THE UNIVERSITY OF WARWICK

AUTHOR: **Richard Gratwick** DEGREE: **Ph.D.**

TITLE: **Singular Minimizers in the Calculus of Variations**

DATE OF DEPOSIT: ....

I agree that this thesis shall be available in accordance with the regulations governing the University of Warwick theses.

I agree that the summary of this thesis may be submitted for publication.

I **agree** that the thesis may be photocopied (single copies for study purposes only).

Theses with no restriction on photocopying will also be made available to the British Library for microfilming. The British Library may supply copies to individuals or libraries. subject to a statement from them that the copy is supplied for non-publishing purposes. All copies supplied by the British Library will carry the following statement:

> "Attention is drawn to the fact that the copyright of this thesis rests with its author. This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's written consent."

AUTHOR'S SIGNATURE: ....

USER'S DECLARATION

1. I undertake not to quote or make use of any information from this thesis without making acknowledgement to the author.
2. I further undertake to allow no-one else to use this thesis while it is in my care.

DATE SIGNATURE ADDRESS

....

....

....

....

....

Singular Minimizers in the Calculus of Variations

by

Richard Gratwick

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Warwick Mathematics Institute

April 2011

THE UNIVERSITY OF WARWICK

Contents

Acknowledgments

I wish to thank Prof. David Preiss for his inspiration and guidance throughout the preparation of this thesis, in numerous enlightening conversations on both mathematical and non-mathematical subjects.

For unwavering personal support I would like to thank both my mother and Melissa.

Finally, I would like to dedicate this thesis to my father, Christopher Gratwick, forever an inspiration.

Declarations

The result of Chapter 2 is based on an article 'A one-dimensional variational problem with continuous Lagrangian and singular minimizer' to appear in *Archive for Rational Mechanics and Analysis*, available online, DOI: 10.1007/s00205-011-0413-3, as joint work of the author and David Preiss. The main idea for non-differentiability of a minimizer at an endpoint, outlined after Lemma 2.22, is due to David Preiss, as is the content of Remark 2.18. The rest of the Chapter is entirely the work of the author.

The result of Chapter 4 is based on an article of the author 'Universal singular sets of superlinear Lagrangians can contain purely unrectifiable F_σ sets', submitted in 2011 to *Zeitschrift für Analysis und ihre Anwendung.*

With the exception of standard results and those results clearly attributed to other authors, all the material in this thesis is new and original work of the author.

No material in this thesis has been submitted for a degree at any other university or institution.

Abstract

This thesis examines the possible failure of regularity for minimizers of one-dimensional variational problems. The direct method of the calculus of variations gives rigorous assurance that minimizers exist, but necessarily admits the possibility that minimizers might not be smooth. Regularity theory seeks to assert some extra smoothness of minimizers.

Tonelli's partial regularity theorem states that any absolutely continuous minimizer has a (possibly infinite) classical derivative everywhere, and this derivative is continuous as a function into the extended real line. We examine the limits of this theorem. We find an example of a reasonable problem where partial regularity fails, and examples where partial regularity holds, but the infinite derivatives of minimizers permitted by the theorem occur very often, in precise senses.

We construct continuous Lagrangians, strictly convex and superlinear in the third variable, such that the associated variational problems have minimizers non-differentiable on dense second category sets. Thus mere continuity is an insufficient smoothness assumption for Tonelli's partial regularity theorem.

Davie showed that any compact null set can occur as the singular set of a minimizer to a problem given via a smooth Lagrangian with quadratic growth. The proof relies on enforcing the occurrence of the Lavrentiev phenomenon. We give a new proof of the result, but constructing also a Lagrangian with arbitrary superlinear growth, and in which the Lavrentiev phenomenon does not occur in the problem.

Universal singular sets record how often a given Lagrangian can have minimizers with infinite derivative. Despite being negligible in terms of both topology and category, they can have dimension two: any compact purely unrectifiable set can lie inside the universal singular set of a Lagrangian with arbitrary superlinearity. We show this also to be true of F_σ purely unrectifiable sets, suggesting a possible characterization of universal singular sets.

Chapter 1

Introduction

1.1 The calculus of variations

The calculus of variations is one of the oldest branches of mathematics, and it remains an active area of research today. It is the study of minimal objects, i.e. objects which, when compared with other "competing" objects, minimize a certain numerical quantity. For example, the straight line between two fixed points is that curve with those fixed endpoints which has the smallest length, likewise the soap bubble formed across a wire loop dipped in soapy water takes the minimum surface area of all shapes that could span that loop.

As demonstrated by these examples, variational problems arise very naturally in real-world contexts: a very old example of a variational problem can even be traced (according to legend and Virgil's *Aeneid*) to Queen Dido of Carthage. Modern study of the calculus of variations is considered to have begun with the works of Newton and Bernoulli in the seventeenth century. Centuries later the subject was still of great interest. Two of Hilbert's famous problems posed to the International Congress of Mathematicians in Paris in 1900 deal with explicit problems in the calculus of variations [see Buttazzo et al., 1998, Introduction]:

Hilbert's 20th problem "Has not every regular variational problem a solution, provided certain assumptions regarding the given boundary conditions are satisfied, and provided also if need be that the notion of a solution shall be suitably extended?"

Hilbert's 19th problem "Are the solutions of regular problems in the calculus of variations always necessarily regular?"

Much of the study of the subject throughout the last century has been on these questions. The developing of the *direct methods of the calculus of variations* has had great impact on these questions, and our understanding of the issues of both existence and regularity of minimizers is not yet complete.

The basic problem of the one-dimensional calculus of variations is to minimize the functional

$$\mathscr{L}(u) = \int_a^b L(x, u(x), u'(x))\, dx \tag{1.1}$$

over some class $\mathscr{A}$ of functions $u\colon [a,b] \to \mathbb{R}$ with fixed boundary conditions, say $u(a) = A$ and $u(b) = B$ for fixed values $A, B \in \mathbb{R}$. Here $[a,b]$ is a fixed closed bounded subinterval of the real line. The function $L\colon [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is known as the *Lagrangian*, and the class $\mathscr{A}$ of functions u which we may consider is known as the *admissible functions*.

A typical example of such a problem is the so-called *brachistochrone*, first formulated by Gallileo, [see Buttazzo et al., 1998, Section 1.3]. The problem is to find, given two fixed points, the curve connecting them which will allow a frictionless mass threaded on the curve to move from one point to the other under the influence of gravity in the shortest time. Modelling the two points as (a, A) and (b, B) in the plane with $a < b$ and $A > B$, and assuming gravity to act in the direction of the negative y-axis, this amounts to minimizing the integral

$$\mathscr{L}(u) = \int_a^b \left(\frac{1 + |u'(x)|^2}{A - u(x)} \right)^{1/2} dx,$$

where each $u\colon [a,b] \to \mathbb{R}$ represents a possible curve connecting the two endpoints, i.e. $u(a) = A$, $u(b) = B$. The correct solution (the curve u must be a cycloid) was found by Johann Bernoulli in 1697.

Also posed by Galileo is the *heavy chain problem*: what shape is formed by a thin, heavy, inextensible chain suspended at its ends? This, when modelled as above, is equivalent to minimizing

$$\int_a^b u(x)(1 + |u'(x)|^2)^{1/2}\, dx$$

under the boundary conditions $u(a) = A$, $u(b) = B$, and the further condition imposed by the inextensibility, that

$$\int_a^b (1 + |u'(x)|^2)^{1/2}\, dx = l,$$

where l is the length of the chain. The solution, found independently by the Bernoulli brothers, Huygens, and Leibniz, is of form $u(x) = \alpha^{-1}(\cosh(\alpha x + \beta) + \gamma)$, where the constants α, β, γ depend on the data.

1.2 Direct methods

Traditionally solutions were sought to minimization problems among smooth functions, e.g. of class C^2, and the approach was to derive and examine so-called *necessary conditions* for functions to be minimizers. Certain conditions were derived which would have to be satisfied by any minimizer, and solutions were then sought among objects which did indeed satisfy these conditions. The best-known example is the *Euler Lagrange equation*

$$\frac{d}{dx} L_p(x, u(x), u'(x)) = L_y(x, u(x), u'(x)),$$

which must be satisfied by any function u furnishing a minimum to (1.1). This already requires that the Lagrangian L is at least of class C^1, say, and the derivation, which proceeds by computing the first variation of the functional $\mathscr{L}$ at a minimizer u, also requires some justification of an interchange of limits and integration.

However, this approach assumes that every minimization problem has a smooth solution: necessary conditions are found precisely by supposing a function to be a minimizer, and on this assumption proving properties of it. Such an assumption was apparently made without comment by, for example, Dirichlet and Riemann. There is, however, no justification for such an assumption in complete generality.

Tonelli realized that there was a need to prove directly that minimizers of variational problems certainly exist. He noticed that the notions of lower semicontinuity discussed by Baire, coupled with appropriate compactness properties of function spaces, would allow minimizers to be found as limits of minimizing sequences. Thus the space in which a minimizer was sought became critical to the validity of the theory.

Tonelli's method is what is now known as the *direct method* of the calculus of variations. Although we are concerned only with the one-dimensional case, the strategy is the same in higher dimensions, for maps u from $\mathbb{R}^n$ to $\mathbb{R}^N$, see e.g. the books of Dacorogna [2008] and Giusti [2003] for a discussion of this general case. The plan is as sketched above. We first take a minimizing sequence, i.e. a sequence of admissible functions u_k such that $\mathscr{L}(u_k) \to \inf_{u \in \mathscr{A}} \mathscr{L}(u)$. Using properties of the Lagrangian and the function space from which our competing objects are drawn, we show that we can extract a convergent subsequence, not re-indexed, $u_k \to u$ say. This function u is our candidate minimizer. To prove it is indeed a minimizer, we again need to use properties of the Lagrangian to show that the functional $\mathscr{L}$ is sequentially lower semicontinuous with respect to the topology in which the u_k converge to u. Then

$$\mathscr{L}(u) \leq \liminf_{k\to\infty} \mathscr{L}(u_k) = \lim_{k\to\infty} \mathscr{L}(u_k) = \inf_{u\in\mathscr{A}} \mathscr{L}(u),$$

since the sequence u_k was precisely chosen to be $\mathscr{L}$-minimizing.

This discussion requires a careful choice of topology in our function space. Our topology needs to be such that we can prove sequential compactness of minimizing sequences, under reasonable assumptions on the Lagrangian, and also sequential lower semicontinuity of the functional $\mathscr{L}$. There is a tension here: the first requirement asks for an abundance of convergent sequences, whereas the latter is more easily satisfied the fewer convergent sequences there are. Happily, this is not a fatal tension. The weak topology on Sobolev spaces is such that this method will in fact succeed.

Sobolev spaces as they are known now had not been defined in Tonelli's time, but the conditions he imposed, working in the space of absolutely continuous functions, are exactly those required when approaching the question from the point of view of weak topologies in Banach spaces. The absolutely continuous functions, introduced by Vitali, are precisely those functions for which the fundamental theorem of calculus holds: i.e. they are (classically) differentiable almost everywhere, and can be written as indefinite integrals of their derivatives. This is a strictly larger space than the spaces of C^1 and Lipschitz functions, in which solutions had previously been sought.

To prove sequential compactness of a minimizing sequence, one needs to impose on the Lagrangian the condition of superlinear growth in the third variable, i.e. that there exists a function $\omega\colon \mathbb{R} \to \mathbb{R}$, satisfying $\omega(p)/|p| \to \infty$ as $|p| \to \infty$, such that for all (x, y, p), $L(x, y, p) \geq \omega(p)$, [see Buttazzo et al., 1998, Theorem 2.13]. This corresponds to the *coercivity* condition seen in multidimensional problems.

The condition required for lower semicontinuity is that the Lagrangian is convex as a function of p, for each fixed $(x, y) \in \mathbb{R}^2$. Convexity also suffices in higher dimensions, but is rather too strong: the more nebulous condition of *quasiconvexity* was shown by Morrey [1952] to be the appropriate condition for weak sequential lower semicontinuity.

Some minimal smoothness assumptions are necessary for the details of the machinery to run, for example that $L \in C^1([a,b] \times \mathbb{R} \times \mathbb{R})$ certainly suffices.

Summarizing, we can state Tonelli's existence theorem, [see Buttazzo et al., 1998, Theorem 3.7].

Theorem 1.1. Let $L\colon [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be such that

- L and L_p are continuous;
- $p \mapsto L(x,y,p)$ is convex for all $(x,y) \in \mathbb{R}^2$; and
- L is superlinear with respect to p for each fixed $(x,y) \in \mathbb{R}^2$.

Then there exists a minimizer of (1.1) over the class of absolutely continuous functions.

This theorem can be extended to a more general class of Lagrangians, in particular those for which the first condition above is replaced with the requirement that L is a *Carathéodory function*, i.e. L is measurable as a function of x for all fixed $(y,p) \in \mathbb{R}^2$, and continuous as a function of (y,p) for almost every $x \in [a,b]$ [see Buttazzo et al., 1998, Section 3.2, Remark 1]. The observation most relevant for our discussion, particularly in light of the result of Chapter 2, is that continuity of the Lagrangian in (x,y,p) is already more than strong enough to guarantee existence of an absolutely continuous minimizer. Tonelli showed [1923; 1934] that the superlinearity condition may also be weakened slightly, e.g. it suffices for L to have superlinear growth in p for all values of (x,y) except those on, for example, the graph of a curve of finite length [see Buttazzo et al., 1998, Section 3.2, Remark 4].

Tonelli's first existence statement for the special case of superlinear growth of form $p \mapsto p^\alpha$ for $\alpha > 1$ can be found in Tonelli [1915] and Tonelli [1923], and for the general case in Tonelli [1934]. We mention also Cesari [1983] as a good reference on this topic. The trick to proving existence results is to enlarge the space of admissible functions to a space with a topology suitable for finding minimizers. The disadvantage is now this: classically we could be sure that any solution we found would be smooth, since we only ever considered smooth functions. Tonelli's existence result tells us that a minimizer exists, but tells us no more than that it is absolutely continuous. For an arbitrary absolutely continuous function, statements such as the Euler-Lagrange equation do not even make much sense: the derivative is not known to exist everywhere (only almost everywhere), and certainly no assertion about continuity of the derivative can be made in general.

This leads to the second main aspect of the direct method of the calculus of variations: *regularity theory*. This is the study of what further smoothness (or *regularity*) properties can be asserted of a minimizer, over an arbitrary admissible function. This usually consists of statements that higher derivatives of the minimizer exist and lie in certain Lipschitz or Hölder spaces, and satisfy certain equations, e.g the Euler-Lagrange equation, in various senses.

Under correspondingly strong conditions on the Lagrangian, it is possible to prove full regularity of a minimizer $u \in \mathrm{AC}(a, b)$. That is, the minimizer is of class C^k when the Lagrangian is of class C^k. The following is essentially Theorem 4.1 in Buttazzo et al. [1998], where one can find the proof.

Theorem 1.2. Suppose $L \in C^k([a, b] \times \mathbb{R} \times \mathbb{R})$, for $k \geq 2$, is such that

- there exist $c_0, c_1 > 0$ and $m > 1$ such that for all (x, y, p)

$$c_0|p|^m \leq L(x, y, p) \leq c_1(1 + |p|)^m;$$

- there exists a function $M\colon [0, \infty) \to (0, \infty)$ such that for all $(x, y) \in [a, b] \times \mathbb{R}$ with $x^2 + y^2 \leq R^2$, we have

$$|L_y(x, y, p)| + |L_p(x, y, p)| \leq M(R)(1 + |p|^2);$$

  and

- for all $(x, y) \in [a, b] \times \mathbb{R}$ we have

$$L_{pp}(x, y, p) > 0.$$

Then any minimizer u over the class of absolutely continuous functions is in fact of class C^k on $[a, b]$. Moreover, if L is real analytic, then u is real analytic.

1.2.1 Partial regularity

However, we should observe that the conditions required to pass to full regularity are notably stronger than those required just to prove existence. It seems reasonable to ask what hope of regularity there can be under minimal assumptions beyond those required for existence. Tonelli examined this question as well, and proved his *partial regularity theorem*, which is the result on which this thesis will concentrate. This statement is Theorem 4.6 in Buttazzo et al. [1998]:

Theorem 1.3. Let $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be of class C^∞ and satisfy $L_{pp}(x, y, p) > 0$ for every $(x, y, p) \in [a, b] \times \mathbb{R} \times \mathbb{R}$. Let u be a (strong local) minimizer of (1.1) over the absolutely continuous functions.

Then the classical derivative u' of u exists everywhere on $[a, b]$, albeit possibly with infinite values, and this derivative is continuous as a map into the extended real line $\mathbb{R} \cup \{\pm\infty\}$. Thus the *singular set* $E := \{x \in [a, b] : |u'(x)| = \infty\}$ is a closed set, necessarily of Lebesgue measure zero, and u is of class C^∞ off E.

Tonelli's original versions of this result can be found in Tonelli [1915], under an additional superlinearity condition, and later without in Tonelli [1923]. The methods of proof differ somewhat, see Clarke and Vinter [1985b].

We make some observations about this result. Firstly, note that the partial regularity assertion makes three statements: existence of the classical derivative, continuity of the derivative, and regularity away from a closed null set, the *singular set*.

The convexity condition required for existence has been strengthened slightly to the requirement that $L_{pp} > 0$, as in the theorem for full regularity. However, the superlinearity condition is not required. Some sort of growth condition is required here: the condition can be relaxed to (non-strict) convexity, if a superlinearity condition is also imposed. Csörnyei et al. [2008] established a version of the result with no convexity assumption at all, hence having to stipulate superlinear growth, and work in a generalized setting to deal with the lack of classical existence results. The only other aspect to discuss is the smoothness assumption made; the situation is not restricted to the C^∞ case: if the Lagrangian is of class C^k, then the minimizer is of class C^k off E. Positive results have been proved for Lagrangians satisfying only various Hölder and Lipschitz conditions. We discuss optimality of the conditions for the partial regularity theorem in greater detail in Chapter 2, collecting known results, and providing a construction schema showing that continuity of the Lagrangian is an insufficient smoothness assumption to prove everywhere differentiability of minimizers.

1.2.2 Lavrentiev phenomenon

Tonelli made no comment as to whether the singular set he defined can be non-empty. Minimizers with infinite derivative were first exhibited by Lavrentiev [1926]. Lavrentiev's work does not address the precise question of non-empty singular set under conditions of partial regularity (see Chapter 3), but he proves something rather more remarkable. Lavrentiev constructs a Lagrangian $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, superlinear and convex in p, such that the infimum of the functional $\mathscr{L}$ over the admissible functions of class C^1 is *strictly larger* than that over all admissible absolutely continuous functions:

$$\inf_{\substack{u \in \mathrm{AC}(a,b) \\ u(a)=A \\ u(b)=B}} \mathscr{L}(u) < \inf_{\substack{u \in C^1([a,b]) \\ u(a)=A \\ u(b)=B}} \mathscr{L}(u).$$

So it is impossible to approximate the minimum value using only smooth functions. This occurrence is known as the *Lavrentiev phenomenon.* Manià [1934] gave an example exhibiting the same phenomenon, but where the Lagrangian is a polynomial: the Lagrangian $L\colon [0,1] \times \mathbb{R} \times \mathbb{R} \to [0,\infty)$ given by

$$L(x,y,p) = (y^3 - x)^2 p^6$$

gives rise to the variational problem of minimizing

$$\mathscr{L}(u) = \int_0^1 ((u(x))^3 - x)^2 (u'(x))^6 \, dx$$

over those absolutely continuous functions u with $u(0) = 0$ and $u(1) = 1$. This clearly has a minimum value of 0 for the function $u(x) = x^{1/3}$. Thus the minimizer is not fully regular. However, Manià proves rather more, showing the existence of a positive number $\eta > 0$ such that $\mathscr{L}(u) \geq \eta > 0$ for any Lipschitz u. Any variational problem exhibiting the Lavrentiev phenomenon clearly has a minimizer which fails to be fully regular.

1.2.3 The singular set

Recall that in the situation of the partial regularity theorem, for a given minimizer u we associate with it a subset E of $[a,b]$ comprising those points $x \in [a,b]$ where $|u'(x)| = \infty$. Tonelli's result on continuity of the derivative tell us that it is closed; and since u' must be integrable, since u is absolutely continuous, we know also that is must be Lebesgue null.

Various people, including Tonelli himself, have produced results guaranteeing that the singular set must in fact be empty, i.e. the minimizer is fully regular. Examples of non-empty singular sets first appeared in Ball and Mizel [1984, 1985], and this direction of examination culminated in Davie's result [Davie, 1988] that any closed null set can appear as the singular set of a minimizer of a problem with smooth Lagrangian L, superlinear in p and with $L_{pp} > 0$. Some of the examples in Ball and Mizel [1985], and Davie's construction, exhibit the Lavrentiev phenomenon. In Chapter 3 we recall the precise information available about the singular set, and give another proof of Davie's result, without incurring a Lavrentiev gap.

1.2.4 Recording singular behaviour

Sychëv [1992] established a connexion between regularity of minimizers u of Hölder-continuous Lagrangians on the interval $[a, b]$ with boundary conditions $u(a) = A$ and $u(b) = B$, and the behaviour of the value function $S\colon \mathbb{R}^4 \to \mathbb{R}$ given by

$$S(a, A, b, B) = \inf\{\mathscr{L}(u) : u \in \mathrm{AC}(a, b),\ u(a) = A,\ u(b) = B\}.$$

Roughly speaking, this function is Lipschitz on a neighbourhood of a point (a, A, b, B) if and only if any minimizer to the corresponding variational problem on $[a, b]$ with boundary data $u(a) = A$ and $u(b) = B$ is regular. Sychev and Mizel [1998] produced related results in the vectorial case $u\colon [a, b] \to \mathbb{R}^n$.

Given a smooth superlinear Lagrangian, it is clear that it is somewhat unexpected for minimizers with infinite derivative to exist, since steep derivatives in u suggest large values of $L(x, u, u')$ exactly by the superlinearity condition. The *universal singular set* of a Lagrangian L, introduced by Ball and Nadirashvili [1993], records where singular minimizers can occur: a point $(x_0, y_0) \in \mathbb{R}^2$ is in the universal singular set of L, hereafter $\mathrm{uss}(L)$, if there is an interval $[a, b]$ in $\mathbb{R}$ containing x_0 and a choice of boundary conditions $u(a) = A$, $u(b) = B$, such that there is a minimizer u of the associated variational problem with $u(x_0) = y_0$ and $|u'(x_0)| = \infty$. It is known that the universal singular set is of the first Baire category [Ball and Nadirashvili, 1993] and that the two-dimensional Hausdorff measure of the universal singular set is zero [Sychëv, 1994]. Thus we know that in terms of both measure and topology, the universal singular set is negligible.

Csörnyei et al. [2008] study the geometry of the universal singular sets, investigating their intersections with rectifiable curves. They show that they are in some sense "almost" purely unrectifiable, and that any set which can be covered by universal singular sets of Lagrangians of arbitrary superlinearity is indeed purely unrectifiable. On the other hand, they also construct Lagrangians with large universal singular sets, in particular showing that any compact purely unrectifiable set can be covered by universal singular sets of arbitrary superlinearity. In particular universal singular sets can have maximum dimension possible for a subset of the plane. In Chapter 4 we discuss the universal singular sets in more detail, and show that F_σ purely unrectifiable sets can lie inside universal singular sets. This seems, as we explain, to come near to one direction of a characterization of universal singular sets.

1.3 Higher dimensions

Finally we note that (partial) regularity questions are very actively pursued in higher dimensions, in the analysis of multi-dimensional variational problems and (nonlinear) elliptic systems. The issue seems rather harder in this context, and it is consequently an area of somewhat greater study. For general introductions to the subject of the calculus of variations in higher dimensions see the books of Morrey [2008] and Dacorogna [2008]. Concentrating more on regularity theory are Giusti [2003] and Giaquinta [1983], and for example the survey of Mingione [2006]. We give a very rough sketch of the situation. The majority of the results are at least first formulated for Lagrangians which are functions only of the derivative Du of the admissible functions u; we restrict our discussion to this situation.

For scalar-valued functions, i.e. when minimizing over functions $u\colon \Omega \to \mathbb{R}$ for some domain $\Omega \subseteq \mathbb{R}^n$, we can prove Hölder continuity of the derivative Du. This is known as De Giorgi-Nash-Moser theory, and is based on the fundamental result of De Giorgi [1957] that solutions of linear elliptic equations in divergence form are $C^{0,\alpha}$. See Giusti [2003] and Giaquinta [1983] for discussions of this. That this fails in the vectorial case, $u\colon \Omega \to \mathbb{R}^N$ for $N \geq 2$, was shown via a number of counterexamples, first from De Giorgi himself [1968], then later also Giusti and Miranda [1968], Nečas [1977], and Šverák and Yan [2000]; Sverák and Yan [2002].

The interest then turned to partial regularity. Partial regularity statements in this context assert existence and (local) Hölder continuity of the derivative on a relatively open set of full measure in the domain. For convex problems positive results were proved by, among others, Morrey [1967/1968], Giusti and Miranda [1968/1969] and Giaquinta and Giusti [1978]. However, as mentioned above, the critical assumption for existence in the vectorial case is not convexity, but quasiconvexity: convexity is an unnecessarily strong condition. Thus arguably the most interesting issue is that of partial regularity under only the assumption of quasiconvexity.

On this topic, the main result is by Evans [1986], generalized later by Acerbi and Fusco [1987], and Giaquinta and Modica [1986].

Theorem 1.4 (Evans [1986]). Suppose $L\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ is of class C^2 and is such that

- there exists $\gamma > 0$ such that

$$\int_{B_r(x)} (L(\xi) + \gamma |D\phi(y)|^2)\, dy \leq \int_{B_r(x)} L(\xi + D\phi(y))\, dy$$

for all $x \in \mathbb{R}^n$, $r > 0$, $\xi \in \mathbb{R}^{N\times n}$, and test functions $\phi \in C_0^1(B_r(x); \mathbb{R}^n)$; and

- $|D^2 L(\xi)| \leq C$ for all $\xi \in \mathbb{R}^{N\times n}$, for some constant $C > 0$.

Let $u \in W^{1,2}(\Omega, \mathbb{R}^N)$ be a minimizer of

$$\mathscr{L}(u) = \int_\Omega L(Du(x))\, dx$$

subject to fixed boundary conditions on $\partial\Omega$.

Then there is an open set $\Omega_0 \subseteq \Omega$ such that $\mathrm{meas}(\Omega\backslash\Omega_0) = 0$ and for all $\alpha \in (0, 1)$ we have $Du \in C^\alpha(\Omega_0; \mathbb{R}^{N\times n})$.

An interesting observation in this higher-dimensional situation is that partial regularity is very much a *variational* phenomenon in the quasiconvex case, rather than a result about extremals (i.e. solutions of the associated Euler-Lagrange equation). Without the stronger assumption of convexity, it is possible to find functionals satisfying Evans' partial regularity theorem, for which there are extremals failing to be of class C^1 in any open subset of the domain [Müller and Šverák, 2003].

The question of low-order partial regularity, discussed in section 4.3 of Mingione [2006], and which we discuss in the one-dimensional case in Chapter 2, has recently been addressed by Foss and Mingione [2008]. They prove a positive result of partial $C^{0,\alpha}$ regularity for solutions of nonlinear elliptic systems, and quasiconvex variational problems, assuming only continuity of the coefficients.

Much of the work sharpening partial regularity results has been on understanding better the singular set. Kristensen and Mingione provide estimates of the Hausdorff dimension in the convex case [2006] and, assuming the minimizer to be Lipschitz, in the quasiconvex case [2007].

1.4 Basic notions and notation

We record here the set-up and notation used for the remainder of the thesis. Further notation and terminology used only within the individual chapters will be introduced in the appropriate chapter.

Our variational problems will always consider real-valued functions on a closed bounded subinterval $[a, b]$ of the real line $\mathbb{R}$. We denote by meas the n-dimensional Lebesgue measure on $\mathbb{R}^n$, and recall that a *Lebesgue null set*, or just *null set*, is a set $E \subseteq \mathbb{R}^n$ for which $\mathrm{meas}(E) = 0$. Any property which holds for all values of a subset of $\mathbb{R}^n$ except perhaps on a null set we say holds almost everywhere.

The 1-dimensional Hausdorff measure (used only in $\mathbb{R}^2$) will be denoted by $\mathscr{H}^1$; for the definition and properties, we recommend Federer [1969].

We let $\|\cdot\|_2$ denote the usual Euclidean norm on $\mathbb{R}^n$, which is the norm used throughout and for the following definitions. For $r > 0$, we will use $B_r(x)$ for the open ball in this Euclidean metric of radius $r > 0$ around the point $x \in \mathbb{R}^n$, similarly $B_r(X)$ denotes the Euclidean r-neighbourhood of a subset $X \subseteq \mathbb{R}^n$. The diameter of a non-empty set $X \subseteq \mathbb{R}$, $\operatorname{diam}(X)$, is defined as

$$\operatorname{diam}(X) = \sup\{|x - y| : x, y \in X\}.$$

The distance between two subsets $X, Y \subseteq \mathbb{R}^n$ shall be denoted $\operatorname{dist}(X, Y)$, and is defined by

$$\operatorname{dist}(X, Y) = \inf\{\|x - y\|_2 : x \in X,\ y \in Y\}.$$

In the case that $X = \{x\}$ for some $x \in \mathbb{R}^n$ we write $\operatorname{dist}(x, Y)$, and we interpret this value as ∞ if one of the sets is empty. Notation $X \Subset Y$ is used when the closure $\overline{X}$ of X is compact and contained in Y. Orthogonal projection from the plane onto the x-axis is denoted by π_X, i.e. $\pi_X \colon \mathbb{R}^2 \to \mathbb{R}$ is given by $\pi_X(x, y) = x$.

Given a function $f \colon \mathbb{R}^n \to \mathbb{R}$, we let $\|f\|_\infty$ denote the supremum norm of the function:

$$\|f\|_\infty = \sup\{|f(x)| : x \in \mathbb{R}^n\}.$$

The *support* of f, $\operatorname{spt}(f)$, is the smallest closed set outside of which f is zero:

$$\operatorname{spt}(f) = \overline{\{x \in \mathbb{R}^n : f(x) \neq 0\}}.$$

The classical derivative of a function $u \colon \mathbb{R} \to \mathbb{R}$ is represented by u'. Partial derivatives shall be denoted by subscripts, e.g. Φ_x, Φ_y, and L_p for functions $\Phi = \Phi(x, y) \colon \mathbb{R}^2 \to \mathbb{R}$ and $L = L(x, y, p) \colon \mathbb{R}^3 \to \mathbb{R}$. We emphasize that all derivatives are understood in the classical sense: there is no use of weak (distributional) derivatives, or of any non-smooth analysis as in the works of Clarke and Vinter. For any function $u \colon \mathbb{R} \to \mathbb{R}$ we let $U \colon \mathbb{R} \to \mathbb{R}^2$ be given by $U(x) = (x, u(x))$. We write

$$\operatorname{Lip}(u) = \sup_{\substack{t,x,\in X \\ t \neq x}} \frac{|u(t) - u(x)|}{|t - x|}.$$

Although of course not true in general, this will always be a finite number in our usage. When this is the case, the function u is said to be *Lipschitz*. For a point

$x \in \mathbb{R}$, $u_-(x)$ and $u_+(x)$ shall denote the left- and right-hand limits respectively:

$$u_-(x) = \lim_{\substack{t\to x\\ t<x}} u(t) \text{ and } u_+(x) = \lim_{\substack{t\to x\\ t>x}} u(t).$$

The upper and lower Dini derivatives of a function $u \in \mathrm{AC}(a,b)$ at a point $x \in [a,b]$ are given respectively by

$$\overline{D}u(x) = \limsup_{\substack{t\to x\\ t\neq x}} \frac{u(t)-u(x)}{t-x}, \text{ and } \underline{D}u(x) = \liminf_{\substack{t\to x\\ t\neq x}} \frac{u(t)-u(x)}{t-x}.$$

For a measurable function $u\colon \mathbb{R}^n \to \mathbb{R}$, and an open subset $X \subseteq \mathbb{R}^n$, the $L^\infty(X)$ norm is given by

$$\|u\|_{L^\infty(X)} = \sup\{c \in \mathbb{R} : \mathrm{meas}(\{x \in X : |u(x)| \geq c\}) > 0\}.$$

For $p \in [1,\infty)$, the $L^p(X)$ norm is defined by

$$\|u\|_{L^p(X)} = \left(\int_X |u|^p\right)^{1/p}.$$

The function u is said to lie in $L^p(X)$ if the number $\|u\|_{L^p(x)}$ is finite, and in $L^p_{\mathrm{loc}}(X)$ if $\|u\|_{L^p(Y)}$ is finite for every compact $Y \subseteq X$. The Sobolev space $W^{1,p}(X)$ is the space of functions $u \in L^p(X)$ with weak derivatives also in $L^p(X)$. We will not need the notion of weak derivative, and shall rarely mention Sobolev spaces. It suffices to observe that in the case that the classical derivative of a function u exists almost everywhere, then this is a weak derivative. For more information on the topic, we recommend Ziemer [1989].

