" reasoning. Tonelli's coercivity condition is clearly not fulfilled.

$$\begin{aligned}(t - x^3)^2 v^{2m} \geq \alpha|v|^r + \beta, \qquad & \forall (t, x, v) \in [a, b] \times \mathbb{R} \times \mathbb{R}.\\ \text{Choose } v_0 \text{ big enough s.t. } & \alpha|v_0|^r + \beta = c > 0 \text{ then}\\ & (t - x^3)^2 v_0^{2m} \geq \alpha|v_0|^r + \beta = c \ \ (> 0).\end{aligned} \tag{4.3}$$

However $(t - x^3)^2$ can be made 0, choose e.g. $t = a, x = a^{1/3}$ which implies $0 \geq c > 0$, a contradiction.

4.1.2 Ball and Mizel example

A setting were Tonelli's coercivity condition is also fulfilled is given by a more complicated example. It is quite recent (1985) and can be found in great detail in [3], a short conclusion of their results can be found in [2].

One ingredient is an important result proved by Tonelli himself (for $n = 1$, directly after his existence result) and later generalized to higher dimensions by Clarke and Vinter. For a proof see [3].

Theorem 4.2. (Tonelli's partial regularity result) *Let $\Lambda \in C^3$ in (t, x, v), $\Lambda_{vv} > 0$ and $x_* \in AC[a, b]$ be a strong minimizer of $J(x) = \int_a^b \Lambda(t, x, x')dt$. Then $x_*' : [a, b] \to \mathbb{R} \cup \{-\infty, \infty\}$ is continuous and the* Tonelli set *$E = \{t \in [a, b] : |x_*'(t)| = \infty\}$ is a closed set of measure 0. Also x_* is of type C^1 outside of E.*

The set E is often empty, the example below is one of the first known where it is non-empty. Ball and Mizel now consider the functional

$$J(x) = \int_0^1 (t^2 - x(t)^3)^2 x'(t)^{14} + \epsilon x'(t)^2 dt, \qquad x(0) = 0, x(1) = k > 0,$$

with its corresponding EL

$$\frac{d}{dt}(7(t^2 - x(t)^3)^2 x'(t)^{13} + \epsilon x'(t) = -3x(t)^2(t^2 - x(t)^3)x'(t)^{14}. \tag{4.4}$$

Now depending on the size of ϵ, there are 2 solutions ($\epsilon < \epsilon_0$) or one solution ($\epsilon = \epsilon_0$) of the form $rx^{2/3}, r > 0$ or no solution at all $\epsilon > \epsilon_0$.
The integrant has a scale invariant property, i.e. $\Lambda(\lambda t, \lambda^{2/3}x, \lambda^{-1/3}v) = \lambda^{-2/3}\Lambda(t, x, v)$, which makes it possible to transform the EL-equation into an autonomous system of two variables q and z. Ball and Mizel then proceed by a rather technical analysis of the phase plots to show correspondence between smooth solutions of (4.4) and trajectories in the phase plot. This leads then to the behaviour and amount of the solutions in AC for different ϵ and k. Note however that the (IEL) does not hold as $\lim_{t\to 0^+} \Lambda_v(t, x, x') = +\infty$, the minimizer is singular at the origin (the Tonelli set is $\{0\}$), by other calculations.

In a second example they consider the functional

$$\int_{-1}^{1} (t^4 - x^6)^2 |x'|^s + \epsilon x'^2 dt, \qquad x(-1) = k_1, x(1) = k_2.$$

Here for $s \geq 27$ and sufficiently small ϵ the Tonelli set E is an interior point $\{0\}$. The AC minimizer (Ball and Mizel actuall work in the Sobolev space $W^{1,p}$) does not fulfill the IEL equation, not even in the sense of distributions (weak IEL). This leads to the fact that the infimum over Lipschitz functions (actually $W^{1,q}, q \geq 3$) is strictly greater than the infimum over the AC functions.

4.1.3 Infinite gap

Take the functional

$$J(x) = \int_0^2 \frac{1}{|x(t)|} dt, \qquad x(0) = x(2) = 0.$$

Let x_n be a sequence of functions in $AC[0,2]$ defined by

$$x_n(t) = \begin{cases} n \cdot \sqrt{t} & \text{if } t \in [0,1], \\ n \cdot \sqrt{2-t} & \text{if } t \in (1,2]. \end{cases}$$

$$J(x_n) = \int_0^2 \frac{1}{|x_n(t)|} dt = \frac{2}{n} \int_0^1 \frac{1}{\sqrt{t}} dt = \frac{4}{n} \xrightarrow{n\to\infty} 0,$$

and the infimum in AC is 0. But for y being any Lipschitz function, $y'(t) \leq M$. So it is possible to bound $|y(t)|$ by the function

$$|y(t)| \leq g(t) := Mt.$$

That implies however that $\frac{1}{g(t)} \leq \frac{1}{|y(t)|}$ and thus

$$J(y) = \int_0^2 \frac{1}{|y(t)|} dt \geq \int_0^2 \frac{1}{g(t)} dt = \int_0^2 \frac{1}{Mt} dt = \frac{1}{M} \cdot \infty = +\infty.$$

Thus an infinite Lavrentiev-like gap occurs, i.e.

$$\inf_{x \in AC[0,2]} J(x) = 0 < +\infty = \inf_{x \in Lip[0,2]} J(x).$$

Remark 4.3. It should be possible to use a similar approach and derive an infinite gap between C^k and C^{k+1}.

4.2 Gaps between higher classes

This setting is not as widely researched in comparison to the *AC* vs *Lip* case, one reason is probably the fact that already comparable "weak" conditions on the Lagrangian prevent gaps to occur. We will elaborate more on that in the next section. The examples are derived from [15], where also a gap between *Lip* and C^1 is shown, however the approach here simplifies the general proof strategy (could also be applied to the situation in [15]).

Remark 4.4. Here we do not have an example of the gap between Lipschitz and C^1 written out. It could be easily constructed by taking $g(t,x) = \frac{x^4-t^4}{x^4+t^4}$ but would have the exact same arguments with Taylor-expansions as the two other cases, i.e. it does not provide particular insight. Using $g(t,x) = \frac{x^4-t^4}{x^4+t^4}\frac{1}{x^2}$ can lead to a infinite gap between *Lip* and C^1.

4.2.1 Gap between C^1 and C^2

Take now the function $g(t,x)$ as

$$g(t,x) = \frac{x^4-t^6}{x^4+t^6},$$

and the following basic problem

$$\text{minimize } J(x) = \int_{-1}^{1} \Lambda(t,x,x')dt, \qquad x(-1) = x(1) = 1,$$

where the Lagrangian $\Lambda(t,x,v)$ is now given by

$$\Lambda(t,x,x') = \left|\frac{d}{dt}g(t,x)\right| = \left|\frac{4t^5x^3(2tx'-3x)}{(x^4+t^6)^2}\right|.$$

Choose now $x_*(t) = |t|^{3/2}$, a function that is in $C^1[-1,1]$ but not in $C^2[-1,1]$ (the second derivative is unbounded at $t=0$). Now $\Lambda(t,x_*,x_*')$ is 0 for all $t \in [-1,1]$ and thereby the whole functional is zero. However the behaviour for $x \in C^2[-1,1]$ looks different. As the integrand in the functional is always nonnegative, we can split up the integral, i.e.

$$\begin{aligned}
J(x) = \int_{-1}^{1}\left|\frac{d}{dt}g(t,x)\right|dt &\geq \lim_{\epsilon\to 0}\left(\int_{-1}^{-\epsilon}\left|\frac{d}{dt}g(t,x)\right|dt + \int_{\epsilon}^{1}\left|\frac{d}{dt}g(t,x)\right|dt\right)\\
&\geq \lim_{\epsilon\to 0}\left(\left|\int_{-1}^{-\epsilon}\frac{d}{dt}g(t,x)dt\right| + \left|\int_{\epsilon}^{1}\frac{d}{dt}g(t,x)dt\right|\right)\\
&= \lim_{\epsilon\to 0}\left(|g(-\epsilon,x(-\epsilon)) - g(-1,x(-1))| + |g(1,x(1)) - g(\epsilon,x(\epsilon))|\right)\\
&= \lim_{\epsilon\to 0}|g(-\epsilon,x(-\epsilon))| + \lim_{\epsilon\to 0}|g(\epsilon,x(\epsilon))|,
\end{aligned}$$

where we have used that $g(1,x(1)) = g(-1,x(-1)) = 0$. Start by writing $x(t)$ close to $t=0$ as

$$x(\epsilon) = x(0) + \epsilon x'(0) + \mathcal{O}(\epsilon^2).$$

This inserted into g leads to

$$g(\epsilon,x(\epsilon)) = \frac{(x(0)+\epsilon x'(0)+\mathcal{O}(\epsilon^2))^4 - \epsilon^6}{(x(0)+\epsilon x'(0)+\mathcal{O}(\epsilon^2))^4 + \epsilon^6}.$$

We want to look at the limit when $\epsilon \to 0^+$. There are three cases to consider.