Definition 1.5 (Absolutely continuous function)**.** We say $u\colon [a,b] \to \mathbb{R}$ is *absolutely continuous* if for every $\epsilon > 0$ there exists $\delta > 0$ such that, for all $n \geq 1$ and all non-overlapping subintervals $\{(a_i,b_i)\}_{i=1}^n$ of $[a,b]$, we have that

$$\sum_{i=1}^n |u(b_i) - u(a_i)| < \epsilon \text{ whenever } \sum_{i=1}^n |b_i - a_i| < \delta.$$

Such functions are [see Buttazzo et al., 1998, Section 2.2] exactly those functions which have almost everywhere a classical derivative, and this derivative is a Lebesgue integrable function, and satisfies the fundamental theorem of calculus:

$$u(y) - u(x) = \int_x^y u'(t)\,dt$$

for all $x, y \in [a, b]$. Moreover, these are essentially the Sobolev functions $W^{1,1}(a, b)$; i.e. those integrable functions which have a weak or distributional derivative which is also integrable. Precisely, every $W^{1,1}(a, b)$ function can be modified on a set of measure zero to be equal to an absolutely continuous function. Since our attention is precisely pointwise properties of derivatives, it seems more natural to frame our discussion in terms of absolutely continuous functions. We shall write $\mathrm{AC}(a, b)$ for the class of absolutely continuous functions on $[a, b]$.

For fixed $[a, b] \subseteq \mathbb{R}$, by a *Lagrangian* we shall mean a Borel measurable function $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Further conditions will be imposed on Lagrangians at various points; since much of the thesis is a discussion of optimality of conditions, we make no further standing assumptions at this stage and refer rather to individual discussions of the various topics. The basic problem of the one-dimensional calculus of variations is that of minimizing the functional $\mathscr{L}\colon \mathrm{AC}(a, b) \to \mathbb{R}$ given by

$$\mathscr{L}(u) = \int_a^b L(x, u(x), u'(x))\, dx \tag{1.2}$$

over those functions $u \in \mathrm{AC}(a, b)$ with boundary conditions $u(a) = A$ and $u(b) = B$, for some $A, B \in \mathbb{R}$. We shall let

$$\mathscr{A} = \mathscr{A}_{A,B} = \{u \in \mathrm{AC}(a, b) : u(a) = A,\ u(b) = B\}$$

denote the collection of *admissible functions* for the problem.

Definition 1.6. Let $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a Lagrangian, and boundary values $A, B \in \mathbb{R}$ be fixed. A function $u \in \mathscr{A}$ is a *minimizer* (or *global minimizer*) for the problem (1.2) if $\mathscr{L}(u) \le \mathscr{L}(v)$ for all $v \in \mathscr{A}$.

Definition 1.7. Let $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a Lagrangian, and boundary values $A, B \in \mathbb{R}$ be fixed. A function $u \in \mathscr{A}$ is a *strong relative minimizer* for the problem (1.2) if there exists $\delta > 0$ such that for any $v \in \mathscr{A}$ with $\|u - v\|_\infty < \delta$, we have $\mathscr{L}(u) \le \mathscr{L}(v)$.

We will rarely make use of this latter notion, and record it here just for reference, since most partial regularity results can be stated for strong local minimizers. Without further comment, a *minimizer* will always refer to a global minimizer.

Often in integrals we will omit the dummy variable of integration, e.g. we write $\int_a^b L(x, u, u')$ for $\int_a^b L(x, u(x), u'(x))\, dx$. We emphasize that variables u, v, w will always represent functions which take such dummy variables as their arguments.

Chapter 2

Optimal conditions for Tonelli's partial regularity theorem

2.1 Introduction

2.1.1 Positive results

The following theorem, stated here as by Ball and Mizel [1985], who also give a proof, is a typical statement of Tonelli's partial regularity theorem.

Theorem 2.1. Let $L\colon \mathbb{R}^3 \to \mathbb{R}$ be of class C^3, be bounded below, and satisfy $L_{pp} > 0$. Suppose u is a strong relative minimizer of (1.2).

Then u is of class C^1 when considered as a map into the extended real line $\mathbb{R} \cup \{\pm\infty\}$.

All the classical statements of Tonelli's theorem require at least that L be continuously differentiable. However, for the most general existence results, even the smoothness assumption of continuity is stronger than necessary (see Chapter 1). The pursuit of partial regularity is motivated by the desire to examine what further regularity we can expect of minimizers, over arbitrary elements of $\mathrm{AC}(a, b)$, under only minimal strengthenings of the "natural" assumptions for the problem, i.e. those that guarantee existence of a minimizer. Thus it is sensible to ask how far the assumptions for partial regularity can be lowered.

In the statement of Theorem 2.1, there are two main assumptions, both, as stated, stronger than those required for existence: the condition that $L_{pp} > 0$, and the smoothness condition. (We also assume above that L is bounded below. This is reasonable in light of the assumptions required for existence, but is also stronger than necessary; see the results below.) Recall that convexity is necessary for the existence results, but that the condition $L_{pp} > 0$ is a stronger strict convexity assumption.

The condition on the second derivative with respect to p cannot be weakened. If we imposed only $L_{pp} \geq 0$, then picking any function $w \in \mathrm{AC}(a,b)$, we could define $L\colon [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$L(x,y,p) = (y - w(x))^2 p^2,$$

which has $L_{pp}(x,y,p) = 2(y-w(x))^2 \geq 0$ for all $(x,y,p) \in [a,b] \times \mathbb{R} \times \mathbb{R}$. The associated functional

$$\mathscr{L}(u) = \int_a^b (u(x) - w(x))^2 (u'(x))^2 \, dx$$

is clearly minimized over $\mathscr{A} = \mathscr{A}_{w(a),w(b)}$ by $w \in \mathrm{AC}(a,b)$. However, $w \in \mathrm{AC}(a,b)$ was arbitrary and therefore has no higher regularity. Imposing just convexity will only furnish higher regularity if combined with a further growth condition, e.g. superlinearity.

So the only condition to examine in the pursuit of optimal conditions is the smoothness assumption. We collect here all the recent work in this direction.

Theorem 2.2 (Clarke and Vinter [1985a])**.** Suppose $L\colon [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is such that

(CVH1) L is locally bounded, measurable as a function of x, and convex as a function of p;

(CVH2) L is locally Lipschitz in (y,p) uniformly in x, i.e. for each bounded $C \subseteq \mathbb{R}^2$, there exists K such that

$$|L(x,y_1,p_1) - L(x,y_2,p_2)| \leq K|(y_1 - y_2, p_1 - p_2)|$$

for all $(y_1,p_1), (y_2,p_2) \in C$, and all $x \in [a,b]$; and

(CVH3) L is superlinear in p.

Then a minimizer $u \in \mathrm{AC}(a,b)$ of (1.2) exists. Furthermore, let $x \in [a,b]$ be such that

$$\liminf_{\substack{s,t \to x \\ a \leq s \leq x \leq t \leq b \\ s \neq t}} \frac{|u(s) - u(t)|}{|s-t|} < \infty.$$

Then

(CVi) On some interval I containing x as an interior point the function u is Lipschitz and satisfies the Euler-Lagrange equation in the sense of nonsmooth analysis, i.e. differential inclusions.

(CVii) If, in addition, for all $t \in [a, b]$ and all $q \in \mathbb{R}$, the function $p \mapsto L(t, u(t), p)$ is strictly convex and the function $s \mapsto L(s, u(t), q)$ is continuous at t, then u is C^1 on I.

(CViii) If in addition to this hypothesis (CVii), for each $t \in [a, b]$ L is of class C^k near $(t, u(t), u'(t))$ and $L_{pp}(t, u(t), u'(t)) > 0$, then u is of class C^k in I, for $k \geq 2$.

Corollary 2.3 (Clarke and Vinter [1985a])**.** Let $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfy the three assumptions (CVH1)–(CVH3). Let $u \in \mathrm{AC}(a, b)$ be a minimizer of (1.2).

Then there is a closed null set $E \subseteq [a, b]$ such that u' is locally bounded off E.

Remark 2.4. Notice for statement (CVi), which asserts that u is locally Lipschitz on an open subset of full measure, that no strict convexity assumption is made. Hence the presence of the superlinearity assumption.

Corollary 2.5 (Clarke and Vinter [1985a])**.** Let $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfy the three assumptions (CVH1)–(CVH3)and the further conditions of (CVii) above.

Then for any $x \in [a, b]$ we have

$$\liminf_{\substack{s,t \to x \\ a \leq s \leq x \leq t \leq b \\ s \neq t}} \frac{|u(s) - u(t)|}{|s - t|} = \limsup_{\substack{s,t \to x \\ a \leq s \leq x \leq t \leq b \\ s \neq t}} \frac{|u(s) - u(t)|}{|s - t|}.$$

Corollary 2.6 (Clarke and Vinter [1985a])**.** Let $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfy the three assumptions (CVH1)–(CVH3) and the further conditions of (CVii) above.

Then for any $x \in [a, b]$ the limit

$$\lim_{\substack{t \to x \\ a \leq t \leq b}} \frac{u(t) - u(x)}{t - x}$$

exists as a finite or infinite value.

Clarke and Vinter also examine a range of conditions to move to full regularity, see Chapter 3 for a discussion of this question. Their setting is in fact the vectorial case, dealing with functions $u\colon [a, b] \to \mathbb{R}^N$. This example of the Tonelli regularity result is a corollary of their vectorial regularity results. A discussion of the nonsmooth analysis required to interpret the Euler-Lagrange equation in this situation would take us on a rather long and otherwise irrelevant diversion, so we instead refer to Clarke [1990].

Sychëv [1991, 1992, 1993] proves versions of the result under the usual strict convexity assumption and the condition that L is (locally) Hölder continuous.

Theorem 2.7 (Sychëv [1993])**.** Suppose $L\colon \mathbb{R}^3 \to \mathbb{R}$ is such that

- L is locally Hölder continuous; and
- $L_{pp} \geq \mu > 0$ and is continuous.

Then any minimizer $u \in \mathrm{AC}(a, b)$ of (1.2) has a (possibly infinite) derivative everywhere, and u' is continuous as a map into the extended real line.

Csörnyei et al. [2008] prove Tonelli's partial regularity result under different weak smoothness assumptions on L, the important condition being a locally uniform Lipschitz condition in y.

Theorem 2.8 (Csörnyei et al. [2008])**.** Suppose $L\colon \mathbb{R}^3 \to \mathbb{R}$ is such that

- L is bounded below and locally bounded above;
- L is superlinear; and
- L is locally Lipschitz in y locally uniformly in (x, p), i.e. for every $R > 0$ there is $C > 0$ such that

$$|L(x, y_1, p) - L(x, y_2, p)| \leq C|y_1 - y_2|, \tag{L}$$

  when $(x, y_i, p) \in \mathbb{R}^3$ are such that $|x|, |y_i|, |p| \leq R$ for $i = 1, 2$.

Then for any (generalized) minimizer $u \in \mathrm{AC}(a, b)$ of (1.2), there exist disjoint closed null sets $E_\pm \subseteq [a, b]$ such that

- u is locally Lipschitz on $[a, b]\backslash E_+ \cup E_-$;
- $\lim_{t\neq x,\ \max\{d(t,E_+),d(x,E_+),|t-x|\}\to 0} \frac{u(x)-u(t)}{x-t} = \infty$; and
- $\lim_{t\neq x,\ \max\{d(t,E_-),d(x,E_-),|t-x|\}\to 0} \frac{u(x)-u(t)}{x-t} = -\infty$.

They work without any kind of convexity assumption, hence the superlinearity assumption for the regularity results. Lack of convexity means a lack of classical existence theory, hence the notion of *generalized* minimizer in this statement. We explain this notion in Chapter 4, but for the moment the word "generalized" can be ignored in this result, and the minimizer taken to be a minimizer in the classical sense.

The most recent result is in Ferriero [2010], and is framed in rather a general setting, so we first require some definitions. We state his result as given (albeit only for scalar-valued maps u), and also a simplified particular version most interesting for us. Throughout $\Sigma_L \subseteq [a,b]$ is a closed set of measure zero on which the Lagrangian L is not defined.

Definition 2.9 (Affine minorized)**.** Lagrangian $L\colon [a,b]\backslash\Sigma_L \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is *affine minorized* in p, locally uniformly in $(x,y) \in ([a,b]\backslash\Sigma_L) \times \mathbb{R})$, if for every compact $K \subseteq ([a,b]\backslash\Sigma_L) \times \mathbb{R})$ there exist $q \in \mathbb{R}$ and $\beta \geq 0$ such that $L(x,y,p) \geq pq - \beta$ for every $(x,y,p) \in K \times \mathbb{R}$.

Definition 2.10 (Bounded intersection property)**.** We say that a Lagrangian $L\colon [a,b]\backslash\Sigma_L \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ has the *bounded intersection property* in p, locally uniformly in $(x,y) \in ([a,b]\backslash\Sigma_L) \times \mathbb{R}$, if for any compact $K \subseteq ([a,b]\backslash\Sigma_L) \times \mathbb{R}$, and for every $p \in \mathbb{R}$, there exists $q(x,y) \in \partial_p L^{**}(x,y,p)$, the subgradient of L^{**} with respect to the third variable at p, such that the set $\{A(x,y,p) : (x,u) \in K\}$ is bounded. Here

$$A(x,y,p) := \{r \in \mathbb{R} : L^{**}(x,y,r) = L^{**}(x,y,p) + q(x,y)(r-p)\},$$

and L^{**} is the usual convexification of L with respect to the third variable, i.e. the maximal function below L which is convex with respect to the third variable.

Definition 2.11. The oscillation of a function $u \in L^\infty(a,b)$ at a point $x_0 \in (a,b)$ is defined by

$$\mathrm{osc}_{x_0}(u) := \lim_{\epsilon\to 0} \sup \{c \in \mathbb{R} : \mathrm{meas}\left(\{(x,t) \in B_\epsilon(x_0) \times B_\epsilon(x_0) : |u(x) - u(t)| > c\}\right) > 0\}$$

We also need two conditions, Ferriero's conditions (H1) and (H2), which we label (FH1) and (FH2):

(FH1) for each $R > 0$ there exists integrable $C_R\colon [a,b] \to [0,\infty)$ such that

$$|L(x,y_1,p) - L(x,y_2,p)| \leq C_R(x)|y_1 - y_2|$$

for almost every $x \in [a,b]$ and all $|y_i|, |p| \leq R$, for $i = 1,2$;

(FH2) the Lagrangian L is invariant under a group of C^1 transformations $(\tau^s(x), \phi^t(y))\colon [a,b] \times \mathbb{R} \to [a,b] \times \mathbb{R}$, with $s,t \in [-1,1]$, such that $(\tau^0(x), \phi^0(y)) = (x,y)$ and $|\partial_s \tau^s(x)|_{s=0}| + |\partial_t \phi^t(y)|_{t=0}| \neq 0$, for every x, y.

That is, for arbitrary $s,t \in [-1,1]$, $x_0, x_1 \in [a,b]$ such that $x_0 < x_1$, and admissible functions $u \in \mathrm{AC}(a,b)$,

$$\int_{\tau^s(x_0)}^{\tau^s(x_1)} L\left(\tau, \phi^t\left(u((\tau^s)^{-1}(\tau))\right), \frac{d}{d\tau}\phi^t\left(u((\tau^s)^{-1}(\tau))\right)\right) d\tau \\ = \int_{x_0}^{x_1} L(x, u(x), u'(x))\, dx.$$

Theorem 2.12 (Ferriero [2010])**.** Suppose $L = L(x,y,p)\colon [a,b]\backslash\Sigma_L \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous with respect to x and y for almost every $p \in \mathbb{R}$, and Borel measurable in p for every $(x,y) \in ([a,b]\backslash\Sigma_L) \times \mathbb{R}$. Suppose further that L is affine minorized and has the bounded intersection property in p locally uniformly in $(x,y) \in ([a,b]\backslash\Sigma_L) \times \mathbb{R}$, and one of the conditions (FH1) or (FH2) holds. Let $u \in \mathrm{AC}(a,b)$ be a minimizer of (1.2).

Then the set

$$\left\{x_0 \in [a,b]\backslash\Sigma_L : \limsup_{\epsilon\to 0} \frac{1}{2\epsilon}\left|\int_{x_0-\epsilon}^{x_0+\epsilon} u'(x)\, dx\right| < \infty\right\}$$

is an open set of full measure on which u' is locally bounded, and for any point x_0 in this open set, there exists $p(x_0) \in \mathbb{R}$, where $p(x_0) = u'(x_0)$ if x_0 is a Lebesgue point of u', such that

$$\mathrm{osc}_{x_0}(u') \leq \mathrm{diam}(A(x_0, u(x_0), p(x_0))).$$

Moreover, there exist disjoint closed subsets E_+ and E_- of $[a,b]\backslash\Sigma_L$ such that for any $x \in E_\pm$, we have $\lim_{t\to x} u'(t) = \pm\infty$.

We have stated this result only for the case which concerns us of real-valued functions. Ferriero's statement is in fact for functions $u\colon [a,b] \to \mathbb{R}^N$ for all $N \geq 1$, in which case the Lagrangian is a function $L\colon [a,b]\backslash\Sigma_L \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$. As remarked at the end of Ferriero [2010], superlinear growth implies the bounded intersection property. It is clear that strict convexity will imply the affine minorization condition. Thus these assumptions give us the following version of Ferriero's result.

Theorem 2.13 (Ferriero [2010], a special case)**.** Suppose $L\colon [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous, superlinear and strictly convex in p, and for each $R > 0$ there exists integrable function $C_R\colon [a,b] \to [0,\infty)$ such that

$$|L(x,y_1,p) - L(x,y_2,p)| \leq C_R(x)|y_1 - y_2|$$

for almost every $x \in [a, b]$ and all $(y_i, p) \in \mathbb{R}^2$ such that $|y_i|, |p| \leq R$, for $i = 1, 2$.

Then for any minimizer $u \in \mathrm{AC}(a, b)$ of (1.2) there is a closed null set $E \subseteq [a, b]$ such that u is locally Lipschitz off E.

It is this "integrable-Lipschitz constant" condition of Ferriero that the examples we construct below most clearly violate, see Remark 2.18 below.

2.1.2 The limit of partial regularity

It is tempting to think that just the assumption of continuity suffices to prove partial regularity. Certainly there is a clear argument in this situation showing that cusp points cannot occur in a minimizer: if both one-sided derivatives exist at a point, then they must be equal. We show how for example the modulus function $|\cdot|\colon [-1, 1] \to \mathbb{R}$ can never be a minimizer with respect to its own boundary conditions of a problem with a Lagrangian of form $L(x, y, p) = \phi(x, y - |x|) + p^2$ for continuous $\phi\colon [-1, 1] \times \mathbb{R} \to [0, \infty)$ with $\phi(x, 0) = 0$ for all $x \in [-1, 1]$. (These Lagrangians are precisely those which we consider later in the chapter to find counter-examples to partial regularity.)

Choose $\epsilon > 0$ such that $0 \leq \phi(x, y) \leq 1/2$ for $(x, y) \in [-\epsilon, \epsilon] \times [-\epsilon, \epsilon]$. Then consider the admissible function $w\colon [-1, 1] \to \mathbb{R}$ defined by

$$w(x) = \begin{cases} |x| & x \notin [-\epsilon, \epsilon] \\ \epsilon & x \in [-\epsilon, \epsilon]. \end{cases}$$

Then by choice of ϵ,

$$\begin{aligned} \mathscr{L}(|\cdot|) - \mathscr{L}(w) &= \int_{-1}^{1} (\phi(x, |x| - |x|) + 1)\, dx - \int_{-1}^{1} \left(\phi(x, w(x) - |x|) + (w'(x))^2\right)\, dx \\ &= \int_{-\epsilon}^{\epsilon} 1 - \phi(x, \epsilon - |x|)\, dx \\ &\geq 2\epsilon - 2\epsilon/2 \\ &> 0, \end{aligned}$$

hence $|\cdot|$ is not a minimizer.

That a function has cusp-points is not, however, the only way in which differentiability can fail.

In this chapter we show that some smoothness assumption stronger than mere continuity (even in all three variables) of L is necessary to obtain partial regularity. We construct continuous Lagrangians, superlinear in p and with $L_{pp} > 0$, such that the associated variational problems have minimizers each violating the partial regularity statement in a different way.

2.2 Construction of continuous Lagrangians with non-differentiable minimizers

2.2.1 Results

We prove two theorems. For any given $T > 0$, we construct Lagrangians $L\colon [-T,T]\times \mathbb{R} \times \mathbb{R} \to [0,\infty)$ and consider the problem of minimizing the functional

$$\mathrm{AC}(-T,T) \ni u \mapsto \mathscr{L}(u) = \int_{-T}^{T} L(x, u(x), u'(x))\, dx \tag{2.1}$$

over those u with prescribed boundary conditions $u(-T) = A$, $u(T) = B$.

Theorem 2.14. Let $T > 0$. Then there exists Lipschitz $w \in \mathrm{AC}(-T,T)$ and continuous $\phi\colon [-T,T]\times\mathbb{R} \to [0,\infty)$ such that defining $L(x,y,p) = \phi(x, y-w(x))+p^2$ gives a continuous Lagrangian $L\colon [-T,T] \times \mathbb{R} \times \mathbb{R} \to [0,\infty)$, superlinear in p and with $L_{pp} > 0$, such that

- w is a minimizer for the problem (2.1) with respect to its own boundary conditions, i.e. with $A = w(-T)$ and $B = w(T)$; but
- for dense G_δ (and hence second category) set $\mathcal{N} \subseteq [-T,T]$, we have $x \in \mathcal{N}$ implies
$$\overline{D}w(x) \geq 1 \text{ and } \underline{D}w(x) \leq -1.$$

Theorem 2.15. Let $T > 0$. Then there exists $w \in \mathrm{AC}(-T,T)$ and continuous $\phi\colon [-T,T] \times \mathbb{R} \to [0,\infty)$ such that defining $L(x,y,p) = \phi(x, y - w(x)) + p^2$ gives a continuous Lagrangian $L\colon [-T,T] \times \mathbb{R} \times \mathbb{R} \to [0,\infty)$, superlinear in p and with $L_{pp} > 0$, such that

- w is a minimizer for the problem (2.1) with respect to its own boundary conditions, i.e. with $A = w(-T)$ and $B = w(T)$; but
- for dense G_δ (and hence second category) set $\mathcal{N} \subseteq [-T,T]$, we have $x \in \mathcal{N}$ implies
$$\overline{D}w(x) = +\infty \text{ and } \underline{D}w(x) = -\infty.$$

These two theorems follow from our main result, which is the following:

Theorem 2.16. Let $T > 0$ and $g, h\colon [-T, T] \to \mathbb{R}$ be such that, writing $f(x) = x^{-1}g(x)$, the following properties hold (we comment later on the role of each assumption):

(2.i) f, h are even functions;

(2.ii) $g \in C([-T, T])$, and $f, h \in C^2([-T, T]\setminus\{0\})$;

(2.iii) f is non-increasing on $[0, T]$ and $f \geq 1$ near 0;

(2.iv) g is strictly increasing on $[0, T]$, and $g(0) = 0$;

(2.v) g is concave on $[0, T]$;

(2.vi) $h(x) \to \infty$ as $x \to 0$;

(2.vii) $g', gh' \in L^2(-T, T)$;

(2.viii) $g(x)(|g'(x)h'(x)| + |g''(x)| + |g(x)(h'(x))^2| + |g(x)h''(x)|) \to 0$ as $0 < |x| \to 0$;

(2.ix) $x \mapsto |xf'(x)| + |xf(x)h'(x)|$ is increasing on $[0, T]$;

(2.x) $x \mapsto |f'(x)| + |f(x)h'(x)|$ is decreasing on $[0, T]$;

(2.xi) there exists some non-negative function $\kappa \in C(0, T)$ with $\kappa(x) \to 0$ as $x \to 0$, such that

$$41g(c)(|cf'(c)| + |cf(c)h'(c)|) + 8\Psi(c) \leq 5\int_0^{g^{-1}(g(c)/5))} \kappa(x)\, dx,$$

where we have defined $\Psi\colon [0, T] \to [0, \infty)$ by

$$\Psi(c) = \int_0^c (\min\{c(|f'|+|fh'|), 2|f|+|xf'|+|xfh'|\})(2|f|+|xf'|+|xfh'|)\, dx.$$

Then there exists a subinterval $[-T_0, T_0]$ of $[-T, T]$, a function $w \in \mathrm{AC}(-T_0, T_0)$ and a continuous function $\phi\colon [-T_0, T_0] \times \mathbb{R} \to [0, \infty)$ such that defining $L(x, y, p) = \phi(x, y - w(x)) + p^2$ gives a continuous Lagrangian $L\colon [-T_0, T_0] \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$, superlinear in p and with $L_{pp} > 0$, such that

- w minimizes the associated variational problem

$$\mathrm{AC}(-T_0, T_0) \ni u \mapsto \mathscr{L}(u) = \int_{-T_0}^{T_0} L(x, u(x), u'(x))\, dx,$$

over those $u \in \mathrm{AC}(-T_0, T_0)$ with $u(\pm T_0) = w(\pm T_0)$; but

- for a dense G_δ (and hence second category) set $\mathcal{N} \subseteq [-T_0, T_0]$, we have $x \in \mathcal{N}$ implies

$$0 \leq g'(0) \leq \overline{D}w(x) \leq 2g'(0),$$

and

$$-2g'(0) \leq \underline{D}w(x) \leq -g'(0) \leq 0.$$

In particular w is non-differentiable on $\mathcal{N}$.

Remark 2.17. Despite the fact that our minimizers are in the sense of differentiability quite badly behaved, we can see quite easily from our construction—since it is essentially a limiting process—that the Lavrentiev phenomenon does not occur in these examples. We prove this fact in Section 2.2.6 below.

Remark 2.18. Even without Sychëv's results it is immediate that the Lagrangians we construct are not locally Hölder: the main ingredient, the function $\tilde{\phi}(x, y)$ defined after the proof of Lemma 2.21, satisfies $|\tilde{\phi}(x, |x|) - \tilde{\phi}(x, 0)| \geq |g(x)||\tilde{w}''(x)|$, which in the explicit examples given tends to zero with speed controlled only by logarithms of $|x|$. A more interesting remark is that the same estimate shows that the (local) Lipschitz constant, say $C(x)$, of the function $\tilde{\phi}(x, \cdot)$ is not integrable (since $|\tilde{w}'(x)|$ cannot be continuous at zero). This is in fact necessary: the positive statement from Ferriero [2010] quoted above precisely shows that integrability of $C(x)$ already implies Tonelli-type partial regularity of the minimizers.

Remark 2.19. It is immediate that the set $\mathcal{N}$ of non-differentiability points cannot be σ-porous, since it is a second category set. We have not made any further study of the set; in particular the question of its possible Hausdorff dimension remains unknown.

In Section 2.2.2, we give the (rather intricate) details of this general construction. In Section 2.2.3 we prove that w is indeed a minimizer, and in Section 2.2.4 we prove that w has the claimed (non-)differentiability properties. We draw together our arguments in Section 2.2.5. We conclude the general argument in Section 2.2.6, where we show that the Lavrentiev phenomenon does not occur in this problem. Finally in Section 2.2.7 we give the explicit proofs and calculations which allow us to infer Theorems 2.14 and 2.15 from Theorem 2.16.

2.2.2 The construction

Suppose T, g, and h are as in Theorem 2.16. We shall occasionally have to distinguish the two cases of whether or not g is Lipschitz, i.e. whether or not $\|g'\|_\infty$ is a finite number. In the non-Lipschitz case, these discussions can be ignored, as everything is satisfied trivially when this value is infinite.

We choose $T_0 < \min\{1/2, T/2\}$ small enough so that for $x \in [-2T_0, 2T_0]$ the following conditions hold:

(2.a) $g(x) \geq x$;

(2.b) $|g(x)| \leq 1$;

(2.c) $|g(x)h'(x)| < \|g'\|_\infty/2$; and

(2.d) $|h'(x)| \geq 1$.

Given any sequence of points in $(-T_0, T_0)$, we can construct a Lagrangian L and minimizer w with the set of non-differentiability points of w containing this sequence. The construction is essentially inductive, and hinges on the fact that a certain function $\tilde{w}$ is non-differentiable at one point, but minimizes a problem with continuous Lagrangian. This basic Lagrangian is of form $(x, y, p) \mapsto \tilde{\phi}(x, y - \tilde{w}(x)) + p^2$ for a "weight function" $\tilde{\phi}\colon [-T_0, T_0] \times \mathbb{R} \to [0, \infty)$ which penalizes functions which stray from $\tilde{w}$, i.e. $\tilde{\phi}(\cdot, 0) = 0$ and $|y| \mapsto \tilde{\phi}(x, y)$ is increasing. This immediately gives us a one-point example of non-differentiability of a minimizer, which already suffices to provide a counter-example to any Tonelli-like partial regularity result. Other points of non-differentiability are included by inserting translated and scaled copies of $\tilde{w}$ into the original $\tilde{w}$, and passing to the limit, w, say. The final Lagrangian is of form $(x, y, p) \mapsto \phi(x, y - w(x)) + p^2$, where ϕ is a sum of suitably modified translated and truncated copies $\tilde{\phi}_n$ of $\tilde{\phi}$, each of which penalizes functions which stray from w in a neighbourhood of x_n. We observe that many of the technicalities of the following proof are related to guaranteeing the existence and appropriate properties of w and L, and are in some sense secondary to the main points of the proof.

Some remarks on the conditions (2.i)–(2.xi) seem appropriate. We remind the reader that we shall conclude this chapter by exhibiting examples of functions which do satisfy these conditions, providing us with the particular cases Theorems 2.14 and 2.15 of the main theorem, so this long list of conditions is not so long as to be unreasonable. We suggest reading the general proof with the specific functions given in Example 2.34 in mind, to help visualize the steps in what is presented here as an unavoidably rather abstract and technical construction.

Condition (2.i) just implies that our basic minimizer $\tilde{w}$ is an odd function; the symmetry here simplifies the technicalities. Condition (2.ii) ensures that $\tilde{w}$ is smooth away from the point of singularity, allowing us to use integration by parts in the proof of minimality. That g is concave on $[0, T]$ guarantees that the convex hull of our minimizer $\tilde{w}$ is given by the graph of $\pm|g|$. Conditions (2.iii), (2.iv), (2.ix), and (2.x) indicate the delicate shape required of this convex hull; the latter two are useful when we estimate the errors made by comparing competing functions with not $\tilde{w}$, but the function obtained by replacing the graph of $\tilde{w}$ with a line segment on a certain well-chosen interval. Condition (2.vi) is required so that we can make $\tilde{w}$ oscillate arbitrarily close to 0 by means of the function $x \mapsto \sin h(x)$. Condition (2.vii) ensures that the energy $\mathscr{L}(\tilde{w})$ of our minimizer $\tilde{w}$ is finite. The two conditions (2.viii) and (2.xi) are the crucial properties we need to ensure that $\tilde{w}$ is a minimizer of a problem with continuous Lagrangian.

Define $\tilde{w}\colon [-T, T] \to \mathbb{R}$ by

$$\tilde{w}(x) = \begin{cases} g(x) \sin h(x) & x \neq 0 \\ 0 & x = 0, \end{cases}$$

so by (2.ii),

$$\tilde{w} \in C^2([-T, T] \backslash \{0\}). \tag{2.2}$$

Note for $x \in [-T, T] \backslash \{0\}$,

$$\begin{aligned} \tilde{w}'(x) &= g'(x) \sin h(x) + g(x)h'(x) \cos h(x) \\ &= (xf'(x) + f(x)) \sin h(x) + xf(x)h'(x) \cos h(x), \end{aligned}$$

and we observe by (2.i) that this is an even function, since the derivative of an odd function is even, that of an even function is odd, and a product of odd functions is even. Also note that for almost every $x \in [-T, T]$,

$$|\tilde{w}'(x)| \leq |g'(x)| + |g(x)h'(x)|, \tag{2.3}$$

and therefore by (2.vii) that

$$\tilde{w}' \in L^2(-T, T). \tag{2.4}$$

Also note that for $x \in [-T, T] \backslash \{0\}$,

$$\begin{aligned} w''(x) = g'(x)h'(x) \cos h(x) + g''(x) \sin h(x) \\ + g(x) \left(-(h'(x))^2 \sin h(x) + h''(x) \cos h(x)\right) + g'(x)h'(x) \cos h(x) \end{aligned}$$

and therefore that

$$\begin{aligned}|\tilde{w}''(x)| &\le |g'(x)h'(x)| + |g''(x)| + |g(x)(h'(x))^2| + |g(x)h''(x)| + |h'(x)g'(x)| \\ &= 2|g'(x)h'(x)| + |g''(x)| + |g(x)(h'(x))^2| + |g(x)h''(x)|.\end{aligned}$$

Hence see by (2.viii) that

$$|g(x)\tilde{w}''(x)| \to 0 \text{ as } 0 < |x| \to 0. \tag{2.5}$$

In particular, since this function is therefore bounded on a neighbourhood of 0, by (2.2) and (2.ii) we have that

$$\|g(x)\tilde{w}''(x)\|_{L^\infty(-T,T)} < \infty. \tag{2.6}$$

The following functions shall give us for each $x \in [-T_0, T_0]$ the exact coefficients we shall eventually need in our weight function $\tilde{\phi}$. Using the function $\kappa \in C(0,T)$ from condition (2.xi), we define $\psi^1, \psi^2 \colon [-T,T] \to [0,\infty)$ by

$$\psi^1(x) = \begin{cases} \frac{\kappa(|x|)}{|g(x)|} & x \neq 0 \\ 0 & x = 0 \end{cases} \quad \text{and} \quad \psi^2(x) = \begin{cases} 3 + 4|\tilde{w}''(x)| & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Note ψ^1 is well-defined by (2.iv). Now we define $\psi \colon [-T,T] \to [0,\infty)$ by $\psi(x) = \psi^1(x) + \psi^2(x)$. Using (2.2), (2.ii), the conditions on κ in (2.xi), and (2.5), we have

(ψ:1) $\psi \in C([-T,T]\backslash\{0\})$;

(ψ:2) $x \mapsto g(x)\psi(x)$ defines a continuous function on $[-T,T]$, with value 0 at 0.