1. **x(0)≠0:** When passing to the limit the only term that remains is $x(0)$, thus

$$\lim_{\epsilon\to 0} g(\epsilon, x(\epsilon)) = \frac{x(0)}{x(0)} = 1.$$

2. **x(0)=0 and x'(0)≠0:** Expanding now both the numerator and the denominator gives

$$\begin{aligned} g(\epsilon, x(\epsilon)) &= \frac{(\epsilon x'(0) + \mathcal{O}(\epsilon^2))^4 - \epsilon^6}{(\epsilon x'(0) + \mathcal{O}(\epsilon^2))^4 + \epsilon^6} \\ &= \frac{\epsilon^4 x'(0)^4 + x'(0)^3\mathcal{O}(\epsilon^5) + x'(0)^2\mathcal{O}(\epsilon^6) + x'(0)\mathcal{O}(\epsilon^7) + \mathcal{O}(\epsilon^8)) - \epsilon^6}{\epsilon^4 x'(0)^4 + x'(0)^3\mathcal{O}(\epsilon^5) + x'(0)^2\mathcal{O}(\epsilon^6) + x'(0)\mathcal{O}(\epsilon^7) + \mathcal{O}(\epsilon^8)) + \epsilon^6}. \end{aligned}$$

   When passing to the limit, the first terms in the numerator and denominator dominate thus

$$\lim_{\epsilon\to 0} g(\epsilon, x(\epsilon)) = \lim_{\epsilon\to 0} \frac{\epsilon^4 x'(0)^4}{\epsilon^4 x'(0)^4} = 1.$$

3. **x(0)=x'(0)=0:** Here

$$\lim_{\epsilon\to 0} g(\epsilon, x(\epsilon)) = \lim_{\epsilon\to 0} \frac{\mathcal{O}(\epsilon^8) - \epsilon^6}{\mathcal{O}(\epsilon^8) + \epsilon^6} = \lim_{\epsilon\to 0} \frac{-\epsilon^6}{+\epsilon^6} = -1.$$

The scenario for $g(-\epsilon, x(-\epsilon))$ is the same, it follows

$$\begin{aligned} \lim_{\epsilon\to 0} |g(-\epsilon, x(-\epsilon))| = \lim_{\epsilon\to 0} |g(\epsilon, x(\epsilon))| = 1, \\ \implies J(x) \geq 1 + 1 = 2. \end{aligned}$$

It follows that $\inf_{x\in C^2} J(x) = 2 > 0 = \inf_{x\in C^1} J(x)$.

4.2.2 Gap between C^k and C^{k+1}

This example is a slight generalization of the previous subsection, but follows the same spirit. Define now $g(t, x)$ as

$$g(t, x) = \frac{x^2 - |t|^{2p}}{x^2 + |t|^{2p}}, \qquad 1 \leq k < p < k + 1,$$

and the following basic problem

$$\text{minimize } J(x) = \int_{-1}^{1} \Lambda(t, x, x')dt, \qquad x(-1) = x(1) = 1.$$

The Lagrangian $\Lambda(t, x, v)$ is now given by

$$\Lambda(t, x, x') = \left|\frac{d}{dt} g(t, x)\right| = \left|\frac{d}{dt}\left(\frac{x^2 - |t|^{2p}}{x^2 + |t|^{2p}}\right)\right| = \left|\frac{4x(t)x'(t)|t|^{2p} - 4p|t|^{2(p-1)}tx(t)^2}{(x(t)^2 + |t|^{2p})^2}\right|,$$

where we have made use of the fact that $\frac{d}{dt}|t|^q = qt|t|^{q-2}$ for $q \geq 2$. We see now that the Lagrangian is zero if we choose $x_*(t) = |t|^p$,

$$4x_*(t)x_*'(t)|t|^{2p} - 4p|t|^{2(p-1)}tx_*(t)^2 = 4|t|^p pt|t|^{p-2}|t|^{2p} - 4p|t|^{2(p-1)}t|t|^{2p} = 0.$$

Now as $k < p < k+1$ it follows that x_* is element of C^k but not of C^{k+1}. Consider now the case where $x \in C^{k+1}$. As previously derived

$$J(x) \geq \lim_{\epsilon \to 0} \left(\left| \int_{-1}^{-\epsilon} \frac{d}{dt} g(t,x)dt \right| + \left| \int_{\epsilon}^{1} \frac{d}{dt} g(t,x)dt \right| \right) \geq \lim_{\epsilon \to 0} |g(-\epsilon, x(-\epsilon))| + \lim_{\epsilon \to 0} |g(\epsilon, x(\epsilon))|. \qquad (\dagger)$$

Again writing $x(t)$ as its Taylor expansion close to 0 leads to

$$x(\epsilon) = x(0) + \epsilon x'(0) + ... + \epsilon^k x^{(k)}(0) + \mathcal{O}(\epsilon^{k+1}),$$

$$g(\epsilon, x(\epsilon)) = \frac{(x(0) + \epsilon x'(0) + ... + \epsilon^k x^{(k)}(0) + \mathcal{O}(\epsilon^{k+1}))^2 - \epsilon^{2p}}{(x(0) + \epsilon x'(0) + ... + \epsilon^k x^{(k)}(0) + \mathcal{O}(\epsilon^{k+1}))^2 + \epsilon^{2p}}.$$

Now when passing to the limit there are several cases to consider. Let now $0 \leq j < k$ be the first index such that $x^{(j)} \neq 0$. Then

$$\lim_{\epsilon \to 0} g(\epsilon, x(\epsilon)) = \lim_{\epsilon \to 0} \frac{(\epsilon^j x^{(j)}(0) + ... + \epsilon^k x^{(k)}(0) + \mathcal{O}(\epsilon^{k+1}))^2 - \epsilon^{2p}}{(\epsilon^j x^{(j)}(0) + ... + \epsilon^k x^{(k)}(0) + \mathcal{O}(\epsilon^{k+1}))^2 + \epsilon^{2p}} = \lim_{\epsilon \to 0} \frac{e^{2j} \left(x^{(j)}(0)\right)^2}{e^{2j} \left(x^{(j)}(0)\right)^2} = 1,$$

because all the other terms with higher index are of higher power than j in ϵ and vanish in the limit. If $x^{(j)} = 0 \; \forall j : 0 \leq j \leq k$, then only the the following remains ($2p < 2(k+1)$).

$$\lim_{\epsilon \to 0} \frac{\mathcal{O}(\epsilon^{k+1})^2 - \epsilon^{2p}}{\mathcal{O}(\epsilon^{k+1}) + \epsilon^{2p}} = \lim_{\epsilon \to 0} \frac{\mathcal{O}(\epsilon^{2k+2}) - \epsilon^{2p}}{\mathcal{O}(\epsilon^{2k+2}) + \epsilon^{2p}} = \lim_{\epsilon \to 0} \frac{-\epsilon^{2p}}{\epsilon^{2p}} = -1.$$

In any case the limit is ± 1, thus from ($\dagger$)

$$J(x) \geq \lim_{\epsilon \to 0} |g(-\epsilon, x(-\epsilon))| + \lim_{\epsilon \to 0} |g(\epsilon, x(\epsilon))| = 1 + 1 = 2,$$

and a Lavrentiev-like phenomenon occurs, that is

$$\inf_{x \in C^k} J(x) = 0 < 2 = \inf_{x \in C^{k+1}} J(x).$$

5 Regularity

In order to obtain an existence theory for a global minimizer we had to extend the class of admissible functions from *Lip*$[a,b]$ to *AC*$[a,b]$. However knowledge about existence alone is not such a big help in finding the actual solution. We would like some necessary condition so that we have a set of candidate-solutions. The "strongest" necessary condition we had so far was (IEL), which deals with Lipschitz functions. Now it is dangerous to hastily generalize the IEL to absolutely continuous functions, as there are functions in *AC* that are not Lipschitz. In the previous section a couple of examples were treated where the solutions in different classes (e.g. *AC* vs *Lip*) were actually different. For being able to use (IEL) (and have access to a vast range of numerical solvers), we want to force the existing solution to lie in a smaller class than *AC*$[a,b]$, i.e. in *Lip*$[a,b]$, by using additional structure of the functional. In general, these procedures are known as regularity conditions. (The chapter is based on [6, Chapter 16.2, 16.3])

5.1 Tonelli'-Morrey theorem - Regularity via growth conditions

The first theorem only lifts the solution from *AC* to *Lip*, but, depending on the integrand, it may be even possible to assert higher regularities using results like 1.8.

Theorem 5.1. (Tonelli-Morrey) *Suppose that Λ and its derivatives in the second and third variable Λ_x, Λ_v are continuous. Suppose also that*

$$|\Lambda_x(t,x,v)| + |\Lambda_v(t,x,v)| \leq c|v| + c|\Lambda(t,x,v)| + d(t) \qquad \forall (t,x,v) \in [a,b] \times S \times \mathbb{R}^n, \tag{!}$$

where $c \in \mathbb{R}$ is a constant, $S \subseteq \mathbb{R}^n$ bounded and $d \in L^1[a,b]$. Then any (weak) local minimizer x_* satisfies (IEL).

Proof. The idea is to show that

$$f(t) = \Lambda_v(t, x_*(t), x_*'(t)) - \int_a^t \Lambda_x(s, x_*(s), x_*'(s)) ds,$$

is constant almost everywhere (i.e. IEL is fulfilled) by using the fundamental lemma of the Calculus of Variations 1.7, once the hypotheses $\int_a^b f(t)\phi'(t)dt = 0 \ \ \forall \phi \in Lip[a,b]$ is established.
Now Λ_v and Λ_x are bounded above by an L^1-function from (!) (for the RHS, $c|x_*'(t)| \in L^1[a,b]$ as $x_* \in AC[a,b]$ and $|\Lambda(t, x_*(t), x_*'(t))| \in L^1[a,b]$), moreover the integral of Λ_x is also in L^1(even continuous), it follows that $f \in L^1[a,b]$.
Take now a subset Y of *Lip*$[a,b]$ given by $Y = \{y \in Lip[a,b] : \|y\| + \|y'\| \leq 1\}$. Take an arbitrary Lipschitz function ϕ; the function values $\phi(t)$ and the derivative $\phi'(t)$ are essentially bounded in $[a,b]$, thus ϕ can be written as a finite sum of functions in Y. It is thus enough to show the hypotheses of the lemma holds for all functions in Y (i.e $\int_a^b f(t)y'(t)dt = 0 \ \forall y \in Y$), it then holds for all functions in *Lip*$[a,b]$. Define now the function

$$g(t,s) = \Lambda(t, x_*(t) + sy(t), x_*'(t) + sy'(t)) - \Lambda(t, x_*(t), x_*'(t)), \qquad s \in [0,1],$$

for all $t \in [a,b]$ whenever $x_*'(t)$ and $y'(t)$ both exist. The definition holds almost everywhere in $[a,b]$ as the combined measure of two measure zero sets (where $x_*'(t)$ or $y'(t)$ respectively do not exist) is again zero.
Now x_* was a local minimizer to begin with and thus

$$\int_a^b g(t,s)dt = J(x_* + sy) - J(x_*) \geq 0 \qquad \text{if } s \text{ sufficiently small.} \tag{†}$$

Consecutively we perform the following estimate,

$$\begin{aligned}\left|\frac{d}{ds}g(t,s)\right| &= |\Lambda_x(t,x_*(t)+sy(t),x_*'(t)+sy'(t))y(t)+\\ &\qquad \Lambda_v(t,x_*(t)+sy(t),x_*'(t)+sy'(t))y'(t)|\\ &\le |\Lambda_x(t,x_*(t)+sy(t),x_*'(t)+sy'(t))|\cdot|y(t)|+\\ &\qquad |\Lambda_v(t,x_*(t)+sy(t),x_*'(t)+sy'(t))|\cdot|y'(t)|\\ &\le |\Lambda_x(t,x_*(t)+sy(t),x_*'(t)+sy'(t))|+|\Lambda_v(t,x_*(t)+sy(t),x_*'(t)+sy'(t))|\\ &\le c\,(|x_*'(t)+sy'(t)|+|\Lambda(t,x_*(t)+sy(t),x_*'(t)+sy'(t))|)+d(t)\\ &\le c\,(1+|x_*'(t)|+|\Lambda(t,x_*(t)+sy(t),x_*'(t)+sy'(t))|)+d(t)\\ &= c\,(1+|x_*'(t)|+|\Lambda(t,x_*(t),x_*'(t))|+|g(t,s)|)+d(t).\end{aligned}$$

We compare the far LHS to the far RHS. $x_*$is in $AC[a,b]$, so $x_*' \in L^1[a,b]$, and $\int_a^b |\Lambda(t,x_*(t),x_*'(t))|$ finite (because x_* is a minimum) so the estimate can be combined into

$$\left|\frac{d}{ds}g(t,s)\right| \le c|g(t,s)|+d_2(t), \qquad d_2\in L^1[a,b].$$

Now a special case of Gronwall's lemma (see [16, p.2]), namely

$$u' \le au+b \implies u(s)\le u(0)e^{as}+\frac{b}{a}(e^{as}-1).$$

Now seeing s as the variable and $d_2(t)$ constant wrt. s,

$$|g(t,s)| \le |g(t,0)e^{cs}|+\frac{d_2(t)}{c}(e^{cs}-1)=\frac{d_2(t)}{c}(e^{cs}-1).$$

Now $0\le s\le 1$, so the function $e^{cs}-1$ is bounded above by the function $h(s)=(e^c-1)s=M_1 s$, thus

$$|g(t,s)| \le \frac{d_2(t)}{c}(e^{cs}-1)\le \frac{d_2(t)}{c}M_1 = sMd_2(t).$$

Using the fact that $g(t,0)=0$ leads to

$$\left|\frac{g(t,s)-g(t,0)}{s}\right| = \frac{|g(t,s)|}{s} \le \frac{sMd_2(t)}{s} = Md_2(t) \quad \text{for almost all } t.$$

But that means the difference quotient is bounded by the integrable function $Md_2(t)$ for every s, so Lebesgue's dominated convergence theorem applies and the limit can be moved inside the integral. For sufficiently small s from (†) it follows

$$\begin{aligned} 0 &\le \int_a^b g(t,s)dt \qquad \forall s\in(0,\epsilon),\\ \implies 0 &\le \int_a^b \frac{g(t,s)}{s}dt = \int_a^b \frac{g(t,s)-g(t,0)}{s}dt \qquad \forall s\in(0,\epsilon).\end{aligned}$$

Now letting $s \to 0^+$ (using dominated convergence theorem to bring the limit inside the integral) leads to

$$\begin{aligned}
0 &\leq \lim_{s\to 0^+} \int_a^b \frac{g(t,s)-g(t,0)}{s} dt \\
&= \int_a^b \lim_{s\to 0^+} \frac{g(t,s)-g(t,0)}{s} dt = \int_a^b \frac{d}{ds} g(t,s)\Big|_{s=0} dt \\
&= \int_a^b \Lambda_x(t, x_*(t), x_*'(t)) \cdot y(t) + \Lambda_v(t, x_*(t), x_*'(t)) \cdot y'(t) dt \\
&= \int_a^b \Lambda_x(t, x_*(t), x_*'(t)) \cdot y(t) dt + \int_a^b \Lambda_v(t, x_*(t), x_*'(t)) \cdot y'(t) dt \\
&= \left[\int_a^t \Lambda_x(u, x_*(u), x_*'(u)) du \cdot y(t) \right]_a^b - \int_a^b \left(\int_a^t \Lambda_x(u, x_*(u), x_*'(u)) du \cdot y'(t) \right) dt \\
&\qquad\qquad + \int_a^b \Lambda_v(t, x_*(t), x_*'(t)) \cdot y(t)' dt \\
&= \int_a^b y(t)' \left(\Lambda_v(t, x_*(t), x_*'(t)) - \int_a^t \Lambda_x(u, x_*(u), x_*'(u)) du \right) dt \\
&= \int_a^b f(t) y'(t) dt.
\end{aligned}$$

Now y can be replaced by $-y$, the argument repeats to result in $0 \leq -\int_a^b f(t)y'(t)dt$, so we have actually equality, $\int_a^b f(t)y'(t)dt = 0 \; \forall y \in Y$ and the fundamental lemma gives directly (IEL). □

The minimizer fulfilling (IEL) is already a really nice property, usually we can find it by solving the IEL, numerically if necessary. However it does not tell much about the solution in itself, we would like to know its class too.

Definition 5.2. Nagumo growth: The Lagrangian is said to have Nagumo growth along x_* if there exists a superlinear function $h : \mathbb{R}_+ \to \mathbb{R}$, i.e. $\lim_{t\to\infty} \frac{h(t)}{t} = +\infty$ that bounds the Lagrangian from below, i.e.

$$\Lambda(t, x_*(t), v) \geq h(|v|) \qquad t \in [a,b], v \in \mathbb{R}^n.$$

Corollary 5.3. *If additionally to the growth condition in 5.1, $\Lambda(t,x,v)$ is also convex in v and has Nagumo growth along x_* then $x_* \in Lip[a,b]$.*

Proof. From the characterization for differentiable convex functions (2.18) $f(y) - f(x) \geq \langle f'(x), y - x\rangle$, as Λ is convex in v (additional assumption)

$$\begin{aligned}
\Lambda(t, x_*(t), 0) - \Lambda(t, x_*(t), x_*'(t)) &\geq \Lambda_v(t, x_*(t), x_*'(t) \cdot (0 - x_*'(t)) \\
&= -\Lambda_v(t, x_*(t), x_*'(t)) \cdot x_*'(t) \qquad \text{a.e.}
\end{aligned}$$

This inequality and the Nagumo growth condition lead to

$$\begin{aligned}
&\Lambda(t, x_*(t), 0) + \Lambda_v(t, x_*(t), x_*'(t)) \cdot x_*'(t) \geq \Lambda(t, x_*, x_*') \geq h(|x_*'(t)|) \quad \text{a.e.} \\
&\implies h(|x_*'(t)|) \leq \max_{t\in[a,b]} |\Lambda_v(t, x_*(t), x_*'(t))||x_*'(t)| + \max_{t\in[a,b]} |\Lambda(t, x_*(t), 0)| = A|x_*'(t)| + B \text{ a.e.} \\
&\implies |x_*'(t)| \leq D \text{ a.e.} \implies x_* \in Lip[a,b].
\end{aligned}$$

The first implication follows from the fact that $\Lambda(t, x_*(t), 0)$ and $\Lambda_v(t, x_*(t), x_*'(t))$ are continuous(by assumption) and thus bounded on $[a, b]$. But h was of superlinear growth, so for h to be bounded by a linear function a.e., its argument has to be bounded almost everywhere as well, that implies $x_* \in Lip$. □

5.2 Regularity for higher classes

Examples in the previous section showed that a gap can also occur between higher classes. Regularity conditions here are also possible. We follow here the outline of [6, Exercise 21.13].