By (ψ:2) we can choose $C \in (0,\infty)$ such that

$$C \ge 1 + 5|g(x)|\psi(x), \text{ for all } x \in [-T,T]. \tag{2.7}$$

We also, by (2.4), define $D \in (1,\infty)$ by $D = \|\tilde{w}'\|_{L^2(-T,T)} + 1$.

Let $\{x_n\}_{n=0}^\infty$ be a sequence of distinct points in $(-T_0, T_0)$. (Singularity at the endpoints requires only minor modifications (largely notational) to the construction given here, which we do not make explicit.) We assume $x_0 = 0$. We choose a decreasing sequence of constants $\{\sigma_n\}_{n=1}^\infty$ such that for $n \ge 1$,

$$0 < \sigma_n \le \min_{0 \le i < n} |x_i - x_n|/2,$$

and define sequences of translated functions $\{\tilde{w}_n\}_{n=0}^{\infty}$, $\{g_n\}_{n=0}^{\infty}$, and $\{\psi_n\}_{n=0}^{\infty}$, where $\tilde{w}_n, g_n \colon [-T_0, T_0] \to \mathbb{R}$ and $\psi_n \colon [-T_0, T_0] \to [0, \infty)$, by $\tilde{w}_n(x) = \tilde{w}(x - x_n)$, $g_n(x) = g(x - x_n)$, and $\psi_n(x) = \psi(x - x_n)$.

We want to construct a sequence $w_n \in W^{1,2}(-T_0, T_0)$, where (up to a constant multiple and the addition of a scalar) $w_n = \tilde{w}_i$ on a neighbourhood of x_i, thus w_n is singular at x_i, for each $0 \le i \le n$. We first construct a decreasing sequence $\{T_n\}_{n=0}^{\infty}$ of numbers $T_n \in (0, 1)$ and thus intervals $Y_n := [x_n - T_n, x_n + T_n]$ as follows. In the inductive construction of w_n we shall modify w_{n-1} only on Y_n.

Define a sequence $\{K_n\}_{n=0}^{\infty}$ of constants $K_n \ge 1$ by setting $K_0 = 1$, and so that for $n \ge 1$, we have

$$K_n \ge 1 + K_{n-1}, \text{ and} \tag{2.8}$$

$$2\sum_{i=0}^{n-1}(|\tilde{w}_i''(x)| + |\tilde{w}_i'(x)| + 1) \le K_n, \text{ whenever } |x_i - x| \ge \sigma_n \text{ for all } 0 \le i \le n-1. \tag{2.9}$$

This is possible for $K_n < \infty$ by (2.2).

Also for $n \ge 0$ define sequence $\{\theta_n\}_{n=0}^{\infty}$ of numbers $\theta_n \ge 2$ by setting $\theta_0 = 2$ and for $n \ge 1$ defining

$$\theta_n = \frac{52 g(T_0)}{\sigma_n} + 12 K_n. \tag{2.10}$$

Write $\tilde{g}_n = \theta_n g_n$. This scaling constant θ_n is an unimportant technicality, and just guarantees later that $|g_m(x) - g_m(x_n)| \le \theta_n |g_n(x)| (= \theta_n |g_n(x) - g_n(x_n)|)$ for all $x \in [-T_0, T_0]$, for all $0 \le m \le n$.

For $n \ge 1$ we inductively define $T_n \in (0, 1)$ small enough such that $Y_n = [x_n - T_n, x_n + T_n] \subseteq Y_0 = [-T_0, T_0]$, and the following conditions hold:

(T:1) $T_n < \sigma_n$;

(T:2) $T_n < T_{n-1}/2$; and

(T:3) $|\tilde{g}_n(x)\psi_n(x)| < 2^{-n}/5$ for $x \in Y_n$.

Note that (T:3) is possible by (ψ:2). Since we modify w_{n-1} only on Y_n to construct w_n, we need to add more weight to our Lagrangian only for $x \in Y_n$. Recalling that we are always working with translations of the same basic function $\tilde{\phi}$ (which we will define explicitly later), we know that we can choose the intervals Y_n small enough so that summing all the extra "weights" we need, we still converge to a continuous function. That the intervals of modification are small enough in this sense is the reason behind conditions (T:2) and (T:3). Since $T_0 < 1$, (T:2) guarantees in particular that

$$T_n < 2^{-n} \text{ for all } n \geq 0. \tag{2.11}$$

Condition (T:1) guarantees that the points in Y_n are far away from the previous x_i:

$$|x_i - x| > \sigma_n \text{ for } 0 \leq i < n, \text{ whenever } x \in Y_n. \tag{2.12}$$

Suppose otherwise, then for some $0 \leq i < n$ and some $x \in Y_n$ we have

$$|x_i - x_n| \leq |x_i - x| + |x - x_n| < \sigma_n + T_n < 2\sigma_n,$$

which contradicts the choice of σ_n. This stops the subintervals we later consider from overlapping.

We emphasize that this sequence is constructed independently of the later constructed w_n; the inductive construction of these functions will require us to pass further down the sequence $\{T_n\}_{n=0}^{\infty}$ than induction would otherwise allow, as we now see. For $n \geq 0$, find $m_n > n$ such that

$$2^{-m_n} < \frac{T_{n+1}^2}{64}. \tag{2.13}$$

Choose open cover $G_n \subseteq [-T_0, T_0]$ of the points $\{x_i\}_{i=0}^{m_n}$ such that, where $C > 0$ is as in (2.7),

$$\text{meas}(G_n) \leq \frac{T_{n+1}^2}{32C}. \tag{2.14}$$

Now, by (ψ:1) we can find $M_n \in (1, \infty)$ such that we have

$$\sum_{i=0}^{m_n} (\max\{\psi_i(x), \psi_i(x_i + T_i)\}) \leq M_n \text{ whenever } x \in [-T_0, T_0] \backslash G_n. \tag{2.15}$$

We note also that by (2.iv), for each $n \geq 0$ there exists $\eta_n \in (0, 1)$ such that for all $0 \leq i < n$,

$$|g_i(x)| \geq \eta_n \text{ whenever } |x_i - x| \geq \sigma_n. \tag{2.16}$$

Let $R_0 = T_0$ and for $n \geq 1$ inductively construct a decreasing sequence $\{R_n\}_{n=0}^{\infty}$ of numbers $R_n \in (0, T_n)$ such that:

(R:1) $\int_{-R_n}^{R_n} |\tilde{w}'|^2 < \frac{T_n^4}{8 \cdot 2048 D^2}$;

(R:2) $R_n < R_{n-1}/2$ and $g(R_n) < \frac{g(R_{n-1})}{2K_n}$;

(R:3) $g(R_n) < \frac{T_n^5 \eta_n}{(584 \cdot 2056) D^2 K_n^2 M_{n-1}}$;

(R:4) $g(R_n) < \frac{2^{-(n+1)}\|g'\|_\infty T_n}{17K_n}$; and

(R:5) $|gh'(x)| < 2^{-(n+3)}\|g'\|_\infty$ for $x \in [-R_n, R_n]$.

Note that (R:1) is possible by (2.4), (R:2)–(R:4) are possible by (2.iv), and (R:5) is possible since (2.viii) implies in particular that $(gh')^2(x) \to 0$ as $x \to 0$. Define subintervals $Z_n := [x_n - R_n, x_n + R_n]$ of Y_n. These intervals are those on which we aim to insert a copy of $\tilde{w}_n$ into w_{n-1}. The Z_n must be a very much smaller subinterval of Y_n to allow the estimates we require to hold; the point of this stage in the construction is that we now let the derivative of w_n oscillate on Z_n, so we have to make the measure of this set very small to have any control over the convergence.

Lemma 2.20. There exists a sequence of $w_n \in W^{1,2}(-T_0, T_0)$ satisfying, for $n \geq 0$:

(2.20.1) $w_n(x) = \lambda_n \tilde{w}_n(x) + \rho_n$ when $x \in [x_n - \tau_n, x_n + \tau_n]$, for some $\tau_n \in (0, R_n]$, some $\lambda_n \in [1, 2)$, and some $\rho_n \in \mathbb{R}$;

(2.20.2) w_n' exists and is locally Lipschitz on $[-T_0, T_0]\backslash\{x_i\}_{i=0}^n$;

(2.20.3) $|w_n(x) - w_n(x_i)| \leq (2 - 2^{-n})|\tilde{g}_i(x)|$ for all $x \in [-T_0, T_0]$ and for all $0 \leq i \leq n$;

(2.20.4) $\sup_{x\in[-T_0,T_0]\backslash\{x_i\}_{i=0}^n} |w_n'(x)| \leq (2 - 2^{-(n+1)})\|g'\|_\infty$;

(2.20.5) $|w_n'(x)| \leq K_{n+1}$ when $|x - x_{n+1}| \leq \sigma_{n+1}$, in particular on Y_{n+1};

(2.20.6) w_n'' exists almost everywhere, and satisfies $|w_n''(x)| \leq K_{n+1}$ for almost every $x \in [-T_0, T_0]$ such that $|x - x_{n+1}| \leq \sigma_{n+1}$, in particular on Y_{n+1};

and for $n \geq 1$:

(2.20.7) $w_n = w_{n-1}$ off Y_n;

(2.20.8) $\|w_n - w_{n-1}\|_\infty < 6K_n g(R_n)$;

(2.20.9) $w_n(x_i) = w_{n-1}(x_i)$ for all $0 \leq i \leq n$;

(2.20.10) $\|w_n' - w_{n-1}'\|_{L^2(-T_0,T_0)} < \frac{T_n^2}{32D}$;

(2.20.11) $|w_n'(x)| < |w_{n-1}'(x)| + 2^{-n}$ for almost every $x \notin [x_n - \tau_n, x_n + \tau_n]$; and

(2.20.12) $|w_n''(x)| < |w_{n-1}''(x)| + 2^{-n}$ for almost every $x \notin [x_n - \tau_n, x_n + \tau_n]$.

Proof. We easily see that defining $w_0 = \tilde{w}_0$ satisfies all the required conditions. That $w_0 \in W^{1,2}(-T_0, T_0)$ follows from (2.ii) and (2.4). Condition (2.20.1) is trivial for $\tau_0 = T_0$, $\lambda_0 = 1$, and $\rho_0 = 0$; (2.20.2) follows from (2.2); and (2.20.3) is evident from the definition of $\tilde{w}$, since $w_0(x_0) = \tilde{w}(0) = 0$ and $\theta_0 \geq 2$. Condition (2.20.4) is given by (2.3) and (2.c). Conditions (2.20.5) and (2.20.6) are given precisely by (2.9), since $|x - x_{n+1}| \leq \sigma_{n+1}$ implies that $|x - x_i| > \sigma_{n+1}$ for all $0 \leq i \leq n$, by choice of σ_{n+1}.

Suppose for $n \geq 1$ we have constructed $w_i \in W^{1,2}(-T_0, T_0)$ as claimed for all $0 \leq i < n$. We demonstrate how to insert a scaled copy of $\tilde{w}_n$ into w_{n-1}. We introduce in this proof a number of variables, e.g. m, which only appear in this inductive step. Although they do of course depend on n, we do not index them as such, since they are only used while n is fixed.

Condition (T:1) implies that $x_i \notin Y_n$ for all $0 \leq i < n$, thus w'_{n-1} exists and is Lipschitz on Y_n by inductive hypothesis (2.20.2). Let $m := w'_{n-1}(x_n)$, so $|m| < K_n$ by (2.20.5). On some yet smaller subinterval $[x_n - \tau_n, x_n + \tau_n]$ of Z_n we aim to replace w_{n-1} with a copy of $\tilde{w}_n$, connecting this with w_{n-1} off Y_n without increasing too much either the first or second derivatives, hence the choice of R_n as very much smaller than T_n. Moreover we want to preserve a continuous first derivative. Hence we displace w_{n-1} by a C^1 function—dealing with either side of x_n separately—so that on either side we approach x_n on an affine function of gradient m (a different function either side, in general), which we then connect up with $\tilde{w}_n$ at a point where $\tilde{w}'_n = m$. Because we need careful control over the first and second derivatives, it is easiest to construct explicitly the cut-off function we in effect use.

A slight first problem is in the case that g and therefore $\tilde{w}$ is Lipschitz, when it is possible that so small might be the interval $[x_n - R_n, x_n + R_n]$ on which we consider $\tilde{w}_n$, the derivative $\tilde{w}'_n$ might never be large enough to allow us to join with an affine function of gradient m: recalling (2.3), and using (2.viii),

$$|\tilde{w}'(x)| \leq \|g'\|_\infty + |g(x)h'(x)| \to \|g'\|_\infty \quad \text{as } 0 < |x| \to 0.$$

It is possible however that $|m| > \|g'\|_\infty$, hence the possible need to scale $\tilde{w}_n$ up slightly by some number $\lambda_n \in (1, 2)$ to ensure we can find points where the derivatives can agree.

Let $m_+ = \sup_{(x_n, x_n+R_n]} \tilde{w}'_n$, and $m_- = \inf_{(x_n, x_n+R_n]} \tilde{w}'_n$. The definition of $\tilde{w}$ and (2.vi) imply that there exist sequences $\{s_k\}_{k=1}^\infty$ and $\{t_k\}_{k=1}^\infty$ of elements $s_k, t_k \in (x_n, x_n+R_n]$, $s_k < t_k \leq s_{k-1}$, such that $s_k, t_k \to x_n$ as $k \to \infty$, and $\tilde{w}_n(s_k) = -g_n(s_k)$ and $\tilde{w}_n(t_k) = g_n(t_k)$. Then since $\tilde{w}_n$ and g_n are C^2 on $(x_n, x_n + R_n]$, the mean value theorem implies that there exist $\zeta_k \in (s_k, t_k)$ and $\xi_k \in (s_k, t_k)$ such that, using also (2.iv),

$$\tilde{w}_n'(\zeta_k) = \frac{\tilde{w}_n(t_k) - \tilde{w}_n(s_k)}{t_k - s_k} = \frac{g_n(t_k) + g_n(s_k)}{t_k - s_k} > \frac{g_n(t_k) - g_n(s_k)}{t_k - s_k} = g_n'(\xi_k).$$

But by concavity, $g_n'(\xi_k) \to g'(0) = \|g'\|_\infty$ as $k \to \infty$, so we have that $m_+ \geq \|g'\|_\infty$. Similarly $|m_-| = -m_- \geq \|g'\|_\infty$.

So if $|m| \leq \|g'\|_\infty$, then it is a trivial consequence of the continuity of $\tilde{w}_n'$ away from x_n and the intermediate value theorem that there exists $\tau_n \in (0, R_n]$ such that $\tilde{w}_n'(x_n - \tau_n) = m = \tilde{w}_n'(x_n + \tau_n)$. So no scaling is required, and we set $\lambda_n = 1$.

If $|m| > \|g'\|_\infty$, in general we might have to scale $\tilde{w}_n$ up slightly. Note by (2.3) and (R:5) that

$$m_+ = \sup_{x \in (0, R_n]} \tilde{w}'(x) \leq \|g'\|_\infty + \sup_{x \in (0, R_n]} |(gh')(x)| \leq (1 + 2^{-(n+3)})\|g'\|_\infty$$

and similarly $m_- \geq -(1 + 2^{-(n+3)})\|g'\|_\infty$.

So we have

$$\|g'\|_\infty \leq \min\{|m_+|, |m_-|\} \leq (1 + 2^{-(n+3)})\|g'\|_\infty. \tag{2.17}$$

Put $\lambda_n = m / \min\{|m_+|, |m_-|, |m|\}$, so using inductive hypothesis (2.20.4) and (2.17) we have

$$|\lambda_n| \leq \frac{(2 - 2^{-n})\|g'\|_\infty}{\|g'\|_\infty} < 2.$$

The values $m_\pm$ are attained, say $\tilde{w}_n'(x_+) = m_+$ and $\tilde{w}_n'(x_-) = m_-$ for points $x_+, x_- \in (x_n, x_n + R_n]$. Evidently the function $|\lambda_n \tilde{w}_n'|$ takes its maximum value over $(x_n, x_n + R_n]$ at x_+ or x_-, and so calculating, using inductive hypothesis (2.20.4), (2.17), and our above bounds on $|\tilde{w}_n'(x\pm)|$, we see

$$\begin{aligned}
|\lambda_n \tilde{w}_n'(x_+)| &< \frac{|m|(1 + 2^{-(n+3)})\|g'\|_\infty}{\min\{|m_+|, |m_-|, |m|\}} \\
&\leq (2 - 2^{-n})(1 + 2^{-(n+3)})\|g'\|_\infty \\
&= (2 + 2^{-n-2} - 2^{-n} - 2^{-2n-3})\|g'\|_\infty \\
&= (2 - 2^{-n}(1 - 2^{-2} + 2^{-n-3}))\|g'\|_\infty \\
&\leq (2 - 2^{-(n+1)})\|g'\|_\infty,
\end{aligned}$$

and similarly for $|\lambda_n \tilde{w}_n'(x_-)|$, we see that $|\lambda_n \tilde{w}_n'(x_-)| < (2 - 2^{-(n+1)})\|g'\|_\infty$ on

$(x_n, x_n + R_n]$, and since this is an even function we have

$$|\lambda_n \tilde{w}_n'(x)| < (2 - 2^{-(n+1)})\|g'\|_\infty \text{ for all } x \in Z_n \backslash \{x_n\}. \tag{2.18}$$

We now show we have indeed scaled $\tilde{w}_n$ to be large enough, despite ensuring this bound holds. If $m \geq 0$ we see that

$$\lambda_n \tilde{w}_n'(x_+) = \frac{m\tilde{w}_n'(x_+)}{\min\{|m_+|, |m_-|, |m|\}} \geq m,$$

and

$$\lambda_n \tilde{w}_n'(x_-) = \frac{m\tilde{w}_n'(x_-)}{\min\{|m_+|, |m_-|, |m|\}} \leq -m \leq m,$$

and if $m \leq 0$ we see that

$$\lambda_n \tilde{w}_n'(x_+) = \frac{m\tilde{w}_n'(x_+)}{\min\{|m_+|, |m_-|, |m|\}} \leq m,$$

and

$$\lambda_n \tilde{w}_n'(x_-) = \frac{m\tilde{w}_n'(x_-)}{\min\{|m_+|, |m_-|, |m|\}} \geq -m \geq m.$$

So in either case, since $\tilde{w}_n'$ is continuous on $(x_n, x_n + R_n]$, we can apply the intermediate value theorem and find $\tau_n \in (0, R_n]$ such that $\lambda_n \tilde{w}_n'(x_n + \tau_n) = m$. Thus also of course $\lambda_n \tilde{w}_n'(x_n - \tau_n) = m$.

We now construct the cut-off functions χ_l and χ_r we use on the left and right of x_n respectively. Additional constants and functions used in the construction are labelled similarly.

Let $\delta_l = m - w_{n-1}'(x_n - R_n)$. So we see by inductive hypothesis (2.20.6), since $Z_n \subseteq Y_n$ that

$$|\delta_l| = |w_{n-1}'(x_n) - w_{n-1}'(x_n - R_n)| \leq \|w_{n-1}''\|_{L^\infty(Z_n)} R_n \leq K_n R_n. \tag{2.19}$$

Define

$$c_l = w_{n-1}(x_n) + \lambda_n \tilde{w}_n(x_n - \tau_n) + m(\tau_n - R_n) - w_{n-1}(x_n - R_n).$$

The point is that the function $x \mapsto m(x - (x_n - R_n)) + w_{n-1}(x_n - R_n) + c_l$ is an affine function with gradient m which takes value $w_{n-1}(x_n - R_n) + c_l$ at $(x_n - R_n)$ and value $w_{n-1}(x_n) + \lambda_n \tilde{w}_n(x_n - \tau_n)$ at $(x_n - \tau_n)$.

Note that by definition of $\tilde{w}$ and inductive hypothesis (2.20.5), we have

$$\begin{aligned}|c_l| &\leq |\lambda_n \tilde{w}_n(x_n - \tau_n)| + |w_{n-1}(x_n) - w_{n-1}(x_n - R_n)| + |m||\tau_n - R_n| \\ &< 2g(\tau_n) + K_n R_n + K_n R_n \\ &< 4K_n g(R_n), \end{aligned} \tag{2.20}$$

using (2.iv), (2.a) and that $K_n \geq 1$ in the simplification to get the last line. Now put $d_l = \frac{4}{T_n}(c_l - \frac{\delta_l}{2}(T_n/2 - R_n)$. Define the piecewise affine function $q_l \colon [-T_0, T_0] \to \mathbb{R}$ by stipulating

$$q_l(x_n - T_n) = 0 = q_l(x_n - T_n/2),\ q_l(x_n - 3T_n/4) = d_l,$$

and

$$q_l(x) = \begin{cases} 0 & x \leq x_n - T_n \\ \delta_l & x \geq x_n - R_n \\ \text{affine} & \text{otherwise.} \end{cases}$$

So by definition of d_l,

$$\int_{-T_0}^{x_n - R_n} q_l(x)\,dx = \int_{x_n - T_n}^{x_n - R_n} q_l(x)\,dx = \frac{1}{2}\left(\frac{T_n d_l}{2} + (T_n/2 - R_n)\delta_l\right) = c_l. \tag{2.21}$$

Now, $\|q_l\|_\infty = \max\{|\delta_l|, |d_l|\}$. We see by (2.20), (2.19), (2.a), and since $T_n < 1$, that

$$\begin{aligned}|d_l| \leq \frac{4}{T_n}\left(|c_l| + \frac{|\delta_l|}{2}(T_n/2 - R_n)\right) &< \frac{4}{T_n}\left(4K_n g(R_n) + \frac{T_n K_n R_n}{4}\right) \\ &= \frac{16 K_n g(R_n)}{T_n} + K_n R_n \\ &< \frac{17 K_n g(R_n)}{T_n}. \end{aligned} \tag{2.22}$$

So, comparing with (2.19) and recalling again (2.a), we have

$$\|q_l\|_\infty \leq \frac{17 K_n g(R_n)}{T_n}. \tag{2.23}$$

Also, q_l' exists almost everywhere and satisfies $\|q_l'\|_{L^\infty(-T_0,T_0)} = \max\left\{\frac{4|d_l|}{T_n}, \frac{|\delta_l|}{T_n/2 - R_n}\right\}$. Note firstly by (2.22) and (R:3) that

$$\frac{4|d_l|}{T_n} < \frac{4}{T_n}\left(\frac{17 K_n g(R_n)}{T_n}\right) = \frac{68 K_n g(R_n)}{T_n^2} < 2^{-(n+1)},$$

and secondly that since (R:3) in particular implies $R_n < T_n/4$, using (2.19) and (R:3) we see that

$$\frac{|\delta_l|}{(T_n/2) - R_n} < \frac{4R_nK_n}{T_n} < 2^{-(n+1)}.$$

Hence

$$\|q_l'\|_{L^\infty(-T_0,T_0)} < 2^{-(n+1)}. \tag{2.24}$$

We can now define $\chi_l\colon [-T_0, T_0] \to \mathbb{R}$ by $\chi_l(x) = \int_{-T_0}^{x} q_l(t)\,dt$. This gives $\chi_l \in C^1(-T_0, T_0)$ such that $\chi_l' = q_l$ everywhere, $\chi_l'' = q_l'$ almost everywhere, and, by (2.21),

$$\chi_l(x_n - T_n) = 0,\ \chi_l(x_n - R_n) = c_l,\ \chi_l'(x_n - R_n) = q_l(x_n - R_n) = \delta_l.$$

We perform a very similar argument on the right of x_n, to construct piecewise affine function $q_r\colon [-T_0, T_0] \to \mathbb{R}$. Define

$$c_r = w_{n-1}(x_n) + \lambda_n\tilde{w}_n(x_n + \tau_n) + m(R_n - \tau_n) - w_{n-1}(x_n + R_n),$$

and $\delta_r = m - w_{n-1}'(x_n + R_n)$, and finally $d_r = \frac{4}{T_n}(c_r + \frac{\delta_r}{2}(T_n/2 - R_n))$. Then again stipulate

$$q_r(x_n + T_n/2) = 0 = q_r(x_n + T_n),\ q_r(x_n + 3T_n/4) = -d_r,$$

and elsewhere

$$q_r(x) = \begin{cases} \delta_r & x \le x_n + R_n \\ 0 & x \ge x_n + T_n \\ \text{affine} & \text{otherwise.} \end{cases}$$

So by definition of d_r, we have

$$\int_{x_n+R_n}^{x_n+T_n} q_r(x)\,dx = \frac{1}{2}\left(\delta_r(T_n/2 - R_n) - \frac{d_rT_n}{2}\right) = -c_r. \tag{2.25}$$

All the numbers c_r, δ_r, d_r satisfy the same bounds as their left-hand counterparts, and thus q_r satisfies the same bounds as q_l above, i.e.

$$\|q_r\|_\infty \le \frac{17K_ng(R_n)}{T_n} \tag{2.26}$$

and

$$\|q_r'\|_{L^\infty(-T_0,T_0)} < 2^{-(n+1)}. \tag{2.27}$$

Defining $\chi_r\colon [-T_0, T_0] \to \mathbb{R}$ by

$$\chi_r(x) = c_r - \delta_r((x_n + R_n) - (-T_0)) + \int_{-T_0}^{x} q_r(t)\, dt$$

gives $\chi_r \in C^1(-T_0, T_0)$ such that $\chi_r' = q_r$ everywhere, $\chi_r'' = q_r'$ almost everywhere, and, by (2.25),

$$\chi_r(x_n + R_n) = c_r,\ \chi_r(x_n + T_n) = 0,\ \chi_r'(x_n + R_n) = q_r(x_n + R_n) = \delta_r.$$

We can now define $w_n\colon [-T_0, T_0] \to \mathbb{R}$ by

$$w_n(x) = \begin{cases} w_{n-1}(x) + \chi_l(x) & x \leq x_n - R_n \\ m(x - (x_n - R_n)) + w_{n-1}(x_n - R_n) + c_l & x_n - R_n < x < x_n - \tau_n \\ \lambda_n \tilde{w}_n(x) + w_{n-1}(x_n) & x_n - \tau_n \leq x \leq x_n + \tau_n \\ m(x - (x_n + R_n)) + w_{n-1}(x_n + R_n) + c_r & x_n + \tau_n < x < x_n + R_n \\ w_{n-1}(x) + \chi_r(x) & x_n + R_n \leq x. \end{cases}$$

We see w_n is continuous by construction. Condition (2.20.1) is immediate, with λ_n and τ_n as defined, and $\rho_n = w_{n-1}(x_n)$. We note that since $\chi_l(x) = 0$ for $x < x_n - T_n$, $\chi_r(x) = 0$ for $x > x_n + T_n$, we have that $w_n = w_{n-1}$ off Y_n, as required for (2.20.7).

We see that w_n' exists off $\{x_i\}_{i=0}^n$ by inductive hypothesis (2.20.2), (2.2), and by construction, recalling the definitions of δ_l, δ_r, and τ_n. It is given by

$$w_n'(x) = \begin{cases} w_{n-1}'(x) + q_l(x) & x \leq x_n - R_n \\ m & x_n - R_n < x < x_n - \tau_n \\ \lambda_n \tilde{w}_n'(x) & x_n - \tau_n \leq x < x_n,\ x_n < x \leq x_n + \tau_n \\ m & x_n + \tau_n < x < x_n + R_n \\ w_{n-1}'(x) + q_r(x) & x_n + R_n \leq x. \end{cases}$$

This is locally Lipschitz on $[-T_0, T_0] \setminus \bigcup_{i=0}^n \{x_i\}$ by inductive hypothesis (2.20.2), (2.2), and since q_l and q_r are Lipschitz. Hence we have (2.20.2). Also we see that indeed $w_n \in W^{1,2}(-T_0, T_0)$, by inductive hypothesis and (2.4).

Now, estimates (2.23) and (2.26) imply, by (R:4), that

$$\|q_l\|_\infty, \|q_r\|_\infty < 2^{-(n+1)} \|g'\|_\infty.$$

So by inductive hypothesis (2.20.4), and (2.18), we have for $x \notin \{x_i\}_{i=0}^{n}$,

$$|w_n'(x)| \leq \begin{cases} |w_{n-1}'(x)| + |q_l(x)| < (2 - 2^{-(n+1)})\|g'\|_\infty & x \leq x_n - R_n \\ |m| < (2 - 2^{-n)})\|g'\|_\infty < (2 - 2^{-(n+1)})\|g'\|_\infty & x_n - R_n < x < x_n - \tau_n \\ |\lambda_n \tilde{w}_n'(x)| < (2 - 2^{-(n+1)})\|g'\|_\infty & x_n - \tau_n < x < x_n + \tau_n \\ |m| < (2 - 2^{-n)})\|g'\|_\infty < (2 - 2^{-(n+1)})\|g'\|_\infty & x_n + \tau_n < x < x_n + R_n \\ |w_{n-1}'(x)| + |q_r(x)| < (2 - 2^{-(n+1)})\|g'\|_\infty & x_n + R_n \leq x. \end{cases}$$

Thus we have (2.20.4).

We see by (2.23) and (R:3) that for $x \leq x_n - R_n$, $x \notin \{x_i\}_{i=0}^{n-1}$,

$$|w_n'(x) - w_{n-1}'(x)| = |q_l(x)| \leq \frac{17K_n g(R_n)}{T_n} < 2^{-n}; \tag{2.28}$$

and similarly for $x \geq x_n + R_n$, $x \notin \{x_i\}_{i=0}^{n-1}$, by (2.26) and (R:3) we have that

$$|w_n'(x) - w_{n-1}'(x)| = |q_r(x)| \leq \frac{17K_n g(R_n)}{T_n} < 2^{-n}. \tag{2.29}$$

For $x_n - R_n < x < x_n - \tau_n$ and $x_n + \tau_n < x < x_n + R_n$, we use inductive hypothesis (2.20.6) and (R:3) to see that

$$|w_n'(x) - w_{n-1}'(x)| = |m - w_{n-1}'(x)| \leq K_n R_n < 2^{-n}. \tag{2.30}$$

Hence (2.20.11) holds. We can now check (2.20.10). First note that, using the definition of w_n and (2.20.7) (which we have checked for n),

$$\begin{aligned}
&\int_{-T_0}^{T_0} |w_n'(x) - w_{n-1}'(x)|^2 \, dx \\
&= \int_{Y_n} |w_n'(x) - w_{n-1}'(x)|^2 \, dx \\
&= \int_{x_n - T_n}^{x_n - R_n} |q_l(x)|^2 \, dx + \int_{(x_n - R_n, x_n - \tau_n) \cup (x_n + \tau_n, x_n + R_n)} |w_{n-1}'(x_n) - w_{n-1}'(x)|^2 \, dx \\
&\quad + \int_{x_n - \tau_n}^{x_n + \tau_n} |\lambda_n \tilde{w}_n'(x) - w_{n-1}'(x)|^2 \, dx + \int_{x_n + R_n}^{x_n + T_n} |q_r(x)|^2 \, dx.
\end{aligned}$$

Now, by (2.23) and (2.b),

$$\begin{aligned}
\int_{x_n-T_n}^{x_n-R_n} |q_l(x)|^2\,dx &\le \int_{x_n-T_n}^{x_n-R_n} \frac{(17K_n g(R_n))^2}{T_n^2}\,dx \\
&\le \frac{289(g(R_n))^2K_n^2}{T_n} \\
&\le \frac{289g(R_n)K_n^2}{T_n}
\end{aligned}$$

and similarly using (2.26),

$$\begin{aligned}
\int_{x_n+R_n}^{x_n+T_n} |q_r(x)|^2\,dx &\le \int_{x_n+R_n}^{x_n+T_n} \frac{(17K_n g(R_n))^2}{T_n^2}\,dx \\
&\le \frac{289g(R_n)K_n^2}{T_n}.
\end{aligned}$$

Further, by inductive hypothesis (2.20.6)

$$\int_{x_n-R_n}^{x_n-\tau_n} |w'_{n-1}(x_n) - w'_{n-1}(x)|^2\,dx \le \int_{x_n-R_n}^{x_n-\tau_n} (K_nR_n)^2 \le R_n(K_nR_n)^2,$$

and similarly

$$\int_{x_n+\tau_n}^{x_n+R_n} |w'_{n-1}(x_n) - w'_{n-1}(x)|^2\,dx \le R_n(K_nR_n)^2.$$

Finally, by inductive hypothesis (2.20.5),

$$\begin{aligned}
\int_{x_n-\tau_n}^{x_n+\tau_n} |\lambda_n\tilde{w}'_n(x) - w'_{n-1}(x)|^2\,dx &\le \int_{x_n-\tau_n}^{x_n+\tau_n} 2(\lambda_n^2|\tilde{w}'_n(x)|^2 + K_n^2)\,dx \\
&\le 8\int_{x_n-\tau_n}^{x_n+\tau_n} |\tilde{w}'_n(x)|^2\,dx + 4K_n^2\tau_n \\
&\le 8\int_{-R_n}^{R_n} |\tilde{w}'(x)|^2\,dx + 4K_n^2R_n.
\end{aligned}$$

Combining these estimates, and using (2.a), (R:1), and (R:3), we see that

$$\begin{aligned}
&\int_{-T_0}^{T_0} |w'_n(x) - w'_{n-1}(x)|^2\,dx \\
&\le \frac{2\cdot 289g(R_n)K_n^2}{T_n} + 2R_n(K_nR_n)^2 + 8\int_{-R_n}^{R_n} |\tilde{w}'(x)|^2\,dx + 4K_n^2R_n \\
&\le \frac{578g(R_n)K_n^2}{T_n} + 6R_nK_n^2 + \frac{T_n^4}{2048D^2}
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{584g(R_n)K_n^2}{T_n} + \frac{T_n^4}{2048D^2} \\
&\leq \frac{T_n^4}{2048D^2} + \frac{T_n^4}{2048D^2} \\
&= \frac{T_n^4}{1024D^2}
\end{aligned}
$$

as required.