Lemma 5.4. *Let $h \in Lip[a, b]$ and $\bar{\epsilon} > 0$. Then there exists a polynomial $g(t)$ s.t. $g(a) = h(a), g(b) = h(b)$ and*

$$\|g'\| \leq \|h'\| + \bar{\epsilon}, \qquad \|g - h\| + \|g' - h'\|_{L^1} < \bar{\epsilon},$$

where $\|\cdot\|$ is the essential supremum norm, i.e. $\inf_{M\in\mathbb{R}}\{|f(t)| \leq M \text{ a.e.}\}$

Proof. .
Step 1 Define the sequence of functions h_n as

$$h_n = \begin{cases} \dfrac{h(x + \frac{1}{n} - \frac{x}{nb}) - h(x - \frac{1}{n} + \frac{x}{na})}{1/n} & \text{if } x \in [a, b], \\ 0 & \text{otherwise.} \end{cases}$$

Note that each h_n is measurable, being the composition and sum of continuous functions. Furthermore h_n converges to h' pointwise almost everywhere so invoking [17, Theorem 1.14] shows that the limit is also measurable. Actually we only have convergence almost everywhere, but that does not change the measurability of the limit. For more see [17, p.29 Section 1.37], where it is shown that for the Lebesgue measure (complete measures in general) sets of measure 0 are negligible. We can now apply Lusin's theorem 2.11 with h' to obtain a continuous function f on $[a, b]$ such that $\|f\| \leq \|h'\|$ and they are equal to each other except on a set of measure ϵ, i.e.

$$f(t) = h'(t) \qquad \forall t \in \mathcal{S} \text{ where } \mu(\mathcal{S}) = b - a - \epsilon.$$

Step 2 By Weierstrass approximation theorem we can find a polynomial p of finite degree that approximates f uniformly, i.e. $\|f - p\| < \epsilon$. Moreover

$$\begin{aligned} \|p - h'\| &= \|p - f + f - h'\| \leq \|p - f\| + \|f - h'\| \leq \epsilon + 0 = \epsilon && \text{on } \mathcal{S}, \\ \|p - h'\| &= \|p - f + f - h'\| \leq \|p - f\| + \|f - h'\| \leq \epsilon + 2\|h'\| && \text{on } [a, b]\backslash\mathcal{S}, \end{aligned}$$

by using $\|f - h'\| \leq \|f\| + \|h'\| \leq 2\|h'\|$.

Step 3 Let $\phi(t) = h(a) + \int_a^t p(s)ds$ and c be a constant, s.t. $|c| \leq \dfrac{\epsilon(b - a + 2\|h'\|)}{b - a}$. Define now $g(t) = \phi(t) + c(t - a)$. Now the following holds:

1. $g(t)$ is a polynomial.
2. $g(a) = \phi(a) = h(a)$.

3. $g(b) = h(a) + \int_a^b p(s)ds + c(b-a)$ and $h(b) = h(a) + \int_a^b h'(s)ds$ so for equality ($g(b) = h(b)$)
$$\int_a^b p(s)ds - \int_a^b h'(s)ds = \int_a^b p(s) - h'(s)ds = -c(b-a). \tag{\dagger}$$
Both integrants p and h' are integrable, thus there exists a constant $c \in \mathbb{R}$ such that the inequality holds. Approximation for c yields the necessary bound,
$$\begin{aligned} |c| &= \left| \int_a^b p(s) - h'(s)ds \right| \cdot \frac{1}{b-a} \leq \int_a^b |p(s) - h'(s)|ds \cdot \frac{1}{b-a} \\ &= \left(\int_{\mathcal{S}} |p(s) - h'(s)|ds + \int_{[a,b]\setminus\mathcal{S}} |p(s) - h'(s)|ds \right) \cdot \frac{1}{b-a} \\ &\leq ((b-a-\epsilon)\epsilon + \epsilon(\epsilon + 2\|h'\|)) \cdot \frac{1}{b-a} = \frac{\epsilon(b-a+2\|h'\|)}{b-a}. \end{aligned}$$
It follows that by choosing c appropriately in (†), $g(b) = h(b)$.

4. $\|p\| = \|p - f + f\| \leq \|p - f\| + \|f\| \leq \epsilon + \|h'\|$ implies that
$$\|g'\| = \|p + c\| \leq \|p\| + |c| \leq \|h'\| + |c| + \epsilon.$$

5. In a way similar to above we approximate
$$\begin{aligned} \|g - h\| &= \left\| h(a) + \int_a^t p(s)ds + c(t-a) - h(t) \right\| = \left\| \int_a^t p(s) - h'(s)ds + |c|(t-a) \right\| \\ &\leq \int_a^b |p(s) - h'(s)|ds + |c|(b-a) \leq \epsilon(b-a+2\|h'\|) + |c|(b-a). \end{aligned}$$

6. Again splitting up the integral and approximating each part yields
$$\begin{aligned} \|g' - h'\|_{L^1} &= \int_a^b |p(s) - h'(s) + c|ds \leq \int_a^b |p(s) - h'(s)|ds + c(b-a) \\ &\leq \epsilon(b-a+2\|h'\|) + (b-a)|c|. \end{aligned}$$

Step 4 Now as $h \in Lip$ whenever it exists, i.e. almost everywhere $h' \leq M_0$, so the above inequalities can be written as the following:

(i) $|c| \leq \frac{\epsilon M}{b-a}$,

(ii) $\|g'\| \leq \|h'\| + \epsilon \left(1 + \frac{M}{b-a}\right)$,

(iii) $\|g - h\| \leq 2\epsilon M$,

(iv) $\|g' - h'\|_{L^1} \leq 2\epsilon M$.

Choosing now $\epsilon = \min\{\frac{\bar{\epsilon}}{4M}, \frac{\bar{\epsilon}(b-a)}{b-a+M}\}$ suffices and the statement from the lemma follows, that is $g \in \mathcal{P}$, $g(a) = h(a)$, $g(b) = h(b)$ and
$$\|g'\| \leq \|h'\| + \bar{\epsilon}, \qquad \|g - h\| + \|g' - h'\|_{L^1} < \bar{\epsilon}.$$

□

Theorem 5.5. *If $\Lambda(t,x,v)$ is locally Lipschitz in v for every $(t,x) \in [a,b] \times \mathbb{R}$ and continuous in (t,x,v), then there is no gap between Lip and $\mathcal{P}$, i.e. for all x admissible*

$$\inf_{x \in Lip[a,b]} J(x) = \inf_{x \in \mathcal{P}[a,b]} J(x).$$

Proof. Let $\epsilon > 0$ be given.
Every $x \in Lip[a,b]$ has its function values bounded by M and its derivative by N. By the lemma there exists a polynomial y with $y(a) = x(a), y(b) = y(b)$, $\|y'\| \leq \|x'\| + \delta$ and $\|y - x\|, \|y' - x'\|_{L^1} < \delta$. It follows that y and y' are also bounded by $M + \delta$ and $N + \delta$ respectively. Furthermore Λ is uniformly continuous in (t,x,v), whenever $(t,x,v) \in [a,b] \times [-M-\delta, M+\delta] \times [-N-\delta, N+\delta]$, a compact set, i.e. choosing δ_0 small enough, $|\Lambda(t,x,v) - \Lambda(t,y,v)| < \epsilon \frac{1}{2(b-a)}$ if $\|(t,x,v) - (t,y,v)\| = |x-y| < \delta_0$.

$\Lambda(t,x,v)$ being locally Lipschitz in v means that for any compact interval $K_1 \subseteq \mathbb{R}$ and fixed (t,x) a Lipschitz condition is fulfilled, i.e.

$$|\Lambda(t,x,v) - \Lambda(t,x,w)| \leq L\|(t,x,v) - (t,x,w)\|, \;\; \forall (t,x,v),(t,x,w) \in [a,b] \times \mathbb{R} \times K_1.$$

Choosing now $K_1 = [-N-\delta, N+\delta]$ leads to

$$\begin{aligned}
|J(x) - J(y)| &= \left| \int_a^b \Lambda(t,x,x')dt - \int_a^b \Lambda(t,y,y')dt \right| \leq \int_a^b |\Lambda(t,x,x') - \Lambda(t,y,y')|dt \\
&= \int_a^b |\Lambda(t,x,x') - \Lambda(t,x,y') + \Lambda(t,x,y') - \Lambda(t,y,y')|dt \\
&\leq \int_a^b |\Lambda(t,x,x') - \Lambda(t,x,y')| + \int_a^b |\Lambda(t,x,y') - \Lambda(t,y,y')|dt \\
&\leq \int_a^b L\|(t,x,x') - (t,x,y')\|dt + \int_a^b |\Lambda(t,x,y') - \Lambda(t,y,y')|dt \\
&< L \int_a^b \sqrt{(x'-y')^2}dt + \int_a^b \frac{\epsilon}{2(b-a)}dt \quad \text{(by unif. continuity of } \Lambda) \\
&= L \int_a^b |x' - y'|dt + \frac{\epsilon}{2} = L\|x' - y'\|_{L^1} + \frac{\epsilon}{2} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,
\end{aligned}$$

where $\delta = \min\{\frac{\epsilon}{2L}, \delta_0\}$ to ensure that both terms are less than $\epsilon/2$ was chosen sufficiently small in the beginning. Therefore, as ϵ was arbitrary, $J(y)$ can be made arbitrary close to $J(x)$, the infima are the same. □

Remark 5.6. Note that we have not shown that the infima are the same if the Lagrangian Λ is merely continuous. So in that case it remains open if there are Lavrentiev-like gaps or not.

5.3 Clarke-Vinter theorem - autonomous Lagrangian

One of the most striking results is due to Clarke and Vinter (found in [6, p.330] or [5]), concerning autonomous (independent of t) Lagrangians. Autonomous Lagrangians are an important class, a lot of problems in classical mechanics are autonomous as the physical laws hold for any moment if time in the past or future. The theorem will be the final and most intricated part of this exposition.

Theorem 5.7. (Clarke-Vinter) *If x_* is a global minimizer for the functional $J(x)$ in the basic problem* (P) *over the class AC and the Lagrangian is autonomous, continuous, convex in v and has Nagumo growth along x_*, then x_* is Lipschitz.*

The proof we will give here will be a mixture of the proofs given in [6, p.330] and [5]. It will be structured into three parts. The first part sets up a certain minimization problem and its solution. Then in the second part, a generalized version of Lagrange-multipliers in combination with results from measure theory are used to approach the minimization problem in a different way. In the last step, by using the previous result, the Lipschitz property follows. Clarke and Vinter however used different methods in their original paper [8].
We will make use of some elementary concepts in non-smooth analysis. Roughly speaking it is about making analysis with functions that are not differentiable anymore but have other properties (usually convexity) that make analysis possible.