Now, w_n'' exists almost everywhere and where it does, is given by

$$
w_n''(x) = \begin{cases} w_{n-1}''(x) + q_l'(x) & x < x_n - R_n \\ 0 & x_n - R_n < x < x_n - \tau_n \\ \lambda_n \tilde{w}_n''(x) & x_n - \tau_n < x < x_n,\ x_n < x < x_n + \tau_n \\ 0 & x_n + \tau_n < x < x_n + R_n \\ w_{n-1}''(x) + q_r'(x) & x_n + R_n < x \end{cases}
$$

and thus by (2.24), for almost every $x < x_n - R_n$ we have

$$
|w_n''(x)| \leq |w_{n-1}''(x)| + |q_l'(x)| < |w_{n-1}''(x)| + 2^{-(n+1)} < |w_{n-1}''(x)| + 2^{-n},
$$

and by (2.27), for almost every $x > x_n + R_n$, we have

$$
|w_n''(x)| \leq |w_{n-1}''(x)| + |q_r'(x)| < |w_{n-1}''(x)| + 2^{-(n+1)} < |w_{n-1}''(x)| + 2^{-n}.
$$

Hence (2.20.12), since $w_n'' = 0$ on $Z_n \backslash [x_n - \tau_n, x_n + \tau_n]$. We now check (2.20.5) and (2.20.6). Suppose $|x - x_{n+1}| \leq \sigma_{n+1}$. Then by definition of σ_{n+1}, necessarily $|x - x_i| > \sigma_{n+1}$ for all $0 \leq i \leq n$, so the inequality in (2.9) holds, in particular

$$
2\sum_{i=0}^{n}(|\tilde{w}_i''(x)| + |\tilde{w}_i'(x)|) \leq K_{n+1}
$$

precisely by choice of K_{n+1}.

Let $0 \leq k \leq n$ be such that $x \in Y_k \backslash \bigcup_{i=k+1}^{n} Y_i$. Then by inductive hypothesis (2.20.7) for $k+1, \ldots, n$ (we have checked this for n), we have that $w_n = w_k$ on a neighbourhood of x, so $w_n'(x) = w_k'(x)$ and $w_n''(x) = w_k''(x)$ where both sides exist, i.e. almost everywhere.

If $x \notin [x_k - \tau_k, x_k + \tau_k]$, then by inductive hypotheses (2.20.11) (we have checked this for $k = n$) and (2.20.5) (since $x \in Y_k$), and by (2.8), we have almost everywhere,

$$|w'_n(x)| = |w'_k(x)| \leq |w'_{k-1}(x)| + 2^{-k} \leq K_k + 1 \leq K_{n+1}$$

as required. Similarly by inductive hypotheses (2.20.12) (we have checked this for $k = n$) and (2.20.6) (since $x \in Y_k$), and by (2.8), we have almost everywhere,

$$|w''_n(x)| = |w''_k(x)| \leq |w''_{k-1}(x)| + 2^{-k} \leq K_k + 1 \leq K_{n+1}$$

as required.

If $x \in (x_k - \tau_k, x_k + \tau_k)$, then by inductive hypothesis (2.20.1) (we have checked this holds for $k = n$), almost everywhere we have, using (2.9) again,

$$|w'_n(x)| = |w'_k(x)| = |\lambda_k \tilde{w}'_k(x)| \leq \sum_{i=0}^{k} |\lambda_i \tilde{w}'_i(x)| \leq 2 \sum_{i=0}^{n} |\tilde{w}'_i(x)| \leq K_{n+1}$$

and

$$|w''_n(x)| = |w''_k(x)| = |\lambda_k \tilde{w}''_k(x)| \leq \sum_{i=0}^{k} |\lambda_i \tilde{w}''_i(x)| \leq 2 \sum_{i=0}^{n} |\tilde{w}''_i(x)| \leq K_{n+1}$$

as required.

Now observe that on $[-T_0, x_n - R_n]$ we have by definition of q_l, and using, (2.22), (2.19), and (2.a), that

$$\begin{aligned}
|\chi_l| &\leq \int_{-T_0}^{x_n - R_n} |q_l| \\
&\leq \frac{1}{2} \left(\frac{T_n}{2} |d_l| + (T_n/2 - R_n)|\delta_l| \right) \\
&\leq \left(\frac{17 K_n g(R_n)}{4} + \frac{T_n K_n R_n}{4} \right) \\
&\leq \frac{18 K_n g(R_n)}{4} \\
&< 5 K_n g(R_n).
\end{aligned}$$

A similar estimate holds for χ_r on $[x_n + R_n, T_0]$, using (2.25):

$$\begin{aligned}
\chi_r(x) &= c_r - \delta_r((x_n + R_n) + T_0) + \int_{-T_0}^{x} q_r(t)\, dt \\
&= c_r + \int_{x_n + R_n}^{x} q_r(t)\, dt
\end{aligned}$$

$$\begin{aligned} &= -\int_{x_n+R_n}^{T_0} q_r(t)\,dt + \int_{x_n+R_n}^{x} q_r(t)\,dt \\ &= -\int_{x}^{T_0} q_r(t)\,dt \end{aligned}$$

and hence, since then $|\chi_r| \leq \int_{x_n+R_n}^{T_0} |q_r|$ on $[x_n + R_n, T_0]$, we can estimate as above. So, for $x_n - T_n \leq x \leq x_n - R_n$, we have

$$|w_n(x) - w_{n-1}(x)| = |\chi_l(x)| \leq 5K_n g(R_n)$$

and similarly for $x_n + R_n \leq x \leq x_n + T_n$ we have

$$|w_n(x) - w_{n-1}(x)| = |\chi_r(x)| \leq 5K_n g(R_n).$$

By inductive hypothesis (2.20.5), (2.20), and (2.a), we have for $x_n - R_n < x < x_n - \tau_n$ that

$$\begin{aligned} |w_n(x) - w_{n-1}(x)| &\leq |m(x - (x_n - R_n))| + |w_{n-1}(x_n - R_n) - w_{n-1}(x)| + |c_l| \\ &< K_n R_n + K_n R_n + 4K_n g(R_n) \\ &\leq 6K_n g(R_n) \end{aligned}$$

and similarly for $x_n + \tau_n < x < x_n + R_n$ we have

$$|w_n(x) - w_{n-1}(x)| \leq |m(x - (x_n + R_n))| + |w_{n-1}(x_n + R_n) - w_{n-1}(x)| + |c_r| < 6K_n g(R_n).$$

Finally for $x_n - \tau_n \leq x \leq x_n + \tau_n$, by definition of $\tilde{w}$, inductive hypothesis (2.20.5) again, (2.a), and (2.iv), we have

$$|w_n(x) - w_{n-1}(x)| \leq |\lambda_n \tilde{w}_n(x)| + |w_{n-1}(x_n) - w_{n-1}(x)| \leq 2g(\tau_n) + K_n \tau_n \leq 3K_n g(R_n).$$

Hence we have, using also (2.20.7) (which we have checked for n),

$$\|w_n - w_{n-1}\|_\infty = \sup_{x \in Y_n} |w_n(x) - w_{n-1}(x)| < 6K_n g(R_n)$$

as required for (2.20.8).

We check (2.20.9). Let $0 \leq i \leq n$. If $i < n$, then $x_i \notin Y_n$ by (T:1), so $w_n(x_i) = w_{n-1}(x_i)$ by (2.20.7). We see directly from the construction that $w_n(x_n) = w_{n-1}(x_n)$ since $\tilde{w}_n(x_n) = 0$, as required for the full result.

We can now check (2.20.3). First consider $0 \leq i \leq n - 1$. The result is immediate by inductive hypothesis if $x \notin Y_n$, by (2.20.9) and (2.20.7). So suppose $x \in Y_n$. Then by (2.12) and (2.16), $|g_i(x)| \geq \eta_n$. Therefore by (2.20.9), inductive hypothesis (2.20.3), (2.20.8), and (R:3), we have

$$\begin{aligned}
|w_n(x) - w_n(x_i)| &\leq |w_n(x) - w_{n-1}(x)| + |w_{n-1}(x) - w_{n-1}(x_i)| \\
&\leq \|w_n - w_{n-1}\|_\infty + (2 - 2^{-(n-1)})|\tilde{g}_i(x)| \\
&< 6K_n g(R_n) + (2 - 2^{-(n-1)})|\tilde{g}_i(x)| \\
&\leq 2^{-n}\eta_n + (2 - 2^{-(n-1)})|\tilde{g}_i(x)| \\
&\leq (2 - 2^{-n})|\tilde{g}_i(x)|.
\end{aligned}$$

It just remains to check (2.20.3) in the case $i = n$. We first show that for all $x \in [-T_0, T_0]$, we have chosen θ_n such that

$$|w_{n-1}(x) - w_{n-1}(x_n)| \leq |\theta_n g_n(x)|/2 = |\tilde{g}_n(x)|/2. \tag{2.31}$$

If $|x - x_n| \leq \sigma_n$, we have by inductive hypothesis (2.20.5), (2.10), and (2.iii) that

$$\begin{aligned}
|w_{n-1}(x) - w_{n-1}(x_n)| &\leq K_n|x - x_n| \\
&\leq f(T)\theta_n|x - x_n|/2 \\
&\leq f(|x - x_n|)\theta_n|x - x_n|/2 \\
&= |\tilde{g}_n(x)|/2.
\end{aligned}$$

If $|x - x_n| \geq \sigma_n$, then by inductive hypothesis (2.20.8), (R:2), (2.iv), (2.10), and (2.iii),

$$\begin{aligned}
|w_{n-1}(x) - w_{n-1}(x_n)| &\leq |w_{n-1}(x) - w_0(x)| + |w_0(x) - w_0(x_n)| + |w_0(x_n) - w_{n-1}(x_n)| \\
&\leq 2\|w_{n-1} - w_0\|_\infty + 2\|w_0\|_\infty \\
&\leq 2\left(\sum_{i=1}^{n-1}(\|w_i - w_{i-1}\|_\infty) + g(T_0)\right) \\
&\leq 2\left(\sum_{i=1}^{n-1}(6K_i g(R_i)) + g(T_0)\right) \\
&\leq 2\left(12g(R_0) + g(T_0)\right) \\
&\leq 26g(T_0) \\
&\leq \sigma_n f(T_0)\theta_n/2 \\
&\leq \|x - x_n|f(x - x_n)\theta_n/2 \\
&= |\tilde{g}_n(x)|/2
\end{aligned}$$

as claimed.

Now, suppose first $x \in [x_n - \tau_n, x_n + \tau_n]$. Then by (2.20.1) and the definition of $\tilde{w}$ we have, since $\theta_n \geq 2$,

$$|w_n(x) - w_n(x_n)| = |\lambda_n(\tilde{w}_n(x) - \tilde{w}_n(x_n))| \leq 2|g_n(x)| \leq (2 - 2^{-n})|\tilde{g}_n(x)|.$$

To deal with the case $x_n - R_n \leq x < x_n - \tau_n$, we note first that the condition is satisfied at the endpoints: that it holds at $x = x_n - \tau_n$ follows from above, and using (2.20.9) (which we have checked), inductive hypothesis (2.20.5), (2.20), (2.a), and (2.10), we see that

$$\begin{aligned} |w_n(x_n - R_n) - w_n(x_n)| &= |w_{n-1}(x_n - R_n) + c_l - w_{n-1}(x_n)| \\ &\leq K_n R_n + |c_l| \\ &\leq 5K_n g(R_n) \\ &\leq (2 - 2^{-n})|\tilde{g}_n(x_n - R_n)|. \end{aligned}$$

Since w_n is defined to be affine between these endpoints, and g_n is concave on $[-T_0, x_n]$ and $[x_n, T_0]$, the result holds for all $x \in [x_n - R_n, x_n - \tau_n]$. Similarly the result holds for $x \in [x_n + \tau_n, x_n + R_n]$. Now we have to consider $x \leq x_n - R_n$. In this case we then have by (2.iv) that $g_n(x) \geq g_n(R_n)$, and so we can argue as follows, using (2.20.9), (2.20.8) (both of which we have checked), (2.31), and (2.10):

$$\begin{aligned} |w_n(x) - w_n(x_n)| &\leq |w_n(x) - w_{n-1}(x)| + |w_{n-1}(x) - w_n(x_n)| \\ &\leq \|w_n - w_{n-1}\|_\infty + |w_{n-1}(x) - w_{n-1}(x_n)| \\ &\leq 6K_n g(R_n) + |\tilde{g}_n(x)|/2 \\ &\leq 6K_n |g_n(x)| + |\tilde{g}_n(x)|/2 \\ &\leq |\tilde{g}_n(x)| \\ &\leq (2 - 2^{-n})|\tilde{g}_n(x)|. \end{aligned}$$

We deal with $x \geq x_n + R_n$ similarly. Thus (2.20.3) holds for all $x \in [-T_0, T_0]$ as claimed.

□

We now show easily that this sequence converges to some absolutely continuous w. This w will be our minimizer.

Lemma 2.21. The sequence $\{w_n\}_{n=0}^\infty$ converges uniformly to some function $w \in W^{1,2}(-T_0, T_0)$ such that, for all $n \geq 0$, (2.21.1) $w(x_i) = w_n(x_i)$ for all $0 \leq i \leq n+1$;

(2.21.2) $\|w - w_n\|_\infty \leq 12K_{n+1}g(R_{n+1})$;

(2.21.3) $\|w' - w_n'\|_{L^2(-T_0,T_0)} \leq \frac{T_{n+1}^2}{16D}$; and

(2.21.4) $|w(x) - w(x_n)| \leq 2|\tilde{g}_n(x)|$ for all $x \in [-T_0, T_0]$.

Proof. Let $n \geq 0$. We use (2.20.8) and (R:2) to see that for $m > n$ we have

$$\begin{aligned}\|w_m - w_n\|_\infty &\leq \|w_m - w_{m-1}\|_\infty + \ldots + \|w_{n+1} - w_n\|_\infty \\ &< 6(K_m g(R_m) + \cdots + K_{n+1}g(R_{n+1})) \\ &\leq 6(2^{-(m-(n+1))} + \cdots + 1)K_{n+1}g(R_{n+1}) \\ &< 12K_{n+1}g(R_{n+1}).\end{aligned}$$

Hence, since (R:3) certainly implies that this tends to 0 as $n \to \infty$, the sequence $\{w_n\}_{n=0}^\infty$ is uniformly Cauchy, and so converges uniformly to some $w \in C(-T_0, T_0)$. Condition (2.21.2) follows immediately, (2.21.1) follows directly from (2.20.9), and (2.21.4) follows from (2.20.3).

Now, by (2.20.10) and (T:2)

$$\|w_m' - w_n'\|_{L^2(-T_0,T_0)} \leq \frac{T_m^2}{32D} + \ldots + \frac{T_{n+1}^2}{32D} \leq \frac{T_{n+1}^2}{16D} \tag{2.32}$$

and hence by (2.11) w_n' is Cauchy in $L^2(-T_0, T_0)$, thus converges in $L^2(-T_0, T_0)$. Since w_n' also converges in $L^1(-T_0, T_0)$, we can easily see that this limit is equal almost everywhere to w': for any $x \in [-T_0, T_0]$,

$$\int_{-T_0}^x \lim_{n\to\infty} w_n'(t)\,dt = \lim_{n\to\infty} \int_{-T_0}^x w_n'(t)\,dt = \lim_{n\to\infty} (w_n(x) - w_n(T_0)) = w(x) - w(T_0).$$

Hence $w' \in L^2(-T_0, T_0)$ and (2.21.3) holds, and indeed $w \in W^{1,2}(-T_0, T_0)$. □

Our basic weight function $\tilde{\phi}\colon [-T_0, T_0] \times \mathbb{R} \to [0, \infty)$ will be given by

$$\tilde{\phi}(x,y) = \begin{cases} 0 & x = 0 \\ 5\psi(x)|g(x)| & |y| \geq 5|g(x)| \\ \psi(x)|y| & |y| \leq 5|g(x)|. \end{cases}$$

We need some bound of form $|\tilde{\phi}(x,y)| \leq c|g(x)|\psi(x)$ to ensure continuity of $\tilde{\phi}$; it turns out (see Lemma 2.23) that sensitive tracking of $|y|$ only for $|y| \leq 5|g(x)|$ suffices in the proof of minimality. Our function $\tilde{w}$ was constructed precisely so that (2.5) and hence (ψ:2) hold, and hence that this $\tilde{\phi}$ is continuous.

We in fact will find it useful to split $\tilde{\phi}$ into the summands by which we defined ψ. More precisely, we define for each $n \geq 0$ our translated weight functions $\tilde{\phi}_n^1$, $\tilde{\phi}_n^2 \colon [-T_0, T_0] \times \mathbb{R} \to [0, \infty)$ as follows. For $n \geq 0$, and for $k = 1, 2$, we recall that we need extra weight only on Y_n, so we define for $(x, y) \in Y_n \times \mathbb{R}$

$$\tilde{\phi}_n^k(x, y) = \begin{cases} 0 & x = x_n \\ 5\psi_n^k(x)\tilde{g}_n(x) & |y| \geq 5\tilde{g}_n(x) \\ \psi_n^k(x)|y| & |y| \leq 5\tilde{g}_n(x) \end{cases}$$

and then extend to $[-T_0, T_0] \times \mathbb{R}$ by defining for $(x, y) \in ([-T_0, T_0] \backslash Y_n) \times \mathbb{R}$

$$\tilde{\phi}_n^k(x, y) = \begin{cases} 5\psi_n^k(x_n + T_n)\tilde{g}_n(x_n + T_n) & |y| \geq 5\tilde{g}_n(x_n + T_n) \\ \psi_n^k(x_n + T_n)|y| & |y| \leq 5\tilde{g}_n(x_n + T_n). \end{cases}$$

We thus define $\tilde{\phi}_n \colon [-T_0, T_0] \times \mathbb{R} \to [0, \infty)$ by $\tilde{\phi}_n(x, y) = \tilde{\phi}_n^1(x, y) + \tilde{\phi}_n^2(x, y)$. By ($\psi$:2) we see that $\tilde{\phi}_n \in C([-T_0, T_0] \times \mathbb{R})$.

We claim for fixed $x \in [-T_0, T_0]$, for all $n \geq 0$ and $k = 1, 2$, that

$$\tilde{\phi}_n^k(x, y) \leq \tilde{\phi}_n^k(x, z) \text{ whenever } |y| \leq |z|;$$
$$\mathrm{Lip}(\tilde{\phi}_n^k(x, \cdot)) \leq \max\{\psi_n^k(x), \psi_n^k(x_n + T_n)\}; \text{ and}$$
$$\tilde{\phi}_n^k(x, 0) = 0.$$

The last result is obvious, as are the other results for $x = x_n$. Suppose $x \in Y_n \backslash \{x_n\}$. First consider case $|y| \leq |z| \leq 5\tilde{g}_n(x)$. Then

$$\tilde{\phi}_n^k(x, z) - \tilde{\phi}_n^k(x, y) = |z|\psi_n^k(x) - |y|\psi_n^k(x) \geq 0;$$

and

$$\left|\tilde{\phi}_n^k(x, z) - \tilde{\phi}_n^k(x, y)\right| = \psi_n^k(x)(|z| - |y|) \leq \psi_n^k(x)|z - y|$$

as required, giving that $\mathrm{Lip}(\tilde{\phi}_n^k(x), \cdot) \leq \psi_n^k(x)$ for such values.

In the case when $5\tilde{g}_n(x) \leq |y| \leq |z|$, we have

$$\tilde{\phi}_n^k(x, y) = 5\tilde{g}_n(x)\psi_n^k(x) = \tilde{\phi}_n^k(x, z)$$

and so both results are immediate. In case $|y| \leq 5\tilde{g}_n(x) \leq |z|$ we have

$$\tilde{\phi}_n^k(x,z) - \tilde{\phi}_n^k(x,y) = 5\tilde{g}_n(x)\psi_n^k(x) - \psi_n^k(x)|y| \geq 0;$$

and so

$$\left|\tilde{\phi}_n^k(x,z) - \tilde{\phi}_n^k(x,y)\right| = \psi_n^k(x)(5\tilde{g}_n(x) - |y|) \leq \psi_n^k(x)(|z| - |y|) \leq \psi_n^k(x)|z-y|.$$

Thus in this case again $\mathrm{Lip}(\tilde{\phi}_n^k(x,\cdot)) \leq \psi_n^k(x)$. Both results follow similarly for $x \notin Y_n$: we obtain instead that $\mathrm{Lip}(\tilde{\phi}_n^k(x,\cdot)) \leq \psi_n^k(x_n + T_n)$, hence the claim. Hence of course for all $x \in [-T_0, T_0]$, $\tilde{\phi}_n(x,\cdot)$ is an increasing function with Lipschitz constant at most $\max\{\psi_n(x), \psi_n(x_n + T_n)\}$, and $\tilde{\phi}(x,0) = 0$.

Defining $\phi_n \colon [-T_0, T_0] \times \mathbb{R} \to [0,\infty)$ by $\phi_n(x,y) = \sum_{i=0}^n \tilde{\phi}_i(x,y)$ gives a sequence of functions $\phi_n \in C([-T_0,T_0] \times \mathbb{R})$ such that for each fixed $x \in [-T_0, T_0]$, for all $n \geq 0$,

$$\phi_n(x,y) \leq \phi_n(x,z) \text{ whenever } |y| \leq |z|; \tag{2.33}$$

$$\mathrm{Lip}(\phi_n(x,\cdot)) \leq \sum_{i=0}^{n} (\max\{\psi_i(x), \psi_i(x_i + T_i)\})\,; \text{ and} \tag{2.34}$$

$$\phi_n(x,0) = 0. \tag{2.35}$$

For $n \geq 1$, by (T:3), we see that for all $(x,y) \in [-T_0,T_0] \times \mathbb{R}$

$$0 \leq \tilde{\phi}_n(x,y) \leq \sup_{x \in Y_n} 5\psi_n(x)\tilde{g}_n(x) \leq 2^{-n}.$$

So defining $\phi(x,y) = \sum_{i=0}^\infty \tilde{\phi}_i(x,y)$ gives $\phi \in C([-T_0,T_0] \times \mathbb{R})$ with, by (2.7),

$$\|\phi\|_\infty \leq \|\tilde{\phi}_0\|_\infty + \sum_{i=1}^\infty \|\tilde{\phi}_i\|_\infty \leq \|\tilde{\phi}_0\|_\infty + \sum_{i=1}^\infty 2^{-i} = \|\tilde{\phi}_0\|_\infty + 1 \leq C, \tag{2.36}$$

and

$$\|\phi - \phi_n\|_\infty \leq \sum_{i=n+1}^\infty \|\tilde{\phi}_i\|_\infty \leq \sum_{i=n+1}^\infty 2^{-i} = 2^{-n}. \tag{2.37}$$

By passing to the limit in the above relations (2.33) and (2.35) we see that for fixed $x \in [-T_0, T_0]$,

$$\phi(x,y) \leq \phi(x,z) \text{ whenever } |y| \leq |z|; \text{ and} \tag{2.38}$$

$$\phi(x,0) = 0. \tag{2.39}$$

We shall write $\phi = \phi^1 + \phi^2$, where $\phi^k = \sum_{i=0}^{\infty} \tilde{\phi}_i^k$ for $k = 1, 2$.

We can now define continuous Lagrangian $L\colon [-T_0, T_0] \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$, $L\colon (x, y, p) \mapsto L(x, y, p)$, superlinear and strictly convex in p, by setting

$$L(x, y, p) = p^2 + \phi(x, y - w(x)).$$

Note in fact that L is differentiable with respect to p and $L_{pp}(x, y, p) = 2 > 0$ for all $(x, y, p) \in [-T_0, T_0] \times \mathbb{R} \times \mathbb{R}$, thus it does satisfy the stronger strict convexity assumption required by Tonelli in his statements of partial regularity. Associated with this is the usual variational problem given by defining functional $\mathscr{L}\colon \mathrm{AC}(-T_0, T_0) \to \mathbb{R}$ by

$$\mathscr{L}(u) = \int_{-T_0}^{T_0} L(x, u(x), u'(x))\, dx$$

and seeking to minimize $\mathscr{L}(u)$ over those functions $u \in \mathrm{AC}(-T_0, T_0)$ with boundary conditions $u(\pm T_0) = w(\pm T_0)$. We shall refer to this set-up as $(\star)$.

2.2.3 Minimality

We shall find the following approximations to our functional $\mathscr{L}$ useful: for $n \geq 0$ define $L_n\colon [-T_0, T_0] \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$ by

$$L_n(x, y, p) = p^2 + \phi(x, y - w_n(x)),$$

and define corresponding functional $\mathscr{L}_n\colon \mathrm{AC}(-T_0, T_0) \to [0, \infty)$ by

$$\mathscr{L}_n(u) = \int_{-T_0}^{T_0} L_n(x, u(x), u'(x))\, dx.$$

Working with these approximations is much easier, since there is only a finite number of singularities in w_n. So it is important to know what error we incur by moving to these approximations. This is shown in the next lemma.

Lemma 2.22. Let $u \in \mathrm{AC}(-T_0, T_0)$ and $n \geq 0$. Then

$$|(\mathscr{L}(u) - \mathscr{L}(w)) - (\mathscr{L}_n(u) - \mathscr{L}_n(w_n))| < \frac{T_{n+1}^2}{4}.$$

Proof. We first estimate $|\mathscr{L}(u) - \mathscr{L}_n(u)|$. Recall our definitions of $m_n > n$, $M_n \geq 0$, and $G_n \supseteq \bigcup_{i=0}^{m_n} \{x_i\}$ from page 29. Let $x \in [-T_0, T_0] \setminus G_n$. We see by (2.34) and (2.15)

$$\mathrm{Lip}(\phi_{m_n}(x, \cdot)) \leq \sum_{i=0}^{m_n} (\max\{\psi_i(x), \psi_i(x_i + T_i)\}) \leq M_n.$$

Then using (2.21.2) and (R:3)

$$
\begin{aligned}
|\phi_{m_n}(x, u-w) - \phi_{m_n}(x, u-w_n)| &\leq \mathrm{Lip}(\phi_{m_n}(x,\cdot))|(u(x)-w(x)) - (u(x)-w_n(x))| \\
&\leq M_n\|w-w_n\|_\infty \\
&\leq 12M_nK_{n+1}g(R_{n+1}) \\
&\leq \frac{T_{n+1}^2}{32}.
\end{aligned}
$$

We then have by (2.37) and (2.13)

$$
\begin{aligned}
|\phi(x, u-w) - \phi(x, u-w_n)| &\leq |\phi(x, u-w) - \phi_{m_n}(x, u-w)| \\
&\quad + |\phi_{m_n}(x, u-w) - \phi_{m_n}(x, u-w_n)| \\
&\quad + |\phi_{m_n}(x, u-w_n) - \phi(x, u-w_n)| \\
&\leq 2\|\phi - \phi_{m_n}\|_\infty + \frac{T_{n+1}^2}{32} \\
&\leq 2\cdot 2^{-m_n} + \frac{T_{n+1}^2}{32} \\
&< \frac{2\cdot T_{n+1}^2}{64} + \frac{T_{n+1}^2}{32} \\
&= \frac{T_{n+1}^2}{16}.
\end{aligned}
$$

Now, using (2.36) and (2.14)

$$
\int_{G_n} |\phi(x, u-w) - \phi(x, u-w_n)| \leq 2\int_{G_n} \|\phi\|_\infty \leq 2C\mathrm{meas}(G_n) \leq \frac{T_{n+1}^2}{16}.
$$

So, recalling that $T_0 < 1/2$,

$$
\begin{aligned}
|\mathscr{L}(u) - \mathscr{L}_n(u)| &= \left| \int_{-T_0}^{T_0} \big((u')^2 + \phi(x, u-w)\big) - \big((u')^2 + \phi(x, u-w_n)\big) \right| \\
&\leq \int_{-T_0}^{T_0} |\phi(x, u-w) - \phi(x, u-w_n)| \\
&= \int_{G_n} |\phi(x, u-w) - \phi(x, u-w_n)| \\
&\quad + \int_{[-T_0,T_0]\setminus G_n} |\phi(x, u-w) - \phi(x, u-w_n)| \\
&< \frac{T_{n+1}^2}{16} + \int_{[-T_0,T_0]\setminus G_n} \frac{T_{n+1}^2}{16} \\
&\leq \frac{T_{n+1}^2}{8}. \qquad (2.40)
\end{aligned}
$$

Now we estimate $|\mathscr{L}(w) - \mathscr{L}_n(w_n)|$. First we note that (2.21.3) and the estimate (2.32) imply that for all $n \geq 0$,

$$\|w'\|_{L^2(-T_0,T_0)} \leq \|\tilde{w}_0'\|_{L^2(-T_0,T_0)} + T_1^2/16D \leq \|\tilde{w}_0'\|_{L^2(-T_0,T_0)} + 1,$$

and

$$\|w_n'\|_{L^2(-T_0,T_0)} \leq \|\tilde{w}_0'\|_{L^2(-T_0,T_0)} + 1,$$

hence by definition of D that

$$\|w' + w_n'\|_{L^2(-T_0,T_0)} \leq 2(\|\tilde{w}_0'\|_{L^2(-T_0,T_0)} + 1) \leq 2D.$$

Thus using (2.39), Cauchy-Schwartz, and (2.21.3), we see

$$\begin{aligned}
|\mathscr{L}(w) - \mathscr{L}_n(w_n)| &\leq \int_{-T_0}^{T_0} |(w')^2 - (w_n')^2| \\
&\leq \|w' - w_n'\|_{L^2(-T_0,T_0)} \|w' + w_n'\|_{L^2(-T_0,T_0)} \\
&\leq \frac{2DT_{n+1}^2}{16D} \\
&= \frac{T_{n+1}^2}{8}. \qquad (2.41)
\end{aligned}$$

Combining these two estimates we see

$$\begin{aligned}
|(\mathscr{L}(u) - \mathscr{L}(w)) - (\mathscr{L}_n(u) - \mathscr{L}_n(w_n))| &\leq |\mathscr{L}(u) - \mathscr{L}_n(u)| + |\mathscr{L}(w) - \mathscr{L}_n(w_n)| \\
&< \frac{T_{n+1}^2}{8} + \frac{T_{n+1}^2}{8} \\
&= \frac{T_{n+1}^2}{4}. \qquad \square
\end{aligned}$$

We now show w is the unique minimizer of $(\star)$. We briefly discuss the main ideas behind the proof, which, as mentioned before, are essentially the proof that $\tilde{w}$ minimizes the variational problem with "basic" Lagrangian

$$(x, y, p) \mapsto \tilde{L}(x, y, p) = \tilde{\phi}(x, y - \tilde{w}(x)) + p^2.$$

So suppose for now that $\tilde{u} \in \mathrm{AC}(-T_0, T_0)$ is a minimizer for this basic problem with Lagrangian $\tilde{L}$. If $\tilde{u}(0) = \tilde{w}(0)$, it suffices to argue separately on $[-T_0, 0]$ and $[0, T_0]$. We consider $[0, T_0]$. But $\tilde{w}$ is C^2 on $(0, T_0)$, so we can make the important step of integrating by parts. Moreover, a simple trick relying on $\tilde{u}$ being a minimizer gives us that $|\tilde{u}(x)| \leq |g(x)|$ (see Lemma 2.23 below for the essence of the argument), so $|\tilde{u}(x) - \tilde{w}(x)| \leq 2|g(x)|$. Note that for any two functions $\bar{u}, \bar{w} \colon [-T_0, T_0] \to \mathbb{R}$, we have

$$(\bar{u})^2 - (\bar{w})^2 = (\bar{u} - \bar{w})^2 + 2(\bar{u} - \bar{w})\bar{w} \geq 2(\bar{u} - \bar{w})\bar{w}. \tag{2.42}$$

So we can argue

$$\begin{aligned}
\int_0^{T_0} \big(\tilde{\phi}(x, \tilde{u} - \tilde{w}) + (\tilde{u}')^2\big) - \int_0^{T_0} (\tilde{w}')^2 &\geq \int_0^{T_0} \big(2(\tilde{u}' - \tilde{w}')\tilde{w}' + \tilde{\phi}(x, \tilde{u} - \tilde{w})\big) \\
&= [2(\tilde{u} - \tilde{w})\tilde{w}']_0^{T_0} \\
&\quad + \int_0^{T_0} \big(\tilde{\phi}(x, \tilde{u} - \tilde{w}) - 2(\tilde{u} - \tilde{w})\tilde{w}''\big) \\
&\geq \int_0^{T_0} \big(\psi(x)|\tilde{u} - \tilde{w}| - 2|\tilde{u} - \tilde{w}||\tilde{w}''(x)|\big)
\end{aligned}$$

and hence it suffices to choose ψ large enough to dominate $\tilde{w}''$, which we can do (this is the role of ψ^2).

This argument cannot be performed in the case when $\tilde{u}(0) \neq \tilde{w}(0)$, and there is no *a priori* reason why this might not occur. In this case, we compare $\tilde{u}$ not with $\tilde{w}$ but with a new function we obtain by replacing $\tilde{w}$ with a linear function on an interval around 0.

This basic idea on $\tilde{w}$ is mimicked locally on w around each x_n; more precisely we in fact argue with w_n and then either show that for some n this suffices to give the result for w, or pass to the limit. The techniques of our proof show in fact that w_n is the unique minimizer of the variational problem

$$\mathrm{AC}(-T_0, T_0) \ni u \mapsto \mathscr{L}_n(u)$$

over those u such that $u(\pm T_0) = w_n(\pm T_0) (= w(\pm T_0))$. Thus in particular we get an example of a one-point non-differentiable minimizer: the conditions of Lemma 2.27 below always hold for $n = 0$, which already shows that Tonelli's theorem cannot hold in the continuous case.

We return to the problem proper. Suppose now $u \in \mathrm{AC}(-T_0, T_0)$ is a minimizer for $(\star)$ and $u \neq w$. Note that a minimizer certainly exists, since L is continuous, and superlinear and convex in p, see Theorem 1.1. We now make a number of estimates, with the eventual aim of showing that

$$\mathscr{L}(u) - \mathscr{L}(w) = \int_{-T_0}^{T_0} (u')^2 + \phi(x, u - w) - (w')^2 > 0,$$

which contradicts the choice of u as a minimizer for $(\star)$. Write $v = u - w$, and $v_n = u - w_n$. If $u(x_n) = w(x_n)$ for all $n \geq 0$, then the proof is an easy application of integration by parts as discussed above on the complement of the closure of the points $\{x_n\}_{n=0}^{\infty}$. (In the case that $\{x_n\}_{n=0}^{\infty}$ forms a dense set in $[-T_0, T_0]$, we should immediately have $u = w$ by continuity, thus concluding the proof of minimality of w without using either the assumption that u was a minimizer or that $u \neq w$.) Should $w(x_n) \neq u(x_n)$ for some $n \geq 0$, further argument is required. The next lemma shows us that *since u is a minimizer*, it cannot be too badly behaved around any point $x \in [-T, T]$ where $u(x) \neq w(x)$.

Lemma 2.23. Let $n \geq 0$ be such that $u(x_n) \neq w(x_n)$. Let $J_n \subseteq [-T_0, T_0]$ be the connected component of $[-T_0, T_0]$ containing x_n of the set of points $x \in [-T_0, T_0]$ such that

$$|u(x) - w(x_n)| > 3|\tilde{g}_n(x)|.$$

Note that $J_n \subsetneq [-T_0, T_0]$ is an open subinterval of $[-T_0, T_0]$ since u and w agree at $\pm T_0$ and so by (2.21.4)

$$|u(\pm T_0) - w(x_n)| = |w(\pm T_0) - w(x_n)| \leq 2\,|\tilde{g}_n(\pm T_0)|\,.$$

So there exist $a_n, b_n > 0$ such that $J_n = (x_n - a_n, x_n + b_n)$ and

$$|u(x_n - a_n) - w(x_n)| = 3\theta_n g(a_n) \text{ and } |u(x_n + b_n) - w(x_n)| = 3\theta_n g(b_n).$$

Similarly we choose $\alpha_n, \beta_n > 0$ such that $(x_n - \alpha_n, x_n + \beta_n)$ is the connected component containing x_n of those points for which $|u(x) - w(x_n)| > 2|\tilde{g}_n(x)|$. So $a_n \leq \alpha_n$ and $b_n \leq \beta_n$, but still $(x_n - \alpha_n, x_n + \beta_n) \subseteq [-T_0, T_0]$.

Then in case $u(x_n) > w(x_n)$, u is convex on $(x_n - \alpha_n, x_n + \beta_n)$ and

$$-2\theta_n g'(-\alpha_n) \leq u' \leq 2\theta_n g'(\beta_n) \text{ almost everywhere on } (x_n - \alpha_n, x_n + \beta_n); \quad (2.43)$$

and in case $u(x_n) < w(x_n)$, u is concave on $(x_n - \alpha_n, x_n + \beta_n)$ and

$$-2\theta_n g'(\beta_n) \leq u' \leq 2\theta_n g'(-\alpha_n) \text{ almost everywhere on } (x_n - \alpha_n, x_n + \beta_n). \quad (2.44)$$

Hence

$$\begin{cases} |v_n(x)| \geq \theta_n g(b_n) \text{ for } x \in [x_n, x_n + b_n] & \text{if } b_n \geq a_n \\ |v_n(x)| \geq \theta_n g(a_n) \text{ for } x \in [x_n - a_n, x_n] & \text{if } a_n \geq b_n. \end{cases} \quad (2.45)$$

Finally

$$|u(x) - w(x_n)| \leq 3|\tilde{g}_n(x)| \text{ for } x \notin J_n. \tag{2.46}$$

Proof. We suppose $u(x_n) > w(x_n)$. The argument for the case $u(x_n) < w(x_n)$ is very similar.

Suppose u is not convex on $(x_n - \alpha_n, x_n + \beta_n)$, so there exist points $t_1, t_2 \in (x_n - \alpha_n, x_n + \beta_n)$, $t_1 < t_2$ say, and $\lambda \in [0, 1]$ such that

$$u(\lambda t_1 + (1-\lambda)t_2) > \lambda u(t_1) + (1-\lambda)u(t_2).$$

Let $z\colon [-T_0, T_0] \to \mathbb{R}$ be the affine function with graph passing through $(t_1, u(t_1))$ and $(t_2, u(t_2))$, so

$$z(x) = \frac{u(t_2) - u(t_1)}{t_2 - t_1} \cdot (x - t_1) + u(t_1).$$

So we have by assumption on t_1, t_2 that

$$z(\lambda t_1 + (1-\lambda)t_2) = \lambda u(t_1) + (1-\lambda)u(t_2) < u(\lambda t_1 + (1-\lambda)t_2).$$

Passing to connected components if necessary, we can assume that $z < u$ on (t_1, t_2). We claim that adding a certain constant value onto the function z gives an affine function $\tilde{z}$ such that on some subinterval $(\tilde{t}_1, \tilde{t}_2)$ of (t_1, t_2), we have

$$w(x_n) + 2|\tilde{g}_n| \leq \tilde{z} < u.$$

We then show this contradicts the choice of u as a minimizer for $(\star)$.

Since z is affine and g_n is concave on $[-T_0, x_n]$ and $[x_n, T_0]$, the equation $z = w(x_n) + 2|\tilde{g}_n|$ can in principle have no or up to three distinct solutions on (t_1, t_2), or can be satisfied identically if g_n is affine on this interval. In this latter case the claim is satisfied trivially for $\tilde{z} = z$. If there is at most one solution, then since $z(t_i) = u(t_i) \geq w(x_n) + 2|g_n(t_i)|$ for $i = 1, 2$, evidently $z \geq w(x_n) + 2|\tilde{g}_n|$ on (t_1, t_2). So again we need not modify z at all to get our required $\tilde{z}$.

The case of three distinct solutions is in fact impossible. Suppose we had three such points $s_1, s_2, s_3 \in (t_1, t_2)$. Again by the elementary properties of $\tilde{g}_n$ and z, all three points cannot lie on one side of x_n. So suppose $s_1 \leq x_n \leq s_2 < s_3$. Then for $t < x_n$, by (2.iv), we have that

$$z' = \frac{2|\tilde{g}_n(s_3)| - 2|\tilde{g}_n(s_2)|}{s_3 - s_2} = \frac{2\tilde{g}_n(s_3) - 2\tilde{g}_n(s_2)}{s_3 - s_2} > 0 > -2\tilde{g}_n'(t).$$

Since $t_1 < s_1 \le x_n$, we have $|\tilde{g}_n(s_1)| = -\tilde{g}_n(s_1)$ and $|\tilde{g}_n(t_1)| = -\tilde{g}_n(t_1)$, so

$$\begin{aligned} z(t_1) &= z(s_1) - \int_{t_1}^{s_1} z'(t)\,dt \\ &< w(x_n) - 2\tilde{g}_n(s_1) - \int_{t_1}^{s_1} (-2\tilde{g}_n'(t))\,dt \\ &= w(x_n) - 2\tilde{g}_n(t_1) \\ &= w(x_n) + 2|\tilde{g}_n(t_1)|. \end{aligned}$$

This is a contradiction since $z(t_1) = u(t_1) > w(x_n) + 2|\tilde{g}_n(t_1)|$. Similarly the case $s_1 < s_2 \le x_n \le s_3$ is dealt with.

So it remains to deal with the case where we have two distinct solutions (s_1, s_2)—this is the case in which we have to possibly add a constant to z. The same considerations as in the preceding paragraph show that we must have both solutions lying to one side of x_n. Suppose $x_n \le s_1 < s_2$. Then by (2.ii), $2|\tilde{g}_n| = 2\tilde{g}_n$ is C^2 on (s_1, s_2), so applying the mean value theorem we see that there is a point $s_0 \in (s_1, s_2)$ such that

$$2\tilde{g}_n'(s_0) = \frac{2\tilde{g}_n(s_2) - 2\tilde{g}_n(s_1)}{s_2 - s_2} = \frac{z(s_2) - z(s_1)}{s_2 - s_1} = z'.$$

Define $\tilde{z}$ by

$$\tilde{z}(x) = z'(x - s_0) + w(x_n) + 2\tilde{g}_n(s_0),$$

the tangent to $w(x_n) + 2\tilde{g}_n$ at s_0, so

$$\tilde{z}(s_0) = w(x_n) + 2\tilde{g}_n(s_0) = w(x_n) + 2|\tilde{g}_n(s_0)| < u(s_0).$$

Let $(\tilde{t}_1, \tilde{t}_2)$ be the connected component containing s_0 such that $u > \tilde{z}$ on $(\tilde{t}_1, \tilde{t}_2)$. Since $s_0 \in (s_1, s_2)$, and $z(s_i) = w(x_n) + 2\tilde{g}_n(s_i)$ for $i = 1, 2$, concavity of g implies $w(x_n) + 2\tilde{g}_n(s_0) \ge z(s_0)$. Since $\tilde{z}(s_0) = w(x_n) + 2\tilde{g}_n(s_0)$ by definition, and $z' = \tilde{z}'$, we have $\tilde{z} \ge z$ everywhere. So $u > \tilde{z}$ implies $u > z$, thus $(\tilde{t}_1, \tilde{t}_2) \subseteq (t_1, t_2)$.

We claim $\tilde{z} \ge w(x_n) + 2|\tilde{g}_n|$ on $(\tilde{t}_1, \tilde{t}_2)$. Since $s_0 > x_n$ and $\tilde{z}(s_0) = w(x_n) + 2|\tilde{g}_n(s_0)|$, by concavity we have $\tilde{z} \ge w(x_n) + 2|\tilde{g}_n|$ on (x_n, T_0). Suppose there existed $s \in (\tilde{t}_1, x_n]$ such that $\tilde{z}(s) < w(x_n) + 2|\tilde{g}_n(s)| = w(x_n) - 2\tilde{g}_n(s)$. Then we see as before that

$$\begin{aligned}\tilde{z}(\tilde{t}_1) &= \tilde{z}(s) - \int_{\tilde{t}_1}^{s} \tilde{z}'(t)\,dt \\ &< w(x_n) - 2\tilde{g}_n(s) - \int_{\tilde{t}_1}^{s} (-2\tilde{g}_n')(t)\,dt \\ &= w(x_n) - 2\tilde{g}_n(\tilde{t}_1) \\ &= w(x_n) + 2|\tilde{g}_n(\tilde{t}_1)|,\end{aligned}$$

which contradicts $\tilde{z}(\tilde{t}_1) = u(\tilde{t}_1) > w(x_n) + 2|\tilde{g}_n(\tilde{t}_1)|$. So $\tilde{z} \geq w(x_n) + 2|\tilde{g}_n|$ on $(\tilde{t}_1, \tilde{t}_2)$ indeed. The case where $s_1 < s_2 \leq x_n$ is similar. So we have constructed an affine $\tilde{z}$ as claimed.

Thus, since by (2.21.4) $w \leq w(x_n) + 2|\tilde{g}_n|$, on $(\tilde{t}_1, \tilde{t}_2)$, we have

$$|u - w| = u - w \geq \tilde{z} - w = |\tilde{z} - w|. \tag{2.47}$$

Since $u > \tilde{z}$ on $(\tilde{t}_1, \tilde{t}_2)$, where $\tilde{z}$ is affine, but $u = \tilde{z}$ at the endpoints, we know u is not affine on $(\tilde{t}_1, \tilde{t}_2)$, so we have strict inequality in Hölder's inequality, thus

$$\begin{aligned}\int_{\tilde{t}_1}^{\tilde{t}_2} (u')^2 &= \frac{1}{\tilde{t}_2 - \tilde{t}_1} \left(\int_{\tilde{t}_1}^{\tilde{t}_2} 1^2\right) \left(\int_{\tilde{t}_1}^{\tilde{t}_2} (u')^2\right) \\ &> \frac{1}{\tilde{t}_2 - \tilde{t}_1} \left(\int_{\tilde{t}_1}^{\tilde{t}_2} u'\right)^2 \\ &= \frac{(u(\tilde{t}_2) - u(\tilde{t}_1))^2}{\tilde{t}_2 - \tilde{t}_1} \\ &= (\tilde{t}_2 - \tilde{t}_1)\left(\frac{z(\tilde{t}_2) - z(\tilde{t}_1)}{\tilde{t}_2 - \tilde{t}_1}\right)^2 \\ &= (\tilde{t}_2 - \tilde{t}_1)(\tilde{z}')^2 \\ &= \int_{\tilde{t}_1}^{\tilde{t}_2} (\tilde{z}')^2.\end{aligned} \tag{2.48}$$

Hence defining $\tilde{u}\colon [-T_0, T_0] \to \mathbb{R}$ by

$$\tilde{u}(x) = \begin{cases} u(x) & x \notin (\tilde{t}_1, \tilde{t}_2) \\ \tilde{z}(x) & x \in (\tilde{t}_1, \tilde{t}_2) \end{cases}$$

gives a function $\tilde{u} \in \mathrm{AC}(-T_0, T_0)$ with $\tilde{u}(\pm T_0) = w(\pm T_0)$, and such that, using (2.48), (2.47), and (2.38),

$$\begin{aligned}
\mathscr{L}(\tilde{u}) &= \int_{-T_0}^{T_0} L(x, \tilde{u}, \tilde{u}') \\
&= \int_{-T_0}^{T_0} \big((\tilde{u}')^2 + \phi(x, \tilde{u} - w)\big) \\
&= \int_{[-T_0,T_0]\setminus(\tilde{t}_1,\tilde{t}_2)} \big((u')^2 + \phi(x, u - w)\big) + \int_{\tilde{t}_1}^{\tilde{t}_2} \big((\tilde{z}')^2 + \phi(x, \tilde{z} - w)\big) \\
&< \int_{[-T_0,T_0]\setminus(\tilde{t}_1,\tilde{t}_2)} \big((u')^2 + \phi(x, u - w)\big) + \int_{\tilde{t}_1}^{\tilde{t}_2} \big((u')^2 + \phi(x, u - w)\big) \\
&= \int_{-T_0}^{T_0} L(x, u, u') \\
&= \mathscr{L}(u),
\end{aligned}$$

which contradicts the choice of u as a minimizer. Hence u is indeed convex on $(x_n - \alpha_n, x_n + \beta_n)$.

It now follows that the graph of u on $(x_n - \alpha_n, x_n + \beta_n)$ lies above the tangents to $w(x_n) + 2|\tilde{g}_n|$ at $(x_n - \alpha_n)$ and $(x_n + \beta_n)$:

$$u(x) \geq 2\theta_n g'(\beta_n)(x - (x_n + \beta_n)) + 2\theta_n g(\beta_n) + w(x_n)$$

and

$$u(x) \geq -2\theta_n g'(-\alpha_n)(x - (x_n - \alpha_n)) + 2\theta_n |g(-\alpha_n)| + w(x_n)$$

for $x \in (x_n - \alpha_n, x_n + \beta_n)$. For suppose the first fails, i.e. that for some $t_0 \in (x_n - \alpha_n, x_n + \beta_n)$ we have

$$u(t_0) < 2\theta_n g'(\beta_n)(t_0 - (x_n + \beta_n)) + 2\theta_n g(\beta_n) + w(x_n).$$

Then by convexity the graph of u lies below the chord between the points $(t_0, u(t_0))$ and $(x_n + \beta_n, u(x_n + \beta_n)) = (x_n + \beta_n, w(x_n) + 2\theta_n g(\beta_n))$, which has slope

$$\frac{w(x_n) + 2\theta_n g(\beta_n) - u(t_0)}{x_n + \beta_n - t_0}.$$

By assumption

$$\frac{w(x_n) + 2\theta_n g(\beta_n) - u(t_0)}{x_n + \beta_n - t_0} > 2\theta_n g'(\beta_n)$$

and so since g' is continuous by (2.ii) we have that

$$2\theta_n g'(t) < \frac{w(x_n) + 2\theta_n g(\beta_n) - u(t_0)}{x_n + \beta_n - t_0}$$

on some left neighbourhood of β_n. So for x in this neighbourhood, we have

$$\begin{aligned} w(x_n) + 2\tilde{g}_n(x) &= w(x_n) + 2\tilde{g}_n(x_n + \beta_n) - \int_x^{x_n+\beta_n} 2\tilde{g}_n'(t)\,dt \\ &> w(x_n) + 2\tilde{g}_n(x_n + \beta_n) - \int_x^{x_n+\beta_n} \frac{w(x_n) + 2\theta_n g(\beta_n) - u(t_0)}{x_n + \beta_n - t_0}\,dt \\ &= w(x_n) + 2\tilde{g}_n(x_n + \beta_n) - \frac{w(x_n) + 2\theta_n g(\beta_n) - u(t_0)}{x_n + \beta_n - t_0}(x_n + \beta_n - x) \\ &= u(x_n + \beta_n) - \frac{w(x_n) + 2\theta_n g(\beta_n) - u(t_0)}{x_n + \beta_n - t_0}(x_n + \beta_n - x) \\ &\geq u(x), \end{aligned}$$

which is a contradiction for $x \in (x_n - \alpha_n, x_n + \beta_n)$. Similarly we prove u lies above the other tangent.

We can now prove the claimed bounds on u'. Suppose there exists a $t_0 \in (x_n - \alpha_n, x_n + \beta_n)$ such that $u'(t_0) > 2\theta_n g'(\beta_n)$. Then we have $u'(x) > 2\theta_n g'(\beta_n)$ for all $x \in (t_0, x_n + \beta_n)$ by convexity. Then we see

$$\begin{aligned} u(x_n + \beta_n) &= u(t_0) + \int_{t_0}^{x_n+\beta_n} u'(t)\,dt \\ &> 2\theta_n g'(\beta_n)(t_0 - (x_n + \beta_n)) + w(x_n) \\ &\quad + 2\theta_n g(\beta_n) + ((x_n + \beta_n) - t_0) 2\theta_n g'(\beta_n) \\ &= w(x_n) + 2\theta_n g(\beta_n), \end{aligned}$$

which is a contradiction since $u(x_n + \beta_n) = 2\theta_n g(\beta_n)$ by choice of β_n. The lower bound for u' is proved similarly.

We now prove the important consequence (2.45) of these derivative estimates. Suppose $b_n \geq a_n$. Then using convexity of u, and the fact that (2.iv) implies in this case that $\tilde{g}(b_n) > \tilde{g}(-a_n) = -\tilde{g}(a_n)$, we see that for $x \in J_n$,

$$\begin{aligned} u(x) &\leq \frac{u(x_n + b_n) - u(x_n - a_n)}{b_n + a_n}(x - (x_n + b_n)) + u(x_n + b_n) \\ &= \frac{3\tilde{g}_n(x_n + b_n) + 3\tilde{g}_n(x_n - a_n)}{b_n + a_n}(x - (x_n + b_n)) + w(x_n) + 3\tilde{g}_n(x_n + b_n) \\ &\leq w(x_n) + 3\tilde{g}_n(x_n + b_n). \end{aligned}$$

Fix $x \in [x_n, x_n + b_n]$, we then have by (2.43), which we have just proved,

$$\begin{aligned} u(x) &= u(x_n + b_n) - \int_x^{x_n+b_n} u'(t)\,dt \\ &\geq w(x_n) + 3\tilde{g}_n(x_n + b_n) - \int_x^{x_n+b_n} 2\tilde{g}_n'(x_n + \beta_n)\,dt \\ &= w(x_n) + 3\tilde{g}_n(x_n + b_n) - 2((x_n + b_n) - x)\tilde{g}_n'(x_n + \beta_n). \end{aligned}$$

Also, since $x \leq x_n + b_n$, we have, using (2.20.3), (2.21.1), and concavity of g,

$$\begin{aligned} w_n(x) &\leq w(x_n) + 2\tilde{g}_n(x) \\ &\leq w(x_n) + 2\tilde{g}_n'(x_n + b_n)(x - (x_n + b_n)) + 2\tilde{g}_n(x_n + b_n) \\ &\leq w(x_n) + 2\tilde{g}_n'(x_n + \beta_n)(x - (x_n + b_n)) + 2\tilde{g}_n(x_n + b_n). \end{aligned}$$

So we have

$$\begin{aligned} u(x) - w_n(x) &\geq \big(w(x_n) + 3\tilde{g}_n(x_n + b_n) - 2((x_n + b_n) - x)\tilde{g}_n'(x_n + \beta_n)\big) \\ &\quad - (2\tilde{g}_n'(x_n + \beta_n)(x - (x_n + b_n)) + w(x_n) + 2\tilde{g}_n(x_n + b_n)) \\ &= \tilde{g}_n(x_n + b_n). \\ &= \theta_n g(b_n). \end{aligned}$$

Similarly we can prove that $u(x) - w_n(x) \geq \theta_n g(a_n)$ for $x \in [x_n - a_n, x_n]$ if $a_n \geq b_n$. In the case that $u(x_n) < w(x_n)$ we can prove in the same way that $u(x) - w_n(x) \leq -\theta_n g(a_n)$ on $[x_n - a_n, x_n]$ if $a_n \geq b_n$, or $u(x) - w_n(x) \leq -\theta_n g(b_n)$ on $[x_n, x_n + b_n]$, hence the full result.

The final statement of the Lemma is proved using the techniques we used above to prove convexity of u on $(x_n - \alpha_n, x_n + \beta_n)$. Suppose there is a $t_0 \in (x_n + b_n, T_0)$ such that $u(t_0) > w(x_n) + 3|\tilde{g}_n(t_0)|$. Defining affine $z\colon [-T_0, T_0] \to \mathbb{R}$ by

$$z(x) = 3\tilde{g}_n'(t_0)(x - t_0) + w(x_n) + 3\tilde{g}_n(t_0),$$

we see that $z(t_0) = w(x_n) + 3\tilde{g}_n(t_0) < u(t_0)$, and, using the concavity of $\tilde{g}_n$, that $z \geq w(x_n) + 3\tilde{g}_n$ on (x_n, T_0). The connected component of $[-T_0, T_0]$ containing t_0 on which $z < u$ on I is a subinterval of $(x_n + b_n, T_0)$, since

$$u(x_n + b_n) = w(x_n) + 3\tilde{g}_n(b_n) \leq z(x_n + b_n),$$

and by (2.21.4),

$$u(T_0) = w(T_0) \leq w(x_n) + 2\tilde{g}_n(T_0) < z(T_0).$$

So we have $u(x) > z(x) \geq w(x_n) + 3\tilde{g}_n(x)$ on some open subinterval of $(x_n + b_n, T_0)$. Hence we can perform the same trick as before, constructing a new function $\tilde{u} \in \mathrm{AC}(-T_0, T_0)$ by replacing u with z on this subinterval, such that $\mathscr{L}(\tilde{u}) < \mathscr{L}(u)$, which again contradicts the choice of u as a minimizer. $\square$

Thus we see that if for some $n \geq 0$, $u(x_n) \neq w(x_n)$, then u must be Lipschitz on a neighbourhood of x_n, and its graph cannot escape the cone bounded by the graphs of $x \mapsto w(x_n) \pm 3|\tilde{g}_n(x)|$ off this neighbourhood. We note that the final statement of the Lemma holds by the same argument even in case $u(x_n) = w(x_n)$ and thus when the set J_n introduced is empty.

For the remainder of the proof, we assume that $u(x_n) \neq w(x_n)$ for all $n \geq 0$. If not one can just perform the argument in the proofs of Lemma 2.27 and Corollary 2.28 on the connected components of $[-T_0, T_0] \backslash \overline{\{x_n : u(x_n) = w(x_n)\}}$. We make remarks in these proofs at those points where an additional argument is required in the general case.

For each $n \geq 0$ we now introduce some definitions and notation. Let $a_n, b_n > 0$ and $J_n = (x_n - a_n, x_n + b_n)$ be as in Lemma 2.23. We let $c_n = \max\{a_n, b_n\}$, and write $\tilde{J}_n = [x_n - c_n, x_n + c_n]$. Fix $n \geq 0$. We note the following immediate corollary of (2.46). For $x \notin J_n$, we have for any $i \geq n$, by (2.21.1) and by (2.20.3) that

$$\begin{aligned} |v_i(x)| &\leq |u(x) - w(x_n)| + |w(x_n) - w_i(x)| \\ &= |u(x) - w(x_n)| + |w_i(x_n) - w_i(x)| \\ &< 3|\tilde{g}_n(x)| + 2|\tilde{g}_n(x)| \\ &= 5|\tilde{g}_n(x)|. \end{aligned} \tag{2.49}$$

The inequalities (2.45) from Lemma 2.23 tell us that the graph of a putative minimizer u cannot get too close to that of w around x_n. As we see next, this lower bound of the distance means we have a certain amount of weight concentrated in our Lagrangian around each x_n. The total weight is of course in general even larger—we took an infinite sum of such non-negative terms—but the important term is the $\tilde{\phi}_n$ term which deals precisely with the oscillations introduced by w_n to get singularity of w at x_n.

Lemma 2.24. Let $n \geq 0$, and suppose $\tilde{J}_n \subseteq Y_n$. Then

$$\int_{\tilde{J}_n} \tilde{\phi}_n^1(x, v_n)\, dx \geq 5\theta_n \int_0^{g^{-1}(g(c_n)/5)} \kappa(x)\, dx,$$

where $\kappa \in C(0, T)$ is as in condition (2.xi).

Proof. Suppose $b_n \geq a_n$, so $c_n = b_n$. The case $a_n > b_n$ differs only in trivial notation. Note that $g(g^{-1}(g(c_n)/5)) = g(c_n)/5 \leq g(c_n)$, so $g^{-1}(g(c_n)/5) \leq c_n$ by (2.iv). So for $x \in [x_n, x_n + g^{-1}(g(c_n)/5)]$ we have by (2.45) and (2.iv)

$$|v_n(x)| \geq \theta_n g(c_n) = 5\theta_n g(g^{-1}(g(c_n)/5)) \geq 5\tilde{g}_n(x),$$

hence by definition (noting our one assumption $\tilde{J}_n \subseteq Y_n$), $\tilde{\phi}_n^1(x, v_n) = 5\tilde{g}_n(x)\psi_n^1(x)$ on $[x_n, x_n + g^{-1}(g(c_n)/5)]$. We can now estimate the integral as follows, recalling the definition of ψ_n^1:

$$\begin{aligned}\int_{\tilde{J}_n} \tilde{\phi}_n^1(x, v_n)\, dx &\geq \int_{x_n}^{x_n + g^{-1}(g(c_n)/5)} \tilde{\phi}_n^1(x, v_n)\, dx \\ &= 5\int_{x_n}^{x_n + g^{-1}(g(c_n)/5)} \tilde{g}_n(x)\psi_n^1(x)\, dx \\ &= 5\theta_n \int_{x_n}^{x_n + g^{-1}(g(c_n)/5)} g_n(x)\psi_n^1(x)\, dx \\ &= 5\theta_n \int_0^{g^{-1}(g(c_n)/5)} \kappa(x)\, dx. \end{aligned}$$ □

For $n \geq 0$ define $H_n \subseteq [-T_0, T_0]$ by

$$H_n := \tilde{J}_n \cap [x_n - \tau_n, x_n + \tau_n] = [x_n - d_n, x_n + d_n], \text{ say,}$$

so $d_n \leq c_n$. Note that by construction and (2.20.1)

$$w_n(x_n \pm d_n) = \lambda_n \tilde{w}_n(x_n \pm d_n) + \rho_n; \text{ and } w_n'(x_n \pm d_n) = \lambda_n \tilde{w}_n'(x_n \pm d_n).$$

We cannot immediately mimic the main principle of the proof and integrate by parts across x_n, since $\tilde{w}_n'$ does not exist at x_n. This singularity is of course the whole point of the example. The main trick of the proof was in making the oscillations of $\tilde{w}_n$ near x_n slow enough so that we can replace this function with a straight line on an interval containing x_n. We can then use integration by parts on each side of this interval, and inside the interval exploit the fact that we have now introduced a function with constant derivative. We incur an error in the boundary terms, of course, as we in general introduce discontinuities of the derivative where the line meets $\tilde{w}_n$, but the function $\tilde{w}_n$ moves slowly enough that this error can be dominated by the weight term in the Lagrangian (the role of ψ_n^1).

So let $\tilde{l}_n\colon [-T_0, T_0] \to \mathbb{R}$ denote the affine function with graph connecting $(x_n - d_n, \tilde{w}_n(x_n - d_n))$ and $(x_n + d_n, \tilde{w}_n(x_n + d_n))$, i.e.

$$\tilde{l}_n(x) = \tilde{l}_n'(x - (x_n - d_n)) + \tilde{w}_n(x_n - d_n),$$

where, by definition of $\tilde{w}_n$,

$$\tilde{l}_n' = \frac{\tilde{w}_n(x_n + d_n) - \tilde{w}_n(x_n - d_n)}{2d_n} = f(d_n)\sin h(d_n).$$

Define $l_n\colon [-T_0, T_0] \to \mathbb{R}$ by

$$l_n(x) = \begin{cases} w_n(x) & x \notin H_n \\ \lambda_n \tilde{l}_n(x) + \rho_n & x \in H_n. \end{cases}$$

Clearly $l_n \in \mathrm{AC}(-T_0, T_0)$.