Definition 5.8. Let $f : X \to \mathbb{R}_\infty$ be a function from a normed space into the extended real numbers. An element $\xi \in X^*$ is called a **subgradient** of f at x if it satisfies

$$f(y) - f(x) \geq \xi \cdot (y - x), \qquad \forall y \in X. \tag{SGI}$$

Definition 5.9. The **subdifferential** denoted by $\partial f(x)$ is the set of all subgradients of f at x.

We can think of the subdifferential as the set of all hyperplanes in $X \times \mathbb{R}$ that lie below the epigraph of f. If f is differentiable at x then $\partial f(x) = f'(x)$. Furthermore, analogous to Fermat's rule, a function f attains its minimum at x only if $0 \in \partial f(x)$.

Theorem 5.10. *If $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, then for any $x \in \mathbb{R}^n$, $\partial f(x)$ is a nonempty convex compact set.*

Example 5.11. Let $X = \mathbb{R}$ and $f(x) = |x|$. Then $f'(0)$ does not exist in the classical sense but $\partial f(0) = \{r \in \mathbb{R} \text{ s.t. } -1 \leq r \leq 1\} = [-1, 1]$.

Proposition 5.12. [6, Exercise 13.23] *Let $\Lambda : \mathbb{R}^n \to \mathbb{R}$ be continuous and convex, $v_* \in \mathbb{R}^n$ and $\mathbb{R} \ni \alpha > 0$. Define $f(\alpha) = \Lambda(v_*/\alpha)\alpha$.*
Then f is convex on $(0, \infty)$ and if $\gamma \in \partial f(1)$ then $\exists \xi \in \partial\Lambda(v_*)$ s.t. $\gamma = \Lambda(v_*) - \xi \cdot v_*$.

Proof. Knowing that $\Lambda(tx_1 + (1-t)x_2) \leq t\Lambda(x_1) + (1-t)\Lambda(x_2)$, it follows

$$\begin{aligned}
f(s\alpha_1 + (1-s)\alpha_2) &= (s\alpha_1 + (1-s)\alpha_2)\Lambda\left(\frac{v_*}{s\alpha_1 + (1-s)\alpha_2}\right) \\
&= (s\alpha_1 + (1-s)\alpha_2)\Lambda\left(\frac{sv_*}{s\alpha_1 + (1-s)\alpha_2} + \frac{(1-s)v_*}{s\alpha_1 + (1-s)\alpha_2}\right) \\
&= (s\alpha_1 + (1-s)\alpha_2)\Lambda\left(\frac{s\alpha_1}{s\alpha_1 + (1-s)\alpha_2}\frac{v_*}{\alpha_1} + \frac{(1-s)\alpha_1}{s\alpha_1 + (1-s)\alpha_2}\frac{v_*}{\alpha_2}\right) \\
&\leq (s\alpha_1 + (1-s)\alpha_2)\left[\frac{s\alpha_1}{s\alpha_1 + (1-s)\alpha_2}\Lambda\left(\frac{v_*}{\alpha_1}\right) + \frac{(1-s)\alpha_1}{s\alpha_1 + (1-s)\alpha_2}\Lambda\left(\frac{v_*}{\alpha_2}\right)\right] \\
&= s\alpha_1\Lambda\left(\frac{v_*}{\alpha_1}\right) + (1-s)\alpha_1\Lambda\left(\frac{v_*}{\alpha_2}\right) = sf(\alpha_1) + (1-s)f(\alpha_2).
\end{aligned}$$

If $\alpha_1, \alpha_2 > 0$ and $0 \leq s \leq 1$, it is assured that $0 \leq \frac{s\alpha_1}{s\alpha_1+(1-s)\alpha_2}, \frac{(1-s)\alpha_1}{s\alpha_1+(1-s)\alpha_2} \leq 1$ and obviously $\frac{s\alpha_1}{s\alpha_1+(1-s)\alpha_2} + \frac{(1-s)\alpha_1}{s\alpha_1+(1-s)\alpha_2} = 1$, the convexity of $\Lambda(x)$ can be used.
For the second part we use the chain rule [6, 10.19] and product rule [6, 10.21] for ∂. (The rules are actually stated for ∂_C, the generalized gradient, but as f, Λ convex and lsc, ∂ and ∂_C are the same [6, 10.8]). Rewrite $f(\alpha)$ and apply first the product, then the chain rule (assuming that $\alpha \neq 0$).

$$\begin{aligned} f(\alpha) &= \Lambda\left(\frac{v_*}{\alpha}\right)\alpha = (\Lambda \circ h)(\alpha) \cdot g(\alpha) \qquad \text{with } g(\alpha) = \alpha \text{ and } h(\alpha) = \frac{v_*}{\alpha}, \\ \partial f(\alpha) &\subseteq (\Lambda \circ h)(\alpha) \cdot \partial g(\alpha) + \alpha \cdot \partial(\Lambda \circ h)(\alpha) = \Lambda(h(\alpha)) + \alpha \cdot \partial(\Lambda \circ h)(\alpha) \\ &\subseteq \Lambda(h(\alpha)) + \alpha \cdot \partial\Lambda(h(\alpha)) \cdot h'(\alpha) = \Lambda(h(\alpha)) - \frac{v_*}{\alpha}\partial\Lambda(h(\alpha)). \\ \text{Thus } &\partial f(1) \subseteq \Lambda(h(1)) - v_*\partial\Lambda(h(1)) = \Lambda(v_*) - v_*\partial\Lambda(v_*). \end{aligned}$$

Given $\gamma \in \partial f(1)$, it is thus always possible to choose a $\xi \in \partial\Lambda(v_*)$ such that $\gamma = \Lambda(v_*) - v_*\xi$. □

Lemma 5.13. [5, Proposition 1.2, Corollary 1.3] *Let $S \subseteq \mathbb{R}$ and $\phi : [a,b] \times S \to \mathbb{R}$ s.t. $\phi(t,p)$ is measurable in t for every fixed p and continuous in p for every fixed t. Let Σ be the non-empty set of all bounded measurable functions $p : [a,b] \to S$ s.t. $\int_a^b \phi(t,p(t))dt$ is well defined. Then*

$$\int_a^b \inf_{p \in S} \phi(t,p)dt = \inf_{p \in \Sigma} \int_a^b \phi(t,p(t))dt.$$

In particular if the function $p_* \in \Sigma$ minimizes $\int_a^b \phi(t,p(t))dt$ over Σ, then for almost every t the point $p_*(t)$ minimizes $\phi(t,\cdot)$ over S. See also Kuratowski–Ryll-Nardzewski measurable selection theorem.

Proof. Start by defining $\sigma(t) := \inf_{p \in S} \phi(t,p)$. Also for a countable dense set $\{d_i\}$ in S we have $\sigma(t) = \inf_i \phi(t,d_i)$, because $\phi(t,\cdot)$ is continuous. As the infimum is measurable if only countable sequences are treated, $\sigma(t)$ is measurable.
For an arbitrary element $p_0 \in \Sigma$, it holds that $\sigma(t) \leq \phi(t,p_0(t))$. Thus the integral $\int_a^b \sigma(t)dt$ is well defined, possibly as $-\infty$. If $\int_a^b \phi(t,p_0(t))dt = -\infty$, then the statement holds trivially, so assume it is finite.
Fix now $\epsilon > 0$ and let for each $t \in [a,b]$ the natural number $i(t)$ denote the first index where $\phi(t,d_i) < \sigma(t) + \epsilon$ ($i\colon [a,b] \to \mathbb{R}_+$, but $i([a,b]) \subseteq \mathbb{N}$). By continuity of ϕ in p and density of d_i in S, $i(t)$ always exists. Now we show that $d_{i(t)}\colon [a,b] \to S$ is measurable.
For that we look at the inverse image of an arbitrary open set $U \in S \subseteq \mathbb{R}$ and show that it is measurable (i.e. lies in the corresponding σ-algebra). Let the index set $J \subseteq \mathbb{N}$ be such that $d_j \in U, \forall j \in J$, then

$$\Gamma_1 := \{t \in [a,b] : d_{i(t)} \in U\} = \bigcup_{j \in J} \{t : i(t) = j\}.$$

So if each term in the countable union is measurable, measurability of $d_{i(t)}$ follows. However using the identity

$$\{t : i(t) = j\} = \bigcap_{k=1}^{j-1} \{t : \phi(t,d_k) \geq \sigma(t) + \epsilon\} \cap \{t : \phi(t,d_j) < \sigma(t) + \epsilon\}. \qquad (\star)$$

it is easy to see that the sets involved are measurable as ϕ is measurable in t by assumption. Also for the function $t \mapsto i(t)$ the inverse image of an open set $V \in \mathbb{R}_+$,

$$\Gamma_2 := \{t \in [a,b] : i(t) \in V\} = \bigcup_{v \in V} \{t : i(t) = v\} = \bigcup_{v \in V \cap \mathbb{N}} \{t : i(t) = v\},$$

is measurable by using ($\star$) again, for $v \notin V \cap \mathbb{N}$, $\{t : i(t) = v\} = \varnothing$).
Now define for $m > 0$,

$$\Omega_m := \{t \in [a,b] : i(t) < m\},$$

thus $\Omega_1 \subseteq \Omega_2 \subseteq ...$ is a nondecreasing sequence of sets and $[a,b] = \bigcup_m^\infty \Omega_m$.
Define furthermore the measurable function $p_m \in \Sigma$ by

$$p_m(t) = \begin{cases} d_{i(t)} & \text{if } t \in \Omega_m, \\ p_0(t) & \text{otherwise }. \end{cases}$$

Moreover

$$\inf_{p(\cdot)\in\Sigma} \int_a^b \phi(t,p(t))dt \le \int_a^b \phi(t,p_m(t))dt \le \int_{\Omega_m} \sigma(t) + \epsilon dt + \int_{\Omega_m^c} \phi(t,p_0(t))dt.$$