We shall find the following notation useful, representing the boundary terms we get as a result of integrating by parts, firstly inside H_n, integrating $l_n' v_n'$, and secondly outside H_n, integrating $w_n' v_n'$:

$$I_{n,l} = \lambda_n \tilde{l}_n' v_n(x_n - d_n),\ \ I_{n,r} = \lambda_n \tilde{l}_n' v_n(x_n + d_n);$$
$$E_{n,l} = w_n'(x_n - d_n) v_n(x_n - d_n),\ \ E_{n,r} = w_n'(x_n + d_n) v_n(x_n + d_n).$$

Note that

$$|I_{n,l} - E_{n,l}| = |\lambda_n||v_n(x_n - d_n)(\tilde{l}_n' - \tilde{w}_n'(x_n - d_n))|;\ \text{and} \tag{2.50}$$
$$|I_{n,r} - E_{n,r}| = |\lambda_n||v_n(x_n + d_n)(\tilde{l}_n' - \tilde{w}_n'(x_n + d_n))|. \tag{2.51}$$

Lemma 2.25. Let $n \geq 0$. Then

$$\int_{H_n} (u')^2 - (w_n')^2 > 2(I_{n,r} - I_{n,l}) - 8\Psi(d_n),$$

where Ψ is as defined in condition (2.xi).

Proof. We want to use the following estimate, replacing w_n with the line l_n and estimating the error:

$$\int_{H_n} (u')^2 - (w_n')^2 = \int_{H_n} \left((u')^2 - (l_n')^2\right) + \int_{H_n} \left((l_n')^2 - (w_n')^2\right)$$
$$\geq \int_{H_n} \left((u')^2 - (l_n')^2\right) - \int_{H_n} |(l_n')^2 - (w_n')^2|. \tag{2.52}$$

We claim that this error term can be bounded by the function Ψ defined in (2.xi). Since $w_n' = \lambda_n \tilde{w}_n'$ and $l_n' = \lambda_n \tilde{l}_n'$ on H_n, a factor of $|\lambda_n^2| \leq 4$ comes out of the second (error) term, so we can just estimate this term in the case $n = 0$; the case of general n is just a translation of this base case. We drop the index 0 from the notation.

Observe that for $s \in [0, d]$ we have

$$\frac{d}{ds}(f(s)\sin h(s)) = f'(s)\sin h(s) + f(s)h'(s)\cos h(s)$$

and so

$$\left|\frac{d}{ds}(f(s)\sin h(s))\right| \leq |f'(s)| + |f(s)h'(s)|.$$

Fix $x \in [0, d]$. Since $f \in C^2(0, d)$, we can use the mean value theorem to see, using also (2.x), that for some $t \in (x, d)$,

$$\begin{aligned} |f(d)\sin h(d) - f(x)\sin h(x)| &= \left| \left(\frac{d}{ds} f(s)\sin h(s) \right) \Big|_{s=t} \right| |d - x| \\ &\leq (|f'(t)| + |f(t)h'(t)|)(d - x) \\ &\leq (|f'(x)| + |f(x)h'(x)|)(d - x). \end{aligned}$$

So, going back to the definitions,

$$\begin{aligned} |\tilde{l}' - \tilde{w}'(x)| &= |f(d)\sin h(d) - ((xf'(x) + f(x))\sin h(x) + xf(x)h'(x)\cos h(x))| \\ &\leq |f(d)\sin h(d) - f(x)\sin h(x)| + |xf'(x)| + |xf(x)h'(x)| \qquad (2.53) \\ &\leq (|f'(x)| + |f(x)h'(x)|)(d - x) + x|f'(x)| + x|f(x)h'(x)| \\ &= d(|f'(x)| + |f(x)h'(x)|). \end{aligned}$$

We immediately also see, using (2.iii), that

$$\begin{aligned} |\tilde{l}' \pm \tilde{w}'(x)| &= |f(d)\sin h(d) \pm \left((xf'(x) + f(x))\sin h(x) + xf(x)h'(x)\cos h(x)\right)| \\ &\leq |f(d)| + |xf'(x)| + |f(x)| + |xf(x)h'(x)| \\ &\leq |f(x)|(2 + x|h'(x)|) + |xf'(x)|. \end{aligned}$$

So on $[0,d]$, we have

$$|\tilde{l} - \tilde{w}'(x)| \leq \min\{d(|f'(x)| + |f(x)h'(x)|), |f(x)|(2 + x|h'(x)|) + |xf'(x)|\}.$$

Hence, since the integrand is an even function, we have, by definition of Ψ,

$$\begin{aligned}
&\int_H |(\tilde{l}')^2 - (\tilde{w}')^2| \, dx \\
&= 2\int_0^d |l' - w'||l' + w'| \, dx \\
&\leq 2\int_0^d (\min\{d(|f'| + |fh'|), 2|f| + |xf'| + |xfh'|\})(2|f| + |xf'| + |xfh'|) \, dx \\
&= 2\Psi(d). \qquad (2.54)
\end{aligned}$$

By (2.42) we have, since $l(\pm d) = w(\pm d)$,

$$\begin{aligned}
\int_H \left((u')^2 - (\tilde{l}')^2\right) &\geq \int_H 2\tilde{l}'(u' - \tilde{l}') \\
&= 2\tilde{l}' \int_H (u' - \tilde{l}') \\
&= 2\tilde{l}'[u - \tilde{l}]_{-d}^{d} \\
&= 2\tilde{l}'[v]_{-d}^{d} \\
&= 2(I_r - I_l).
\end{aligned}$$

Putting this and (2.54) into (2.52) gives the result. □

An estimate established in the preceding proof also gives easily the following important result. The errors we incur in our boundary terms by introducing a jump discontinuity in the derivative of our new function l_n are sufficiently small; they can be controlled by the integral over $H_n = [x_n - d_n, x_n + d_n]$ of a continuous function in $c_n \geq d_n$ taking value 0 at x_n (e.g. a translate of κ).

Lemma 2.26. Let $n \geq 0$. Then

$$|I_{n,r} - E_{n,r}| + |I_{n,l} - E_{n,l}| < 20\theta_n g(c_n)(|c_n f'(c_n)| + |c_n f(c_n)h'(c_n)|).$$

Proof. We just have to estimate $|v_n(x_n \pm d_n)|$. Suppose $u(x_n) > w(x_n)$; the argument for $u(x_n) < w(x_n)$ is similar. Suppose also $b_n \geq a_n$, so $c_n = b_n$. The case $a_n > b_n$ is similar. Then $u(x) \leq u(x_n + b_n)$ by convexity of u, for all $x \in J_n$.

If $x_n - d_n \notin J_n$, then (2.49) gives us the immediate estimate

$$|v_n(x_n - d_n)| \le 5\theta_n g(d_n) \le 5\theta_n g(b_n),$$

since $d_n \le b_n$ and by (2.iv).

If $x_n - d_n \in J_n$, then since certainly $x_n + d_n \in J_n$, we can argue that, by definition of J_n,

$$w(x_n) < w(x_n) + 3|\tilde{g}_n(x_n \pm d)| \le u(x_n \pm d_n) \le u(x_n + b_n) = w(x_n) + 3\tilde{g}_n(x_n + b_n)$$

thus

$$0 < u(x_n \pm d_n) - w(x_n) \le 3\theta_n g(b_n).$$

Hence using also (2.21.1), (2.20.3), and (2.iv), since $d_n \le b_n$,

$$\begin{aligned}
|v_n(x_n \pm d_n)| &\le |u(x_n \pm d_n) - w(x_n)| + |w(x_n) - w_n(x_n \pm d_n)| \\
&\le |u(x_n \pm d_n) - w(x_n)| + |w_n(x_n) - w_n(x_n \pm d_n)| \\
&< 3\theta_n g(b_n) + 2\theta_n g(d_n) \\
&\le 5\theta_n g(b_n).
\end{aligned}$$

Hence in both cases $|v_n(x_n \pm d_n)| \le 5\theta_n g(b_n)$. The result then follows by using (2.53) in (2.50) and (2.51), and by (2.ix), since $d_n \le c_n$, and $|\lambda_n| < 2$. □

We now combine our estimates for $\mathscr{L}_n$ across the whole domain $[-T_0, T_0]$, integrating by parts off $\bigcup_{i=1}^n H_i$ and using the above estimate on each H_i. We work with simplifying assumptions implying the relevant intervals do not overlap. We discuss later how to deal with the failure of these assumptions.

Lemma 2.27. Suppose $n \ge 0$ is such that for all $0 \le j \le n$,

$$\tilde{J}_k \cap Y_j = \emptyset \text{ for all } 0 \le k < j; \text{ and} \tag{2.55}$$

$$\tilde{J}_j \subseteq Y_j. \tag{2.56}$$

Then

$$\mathscr{L}_n(u) - \mathscr{L}_n(w_n) \ge \sum_{i=0}^{n} \theta_i g(c_i)(|c_i f'(c_i)| + |c_i f(c_i) h'(c_i)|) + \int_{[-T_0, T_0] \setminus \bigcup_{i=0}^n H_i} |v_n|.$$

Proof. By (2.20.7) and assumption (2.55) we have $w_j = w_k$ on $\tilde{J}_k$ for all

$0 \leq k < j \leq n$, in particular

$$w_n = w_k,\ w_n' = w_k' \text{ and } w_n'' = w_k'' \text{ (whenever both sides exist) on } \tilde{J}_k, \text{ for } 0 \leq k \leq n. \tag{2.57}$$

Also, by assumptions (2.56) and (2.55) together we have that for $0 \leq k < j \leq n$

$$\tilde{J}_k \cap \tilde{J}_j \subseteq \tilde{J}_k \cap Y_j = \emptyset,$$

i.e. the $\{\tilde{J}_j\}_{j=0}^n$ are pairwise disjoint.

Now, let $0 \leq i \leq n$. We see, using (2.42), that

$$\begin{aligned}
&\int_{\tilde{J}_i} \big((u')^2 + \phi(x, v_i) - (w_i')^2\big) \\
&= \int_{\tilde{J}_i} \phi(x, v_i) + \int_{\tilde{J}_i \setminus H_i} \big((u')^2 - (w_i')^2\big) + \int_{H_i} \big((u')^2 - (w_i')^2\big) \\
&\geq \int_{\tilde{J}_i} (\phi^1(x, v_i) + \phi^2(x, v_i)) + \int_{\tilde{J}_i \setminus H_i} 2v_i'w_i' + \int_{H_i} \big((u')^2 - (w_i')^2\big) \\
&\geq \int_{\tilde{J}_i \setminus H_i} (\phi^2(x, v_i) + 2v_i'w_i') + \int_{\tilde{J}_i} \phi^1(x, v_i) + \int_{H_i} \big((u')^2 - (w_i')^2\big).
\end{aligned}$$

Now, by Lemma 2.24 (note this applies by assumption (2.56)) and Lemma 2.25, and since $c_i \geq d_i$,

$$\begin{aligned}
\int_{\tilde{J}_i} \phi^1(x, v_i) + \int_{H_i} ((u')^2 - (w_i')^2) &\geq \int_{\tilde{J}_i} \tilde{\phi}_i^1(x, v_i) + \int_{H_i} ((u')^2 - (w_i')^2) \\
&\geq 5\theta_i \int_0^{g^{-1}(g(c_i)/5)} \kappa(x)\, dx + 2(I_{i,r} - I_{i,l}) - 8\Psi(c_i).
\end{aligned}$$

So combining we have

$$\begin{aligned}
\int_{\tilde{J}_i} \big((u')^2 + \phi(x, v_i) - (w_i')^2\big) \geq 5\theta_i &\int_0^{g^{-1}(g(c_i)/5)} \kappa(x)\, dx + 2(I_{i,r} - I_{i,l}) - 8\Psi(c_i) \\
&+ \int_{\tilde{J}_i \setminus H_i} (\phi^2(x, v_i) + 2v_i'w_i').
\end{aligned} \tag{2.58}$$

Now, for any $x \in [-T_0, T_0]$, write $\mathcal{I}_n(x) = \{j = 0, \ldots, n : x \in Y_j\}$. We show by an easy induction that

$$\sum_{j \in \mathcal{I}_n(x)} \psi_j^2(x) \geq 2|w_n''(x)| + 1 + 2^{-(n-1)} \tag{2.59}$$

for almost every $x \in [-T_0, T_0]$.

For $n = 0$, we have by definition that for all $x \neq x_0$, $\psi_0^2(x) = 3 + 4|w_0''(x)|$ as required. Suppose the result holds for all $0 \leq i \leq n-1$, where $n \geq 1$. Let $i = i(n,x) \leq n$ denote the greatest index in $\mathcal{I}_n(x)$, i.e. the greatest index $i \leq n$ such that $x \in Y_i$. By (2.20.7) we have $w_n''(x) = w_i''(x)$ whenever both sides exist, i.e. almost everywhere. If $x \in (x_i - \tau_i, x_i + \tau_i)$, then $w_i''(x) = \lambda_i \tilde{w}_i''(x)$ by (2.20.1), and by definition, for $x \neq x_i$,

$$\begin{aligned}\sum_{j\in\mathcal{I}_n(x)} \psi_j^2(x) &\geq \psi_i^2(x) \\ &= 3 + 4|\tilde{w}_i''(x)| \\ &\geq 1 + 2^{-(n-1)} + 2|\lambda_i||\tilde{w}_i''(x)| \\ &= 1 + 2^{-(n-1)} + 2|w_i''(x)|\end{aligned}$$

as required. If $x \notin [x_i - \tau_i, x_i + \tau_i]$, then almost everywhere by (2.20.12)

$$|w_i''(x)| \leq |w_{i-1}''(x)| + 2^{-i}$$

so by inductive hypothesis

$$\begin{aligned}\sum_{j\in\mathcal{I}_n(x)} \psi_j^2(x) &\geq \sum_{j\in\mathcal{I}_{i-1}(x)} \psi_j^2(x) \\ &\geq 2|w_{i-1}''(x)| + 1 + 2^{-((i-1)-1)} \\ &\geq 2|w_i''(x)| - 2\cdot 2^{-i} + 1 + 2^{-((i-1)-1)} \\ &= 2|w_i''(x)| + 1 + 2^{-(i-1)} \\ &\geq 2|w_n''(x)| + 1 + 2^{-(n-1)}\end{aligned}$$

as required for (2.59).

Given this, now consider $x \notin \bigcup_{i=0}^n \tilde{J}_i$. Then since $\tilde{J}_i \supseteq J_i$ for all $i \geq 0$, (2.49) gives that $|v_n(x)| \leq 5\tilde{g}_i(x)$ for all $0 \leq i \leq n$. Therefore $\tilde{\phi}_i^2(x, v_n) = |v_n|\psi_i^2(x)$ by definition for $i \in \mathcal{I}_n(x)$. Thus almost everywhere, we have by (2.59) that

$$\begin{aligned}\phi^2(x, v_n(x)) - 2v_n(x)w_n''(x) &\geq \sum_{i\in\mathcal{I}_n(x)} (\tilde{\phi}_i^2(x, v_n(x))) - 2|v_n(x)||w_n''(x)| \\ &= \sum_{i\in\mathcal{I}_n(x)} (\psi_i^2(x)|v_n(x)|) - 2|v_n(x)||w_n''(x)| \\ &= |v_n(x)|\left(\sum_{i\in\mathcal{I}_n(x)} (\psi_i^2(x)) - 2|w_n''(x)|\right) > |v_n(x)|.\end{aligned}$$

Now, let $x \in \tilde{J}_i \backslash H_i$. Note that we must have $i \geq 1$, since $\tau_0 = T_0$. Since $\{\tilde{J}_j\}_{j=0}^n$ are pairwise disjoint, we have that $x \notin \tilde{J}_j$ for $j < i$. Hence, again by (2.49), $|v_i(x)| \leq 5|\tilde{g}_j(x)|$ for all $j < i$, so by definition $\tilde{\phi}_j^2(x, v_i) = \psi_j^2(x)|v_i|$, for $j \in \mathcal{I}_{i-1}(x)$. Since $x \notin H_i$, we have $x \notin [x_i - \tau_i, x_i +, \tau_i]$, and hence that $|w_i''(x)| \leq |w_{i-1}''(x)| + 2^{-i}$ almost everywhere by (2.20.12). Hence by (2.59) we have almost everywhere

$$\begin{aligned}\sum_{j \in \mathcal{I}_{i-1}(x)} \psi_j^2(x) &\geq 1 + 2|w_{i-1}''(x)| + 2^{-(i-2)} \\ &\geq 1 + 2|w_i''(x)| - 2^{-(i-1)} + 2^{-(i-2)} \\ &> 1 + 2|w_i''(x)|,\end{aligned}$$

and so

$$\phi^2(x, v_i) - 2v_i w_i'' \geq \sum_{j \in \mathcal{I}_{i-1}(x)} (\tilde{\phi}_j^2(x, v_i)) - 2|v_i||w_i''| = \sum_{j \in \mathcal{I}_{i-1}(x)} (\psi_j^2(x)|v_i|) - 2|v_i||w_i''| > |v_i|.$$

Thus we have for almost every $x \notin \bigcup_{i=0}^n H_i$, noting the argument on $\tilde{J}_i \backslash H_i$ above applies by (2.57), that

$$\phi^2(x, v_n) - 2v_n w_n'' > |v_n|,$$

and hence

$$\int_{[-T_0, T_0] \backslash \bigcup_{i=0}^n H_i} \left(\phi^2(x, v_n) - 2v_n w_n''\right) \geq \int_{[-T_0, T_0] \backslash \bigcup_{i=0}^n H_i} |v_n|. \tag{2.60}$$

The reason for making this estimate is that we want to integrate $v_n' w_n'$ by parts on $[-T_0, T_0] \backslash \bigcup_{i=0}^n H_i$. Under our standing assumption that $u(x_i) \neq w(x_i)$ for all $i \geq 0$, we see immediately that this is possible, since v_n and w_n' are bounded and absolutely continuous on $[-T_0, T_0] \backslash \bigcup_{i=0}^n H_i$ by (2.20.2), and thus $v_n w_n'$ is absolutely continuous on $[-T_0, T_0] \backslash \bigcup_{i=0}^n H_i$. However, in the general case that $w(x_j) = u(x_j)$ for some $0 \leq j \leq n$, and thus that $w_n(x_j) = u(x_j)$, we have to argue a little more.

We claim that even in this general case the parts formula is still valid on $[-T_0, T_0] \backslash \bigcup_{i=0}^n H_i$; this is the assertion that $v_n w_n'$ can be written as an indefinite integral on $[-T_0, T_0] \backslash \bigcup_{i=0}^n H_i$. The argument of the preceding paragraph gives us that $v_n w_n'$ is absolutely continuous on subintervals bounded away from all x_j with $u(x_j) = w(x_j)$. Fix such an index $0 \leq j \leq n$.

Let $t_j = t_{j,n} = \min\{\sigma_n, \tau_j\}$. By (2.12), and since $\{\sigma_n\}_{n=1}^\infty$ is decreasing, we know $[x_j - \sigma_n, x_j + \sigma_n] \cap Y_m = \emptyset$ for all $j < m \leq n$. So by (2.20.7) and (2.20.1), $w_n = \lambda_j \tilde{w}_j + \rho_j$ on $[x_j - t_j, x_j + t_j]$. It suffices to check that $v_n w_n'$ can be written as an indefinite integral on $(x_j - t_j, x_j + t_j)$. We check that

$$\int_{x_j - t_j}^{x_j} (v_n w_n')'(t)\, dt = -(v_n w_n')(x_j - t_j),$$

the corresponding equality on the right of x_j follows similarly (recall $v_n(x_j) = u(x_j) - w_n(x_j) = 0$).

We know that on those subintervals of $(x_j - t_j, x_j + t_j)$ bounded away from x_j, $v_n w_n'$ is absolutely continuous. We claim that $(v_n w_n')' \in L^1(x_j - t_j, x_j + t_j)$. Given this, we can use the dominated convergence theorem to get the required result as follows.

Since $J_j = \emptyset$, we see by (2.49) that $|v_n(x)| \leq 5|\tilde{g}_j(x)|$ on $(x_j - t_j, x_j + t_j)$. Thus, using (2.3), (2.d), and (2.viii), we see

$$\begin{aligned} |v_n(x) w_n'(x)| &\leq 5|\tilde{g}_j(x)||\lambda_j \tilde{w}_j'(x)| \\ &\leq 10\theta_j |g(x - x_j)|(|g'(x - x_j)| + |g(x - x_j) h'(x - x_j)|) \\ &\leq 10\theta_j |g(x - x_j)|(|g'(x - x_j)||h'(x - x_j)| + |g(x - x_j)(h'(x - x_j))^2|) \\ &\to 0 \text{ as } x \to x_j. \end{aligned}$$

So now, given that the dominated convergence theorem can be applied, we see that

$$\begin{aligned} -(v_n w_n')(x_j - t_j) &= \lim_{\substack{x \to x_j \\ x \neq x_j}} ((v_n w_n')(x) - (v_n w_n')(x_j - t_j)) \\ &= \lim_{\substack{x \to x_j \\ x \neq x_j}} \int_{x_j - t_j}^{x} (v_n w_n')'(t)\, dt \\ &= \int_{x_j - t_j}^{x_j} (v_n w_n')'(t)\, dt. \end{aligned}$$

To see $(v_n w_n')' \in L^1(x_j - t_j, x_j + t_j)$, note that since u is by choice a minimizer for $(\star)$, we have, since $w \in W^{1,2}(-T_0, T_0)$,

$$\int_{-T_0}^{T_0} (u')^2 \leq \mathscr{L}(u) \leq \mathscr{L}(w) = \int_{-T_0}^{T_0} (w')^2 < \infty.$$

Again noting (2.49) still holds, we have, using (2.20.1) and Cauchy-Schwartz, that

$$\begin{aligned} \int_{x_j - t_j}^{x_j + t_j} |(v_n w_n')'| &= \int_{x_j - t_j}^{x_j + t_j} |(v_j w_j')'| \\ &\leq |\lambda_j| \int_{x_j - t_j}^{x_j + t_j} |v_j \tilde{w}_j''| + |\lambda_j| \int_{x_j - t_j}^{x_j + t_j} |v_j' \tilde{w}_j'| \end{aligned}$$

$$\leq 2\int_{x_j-t_j}^{x_j+t_j} |5\tilde{g}_j\tilde{w}_j''| + 2\int_{x_j-t_j}^{x_j+t_j} |u'\tilde{w}_j'| + 4\int_{x_j-t_j}^{x_j+t_j} |\tilde{w}_j'\tilde{w}_j'|$$

$$\leq 2\cdot 5\int_{x_j-t_j}^{x_j+t_j} |\tilde{g}_j\tilde{w}_j''|$$

$$+ 2\left(\int_{x_j-t_j}^{x_j+t_j} |\tilde{w}_j'|^2\right)^{1/2}\left(\left(\int_{x_j-t_j}^{x_j+t_j} |u'|^2\right)^{1/2} + 2\left(\int_{x_j-t_j}^{x_j+t_j} |\tilde{w}_j'|^2\right)^{1/2}\right)$$

$$\leq 10\|\tilde{g}_j\tilde{w}_j''\|_{L^\infty(x_j-t_j,x_j+t_j)}$$

$$+ 2\|\tilde{w}_j'\|_{L^2(x_j-t_j,x_j+t_j)}(\|u'\|_{L^2(x_j-t_j,x_j+t_j)} + 2\|\tilde{w}_j'\|_{L^2(x_j-t_j,x_j+t_j)}).$$

This right hand side is finite by (2.6), (2.4), and the above note.

So, using (2.42), and recalling that $v_n(\pm T_0) = 0$, and using (2.60) (recalling $H_i \subseteq \tilde{J}_i$), we have, integrating by parts as we now know we can do, that

$$\int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} \left(\phi^2(x,v_n) + 2v_n'w_n'\right)$$

$$= 2[v_n w_n']_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} + \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} \left(\phi^2(x,v_n) - 2v_n w_n''\right)$$

$$= -2\sum_{i=0}^n [v_i w_i']_{x_i-d_i}^{x_i+d_i} + \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} \left(\phi^2(x,v_n) - 2v_n w_n''\right)$$

$$\geq -2\sum_{i=0}^n (E_{i,r} - E_{i,l}) + \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} |v_n|. \tag{2.61}$$

So since $\{\tilde{J}_i\}_{i=0}^n$ are pairwise disjoint, we collect all our estimates together and see, using (2.39), (2.57), (2.58), (2.42), (2.61), Lemma 2.26, and properties of κ from (2.xi), that

$$\mathscr{L}_n(u) - \mathscr{L}_n(w_n)$$

$$= \int_{[-T_0,T_0]} \left((u')^2 + \phi(x,v_n) - (w_n')^2\right)$$

$$= \sum_{i=0}^n \int_{\tilde{J}_i} \left((u')^2 + \phi(x,v_i) - (w_i')^2\right) + \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n \tilde{J}_i} \left((u')^2 + \phi(x,v_n) - (w_n')^2\right)$$

$$\geq \sum_{i=0}^n \left(5\theta_i \int_0^{g^{-1}(g(c_i)/5)} \kappa(x)\,dx + 2(I_{i,r} - I_{i,l}) - 8\Psi(c_i) + \int_{\tilde{J}_i\setminus H_i} (\phi^2(x,v_i) + 2v_i'w_i')\right)$$

$$+ \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n \tilde{J}_i} \left(\phi^2(x,v_n) + 2v_n'w_n'\right)$$

$$\geq \sum_{i=0}^n \left(5\theta_i \int_0^{g^{-1}(g(c_i)/5)} \kappa(x)\,dx + 2(I_{i,r} - I_{i,l}) - 8\Psi(c_i)\right)$$

$$
\begin{aligned}
&+\int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} \left(\phi^2(x,v_n)+2v_n'w_n'\right)\\
&\geq \sum_{i=0}^n \left(5\theta_i \int_0^{g^{-1}(g(c_i)/5)} \kappa(x)\,dx + 2((I_{i,r}-I_{i,l})-(E_{i,r}-E_{i,l})) - 8\Psi(c_i)\right)\\
&+\int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} |v_n|\\
&\geq \sum_{i=0}^n \left(5\theta_i \int_0^{g^{-1}(g(c_i)/5)} \kappa(x)\,dx - 2(|I_{i,r}-E_{i,r}|+|I_{i,l}-E_{i,l}|) - 8\Psi(c_i)\right)\\
&+\int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} |v_n|\\
&\geq \sum_{i=0}^n \left(5\theta_i \int_0^{g^{-1}(g(c_i)/5)} \kappa(x)\,dx - 40\theta_i g(c_i)(|c_i f'(c_i)|+|c_i f(c_i)h'(c_i)|) - 8\theta_i\Psi(c_i)\right)\\
&+\int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} |v_n|\\
&\geq \sum_{i=0}^n \theta_i \left(g(c_i)(|c_i f'(c_i)|+|c_i f(c_i)h'(c_i)|)\right) + \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i} |v_n|. \qquad \square
\end{aligned}
$$

Corollary 2.28. Suppose for all $n \geq 0$ our assumptions (2.55) and (2.56) hold. Then

$$
\mathscr{L}(u)-\mathscr{L}(w) \geq \sum_{i=0}^\infty \theta_i \left(g(c_i)(|c_i f'(c_i)|+|c_i f(c_i)h'(c_i)|)\right) + \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^\infty H_i} |v| > 0.
$$

Proof. This follows by the preceding Lemma and the dominated convergence theorem as follows. Writing $\mathbb{1}_X$ for the characteristic function of a set $X \subseteq [-T_0,T_0]$, it is straightforward to see that

$$
\lim_{n\to\infty} \left(|v_n|\mathbb{1}_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i}\right)(x) = \left(|v|\mathbb{1}_{[-T_0,T_0]\setminus\bigcup_{i=0}^\infty H_i}\right)(x)
$$

for all $x \in [-T_0,T_0]$: for $x \in H_k$ for some $k \geq 0$, eventually both sides are 0; for $x \notin \bigcup_{i=0}^\infty H_i$, we see

$$
\begin{aligned}
\left|\mathbb{1}_{[-T_0,T_0]\setminus\bigcup_{i=0}^n H_i}(x)|v_n(x)| - \mathbb{1}_{[-T_0,T_0]\setminus\bigcup_{i=0}^\infty H_i}(x)|v(x)|\right| &= ||v_n(x)|-|v(x)||\\
&\leq |v_n(x)-v(x)|\\
&= |w_n(x)-w(x)|\\
&\to 0 \quad \text{as } n\to\infty.
\end{aligned}
$$

Moreover, since $w_n \to w$ uniformly, we have that

$$\sup_{n\geq 0} \left\| |v_n| \mathbb{1}_{[-T_0,T_0]\setminus\bigcup_{i=0}^{n} H_i} \right\|_\infty \leq \sup_{n\geq 0} \|v_n\|_\infty < \infty.$$

So the dominated convergence theorem implies

$$\begin{aligned}
\lim_{n\to\infty} \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^{n} H_i} |v_n| &= \lim_{n\to\infty} \int_{-T_0}^{T_0} \left(|v| \mathbb{1}_{[-T_0,T_0]\setminus\bigcup_{i=0}^{n} H_i} \right) \\
&= \int_{-T_0}^{T_0} \lim_{n\to\infty} \left(|v_n| \mathbb{1}_{[-T_0,T_0]\setminus\bigcup_{i=0}^{n} H_i} \right) \\
&= \int_{-T_0}^{T_0} \left(|v| \mathbb{1}_{[-T_0,T_0]\setminus\bigcup_{i=0}^{\infty} H_i} \right) \\
&= \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^{\infty} H_i} |v|.
\end{aligned}$$

Lemma 2.22 and (2.11) give that

$$\lim_{n\to\infty} (\mathscr{L}_n(u) - \mathscr{L}_n(w_n)) = \mathscr{L}(u) - \mathscr{L}(w).$$

So since by assumption Lemma 2.27 holds for all $n \geq 0$, we can pass to the limit on each side of the inequality in the conclusion of the Lemma to get the required result.

We note that in the general case we do indeed have strict inequality, as is necessary for the contradiction proof. If $u(x_n) \neq w(x_n)$ for some $n \geq 0$, then $c_n > 0$ and so the infinite sum is strictly positive. If $u(x_n) = w(x_n)$ for all $n \geq 0$, then $[-T_0, T_0]\setminus\bigcup_{i=0}^{\infty} H_i = [-T_0, T_0]$, so on the assumption that $u \neq w$, where both are continuous functions, the integral term must be strictly positive. $\square$

The arguments of the previous lemma and its corollary relied on the intervals we have to give special attention, the $\tilde{J}_j$, being small enough that they did not escape Y_j, or overlap with later Y_k and hence possibly $\tilde{J}_k$. The trick is now that should one of these assumptions fail, thus apparently making the proof more complicated, in fact this means that we can ignore the modifications we made at stage j and beyond. That one of our assumptions fails for j means that $\tilde{J}_j$ is too large, which by the very definition of $\tilde{J}_j$ implies the graph of u is far away from that of w on a set of large measure around x_j. We have chosen our constants so that this large difference between u and w around x_j gives enough weight to our Lagrangian that we can discard all modifications we made to w_{j-1} and hence to L_{j-1} and work just with these instead; the error so incurred is small enough that it is absorbed into this extra weight. Very roughly, if u misses w at x_j by an inconveniently large amount, then we don't have to worry about the fine detail of our variational problem at and beyond the scale j.