As $m \to \infty$, by continuity of measure(μ denotes now ordinary Lebesgue measure) from below $\lim_{m\to\infty} \mu(\Omega_m) = \mu([a,b])$ and by continuity of measure from above $\lim_{m\to\infty} \mu(\Omega_m^c) = \mu(\varnothing) = 0$, the RHS then tends to

$$\int_a^b \sigma(t)dt + (b-a)\epsilon.$$

But since ϵ was arbitrary, it then follows

$$\inf_{p(\cdot)\in\Sigma} \int_a^b \phi(t,p(t))dt \le \int_a^b \sigma(t)dt = \int_a^b \inf_{p\in S} \phi(t,p)dt.$$

For the opposite inequality, $\inf_{p\in S} \phi(t,p)$ gives a function $\psi(t)$ with the property $\psi(t) \le \phi(t,p(t))$ for any function $p(\cdot)$, so certainly for $p(\cdot) \in \Sigma$. This inequality is preserved by the integral and concluding from both inequalities the equality in the result follows
Additionally, if the finite infimum over Σ is attained by p_*,

$$\inf_{p(\cdot)\in\Sigma} \int_a^b \phi(t,p(t))dt = \int_a^b \phi(t,p_*(t))dt = \int_a^b \inf_{p\in S} \phi(t,p)dt,$$

so it follows that $\phi(t,p_*(t)) = \inf_{p\in S} \phi(t,p)$ almost everywhere. $\square$

Proof. (Clarke-Vinter theorem)
Step 1 Consider a measurable function $\alpha : [a,b] \to [1/2, 3/2]$ satisfying $h(\alpha) = \int_b^a \alpha(t)dt = b - a$. (Here any sufficiently small closed positive interval containing 1 would also work.) Define now (for a later change of variable)

$$\tau(t) = a + \int_a^t \alpha(s)ds.$$

As $|\tau(t)'| = |\alpha(t)| \le 3/2$, the derivative is bounded and so $\tau \in Lip[a,b]$. Moreover as $\alpha(s) > 0$, $\tau(t)$ is a strictly increasing function with range $[a,b]$, thus bijective. We proceed by looking at the derivative of the inverse. If τ were in C^1, we could invoke the classical inverse function theorem, but τ being merely Lipschitz needs some justification. In [11, Theorem 5.3] a direct result ($(F^{-1})'(y) = 1/F'(x)$ a.e.) is given, provided that for each p in the co-domain, the pre-image $\tau^{-1}(p)$ is connected. Here τ is bijective, the pre-images are only singleton sets which leads to

$$t'(\tau) = \frac{1}{\tau'(t)} = \frac{1}{\alpha(t)} \text{ a.e.}$$

Reparametrize now the arc $x_*(t)$ by $y(\tau) = x_*(t(\tau))$. The boundary conditions do not change giving

$$J(y) = \int_a^b \Lambda(y(\tau), y'(\tau))d\tau \geq J(x_*).$$

Applying now the change of variable $\tau = \tau(t)$, ($y'(\tau) = x'(t(\tau))/\alpha(t(\tau))$ and $d\tau = \alpha(t)dt$) gives

$$\int_a^b \Lambda\left(x_*(t), \frac{x_*'(t)}{\alpha(t)}\right)\alpha(t)dt \geq J(x_*),$$

where equality holds if $\alpha(t) = \alpha_*(t) \equiv 1$, i.e. α_* solves a certain constrained minimization problem which can be written explicitly as

$$\text{minimize }_{\alpha\in\Omega} f(\alpha) = \int_a^b \Phi(t,\alpha(t))dt = \int_a^b \Lambda\left(x_*(t), \frac{x_*'(t)}{\alpha(t)}\right)\alpha(t)dt,$$

where $\alpha(t)$ is a measurable element of

$$\Omega = \{x \in L^\infty[a,b] : x(t) \in [1/2, 3/2]\},$$

satisfying the constraint

$$h(\alpha) = \int_a^b \alpha(t)dt = b - a.$$

Ω is now a convex set. To show this claim, check if $z(t) = sx(t) + (1-s)y(t), 0 \leq s \leq 1$ are in Ω.

$$\frac{1}{2} = s\frac{1}{2} + (1-s)\frac{1}{2} \leq sx(t) + (1-s)y(t) \leq s\frac{3}{2} + (1-s)\frac{3}{2} = \frac{3}{2}.$$

From the first part of the proposition 5.12, $\Phi(t,\alpha)$ is convex in α for $0 < \alpha < \infty$. Composition of a measurable function and a continuous function is measurable, moreover measurable functions are closed under multiplication, it follows that $\Phi(t)$ is measurable and $f(\alpha)$ is well defined because it stems from a change of variable in $J(y)$. Convexity is preserved by the integral, thus the existence of a minimum is assured. The minimum α_* is now the minimum of a convex function over a convex set, a desirable setting in optimization.

Step 2 This problem is reminiscent of constrained optimization with Lagrange multipliers. Note here the difference between Λ which denotes the Lagrangian, a multivariable function and λ which is merely a scalar multiplier. We now use the finite case of the Hahn-Banach separation theorem (2.24) with the underlying space $\mathbb{R}^2$, i.e. if x_2 is a point not lying in the interior of a convex set K_2, then there exists a nontrivial $y \in (\mathbb{R}^2)^* = \mathbb{R}^2$ s.t. $y \cdot x_2 \leq y \cdot x, \forall x \in K_2$. We apply this to the convex set $K_2 \subseteq \mathbb{R}^2$ defined by

$$K_2 := \{(r, h(\alpha)) : \alpha \in \Omega, r \geq f(\alpha)\}.$$

and the point $x_2 = (f(\alpha_*), b-a)$. Thus we get a non-zero vector $(\lambda_0, \lambda) \in \mathbb{R}^2$ such that

$$\begin{pmatrix}\lambda_0\\ \lambda\end{pmatrix} \cdot \begin{pmatrix}f(\alpha_*)\\ b-a\end{pmatrix} \leq \begin{pmatrix}\lambda_0\\ \lambda\end{pmatrix} \cdot \begin{pmatrix}r\\ h(\alpha)\end{pmatrix},$$
$$\lambda_0 f(\alpha_*) + \lambda(b-a) \leq \lambda_0 r + \lambda h(\alpha) \qquad \text{for } \alpha \in \Omega \text{ and } r \geq f(\alpha).$$

Now λ_0 is non-negative, otherwise the inequality does not hold when $r \to \infty$, which are points still in K_2. Suppose $\lambda_0 = 0$ and λ positive, then $b - a \leq h(\alpha)$. Choosing now $\Omega \ni \alpha(t) \equiv 1/2$ leads to the contradiction $1 \leq 1/2$. In the same sense, if λ negative, choosing $\alpha(t) \equiv 3/2$ gives the contradiction $1 \geq 3/2$. Hence $\lambda_0 \neq 0$, we can thus normalize and take $\lambda_0 = 1$. It then follows

$$f(\alpha_*) + \lambda(b-a) \leq r + \lambda h(\alpha),$$
$$\int_a^b \Lambda(x_*(t), x_*'(t)) + \lambda dt \leq \int_a^b \Lambda\left(x_*(t), \frac{x_*'(t)}{\alpha(t)}\right)\alpha(t) + \lambda\alpha(t)dt.$$

Instead of now minimizing $f(\alpha)$ over Ω (minimum at $\alpha_*(t) \equiv 1$) we can use 5.13 to minimize the function
$$\theta_t(\alpha) := \Lambda\left(x_*(t), \frac{x_*'(t)}{\alpha}\right)\alpha + \lambda\alpha$$
over $\alpha \in [1/2, 3/2]$ i.e. $\theta_t(\alpha)$ attains a minimum on $[1/2, 3/2]$ at the point $\alpha = 1$. Thus 0 belongs to the subdifferential $\partial\theta_t(1)$. Now $\Lambda(x_*(t), v) + \lambda$ is convex in v for every fixed $x_*(t)$, so by the second part of Proposition 5.12 ($\gamma = 0$)

$$\Lambda(x_*(t), x_*'(t)) - \xi(t)x_*'(t) = -\lambda \quad a.e., \tag{$\dagger$}$$

where $\xi(t)$ is an element of the subdifferential of $v \to \Lambda(x_*(t), v)$ at the point $x_*'(t)$.

Step 3 Finally we want to show that $x_*'(t)$ is bounded. Let now t be such that $x_*'(t)$ exists (a.e., being in AC) and ($\dagger$) holds. Then using [Definition 5.8](SGI) (Λ is convex in v) gives ($\xi(t) \in \partial_v\Lambda(x_*(t), x_*'(t))$)

$$\begin{aligned}
&\Lambda\left(x_*(t), \frac{x_*'(t)}{1+|x_*'(t)|}\right) - \Lambda(x_*(t), x_*'(t)) \\
&\geq \xi(t)\left(\frac{x_*'(t)}{1+|x_*'(t)|} - x_*'(t)\right) = \left(\frac{1}{1+|x_*'(t)|} - 1\right)(\xi(t)x_*'(t)) \\
&= \left(\frac{1}{1+|x_*'(t)|} - 1\right)(\Lambda(x_*(t), x_*'(t)) + \lambda).
\end{aligned}$$

Moving terms gives the following

$$\frac{\Lambda(x_*(t), x_*'(t)) + \lambda}{1+|x_*'(t)|} - \lambda \leq \Lambda\left(x_*(t), \frac{x_*'(t)}{1+|x_*'(t)|}\right).$$

Let now $M = \max_{t\in[a,b], \|w\|\leq 1} \Lambda(x_*(t), w)$, which exists as Λ is continuous and $(t, w) \in [a, b] \times [-1, 1]$. Thus

$$\begin{aligned}
&\frac{\Lambda(x_*(t), x_*'(t)) + \lambda}{1+|x_*'(t)|} - \lambda \leq M \iff \\
&\Lambda(x_*(t), x_*'(t)) \leq -\lambda + (M+\lambda)(1+|x_*'(t)|) = M + (M+\lambda)|x_*'(t)| \\
&\implies \Lambda(x_*(t), x_*'(t)) \leq M + (M+|\lambda|)|x_*'(t)|.
\end{aligned}$$

By the Nagumo growth condition $\Lambda(x_*(t), v) \geq h_1(|v|)$ for some superlinear function h_1 which leads in view of the previous inequality to

$$h_1(|x_*'(t)|) \leq M + (M+|\lambda|)|x_*'(t)|$$
$$\frac{h(|x_*'(t)|)}{|x_*'(t)|} \leq \frac{M}{|x_*'(t)|} + (M+|\lambda|) \qquad \text{if } x_*'(t) \neq 0.$$

But $h(v)$ is a superlinear function, so if $|x_*'(t)|$ would go to infinity the LHS would go to infinity whereas the RHS would be bounded. That entails that $|x_*'(t)|$ does not go to infinity, the derivative is thus bounded almost everywhere, which is equivalent to a Lipschitz condition on x_*. □

5.4 An application of regularity to boundary value problems

An interesting way of using existence and regularity results are the study of boundary value problems in ODE-theory. The idea is to find a functional/problem that has a certain ODE as its Euler-Lagrange equation and then use the rich theory of existence and regularity results that was derived in the last chapters.

Example 5.14. Here we look at [6, Exercise 16.22], i.e. show the existence of a solution to the following boundary value problem.

$$x''(t) = \frac{1 - x'(t)^2}{1 + x'(t)^2}, \qquad x \in C^\infty[0,1], x(0) = 0, x(1) = 1. \tag{5.1}$$

In order to do so, consider the following variational problem

$$\text{minimize } J(x) = \int_0^1 e^{\alpha x(t) + \beta x'(t)^2} dt, \qquad x \in AC[0,1], x(0) = 0, x(1) = 1. \tag{P}$$

It is easy to see that ($e^{\alpha x} \geq 1 + \alpha x$ and $e^{\beta v^2} \geq 1 + \beta v^2$)

$$e^{\alpha x + \beta v^2} \geq \alpha x + \beta v^2 + 1 \geq \beta |v|^2 - \alpha |x| + 1,$$

and we are in the case of 3.8 ($r = 2, s = 1$), the extension of Tonelli's theorem, thus (P) has a solution x_* in $AC[0,1]$.

Another important fact is the strict convexity of the Lagrangian. Consider the eigenvalues of the Hessian (positive definite).

$$H_\Lambda = \begin{pmatrix} \alpha^2 e^{\alpha x + \beta v^2} & 2\alpha\beta v e^{\alpha x + \beta v^2} \\ 2\alpha\beta v e^{\alpha x + \beta v^2} & 2\beta e^{\alpha x + \beta v^2} + 4\beta^2 v^2 e^{\alpha x + \beta v^2} \end{pmatrix} \sim \begin{pmatrix} \alpha^2 & 2\alpha\beta v \\ 2\alpha\beta v & 2\beta + 4\beta^2 v^2 \end{pmatrix}.$$

Instead of computing the eigenvalues directly, use the following facts. $det(H_\Lambda) = \lambda_1 \lambda_2$ and $2tr(H_\Lambda) = \lambda_1 + \lambda_2$. For both eigenvalues to be positive it is required that $det, tr > 0$. Thus

$$\begin{aligned} det(H_\Lambda) &= \alpha^2(2\beta + 4\beta^2 v^2) - 4\alpha^2\beta^2 v^2 = 2\beta\alpha^2 > 0 \\ &\implies \beta > 0, \alpha \neq 0 \\ &\implies tr(H_\Lambda) = \alpha^2 + 2\beta + 4\beta^2 v^2 > 0. \end{aligned}$$

Thus Λ is strictly convex in (x, v) if the constants α and β are chosen appropriately. In other words

$$\Lambda(sx_1 + (1-s)x_2, sv_1 + (1-s)v_2) < s\Lambda(x_1, v_1) + (1-s)\Lambda(x_2, v_2) \qquad \text{if } s \in (0,1),$$

$$\begin{aligned} \implies J(sx + (1-s)y) &= \int_a^b \Lambda(sx + (1-s)y, sx' + (1-s)y')dt \\ &< \int_a^b s\Lambda(x, x') + (1-s)\Lambda(y, y')dt = sJ(x) + (1-s))J(y). \end{aligned}$$

That is the functional itself is strictly convex and thereby the solution x_* is unique.

For regularity of the solution, try first the Tonelli-Morrey theorem 5.1. The structural hypothesis is

$$\begin{aligned} |\alpha|e^{\alpha x+\beta v^2} + |2v\beta|e^{\alpha x+\beta v^2} &\leq c(|v| + e^{\alpha x+\beta v^2}) + d(t) \qquad \forall (t,x,v) \in [a,b] \times [k_1,k_2] \times \mathbb{R}, \\ e^{\alpha x+\beta v^2}(|\alpha| + |2\beta|v - c) &\leq c|v| + d(t) \\ e^{\alpha x}(|\alpha| + |2\beta v| - c) &\leq e^{\alpha x+\beta v^2}(|\alpha| + |2\beta v| - c) \leq c|v| + d(t) \\ |v|(2\beta e^{\alpha x} - c) - e^{\alpha x}(c-\alpha) &\leq d(t). \end{aligned}$$

Yet the last inequality does not hold anymore if the bounded set is chosen to be e.g. $S = [k_1, k_2] = [\frac{1}{\alpha}\ln\frac{|c|}{2\beta}+1, \frac{1}{\alpha}\ln\frac{|c|}{2\beta}+2]$. The RHS is the summable function $d(t)$, which is finite (i.e. $< M$) for almost all t). But the LHS is essentially of the form $A|v|+B$ and thereby unbounded as $v \to \infty$; the structural hypothesis is broken at some point and Tonelli-Morrey can not be invoked.

The right way to go is the Clarke-Vinter theorem 5.7. From the previous convexity in (x, v), convexity in v follows. Define now the Nagumo function $\theta(t) := e^{\alpha m}e^{\beta t^2}$ where m is the minimum of $x_*(t)$ over $[a, b]$. It is clear that $\theta(t)$ has superlinear growth. Now

$$\Lambda(t, x_*(t), v) = e^{\alpha x_*(t)}e^{\beta v^2} \geq e^{\alpha m}e^{\beta|v|^2} = \theta(|v|).$$

This holds for all $t \in [a, b]$ and $v \in \mathbb{R}$, the Nagumo growth condition is fulfilled and x_* is Lipschitz. Now we can use 1.10, the Lagrangian is C^∞ and Λ_{vv} (the (2,2) entry in H_Λ) is strictly positive, to lift the solution x_* into C^∞.

Look now at the EL equation to the problem (P).

$$\begin{aligned} \Lambda_x - \frac{d}{dt}\Lambda_v = \alpha e^{\alpha x(t)+\beta x'(t)^2} - \frac{d}{dt}\left(2\beta x'(t)e^{\alpha x(t)+\beta x'(t)^2}\right) &= 0 \\ \alpha e^{\alpha x(t)+\beta x'(t)^2} - \Big(2\beta x''(t)e^{\alpha x(t)+\beta x'(t)^2} + 2\beta x'(t)e^{\alpha x(t)+\beta x'(t)^2}(\alpha x'(t) + 2\beta x'(t)x''(t))\Big) &= 0 \\ \alpha - (2\beta x''(t) + 2\alpha\beta x'(t)^2 + 4\beta^2 x'(t)^2 x''(t)) &= 0 \\ x''(t)\left(4\beta^2 x'(t)^2 - 2\beta\right) + \left(\alpha - 2\alpha\beta x'(t)^2\right) &= 0 \\ x''(t) &= \frac{\alpha - 2\alpha\beta x'(t)^2}{2\beta - 4\beta^2 x'(t)^2}, \end{aligned}$$

where $\alpha \neq 0$, $\beta > 0$ are constant. Choosing $\alpha = 1, \beta = 1/2$ gives back the exact example (5.1) from Exercise 16.22. where we now have shown that there is a solution (existence of minimum) and as Λ strictly convex in (x, v), it follows that the solution of the EL is the unique minimum in C^∞ (by regularity conditions).

Example 5.15. Here we consider ODE-problems of the form $x'' = f(x)$ where $f \in C^\infty$. Try first an autonomous Lagrangian (Lagrangian does not necessarily need to be autonomous, the t-terms could chancel out in the EL-equation). We are looking thus for a $\Lambda(x, v)$ s.t.

$$\Lambda_x(x, v) - \Lambda_{vx}(x, v)v - \Lambda_{vv}(x, v)w = 0 = f(x) - w.$$

Assuming a simplified Lagrangian of the form $\Lambda(x,v) = g(x) + h(v)$, gives instead

$$\Lambda_x(x,v) - \Lambda_{vx}(x,v)v - \Lambda_{vv}(x,v)w = g'(x) - h''(v)w = f(x) - w,$$

which leads to $g(x) = F(x) = c + \int_a^x f(s)ds$ and $h(v) = \frac{v^2}{2}$. If we want to make use of Tonelli existence theorem it is necessary to impose further restrictions on f. For the original case Theorem (3.5) the condition on $f(x)$ is that its antiderivative $F(x)$ is bounded from below. If using the extension Theorem (3.8) unbounded $F(x)$ can be allowed, as long as it decreases slower than $-|x|^s, s < 2$ that is $F(x) \geq -C|x|^s + D$. Once existence is clarified, use the Clarke-Vinter theorem to assert that the solution is Lipschitz. By the smoothness of g and thereby $\Lambda(x,v)$ and the strict convexity in v 1.10 gives a C^∞ solution of the corresponding EL-equation.