Lemma 2.29. Let $n \geq 1$ be such that assumptions (2.55) and (2.56) hold for $n-1$, but for some $0 \leq k < n$ we have $\tilde{J}_k \cap Y_n \neq \emptyset$, i.e. (2.55) fails for n. Then

$$\mathscr{L}_{n-1}(u) - \mathscr{L}_{n-1}(w_{n-1}) \geq T_n^2.$$

Proof. That (2.55) fails for n implies that $c_k \geq T_n$, otherwise choosing $x \in \tilde{J}_k \cap Y_n$ we would have, by (T:1) that

$$|x_n - x_k| \leq |x_n - x| + |x - x_k| \leq T_n + c_k < 2T_n < |x_n - x_k|.$$

So, applying Lemma 2.27 to $n-1$ we see, using this fact, that $\theta_k \geq 1$, (2.a), and (2.d), that

$$\begin{aligned}
\mathscr{L}_{n-1}(u) - \mathscr{L}_{n-1}(w_{n-1}) &\geq \sum_{i=0}^{n-1} \theta_i g(c_i)(|c_i f'(c_i)| + |c_i f(c_i) h'(c_i)|) \\
&\quad + \int_{[-T_0,T_0] \setminus \bigcup_{i=0}^{n-1} H_i} |v_{n-1}| \\
&\geq \theta_k g(c_k)(|c_k f'(c_k)| + |c_k f(c_k) h'(c_k)|) \\
&\geq (g(c_k))^2 |h'(c_k)| \\
&\geq c_k^2 \\
&\geq T_n^2.
\end{aligned}$$

□

Lemma 2.30. Let $n \geq 1$ be such that assumption (2.55) holds for n, assumption (2.56) holds for $n-1$, but $\tilde{J}_n \not\subseteq Y_n$, i.e. (2.56) fails for n. Then

$$\mathscr{L}_{n-1}(u) - \mathscr{L}_{n-1}(w_{n-1}) \geq T_n^2/2.$$

Proof. We suppose that $c_n = b_n$. The case $a_n > b_n$ differs only in trivial notation. That (2.56) fails for n implies that $b_n \geq T_n$. That (2.55) holds for n implies in particular that $Y_n \cap \bigcup_{i=0}^{n-1} \tilde{J}_i = \emptyset$. Thus by Lemma 2.27 for $n-1$, since $H_i \subseteq \tilde{J}_i$ by definition,

$$\mathscr{L}_{n-1}(u) - \mathscr{L}_{n-1}(w_{n-1}) \geq \sum_{i=0}^{n-1} \left(\theta_i g(c_i)(|c_i f'(c_i)| + |c_i f(c_i) h'(c_i)|)\right)$$

$$
\begin{aligned}
&+\int_{[-T_0,T_0]\setminus\bigcup_{i=0}^{n-1} H_i} |v_{n-1}| \\
&\geq \int_{[-T_0,T_0]\setminus\bigcup_{i=0}^{n-1} \tilde{J}_i} |v_{n-1}| \\
&\geq \int_{Y_n} |v_{n-1}| \\
&\geq \int_{x_n}^{x_n+T_n} |v_{n-1}|.
\end{aligned}
$$

But, using (2.45), that $b_n \geq T_n$, and (2.iv), and also using (2.20.8) and (R:3), we know that for $x \in [x_n, x_n + b_n]$ we have

$$
\begin{aligned}
|v_{n-1}(x)| &\geq |v_n(x)| - |w_n(x) - w_{n-1}(x)| \\
&\geq \theta_n g(b_n) - \|w_n - w_{n-1}\|_\infty \\
&> g(T_n) - 6K_n g(R_n) \\
&\geq g(T_n)/2.
\end{aligned}
$$

So we see, by (2.a), that

$$
\mathscr{L}_{n-1}(u) - \mathscr{L}_{n-1}(w_{n-1}) \geq \int_{x_n}^{x_n+T_n} g(T_n)/2 = T_n g(T_n)/2 \geq T_n^2/2. \qquad \square
$$

We can now conclude our proof that w is the unique minimizer of $(\star)$. Choose the least $n \geq 0$ such that one of our crucial assumptions (2.55) or (2.56) fails. We observe that then $n \geq 1$ necessarily, since certainly $\tilde{J}_0 \subseteq [-T_0, T_0]$. If no such n exists, we are in the case of Corollary 2.28 and we are done.

Suppose $n \geq 1$ is such that (2.55) fails for n. Then we are in the case of Lemma 2.29 and we see by Lemma 2.22 that

$$
\mathscr{L}(u) - \mathscr{L}(w) > \mathscr{L}_{n-1}(u) - \mathscr{L}_{n-1}(w_{n-1}) - \frac{T_n^2}{4} \geq \frac{3T_n^2}{4} > 0.
$$

Suppose $n \geq 0$ is such that (2.55) holds for n but (2.56) fails. Then we are in the case of Lemma 2.30 and we see again by Lemma 2.22 that

$$
\mathscr{L}(u) - \mathscr{L}(w) > \mathscr{L}_{n-1}(u) - \mathscr{L}_{n-1}(w_{n-1}) - \frac{T_n^2}{4} \geq \frac{T_n^2}{4} > 0.
$$

This contradicts the choice of u as a minimizer for $(\star)$, so we know that no minimizer $u \neq w$ exists.

2.2.4 Singularity

The extra oscillations we added in to w_n were small enough in magnitude and far enough from x_n to preserve the behaviour of w as being like that of w_n and hence $\tilde{w}_n$ around x_n. In particular, non-differentiability of $\tilde{w}$ at 0 implies non-differentiability of w at x_n for each $n \geq 0$.

Proposition 2.31. Let $n \geq 0$. Then

$$2g'(0) \geq \overline{D}w(x_n) \geq g'(0)$$

and

$$-2g'(0) \leq \underline{D}w(x_n) \leq -g'(0).$$

In particular, $w'(x_n)$ exists if and only if $g'(0) = 0$. Thus, since under the assumptions that g is concave and strictly increasing on $[0, T_0]$, we know w is not differentiable at x_n.

Proof. Let $x \in [-T_0, T_0]$, and let $m > n$. Note that if $x \in Y_i$ for $i > n$, we have by (T:1)

$$|x_n - x_i| \leq |x_n - x| + |x - x_i| \leq |x_n - x| + T_i < |x_n - x| + |x_n - x_i|/2$$

and hence, again by condition (T:1),

$$T_i < |x_n - x_i|/2 < |x_n - x|. \tag{2.62}$$

Now let $x \in [-T_0, T_0]$ be such that $|x - x_n| < T_m$. Then for $n < i \leq m$, again by (T:1) and since the T_i are decreasing,

$$|x - x_i| \geq |x_i - x_n| - |x - x_n| > 2T_i - T_m \geq 2T_i - T_i = T_i,$$

so $x \notin Y_i$ for all $n < i \leq m$. If $x \notin Y_i$ for any $i > n$ then $w(x) = w_n(x)$ by (2.20.7), and the following argument is trivial. Otherwise choose least $i > n$ such that $x \in Y_i$, so $w_n(x) = w_{i-1}(x)$. Then by the above argument we must have $i > m$, and so by (2.21.2), (R:3), and (2.62),

$$|w(x)-w_n(x)| = |w(x)-w_{i-1}(x)| \leq \|w-w_{i-1}\|_\infty \leq 12K_i g(R_i) < 2^{-i}\,T_i < 2^{-i}|x-x_n|.$$

Hence we have by (2.21.1), and since $i > m$,

$$\left| \frac{w(x) - w(x_n)}{x - x_n} - \frac{w_n(x) - w_n(x_n)}{x - x_n} \right| = \left| \frac{w(x) - w_n(x)}{x - x_n} \right| \leq \frac{2^{-i}|x - x_n|}{|x - x_n|} < 2^{-m}.$$

As $x \to x_n$, we may choose $m \to \infty$. Hence by (2.20.1) and definition of $\tilde{w}_n$,

$$\overline{D}w(x) = \lambda_n \overline{D}\tilde{w}_n(x_n) = \lambda_n g'(0)$$

and

$$\underline{D}w(x) = \lambda_n \underline{D}\tilde{w}_n(x_n) = -\lambda_n g'(0).$$

Since $1 \leq \lambda_n < 2$, we get the result. □

2.2.5 Conclusion

We can now obtain the precise statement of Theorem 2.16.

Proof of Theorem 2.16. Let our sequence $\{x_n\}_{n=0}^{\infty}$ be an enumeration of the rationals in $(-T_0, T_0)$. Define

$$\mathcal{N} = \{x \in (-T_0, T_0) : \overline{D}w(x) \geq g'(0) \text{ and } \underline{D}w(x) \leq -g'(0)\}.$$

Then density of $\mathcal{N}$ is immediate by Proposition 2.31. Since $g'(0) \neq 0$, it is G_δ: $\mathcal{N} = \bigcap_{k=1}^{\infty}(\mathcal{N}_k^+ \cap \mathcal{N}_k^-)$ where

$$\mathcal{N}_k^+ = \Bigg\{ x \in [-T_0, T_0] : \frac{w(t) - w(x)}{t - x} > g'(0) - 1/k \quad \text{for some } t \in [-T_0, T_0] \text{ such that } |t - x| < 1/k \Bigg\}$$

and

$$\mathcal{N}_k^- = \Bigg\{ x \in [-T_0, T_0] : \frac{w(t) - w(x)}{t - x} < -g'(0) + 1/k \quad \text{for some } t \in [-T_0, T_0] \text{ such that } |t - x| < 1/k \Bigg\}$$

are open sets. That $\mathcal{N}$ is therefore second category follows by density and Baire's theorem. □

2.2.6 Non-occurrence of the Lavrentiev phenomenon

Our construction of a problem with continuous Lagrangian and non-differentiable minimizer does *not* exhibit the Lavrentiev phenomenon.

Proposition 2.32. Let $w \in \mathrm{AC}(-T_0, T_0)$ and $\phi \in C([-T_0, T_0] \times \mathbb{R})$ be as constructed above. Then there exists a sequence $\{u_n\}_{n=1}^{\infty}$ of admissible functions $u_n \in C^1(-T_0, T_0)$ such that

$$|\mathscr{L}(u_n) - \mathscr{L}(w)| \to 0 \quad \text{as } n \to \infty.$$

Proof. Let $n \geq 1$. First we note that, as argued on page 66, by construction $w_n = \lambda_i \tilde{w}_i + \rho_i$ on $[x_i - \sigma_n, x_i + \sigma_n] \cap [x_i - \tau_i, x_i + \tau_i]$. Choose $\epsilon_n > 0$ such that

$$\epsilon_n < \frac{1}{2} \min \left\{ \min_{i=0,\ldots n} \tau_i, \ \sigma_n, \ 1/5n(n+1) \right\}.$$

and define $A_{n,i} := [x_i - 2\epsilon_n, x_i + 2\epsilon_n]$. Then

- $w_n = \lambda_i \tilde{w}_i + \rho_i$ on $A_{n,i}$ for all $0 \leq i \leq n$;
- $$\int_{A_{n,i}} |w_n'(x)|^2 \, dx < 1/974n(n+1)$$
  for all $0 \leq i \leq n$ by (R:1); and
- the intervals $\{A_{n,i}\}_{i=0}^{n}$ are pairwise disjoint, since if $x \in A_{n,i} \cap A_{n,j}$ for some $0 \leq i < j \leq n$, then
  $$|x_i - x_j| \leq |x_i - x| + |x - x_j| \leq 4\epsilon_n < 2\sigma_n,$$
  whereas by choice of σ_n, and since $\{\sigma_n\}_{n=1}^{\infty}$ is decreasing,
  $$|x_i - x_j| \geq 2\sigma_j \geq 2\sigma_n.$$

Let $\tilde{A}_{n,i} := [x_i - \epsilon_n, x_i + \epsilon_n]$. Let $u_{n,i} \colon [-T_0, T_0] \to \mathbb{R}$ denote the affine function with graph connecting the points $(x_i - \epsilon_n, w_n(x_i - \epsilon_n))$ and $(x_i + \epsilon_n, w_n(x_i + \epsilon_n))$.

Fix $x, y \in A_{n,i}$ such that $x < y$. Then, using Cauchy-Schwartz,

$$\begin{aligned}
|w_n(x) - w_n(y)| &\leq \int_y^x |w_n'(t)|\,dt \\
&\leq (x-y)^{1/2} \left(\int_y^x |w_n'(t)|^2\,dt \right)^{1/2} \\
&\leq (4\epsilon_n)^{1/2} \left(\int_{x_i - 2\epsilon_n}^{x_i + 2\epsilon_n} |w_n'(t)|^2\,dt \right)^{1/2} \\
&= 2\epsilon_n^{1/2} \|w_n'\|_{L^2(A_{n,i})}. \qquad (2.63)
\end{aligned}$$

In particular

$$|u_{n,i}'| = \frac{|w_n(x_i + \epsilon_n) - w_n(x_i - \epsilon_n)|}{2\epsilon_n} \leq \epsilon_n^{-1/2} \|w_n'\|_{L^2(A_{n,i})}. \qquad (2.64)$$

Note that w_n' exists and is continuous (in fact Lipschitz) off $\bigcup_{i=0}^n \tilde{A}_{n,i}$ by (2.20.2), since $\{x_i\}_{i=0}^n$ are interior points of this set. For each $0 \leq i \leq n$ choose a cut-off function $\chi_{n,i} \in C^1(-T_0, T_0)$ such that $\chi_{n,i} = 0$ off $A_{n,i}$, $\chi_{n,i} = 1$ on $\tilde{A}_{n,i}$, and $\|\chi_{n,i}'\|_\infty \leq 2/\epsilon_n$. Then we can define $u_n \in C^1(-T_0, T_0)$ by

$$u_n(x) = w_n(x) + \sum_{i=0}^n \chi_{n,i}(x)(u_{n,i}(x) - w_n(x)).$$

Note $u_n(\pm T_0) = w_n(\pm T_0) = w(\pm T_0)$ so u_n are admissible functions in our minimization problem.

If $x \notin \bigcup_{i=1}^n A_{n,i}$, then $u_n(x) = w_n(x)$ and $u_n'(x) = w_n'(x)$.

If $x \in A_{n,i}$ for some $0 \leq i \leq n$, then since the $\{A_{n,i}\}_{i=0}^n$ are pairwise disjoint, we see that

$$\begin{aligned}
u_n'(x) &= w_n'(x) + \chi_{n,i}'(x)(u_{n,i}(x) - w_n(x)) + \chi_{n,i}(x)(u_{n,i}' - w_n'(x)) \\
&= w_n'(x)(1 - \chi_{n,i}(x)) + u_{n,i}'\chi_{n,i}(x) + \chi_{n,i}'(x)(u_{n,i}(x) - w_n(x)).
\end{aligned}$$

Now, by (2.64),

$$|w_n'(x)(1 - \chi_{n,i}(x)) + u_{n,i}'\chi_{n,i}(x))| \leq |w_n'(x)| + \epsilon_n^{-1/2} \|w_n'\|_{L^2(A_{n,i})}$$

and, since by choice of $u_{n,i}$ we have $u_{n,i}(x_i - \epsilon_n) = w_n(x_i - \epsilon_n)$, we can apply (2.63) and (2.64) to see

$$\begin{aligned}|\chi'_{n,i}(x)(u_{n,i}(x)-w_n(x))| &\le 2\epsilon_n^{-1}|u_{n,i}(x)-u_{n,i}(x_i-\epsilon_n)+w_n(x_i-\epsilon_n)-w_n(x)|\\ &\le 2\epsilon_n^{-1}\left(|u'_{n,i}(x-(x_i-\epsilon_n)|+2\epsilon_n^{1/2}\|w'_n\|_{L^2(A_{n,i})}\right)\\ &\le 2\epsilon_n^{-1}\left(\epsilon_n^{-1/2}\|w'_n\|_{L^2(A_{n,i})}(3\epsilon_n)+2\epsilon_n^{1/2}\|w'_n\|_{L^2(A_{n,i})}\right)\\ &\le 10\epsilon_n^{-1/2}\|w'_n\|_{L^2(A_{n,i})}.\end{aligned}$$

So

$$|u'_n(x)| \le |w'_n(x)| + 11\epsilon_n^{-1/2}\|w'_n\|_{L^2(A_{n,i})}.$$

We then get the following important estimate, again using (2.64), and by choice of ϵ_n,

$$\begin{aligned}\int_{A_{n,i}} |(u'_{n,i})^2-(u'_n(x))^2|\,dx &\le \int_{A_{n,i}} |(u'_{n,i})^2|+|(u'_n(x))^2|\,dx\\ &\le \int_{A_{n,i}} \left(\epsilon_n^{-1/2}\|w'_n\|_{L^2(A_{n,i})}\right)^2\\ &\quad+\left(|w'_n(x)|+11\epsilon_n^{-1/2}\|w'_n\|_{L^2(A_{n,i})}\right)^2\,dx\\ &\le 4\epsilon_n(\epsilon_n^{-1}\|w'_n\|^2_{L^2(A_{n,i})}(1+242))+\int_{A_{n,i}}2|w'_n(x)|^2\,dx\\ &= 972\|w'_n\|^2_{L^2(A_{n,i})}+2\|w'_n\|^2_{L^2(A_{n,i})}\\ &= 974\|w'_n\|^2_{L^2(A_{n,i})}\\ &\le 1/n(n+1). \qquad (2.65)\end{aligned}$$

We now note that the estimate (2.54) about the affine function l_n in the minimization proof also applies to the affine $u_{n,i}$ for each $0 \le i \le n$, since $w_n = \lambda_i\tilde{w}_i+\rho_i$ on $A_{n,i}$. Applying this (recalling that a scalar fact of $|\lambda_i|^2 \le 4$ factorizes out of the expression), and using the continuous function κ from assumption (2.xi), we see that, using also (2.iv),

$$\begin{aligned}\int_{A_{n,i}} |(u'_{n,i})^2-(w'_n)^2| \le 8\Psi(2\epsilon_n) &\le 5\int_0^{g^{-1}(g(2\epsilon_n)/5)}\kappa(x)\,dx\\ &\le 5\int_0^{2\epsilon_n}\kappa(x)\,dx\\ &\le 10\epsilon_n\\ &\le 1/n(n+1),\end{aligned}$$

assuming $n \geq 1$ is large enough such that $0 \leq \kappa(x) \leq 1$ for all $x \in (0, \epsilon_n)$.

Now, since $u_n = w_n$ off $\bigcup_{i=0}^n A_{n,i}$, we have, using (2.39) and this estimate, (2.36), (2.65), and the choice of $A_{n,i}$, that, for large $n \geq 1$,

$$
\begin{aligned}
|\mathscr{L}_n(u_n) - \mathscr{L}_n(w_n)| &\leq \left| \int_{-T_0}^{T_0} \phi(x, u_n - w_n) + (u_n')^2 - (w_n')^2 \right| \\
&\leq \sum_{i=0}^n \int_{A_{n,i}} \left(|\phi(x, u_n - w_n)| + |(u_n')^2 - (u_{n,i}')^2| + |(u_{n,i}')^2 - (w_n')^2| \right) \\
&\leq \sum_{i=0}^n (4C\epsilon_n + 2/n(n+1)) \\
&\leq \sum_{i=0}^n (C+2)/n(n+1) \\
&= (C+2)/n.
\end{aligned}
$$

Using this and the estimates (2.40) and (2.41) from Lemma 2.22 we see, since we know $T_n \to 0$ as $n \to \infty$ by (2.11), that

$$
\begin{aligned}
|\mathscr{L}(u_n) - \mathscr{L}(w)| &\leq |\mathscr{L}(u_n) - \mathscr{L}_n(u_n)| + |\mathscr{L}_n(u_n) - \mathscr{L}_n(w_n)| + |\mathscr{L}_n(w_n) - \mathscr{L}(w)| \\
&\leq \frac{T_{n+1}^2}{8} + \frac{(C+2)}{n} + \frac{T_{n+1}^2}{8} \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}
$$

□

2.2.7 Examples

We prove the two specific theorems 2.14 and 2.15 we referred to at the beginning of this chapter, by applying our general result, Theorem 2.16, to certain functions g and h.

We first observe that in each case we choose a specific value of $T > 0$ for which certain inequalities hold which allow us to prove that the conditions of Theorem 2.16 are satisfied. The results do, however, hold for any given arbitrary $T > 0$, by a simple rescaling argument, which we present in the following lemma.

Lemma 2.33. Let $T_0 > 0$, $\phi \in C([-T_0, T_0] \times \mathbb{R})$, and $w \in \mathrm{AC}(-T_0, T_0)$. Define $L\colon [-T_0, T_0] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by $L(x, y, p) = \phi(x, y - w(x)) + p^2$. Let $T > 0$.

Then there exist $\phi_T \in C([-T, T] \times \mathbb{R})$ and $w_T \in \mathrm{AC}(-T, T)$ such that, defining $L_T\colon [-T, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by $L_T(x, y, p) = \phi_T(x, y - w_T(x)) + p^2$,

- w_T is a minimizer of (2.1) with Lagrangian L_T on $(-T, T)$ if and only if w is a minimizer of (2.1) with Lagrangian L on $(-T_0, T_0)$; and

- a set $\mathcal{N}_T \subseteq [-T, T]$ exists as in the conclusion of Theorem 2.16 for w_T and a given g if and only if such a set $\mathcal{N} \subseteq [-T_0, T_0]$ exists for w and this same g.

Proof. This is a straightforward rescaling argument. Define $\mu = T_0/T$, and $w_T \in \mathrm{AC}(-T, T)$ and $\phi_T \in C([-T, T] \times \mathbb{R})$ by

$$w_T(x) = \mu^{-1} w(\mu x) \text{ and } \phi_T(x, y) = \phi(\mu x, \mu y).$$

So

$$\begin{aligned} L_T(x, y, p) &= \phi_T(x, y - w_T(x)) + p^2 \\ &= \phi(\mu x, \mu y - \mu w_T(x)) + p^2 \\ &= \phi(\mu x, \mu y - w(\mu x)) + p^2, \end{aligned}$$

and $\frac{d}{dx} w_T(x) = \mu^{-1} \mu w'(\mu x) = w'(\mu x)$ by chain rule, and for any $u \in \mathrm{AC}(-T, T)$, $\mu u(\mu^{-1} \cdot)$ defines a function in $\mathrm{AC}(-T_0, T_0)$. The correspondence so defined between $\mathrm{AC}(-T, T)$ and $\mathrm{AC}(-T_0, T_0)$ is evidently a bijection. Moreover $u(\pm T) = w_T(\pm T)$ if and only if $\mu u(\mu^{-1} \cdot) \in \mathrm{AC}(-T_0, T_0)$ has $\mu u(\pm \mu^{-1} T_0) = \mu u(\pm T) = \mu w_T(\pm T) = \mu \mu^{-1} w(\pm \mu T) = w(\pm T_0)$. Let $u \in \mathrm{AC}(-T, T)$. Then since w is a minimizer over $\mathrm{AC}(-T_0, T_0)$, we see that

$$\begin{aligned} \int_{-T}^{T} L_T(x, w_T(x), w_T'(x))\, dx &= \int_{-T}^{T} \phi_T(x, w_T(x) - w_T(x)) + (w_T'(x))^2\, dx \\ &= \int_{-T}^{T} \phi(\mu x, \mu w_T(x) - \mu w_T(x)) + (w_T'(x))^2\, dx \\ &= \mu^{-1} \int_{-T_0}^{T_0} \phi(t, w(t) - w(t)) + (w'(t))^2\, dt \\ &\leq \mu^{-1} \int_{-T_0}^{T_0} \phi(t, \mu u(\mu^{-1} t) - w(t)) + \left(\frac{d}{dt} (\mu u(\mu^{-1} t)) \right)^2 dt \\ &= \int_{-T}^{T} \phi(\mu x, \mu u(x) - \mu w_T(x)) + (u'(x))^2\, dx \\ &= \int_{-T}^{T} \phi_T(x, u(x) - w_T(x)) + (u'(x))^2\, dx \\ &= \int_{-T}^{T} L_T(x, u(x), u'(x))\, dx, \end{aligned}$$

thus w_T is a minimizer over $\mathrm{AC}(-T, T)$.

The reverse implication follows by repeating the same argument with the roles of T_0 and T reversed.

We also note that

$$\frac{w_T(x) - w_T(x_n)}{x - x_n} = \frac{\mu^{-1}(w(\mu x) - w(\mu x_n))}{\mu^{-1}(\mu x - \mu x_n)} = \frac{w(\mu x) - w(\mu x_n)}{\mu x - \mu x_n}$$

and thus the difference quotients of w_T at points x_n behave exactly like the difference quotients of w at μx_n; thus since scaling by a constant non-zero factor does not change the topological properties of a subset of the real line, we see that $\mathcal{N}_T$ is as in the theorem for w_T if and only if $\mathcal{N} = \mu\mathcal{N}_T$ is as in the theorem for w. □

Example 2.34 (Lipschitz minimizer)**.** This example is the one to keep in mind throughout the general construction. In this case many of the estimates involving the derivatives of w_n can be made more easily since we have $\mathrm{Lip}(w_n),\ \mathrm{Lip}(w) \leq 2$. Let $g, h\colon [-e^{-e}, e^{-e}] \to \mathbb{R}$ be given by

$$g(x) = x \text{ for all } x \in [-e^{-e}, e^{-e}], \text{ and } h(x) = \begin{cases} \log\log\log 1/|x| & x \neq 0 \\ 0 & x = 0. \end{cases}$$

So

$$\tilde{w}(x) = \begin{cases} x \sin\log\log\log 1/|x| & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Let $T = e^{-e}/5$ and consider g, h restricted to $[-T, T]$. We check the conditions (2.i)–(2.xi) are satisfied. Conditions (2.i)–(2.vi) are clear. It suffices to compute derivatives for $x \in (0, T)$. Computing the first and second derivatives of h gives

$$\begin{aligned} h'(x) &= \frac{-1}{x(\log\log 1/x)(\log 1/x)}, \text{ and} \\ h''(x) &= \frac{(\log\log 1/x)(\log 1/x) - (\log\log 1/x) - 1}{x^2(\log\log 1/x)^2(\log 1/x)^2}. \end{aligned}$$

Thus $|g(x)h'(x)| = \frac{1}{(\log\log 1/x)(\log 1/x)}$, so (2.vii) follows. Moreover

$$\begin{aligned} &|g(x)| \left(|g'(x)h'(x)| + |g''(x)| + |g(x)(h'(x))^2| + |g(x)h''(x)|\right) \\ &\leq \frac{1}{(\log\log 1/x)(\log 1/x)} \Bigg(1 + \frac{1}{(\log\log 1/x)(\log 1/x)} \\ &\quad + \frac{(\log\log 1/x) + 1 + (\log\log 1/x)(\log 1/x)}{(\log\log 1/x)(\log 1/x)}\Bigg) \\ &\leq \frac{5}{(\log\log 1/x)(\log 1/x)} \end{aligned}$$

hence (2.viii) follows. Condition (2.ix) is clear from the expression for h'. Condition (2.x) follows since $h'' \geq 0$ on $[0,T]$, hence

$$x \mapsto |f'(x)| + |f(x)h'(x)| = |h'(x)| = -h'(x)$$

is decreasing. Condition (2.xi) requires a little calculation. First note that

$$2|f(x)| + |xf'(x)| + |xf(x)h'(x)| \leq 2 + \frac{1}{(\log\log 1/x)(\log 1/x)} \leq 3.$$

Now let $c \in [0,T]$, and note that for $x \in (\frac{c}{\log 1/c}, c)$, we have

$$\begin{aligned} c(|f'(x)| + |f(x)h'(x)|) = \frac{c}{x(\log\log 1/x)(\log 1/x)} &< \frac{c}{x(\log\log 1/c)(\log 1/c)} \\ &\leq \frac{1}{\log\log 1/c}. \end{aligned}$$

We use this to estimate

$$\begin{aligned} \Psi(c) &= \int_0^c (\min\{c(|f'| + |fh'|), 2|f| + |xf'| + |xfh'|\})(2|f| + |xf'| + |xfh'|)\, dx \\ &\leq 3\int_0^c (\min\{c(|f'| + |fh'|), 2|f| + |xf'| + |xfh'|\})\, dx \\ &\leq 3\int_0^{\frac{c}{\log 1/c}} (2|f| + |xf'| + |xfh'|)\, dx + \int_{\frac{c}{\log 1/c}}^c c(|f'| + |fh'|)\, dx \\ &\leq \frac{9c}{\log 1/c} + \int_{\frac{c}{\log 1/c}}^c \frac{1}{\log\log 1/c} \\ &\leq \frac{9c}{\log 1/c} + \frac{c}{\log\log 1/c} \\ &\leq \frac{10c}{\log\log 1/c}. \end{aligned}$$

We also see that

$$41g(c)(|cf'(c)| + |cf(c)h'(c)|) = \frac{41c}{(\log\log 1/c)(\log 1/c)} \leq \frac{41c}{\log\log 1/c}.$$

So if we define $\kappa\colon [-T,T] \to \mathbb{R}$ by

$$\kappa(x) = \begin{cases} \frac{242}{\log\log 1/5|x|} & x \neq 0 \\ 0 & x = 0, \end{cases}$$

we note that $g^{-1}(g(c)/5) = c/5$, and use the fact that κ is concave to estimate the integral as follows:

$$5\int_0^{g^{-1}(g(c)/5)} \kappa(x)\,dx = 5\int_0^{c/5} \frac{242}{\log\log 1/5x}\,dx$$
$$\geq \frac{5}{2}\frac{c}{5}\frac{242}{\log\log 1/c}$$
$$= \frac{121}{\log\log 1/c}.$$

Hence κ is as required.

Our theorem then gives us a Lipschitz minimizer $w\colon [-T_0, T_0] \to \mathbb{R}$ such that for a dense G_δ set $\mathcal{N}$ we have $x \in \mathcal{N}$ implies

$$\overline{D}w(x) \geq 1 \text{ and } \underline{D}w(x) \leq -1.$$

Example 2.35 (Non-Lipschitz minimizer)**.** This example, evidently inspired by the previous one, is introduced just to demonstrate that differentiability can fail as strongly as one might wish it: the upper and lower Dini derivatives are plus and minus infinity at each x_n; hence of course this minimizer is not Lipschitz.

We let $T > 0$ be chosen small enough such that for $x \in [0, T]$,

$$(x)^{1/2} \log\log 1/x \leq 1; \tag{2.66}$$
$$(\log 1/x)^{-1/3} \log\log 1/x \leq 1; \tag{2.67}$$
$$\log 1/x \geq \log\log 1/x \geq 3; \text{ and} \tag{2.68}$$
$$x(\log\log 1/x) \leq (\log\log(625))/125. \tag{2.69}$$

Let $g, h\colon [-T, T] \to \mathbb{R}$ be given by

$$g(x) = \begin{cases} x \log\log 1/|x| & x \neq 0 \\ 0 & x = 0, \end{cases} \quad \text{and } h(x) = \begin{cases} \log\log\log 1/|x| & x \neq 0 \\ 0 & x = 0. \end{cases}$$

So $w(x) = x(\log\log 1/|x|) \sin\log\log\log 1/|x|$ for $x \neq 0$. Again, conditions (2.i)–(2.vi) are clear.

We calculate the derivatives, again only for $x \in (0, T)$:

$$f'(x) = \frac{-1}{x(\log 1/x)}$$
$$g'(x) = (\log\log 1/x) - \frac{1}{\log 1/x}$$

$$\begin{aligned}
g''(x) &= \frac{-1}{x(\log 1/x)}\left(\frac{1}{\log 1/x}+1\right)\\
h'(x) &= \frac{-1}{x(\log\log 1/x)(\log 1/x)}\\
h''(x) &= \frac{(\log\log 1/x)(\log 1/x)-(\log\log 1/x)-1}{x^2(\log\log 1/x)^2(\log 1/x)^2}.
\end{aligned}$$

So

$$|g(x)h'(x)| \le \frac{1}{\log 1/x},$$

and so condition (2.vii) is clear. Calculating with the derivatives, and using the estimate (2.68):

$$\begin{aligned}
&|g'(x)h'(x)|+|g''(x)|+|g(x)(h'(x))^2|+|g(x)h''(x)|\\
&\le \frac{1}{x\log 1/x}+\frac{1}{x(\log 1/x)^2(\log\log 1/x)}+\frac{1}{x(\log 1/x)}\left(\frac{1}{\log 1/x}+1\right)\\
&\quad+\frac{1}{x(\log 1/x)^2(\log\log 1/x)}+\frac{(\log\log 1/x)+(\log 1/x)+1}{x(\log 1/x)^2(\log\log 1/x)}\\
&\le \frac{1}{x(\log 1/x)}\left(1+\frac{1}{(\log 1/x)(\log\log 1/x)}+\frac{1}{\log 1/x}+1\right.\\
&\quad\left.+\frac{1}{(\log 1/x)(\log\log 1/x)}+\frac{(\log\log 1/x)+(\log 1/x)+1}{(\log 1/x)(\log\log 1/x)}\right)\\
&= \frac{1}{x(\log 1/x)}\left(2+\frac{2}{\log 1/x}+\frac{3}{(\log 1/x)(\log\log 1/x)}+\frac{1}{(\log\log 1/x)}\right)\\
&\le \frac{4}{x(\log 1/x)},
\end{aligned}$$

which gives that

$$\begin{aligned}
g(x)\left(|g'(x)h'(x)|+|g''(x)|+|g(x)(h'(x))^2|+|g(x)h''(x)|\right) &\le \frac{4x\log\log 1/x}{x\log 1/x}\\
&\le \frac{4\log\log 1/x}{\log 1/x}\\
&\to 0 \quad \text{as } 0<x\to 0,
\end{aligned}$$

as required for (2.viii).