Example 5.16. [6, Exercise 21.17].
Let the given ODE be

$$x''(t) = x^3(t) + bx^2(t) + cx(t) + d\sin(t), \qquad 0 \leq t \leq T, x(0) = x(T) = 0. \tag{*}$$

Now we need a Lagrangian of the form $\Lambda(t,x,v)$ thus

$$\Lambda_x(t,x,v) - \Lambda_{vt}(t,x,v) - \Lambda_{vx}(t,x,v)v - \Lambda_{vv}(t,x,v)w = 0 = x(t)^3 + bx(t)^2 + cx(t) + d\sin(t) - w.$$

Assuming now a Lagrangian of the form $\Lambda(t,x,v) = f(x) + g(t,v)$, the above becomes

$$f'(x) - g_{vt}(t,v) - g_{vv}(t,v)w = x(t)^3 + bx(t)^2 + cx(t) + d\sin(t) - w.$$

That in turn leads to $f(x) = \frac{x^4}{4} + \frac{bx^3}{3} + \frac{cx^2}{2} + D$ and $g(t,v) = \frac{v^2}{2} + vd\cos(t)$. Firstly, using Tonelli's theorem ,

$$\begin{aligned}
\Lambda(t,x,v) &= \frac{x^4}{4} + \frac{bx^3}{3} + \frac{cx^2}{2} + D + \frac{v^2}{2} + vd\cos(t) \\
&\geq \min_x\{\frac{x^4}{4} + \frac{bx^3}{3} + \frac{cx^2}{2} + D\} + \frac{v^2}{2} + vd\cos(t) \\
&= L + \frac{v^2}{2} + vd\cos(t) \geq \alpha|v|^r + \beta, \\
\Leftrightarrow \frac{v^2}{2} - \alpha|v|^r + vd\cos(t) &\geq \frac{v^2}{2} - \alpha|v|^r - |dv| \geq \beta - L,
\end{aligned}$$

there exist a solution in AC. For that let e.g. $\alpha = 1$ and $r = 1.5$, then choose $\beta = \min\{v^2/2 - |v|^r - |dv|\}$ which clearly exists as the minimum of a continuous, coercive function.
Clearly the Clarke-Vinter theorem can not be used but Tonelli-Morray can. The structural condition for $(t,x,v) \in [a,b] \times [-M,M] \times \mathbb{R}$,

$$\begin{aligned}
c|v| + c|\Lambda(t,x,v)| + p(t) &= c|v| + c\left|\frac{x^4}{4} + \frac{bx^3}{3} + \frac{cx^2}{2} + D + \frac{v^2}{2} + vdcos(t)\right| + p(t) \\
&\geq c|v| + h(t) \\
&= |v| + p(t) \qquad \text{(by choosing } c = 1) \\
&\geq |v| + K + |d| \qquad \text{(by choosing } p(t) \in L^1[a,b]) \text{ appropriately)} \\
&\geq |v| + |d\cos(t)| + K \\
&\geq |v + d\cos(t)| + \max_{x\in[-M,M]}\{|x^3 + bx^2 + cx|\} \\
&\geq |v + d\cos(t)| + |x^3 + bx^2 + cx| \\
&= |\Lambda_v(t,x,v)| + |\Lambda_x(t,x,v)|,
\end{aligned}$$

is fulfilled. For the Nagumo growth condition choose $h(v) = v^{1.5} + B$, where B is sufficiently small, then

$$\begin{aligned}\Lambda(t, x_*(t), v) &= \frac{x_*(t)^4}{4} + \frac{bx_*(t)^3}{3} + \frac{cx_*(t)^2}{2} + D + \frac{v^2}{2} + vdcos(t) \\ &\geq \frac{x_*(t)^4}{4} + \frac{bx_*(t)^3}{3} + \frac{cx_*(t)^2}{2} + D + \frac{v^2}{2} - |vd| \\ &\geq L + \frac{v^2}{2} - |vd| \geq |v|^{1.5} + B = h(|v|).\end{aligned}$$

Now $\Lambda_{vv} = g_{vv} = 1 > 0$ and the smoothness of Λ assure that the solution is of class C^∞.
Observe also that for fixed t the Hessian of $\bar{\Lambda}_t(x, v) = \Lambda(t, x, v)$,

$$H_{\bar{\Lambda}} = \begin{pmatrix} 3x^2 + 2bx + c & 0 \\ 0 & 1 \end{pmatrix},$$

has non-negative eigenvalues as long as $b^2 \leq 3c$ (the parabola $3x^2 + 2bx + c$ lies then above the x-axis). Hence $\Lambda(t, x, v)$ is convex in (x, v) and the solution of (*)(the EL) corresponds to the minimum.
For the uniqueness of the solution, note the additional strict convexity of $\Lambda(t, x, v) = f(x) + g(t, v)$ in v. Assume now there exists more than one minimizer, i.e. $J(x) = J(y) = m$. Then $z = \frac{x+y}{2}$ is also a minimizer (follows from convexity of Λ). Thus

$$\int_a^b \frac{\Lambda(t, x, x')}{2} + \frac{\Lambda(t, y, y')}{2} - \Lambda\left(t, \frac{x+y}{2}, \frac{x'+y'}{2}\right) dt = 0.$$

The convexity of Λ in (x, v) implies that the integrand is 0 almost everywhere, in particular for $\Lambda(t, x, v) = f(x) + g(t, v)$, assuming $x' \neq y'$,

$$\begin{aligned}\zeta = \frac{f(x) + g(t, x')}{2} + \frac{f(y) + g(t, y')}{2} &= f\left(\frac{x+y}{2}\right) + g\left(t, \frac{x'+y'}{2}\right) \\ &< f\left(\frac{x+y}{2}\right) + \frac{g(t, x') + g(t, y')}{2} \\ &\leq \frac{f(x) + f(y)}{2} + \frac{g(t, x') + g(t, y')}{2} = \zeta.\end{aligned}$$

But the LHS and the RHS are the same expression, $\zeta < \zeta$ is a contradiction and thus $x' = y'$, which makes the minimizer unique. (Endpoint conditions and derivative determine function completely). For a nice result that generalizes the uniqueness where $\Lambda(t, x, v)$ is convex and (x, v) and strictly convex in x or v, see [15].
Interesting to observe now, if $d = 0$, the trivial solution $x_*(t) \equiv 0$ is the only solution to (*) under the condition $b^2 \leq 3c$. If $d \neq 0$, there exist a unique nontrivial solution to (*), as $x_*(t) \equiv 0$ is not a solution anymore.

References

[1] Cantor function. `https://en.wikipedia.org/wiki/Cantor_function`.

[2] J.M. Ball and V.J. Mizel. One-dimensional variational problems whose minimizers do not satisfy the euler-lagrange equation. *Bulletin of The American Mathematical Society, Volume 11, Number 1*, 1984.

[3] J.M. Ball and V.J. Mizel. One-dimensional variational problems whose minimizers do not satisfy the euler-lagrange equation. *rch. Rational Mech. Anal. 90, 325–388 (1985)*, 1985.

[4] G. Buttazzo, M. Giaquinta, and S. Hildebrandt. *One-dimensional Variational Problems: An Introduction*. Oxford University Press, 1998.

[5] Francis Clarke. A lipschitz regularity theorem. *Ergod. Th. & Dynam. Sys. (2007), 27, 1–6*, 2007.

[6] Francis Clarke. *Functional Analysis, Calculus of Variations and Optimal Contral*. Springer, London, 2013.

[7] Bernard Dacorogna. *Direct Methods in the Calculus of Variations*. Springer-Verlag, Berlin-Heidelberg, 1989.

[8] F.H.Clarke and R.B. Vinter. Regularity properties of solutions to the basic problem in the calculus of variations. *Trans. Amer. Math. Soc. 289*, pages 73 – 98, 1985.

[9] Herman H. Goldstine. *A history of the calculus of variations from the 17th through the 19th century*. Springer, New York, 1980.

[10] David Hilbert. Mathematical problems. *Bull. Amer. Math. Soc. 8 (1902)*, 1902.

[11] Ralph Howard. Alexandrov's theorem on the second derivatives of convex functions via rademacher's theorem on the first derivatives of lipschitz functions. `https://people.math.sc.edu/howard/Notes/alex.pdf`.

[12] Mark Kot. *A First Course in the Calculus of Variations*. American Mathematical Society, Providence, Rhode Island, 2014.

[13] Marek Kuczma. *An Introduction to the Theory of Functional Equations and Inequalities, 2nd Edition*. Birkhäuser, Basel, Boston, Berlin, 2009.

[14] W. Hrusa M. Foss and V.Mizel. The lavrentiev gap phenomenon in nonlinear elasticity. *Arch. Rational Mech. Anal. 167, 337–365 (2003)*, 2003.

[15] Carlos Mantegazza. Some elementary questions in the calculus of variations. *Rend. Sem. Mat. Univ. Padova, , 1–7*.

[16] Stéphane Mitchel. An introduction to evolution pde's, chapter 1 - on the gronwall lemma. `https://www.ceremade.dauphine.fr/~mischler/Enseignements/M2evol2020/ch1.pdf`.

[17] Walter Rudin. *Real and Complex Analysis (2nd edition)*. McGraw-Hill, 1974.

[18] Richard L. Wheeden and Antoni Zygmund. *Measure and Integration: An Introduction to Real Analysis, 2nd edition*. CRC Press, 2015.

[19] Kosaku Yosida. *Functional Analysis (4th edition)*. Springer-Verlag, 1970.