For condition (2.ix), observe that

$$|xf'(x)|+|xf(x)h'(x)| = \left|\frac{-1}{\log 1/x}\right|+\left|\frac{-x(\log\log 1/x)}{x(\log 1/x)(\log\log 1/x)}\right| = \frac{2}{\log 1/x} \qquad (2.70)$$

and

$$|f'(x)|+|f(x)h'(x)| = \left|\frac{-1}{x(\log 1/x)}\right|+\left|\frac{-(\log\log 1/x)}{x(\log 1/x)(\log\log 1/x)}\right| = \frac{2}{x(\log 1/x)}; \quad (2.71)$$

the former is clearly increasing on $(0, T)$, while the latter we can check is decreasing:

$$\frac{d}{dx}\left(\frac{2}{x(\log 1/x)}\right) = \frac{-2}{(x(\log 1/x))^2}\left((\log 1/x) - 1\right) \leq 0.$$

It just remains to deal with (2.xi). So fix $c \in [0, T]$. Firstly we claim that

$$\int_0^c (\log\log 1/x)^2\, dx \leq 3c(\log\log 1/c)^2. \quad (2.72)$$

This is the result of using integration by parts several times. First note that, using the substitution $y = \log 1/x$, and integrating by parts,

$$\begin{aligned}
\int_0^c (\log\log 1/x)^2\, dx &= \int_0^c (\log(-\log x))^2\, dx \\
&= -\int_{\log 1/c}^{\infty} (\log y)^2 \cdot x\, dy \\
&= -\int_{\log 1/c}^{\infty} (\log y)^2 \cdot e^{-y}\, dy \\
&= -\left([-e^{-y}(\log y)^2]_{\log 1/c}^{\infty} + \int_{\log 1/c}^{\infty} \frac{2(\log y)\cdot e^{-y}}{y}\, dy\right) \\
&\leq c(\log\log 1/c)^2 + \int_{\log 1/c}^{\infty} \frac{2(\log y)\cdot e^{-y}}{y}\, dy.
\end{aligned}$$

Examining the second summand, we use Cauchy-Schwartz, and integration by parts twice more to see

$$\begin{aligned}
&\int_{\log 1/c}^{\infty} \frac{2(\log y)\cdot e^{-y}}{y}\, dy \\
&\leq \left(\int_{\log 1/c}^{\infty} e^{-2y}\, dy\right)^{1/2} \left(\int_{\log 1/c}^{\infty} \frac{(2\log y)^2}{y^2}\, dy\right)^{1/2} \\
&= 2\left(\left[\left(\frac{-e^{-2y}}{2}\right)^{1/2}\right]_{\log 1/c}^{\infty}\right)\left(\left[\frac{-(\log y)^2}{y}\right]_{\log 1/c}^{\infty} - \int_{\log 1/c}^{\infty} \frac{-2\log y}{y^2}\, dy\right)^{1/2} \\
&\leq 2^{1/2} c\left(\frac{(\log\log 1/c)^2}{\log 1/c} - \left(\left[\frac{2\log y}{y}\right]_{\log 1/c}^{\infty} - \int_{\log 1/c}^{\infty} \frac{2}{y^2}\, dy\right)\right)^{1/2}
\end{aligned}$$

$$
\begin{aligned}
&= 2^{1/2}c\left(\frac{(\log\log 1/c)^2}{\log 1/c} - \left(\frac{-2\log\log 1/c}{\log 1/c} - \left[\frac{-2}{y}\right]_{\log 1/c}^{\infty}\right)\right)^{1/2} \\
&= 2^{1/2}c\left(\frac{(\log\log 1/c)^2}{\log 1/c} + \frac{2\log\log 1/c}{\log 1/c} + \frac{2}{\log 1/c}\right)^{1/2} \\
&\leq 2^{1/2}c\left(\frac{2(\log\log 1/c)^2}{\log 1/c}\right)^{1/2} \\
&= \frac{2c(\log\log 1/c)}{(\log 1/c)^{1/2}}.
\end{aligned}
$$

Combining with the original expression, we have

$$
\begin{aligned}
\int_0^c (\log\log 1/x)^2\,dx &\leq c(\log\log 1/c)\left((\log\log 1/c) + \frac{2}{(\log 1/c)^{1/2}}\right) \\
&\leq 3c(\log\log 1/c)^2,
\end{aligned}
$$

as claimed.

Now, let

$$
\gamma(c) = \frac{c}{(\log 1/c)^{2/3}}, \tag{2.73}
$$

so by (2.68) $\gamma(c) \leq c$. For $x \in [\gamma(c), c]$, we then have by definition of $\gamma(c)$,

$$
\frac{c}{x\log 1/x} \leq \frac{c}{\gamma(c)(\log 1/c)} = \frac{(\log 1/c)^{2/3}}{\log 1/c} = \frac{1}{(\log 1/c)^{1/3}}. \tag{2.74}
$$

Also note by (2.70) and (2.68) that

$$
2|f(x)| + |xf'(x)| + |xf(x)h'(x)| = 2(\log\log 1/x) + \frac{2}{\log 1/x} \leq 4\log\log 1/x.
$$

We then estimate Ψ in the following way, by splitting the domain of integration:

$$
\begin{aligned}
\Psi(c) &= \int_0^c (\min\{c(|f'| + |fh'|), 2|f| + |xf'| + |xfh'|\})(2|f| + |xf'| + |xfh'|)\,dx \\
&\leq \int_0^{\gamma(c)} (2|f| + |xf'| + |xfh'|)^2\,dx + \int_{\gamma(c)}^c c(|f'| + |fh'|)(2|f| + |xf'| + |fh'|)\,dx.
\end{aligned}
$$

Dealing with the first summand, using the definition of γ and (2.68), we have

$$
\log 1/\gamma(c) = \log\left(\frac{(\log 1/c)^{2/3}}{c}\right) = \frac{2}{3}\log\log 1/c + \log 1/c \leq 2\log 1/c \leq (\log 1/c)^2,
$$

and so

$$\log\log 1/\gamma(c) \le \log(\log 1/c)^2 = 2\log\log 1/c. \tag{2.75}$$

Then, using (2.72), the definition of γ, and (2.67), we see

$$\begin{aligned}
\int_0^{\gamma(c)} (2|f| + |xf'| + |xfh'|)^2\,dx &\le 16\int_0^{\gamma(c)} (\log\log 1/x)^2 \\
&\le 48\gamma(c)(\log\log 1/\gamma(c))^2 \\
&\le 192\frac{c(\log\log 1/c)^2}{(\log 1/c)^{2/3}} \\
&= 192c(\log\log 1/c)\frac{\log\log 1/c}{(\log 1/c)^{2/3}} \\
&\le 192g(c)(\log 1/c)^{-1/3}.
\end{aligned}$$

We use Cauchy-Schwartz on the second summand to see, using (2.71), (2.74), and (2.72), that

$$\begin{aligned}
&\int_{\gamma(c)}^{c} (c|f'| + |fh'|)(2|f| + |xf'| + |fh'|)\,dx \\
&\le \left(\int_{\gamma(c)}^{c} (c(|f'| + |fh'|))^2\,dx\right)^{1/2} \left(\int_{\gamma(c)}^{c} (2|f| + |xf'| + |fh'|)^2\,dx\right)^{1/2} \\
&\le \left(\int_{\gamma(c)}^{c} \left(\frac{c}{x\log 1/x}\right)^2 dx\right)^{1/2} \left(\int_{\gamma(c)}^{c} (4\log\log 1/x)^2\,dx\right)^{1/2} \\
&\le \left(\int_{\gamma(c)}^{c} \left(\frac{1}{(\log 1/c)^{1/3}}\right)^2 dx\right)^{1/2} (4c^{1/2}3^{1/2}(\log\log 1/c)) \\
&\le \frac{c^{1/2}}{(\log 1/c)^{1/3}} \cdot (8c^{1/2}(\log\log 1/c)) \\
&= \frac{8g(c)}{(\log 1/c)^{1/3}}.
\end{aligned}$$

So combining we see that

$$\Psi(c) \le \frac{192g(c)}{(\log 1/c)^{1/3}} + \frac{8g(c)}{(\log 1/c)^{1/3}} = \frac{200g(c)}{(\log 1/c)^{1/3}}. \tag{2.76}$$

So if we define $\kappa\colon (0,T) \to \mathbb{R}$ by

$$\kappa(x) = \frac{(546,448\log\log 1/x)}{(\log 1/x)^{1/3}},$$

we see κ is continuous and $\kappa(x) \to 0$ as $0 < x \to 0$. Note also that this function is concave on $(0, T)$:

$$\frac{d}{dx}\left(\frac{\log\log 1/x}{(\log 1/x)^{1/3}}\right) = \frac{1}{(\log 1/x)^{2/3}}\left(\frac{-1}{x(\log 1/x)^{2/3}} + \frac{\log\log 1/x}{3x(\log 1/x)^{2/3}}\right)$$
$$= \frac{(\log\log 1/x) - 3}{3x(\log 1/x)^{4/3}}$$

and so

$$\frac{d^2}{dx^2}\left(\frac{\log\log 1/x}{(\log 1/x)^{1/3}}\right) = \frac{1}{3x^2(\log 1/x)^{8/3}}\Big(- 3(\log 1/x)^{1/3}$$
$$- ((\log\log 1/x) - 3)\left((\log 1/x)^{4/3} - 4(\log 1/x)^{1/3}/3\right)\Big)$$
$$= \frac{-1}{9x^2(\log 1/x)^{7/3}}\left(9 + ((\log\log 1/x) - 3)(3\log 1/x - 4)\right),$$

which is negative on $(0, T)$ since $\log\log 1/x \geq 3$.

For our fixed $c \in [0, T_0]$, we note that if $x^{1/2} \leq c/5$, then by (2.66) and (2.68) we have

$$g(x) = x\log\log 1/x \leq x^{1/2} \leq c/5 \leq (c\log\log 1/c)/5 = g(c)/5,$$

hence we have lower bound $g^{-1}(g(c)/5) \geq c^2/25$, and thus inequality

$$\log 1/c \geq \log 1/(25g^{-1}(g(c)/5))^{1/2} = (\log 1/25g^{-1}(g(c)/5))/2.$$

Observe that our domain $[-T, T]$ is small enough to ensure $(g^{-1}(g(c)/5))^{1/2} \leq 1/25$: condition (2.69) implies that $g(c)/5 < g(1/625)$. So we have, multiplying by $25(g^{-1}(g(c)/5))^{1/2}$ that

$$25g^{-1}(g(c)/5) \leq (g^{-1}(g(c)/5))^{1/2}$$

and hence that

$$1/(25g^{-1}(g(c)/5)) \geq (1/g^{-1}(g(c)/5))^{1/2}.$$

Therefore

$$\log 1/c \geq (\log(1/g^{-1}(g(c)/5))^{1/2})/2 = (\log 1/g^{-1}(g(c)/5))/4,$$

the ultimate point being that

$$\frac{1}{(\log 1/c)^{1/3}} \leq \frac{4^{1/3}}{(\log 1/(g^{-1}(g(c)/5)))^{1/3}} \leq \frac{4}{(\log 1/(g^{-1}(g(c)/5)))^{1/3}}.$$

So, estimating by the triangle under the graph, since κ is concave, we see that

$$\begin{aligned}\int_0^{g^{-1}(g(c)/5)} \kappa(x)\,dx &= \int_0^{g^{-1}(g(c)/5)} \frac{525,456 \log\log 1/x}{(\log 1/x)^{1/3}}\,dx \\ &\geq \frac{525,456 g^{-1}(g(c)/5) \log\log 1/(g^{-1}(g(c)/5))}{2(\log 1/(g^{-1}(g(c)/5)))^{1/3}} \\ &= \frac{525,456 g(g^{-1}(g(c)/5))}{2(\log 1/(g^{-1}(g(c)/5)))^{1/3}} \\ &= \frac{525,456 c \log\log 1/c}{10(\log 1/(g^{-1}(g(c)/5)))^{1/3}} \\ &\geq \frac{525,456 c \log\log 1/c}{40(\log 1/c)^{1/3}} \\ &= \frac{65,682 c \log\log 1/c}{5(\log 1/c)^{1/3}}.\end{aligned}$$

So, recalling (2.70) and (2.76), we have

$$\begin{aligned}41g(c)(|cf'(c)| + |cf(c)h'(c)|) + 8\Psi(c) &\leq 41(c\log\log 1/c)\left(\frac{2}{\log 1/c} + 8\cdot\frac{200g(c)}{(\log 1/c)^{1/3}}\right) \\ &\leq \frac{41(2 + 8\cdot 200)c\log\log 1/c}{(\log 1/c)^{1/3}} \\ &= \frac{65,682c\log\log 1/c}{(\log 1/c)^{1/3}} \\ &\leq 5\int_0^{g^{-1}(g(c/5))} \kappa(x)\,dx,\end{aligned}$$

as required. Our theorem then gives us an absolutely continuous minimizer $w\colon [-T_0, T_0] \to \mathbb{R}$ and dense G_δ set $\mathcal{N}$ such that for $x \in \mathcal{N}$ we have

$$\overline{D}w(x) = +\infty \text{ and } \underline{D}w(x) = -\infty.$$

Chapter 3

The singular set

3.1 Introduction

Having established optimality of the conditions under which Tonelli's partial regularity theorem holds, from now on we shall consider situations in which the theorem does hold. In particular we shall now only be interested in smooth Lagrangians.

To any minimizer $u \in \mathrm{AC}(a, b)$ of a variational problem (1.2) there is associated a subset of the domain $[a, b]$ which records where the minimizer has infinite derivative.

Definition 3.1. Let $u \in \mathrm{AC}(a, b)$ be a minimizer of (1.2). The *singular set* of u, denoted E, is defined to be the set of points where the derivative is infinite in modulus. That is

$$E = \{x \in [a, b] : |u'(x)| = \infty\}.$$

There is no ambiguity in this definition (e.g. with functions equal almost everywhere) since the partial regularity theorem tells us that for smooth Lagrangians the classical derivative of the minimizer, i.e. the limit of difference quotients, exists everywhere.

The partial regularity theorem, Theorem 1.3, tells us that the singular set is closed. Moreover, since $u \in \mathrm{AC}(a, b)$, we immediately know also that it is a null set. Thus u is locally Lipschitz on a relatively open set of (a, b) of full measure. Tonelli apparently had no information about whether anything further can be said about E, assuming no other conditions.

3.1.1 Moving to full regularity

The first instinct is to find out under what circumstances the singular set is empty. Some work has been done, first by Tonelli himself, showing that, when certain extra conditions are imposed, the singular set must indeed by empty, i.e. the minimizer will be fully regular. The most important result in this direction is that if the minimizer is known to be Lipschitz, then necessarily it is smooth [see Ball and Mizel, 1985, Theorem 2.6]. A proof can be found in Cesari [1983].

Theorem 3.2. Let $L\colon [a,b] \times \mathbb{R} \times \mathbb{R} \to [0,\infty)$ be of class C^3 and let $u \in \mathrm{AC}(a,b)$ be a minimizer of (1.2), moreover such that u is Lipschitz. Suppose further that $L_{pp}(x,u(x),p) > 0$ for all $x \in [a,b]$ and $p \in \mathbb{R}$.

Then $u \in C^3([a,b])$ and satisfies the Euler-Lagrange equation.

Clarke and Vinter [1985a] give some conditions for full regularity, for example the following, stated here for smooth Lagrangians (whereas they work in greater generality, using definitions and techniques of nonsmooth analysis). Here and in some following results we recall the conditions (CVH1)–(CVH3) stated in Chapter 2.

Theorem 3.3 (Clarke and Vinter [1985a])**.** Let $L\colon [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be of class C^1 and satisfy the conditions (CVH1)–(CVH3). Let $u \in \mathrm{AC}(a,b)$ be a minimizer of (1.2). Then

- if there exists $\gamma \in L^1(a,b)$ such that

$$\gamma(x) \leq L_x(x,u(x),u'(x))$$

  for almost every $x \in [a,b]$, then $E \subseteq \{a\}$;

- if there exists $\gamma \in L^1(a,b)$ such that

$$\gamma(x) \geq L_x(x,u(x),u'(x))$$

  for almost every $x \in [a,b]$, then $E \subseteq \{b\}$; and

- if there exists $\gamma \in L^1(a,b)$ such that

$$\gamma(x) \geq |L_x(x,u(x),u'(x))|$$

  for almost every $x \in [a,b]$, then $E = \emptyset$.

Morrey [2008] gives the following criterion, in terms of integrability of the other derivatives of L. It applies also to minimization problems dealing with vector-valued functions $u\colon [a,b] \to \mathbb{R}^N$ for $N \geq 1$.

Theorem 3.4 (Morrey [2008])**.** Let $L\colon [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be of class C^2 and such that for some $m > 1$

- $L(x, y, p) \geq c_1|p|^m - c_2$ for some constants $c_1, c_2 > 0$; and
- there exists $M\colon [0, \infty) \to [0, \infty)$ such that

$$|L_y(x, y, p)|,\ |L_p(x, y, p)| \leq M(R)(1 + |p|^m)$$

  whenever $|x^2| + |y^2| \leq R^2$.

Let $u \in \mathrm{AC}(a, b)$ be a minimizer of (1.2).

Then $u \in C^2([a, b])$ and u satisfies the Euler-Lagrange equation.

Clarke and Vinter [1985a] give a more general version of this result, which again applies also in the vector-valued case. Tonelli [1923] was the first to realize that for scalar-valued functions the integrability of L_p could be discarded.

Theorem 3.5. Let $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be of class C^1 and satisfy the conditions (CVH1)–(CVH3). Let $u \in \mathrm{AC}(a, b)$ be a minimizer of (1.2). Suppose there exists $\gamma \in L^1(a, b)$ such that

$$|L_y(x, u(x), u'(x))| \leq |L_p(x, u(x), u'(x))| + \gamma(x)$$

for almost every $x \in [a, b]$.

Then $E = \emptyset$.

These results can be roughly summarized by the following theorem of Ball and Mizel.

Theorem 3.6 (Ball and Mizel [1985])**.** Let $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$ be of class C^3, superlinear in p for each fixed $(x, y) \in [a, b] \times \mathbb{R}$, and satisfy $L_{pp} > 0$. Let $u \in \mathrm{AC}(a, b)$ be a minimizer for (1.2). Suppose further that either $L_y(\cdot, u(\cdot), u'(\cdot)) \in L^1(a, b)$ or $L_x(\cdot, u(\cdot), u'(\cdot)) \in L^1(a, b)$.

Then $u \in C^3([a, b])$ and satisfies the Euler-Lagrange equation on $[a, b]$.

This leads to full regularity in the autonomous case, i.e. when $L = L(y, p)$.

Theorem 3.7 (Ball and Mizel [1985])**.** Let $L = L(y, p)\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be of class C^3, superlinear in p for each fixed $y \in \mathbb{R}$, and satisfy $L_{pp} > 0$. Let $u \in \mathrm{AC}(a, b)$ be a minimizer for (1.2).

Then $u \in C^3([a, b])$ and u satisfies the Euler-Lagrange equation on $[a, b]$.

Finally we mention the work of Bernstein [1912] on solvability of the Euler-Lagrange equation, which uses a growth condition on the function one gets by differentiating the usual expression of the Euler-Lagrange equation. Tonelli [1923] was again the first to apply this to deducing full regularity of scalar-valued minimizers. We give here a statement of Clarke and Vinter, which again in fact holds for vector-valued functions.

Theorem 3.8 (Clarke and Vinter [1985a])**.** Let $L\colon [a,b]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ be of class C^2, and satisfy the conditions (CVH1)–(CVH3). Let $u\in \mathrm{AC}(a,b)$ be a minimizer of (1.2), such that $L_{pp}(x,u(x),u'(x))>0$ for almost every $x\in[a,b]$. Suppose there exists $\gamma\in L^1(a,b)$ such that

$$\left|\left(\frac{L_y-L_{px}-L_{py}u'}{L_{pp}}\right)\bigg|_{(x,u(x),u'(x))}\right|\le\gamma(x)(|u'(x)|+1)$$

for almost every $x\in[a,b]$.

Then $E=\emptyset$.

3.1.2 Non-empty singular set

That minimizers of variational problems can have infinite derivatives has been known since the paper of Lavrentiev [1926], which presented the celebrated *Lavrentiev phenomenon*, whereby when restricting the above minimization problem to even a dense subclass of the absolutely continuous functions (e.g. C^1 functions), the minimum value is *strictly larger* than that minimum value taken over all absolutely continuous functions. Manià [1934] gave an example of a polynomial Lagrangian which exhibits the same phenomenon. In such examples, the minimizer over the absolutely continuous functions has non-empty singular set E; Manià's example has minimizer $x^{1/3}$ over domain $[0,1]$, thus $E=\{0\}$. However, these examples do not satisfy the precise assumptions of the Tonelli partial regularity theorem, since the condition $L_{pp}>0$ on the Lagrangian L is violated (both the Lavrentiev and Manià examples have only $L_{pp}\ge 0$). Thus the question of whether under the exact original conditions of the theorem, the set E can be non-empty, is not answered by these examples.

Many examples of the failure of full regularity also violate conditions traditionally regarded as necessary conditions for minimizers, for example the Euler-Lagrange equation mentioned in Chapter 1. Assumptions beyond those required for existence are required to derive these so-called necessary conditions. This fact was over-looked for a long time because derivation of necessary conditions was motivated by the search for *smooth* minimizers, and so worries over failure of regularity were not entertained.

Ball and Mizel [1984, 1985] were the first to give a comprehensive examination of this situation, giving examples of smooth Lagrangians satisfying all the conditions for partial regularity, with minimizers exhibiting a lack of full regularity in a number of senses. Minimizers are given which variously have non-empty singular set, fail to satisfy versions of classical necessary conditions, and exhibit the Lavrentiev phenomenon.

We record here the versions of the Euler-Lagrange equation which are referred to in Ball and Mizel [1985].

Definition 3.9 (Euler-Lagrange Equation)**.** Let $L\colon [a,b]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ be of class C^1. Function $u \in \mathrm{AC}(a,b)$ is a solution of the Euler-Lagrange equation if

$$\frac{d}{dx}L_p(x,u(x),u'(x)) = L_y(x,u(x),u'(x)) \tag{EL}$$

for all $x\in[a,b]$.

Definition 3.10 (Weak Euler-Lagrange Equation)**.** Let $L\colon [a,b]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$. Function $u\in\mathrm{AC}(a,b)$ satisfies the weak form of the Euler-Lagrange equation if $L_y(\cdot,u(\cdot),u'(\cdot)),\ L_p(\cdot,u(\cdot),u'(\cdot))\in L^1_{\text{loc}}(a,b)$ and the Euler-Lagrange equation holds in the sense of distributions, i.e.

$$\int_a^b L_p(x,u(x),u'(x))\phi'(x) + L_y(x,u(x),u'(x))\phi(x)\,dx = 0 \tag{WEL}$$

for all test functions $\phi\in C_0^\infty(a,b)$.

Definition 3.11 (Integrated Euler-Lagrange Equation)**.** Let $L\colon [a,b]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$. Function $u\in\mathrm{AC}(a,b)$ satisfies the integrated form of the Euler-Lagrange equation if $L_y(\cdot,u(\cdot),u'(\cdot))\in L^1(a,b)$ and

$$L_p(x,u(x),u'(x)) = \int_a^x L_y(t,u(t),u'(t))\,dt + c \tag{IEL}$$

for some constant $c\in\mathbb{R}$, for almost every $x\in[a,b]$.

This last version is strictly stronger than the statement (EL), since in general it is not possible to integrate the equation (EL). The argument to derive (IEL) needs extra conditions imposed on the derivatives of L. These are in fact necessary: Ball and Mizel modify the example of Manià to show that $L\colon[0,1]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ defined by

$$L(x,y,p) = (x^2-y^3)^2p^{14}+\epsilon p^2$$

is such that for certain values of k and ϵ,

- problem (1.2) has minimizer $u = u_{\epsilon,k} \in \mathrm{AC}(a,b)$ when the boundary conditions are given by $A = 0$ and $B = k$ where $L_y(\cdot, u(\cdot), u'(\cdot)) \notin L^1(0,1)$ and so (IEL) does not hold; and
- no smooth solution of the Euler-Lagrange can satisfy the boundary conditions; but of course a minimizer does exist, thus the minimizer does not satisfy the Euler-Lagrange equation.

In particular, L satisfies the partial regularity theorem, i.e. $L_{pp} > 0$, and, since in fact $u(x) = \bar{k}x^{2/3}$ for some constant $\bar{k}$, we have $E = \{0\}$, an endpoint of the domain. Clarke and Vinter [1984] give an alternative analysis of this example.

Ball and Mizel give another example with a singular set comprising an interior point of the domain, and moreover where the Lavrentiev phenomenon occurs. They show that the function $L\colon [-1,1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by, for $\epsilon > 0$ and $s > 3$,

$$L(x,y,p) = (x^4 - y^6)^2|p|^s + \epsilon p^2$$

is such that for a certain choice of boundary conditions $u(-1) = A$ and $u(1) = B$ and $\epsilon > 0$, each minimizer has singular set $E = \{0\}$, fails to satisfy the Euler-Lagrange equation in weak or integrated form, and moreover the problem exhibits the Lavrentiev phenomenon:

$$\inf\left\{\mathscr{L}(u) : u \in W^{1,q}(-1,1) \cap \mathscr{A}_{A,B}\right\} > \inf\left\{\mathscr{L}(u) : u \in \mathrm{AC}(-1,1) \cap \mathscr{A}_{A,B}\right\}$$

for all $3 \le q \le \infty$. Note that when s is an even integer, L is a polynomial.

We recall from Theorem 3.7 that in the autonomous case, full regularity seemed easier to achieve, with a superlinearity condition in p enforced. However, the next results shows that failure of the superlinear growth condition just at one value of $y \in \mathbb{R}$ suffices to allow a failure of regularity.

Theorem 3.12 (Ball and Mizel [1985])**.** There exists $L \in C^\infty(\mathbb{R}^2)$, $L = L(y,p)$ with $L_{pp} > 0$ and superlinear growth in p for each fixed $y \in \mathbb{R}\backslash\{0\}$, and a choice of boundary conditions $u(-1) = A$ and $u(1) = B$ such that for the variational problem given by minimizing

$$\mathscr{L}(u) = \int_{-1}^{1} L(u(x), u'(x))\, dx$$

over those $u \in \mathrm{AC}(-1,1)$ with the given boundary conditions, there is a unique minimizer $u \in \mathrm{AC}(a,b)$ such that $E = \{x_0\}$ for some $x_0 \in (-1,1)$ and $L_y(u(\cdot), u'(\cdot)) \notin L^1_{\mathrm{loc}}(-1,1)$ (and so u does not satisfy the Euler-Lagrange equation in integrated or weak form).

Ball and Mizel take this idea further, showing that in the autonomous case, the immediate information about the singular set given by Tonelli is optimal.

Theorem 3.13 (Ball and Mizel [1985]). Let $E \subseteq [-1, 1]$ be a closed null set.

Then there exists $L \in C^{\infty}(\mathbb{R}^2)$, with $L_{pp} > 0$ and superlinear growth in p for each fixed $y \in \mathbb{R}\backslash F$ for some null set F, and choice of boundary conditions $u(-1) = A$ and $u(1) = B$ such that for the variational problem given by minimizing

$$\mathscr{L}(u) = \int_{-1}^{1} L(u(x), u'(x))\, dx$$

over those $u \in \mathrm{AC}(-1, 1)$ with the given boundary conditions, there is a unique minimizer $u \in \mathrm{AC}(a, b)$ which is strictly increasing and has singular set exactly E. Moreover, $L_p(u(\cdot), u'(\cdot)) \notin L^1_{\mathrm{loc}}(-1, 1)$, so u does not satisfy the Euler-Lagrange equation in integrated or weak form.

3.1.3 Further information

Clarke and Vinter [1986] tell us that for polynomial Lagrangians the singular set is understood rather more precisely.

Theorem 3.14 (Clarke and Vinter [1986]). Suppose $L \colon [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is such that $L_{pp} > 0$ and is of form

$$L(x, y, p) = \sum_{i=0}^{n} a_i(x, y) p^i$$

where $a_0(x, y)$ is a non-trivial polynomial in x and y, and $a_1(x, y), \ldots, a_n(x, y)$ are of class C^2. Let $u \in \mathrm{AC}(a, b)$ be a minimizer of (1.2).

Then there is a closed null set $E \subseteq [a, b]$ such that

- u is of class C^2 on $(a, b)\backslash E$ and the Euler-Lagrange equation (EL) holds; and
- the classical derivative u' of u exists and has $|u'| = \infty$ at all points in E.

Moreover, the set E is at most countable and contains only finitely many accumulation points.

3.1.4 Characterization of the singular set

Davie [1988] continued the work of Ball and Mizel, extending their result on the characterization of the singular set in the autonomous case to the full general case, showing that nothing more can be said about E in general other than the immediate information given by Tonelli. Given an arbitrary closed null set E, Davie constructs a C^∞ Lagrangian L, superlinear in p and with $L_{pp} > 0$, such that any minimizer (and at least one minimizer exists by Tonelli's existence result) has singular set precisely E.

Theorem 3.15 (Davie [1988])**.** Let $E \subseteq [a, b]$ be a closed set of measure zero.

Then there exist admissible function $v \in \mathscr{A}_{0,1}$, smooth functions $\phi, \psi \in C^\infty(\mathbb{R})$, and $\epsilon > 0$, such that $\psi \geq 0$, $\psi'' \geq 0$, $\psi \circ v \in C^\infty([a, b])$, and defining $L\colon [a, b] \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$ by

$$L(x, y, p) = (\phi(y) - \phi(v(x)))^2 \psi(p) + \epsilon p^2$$

gives a variational problem (1.2) such that any minimizer u over $\mathscr{A}_{0,1}$ has singular set exactly E.

In the same paper Davie proves stronger characterization results (which we shall not record), including the prescribing of the singular minimizer and the p-derivative of the Lagrangian, in the cases where L does not depend on all three variables.

For the proof of Theorem 3.15, Davie constructs an admissible function $v \in \mathrm{AC}(a, b)$ and a Lagrangian L so that there exists a constant (in his notation) $(8\alpha)^{-1} > 0$ such that $\mathscr{L}(v) < (8\alpha)^{-1}$, but for any admissible function $u \in \mathrm{AC}(a, b)$, if for some $c \in E$ we have that $u'(c)$ exists and is finite, then $\mathscr{L}(u) \geq (8\alpha)^{-1}$. Therefore any minimizer (and at least one exists) must have infinite derivative on the set E. Thus the proof rests on the fact that the energy of C^1 functions is bounded away from the infimum of the energy over all $\mathrm{AC}(a, b)$ functions, i.e. that the Lavrentiev phenomenon occurs.

This raises the question of the exact relationship between the singular set and the occurrence of the Lavrentiev phenomenon. If a problem exhibits the Lavrentiev phenomenon, then certainly the singular set of any minimizer over $\mathrm{AC}(a, b)$ must be non-empty, although as discussed the first examples from Lavrentiev and Manià do not satisfy the $L_{pp} > 0$ condition required for classical partial regularity statements. That a minimizer has a non-empty singular set does not, of course, in general imply the occurrence of a Lavrentiev gap. Quite the reverse is in fact the case: one usually has to go to some effort to prove that a Lavrentiev gap does occur. However, it might be conjectured that if a minimizer has a *large* singular set, then a gap must occur. Thus the question is: can one prove Davie's result without inducing a Lavrentiev gap? We show, using the methods which Csörnyei et al. [2008] introduced in the context of universal singular sets (see Chapter 4), that this is indeed possible, i.e.

that the existence of a large singular set does *not* imply occurrence of the Lavrentiev phenomenon. Conversely, knowing that the Lavrentiev phenomenon does not occur does *not* tell us that the minimizer has small singular set.

The methods of Csörnyei et al. also naturally allow us to construct a Lagrangian giving this result which has arbitrary given superlinear growth, so this result is a generalization of Davie's result even without the further result preventing a Lavrentiev gap.

3.2 Full singular set without a Lavrentiev gap

We prove the following theorem.

Theorem 3.16. Let $[a, b]$ be a closed bounded subinterval of the real line, and let $E \subseteq [a, b]$ be closed and Lebesgue null. Let $\omega \in C^\infty(\mathbb{R})$ be strictly convex, such that $\omega(p) \geq \omega(0) = 0$ for all $p \in \mathbb{R}$, and $\omega(p)/|p| \to \infty$ as $|p| \to \infty$ (i.e. ω is superlinear).

Then there exists $L \in C^\infty(\mathbb{R}^3)$, $L = L(x, y, p)$, strictly convex in p and such that $L(x, y, p) \geq \omega(p)$ for all $(x, y, p) \in \mathbb{R}^3$, and function $u \in \mathrm{AC}(a, b)$ such that

- u is the unique minimizer of the functional (1.2) with boundary conditions $A = u(a)$ and $B = u(b)$;
- the singular set of u is precisely E; and
- there exist admissible functions $u_k \in C^\infty([a, b])$ (i.e. $u_k(a) = u(a)$ and $u_k(b) = u(b)$) such that $u_k \to u$ uniformly and $\mathscr{L}(u_k) \to \mathscr{L}(u)$, so the Lavrentiev phenomenon does *not* occur.

We first note that it suffices to prove the result assuming that $E \subseteq (a, b)$, i.e. that E does not contain an endpoint of our domain. This assumption simplifies some technical points in the proof. In the general case, we can expand our domain slightly, and then consider the restriction to the original domain of the function u we construct. This suffices since the restriction to a subinterval of a minimizer is a minimizer of the problem on that subinterval.

For the remainder of the chapter we shall assume $[a, b]$, $\emptyset \neq E \subseteq (a, b)$, and ω are fixed as in Theorem 3.16.

3.2.1 Calibration

Our approach to the construction of minimizers with infinite derivatives is inspired by that in Csörnyei et al. [2008]. We use a calibration argument to prove that functions with a specified derivative are minimizers of (1.2) where the Lagrangian L is constructed via a potential defined on $\mathbb{R}^2$. The original context of this method was the study of universal singular sets, specifically the construction of a Lagrangian with universal singular set containing a certain subset S of the plane. Thus Csörnyei et al. constructed the potential to have singular behaviour at these points S. For each point in S a minimizer was constructed with derivative given via the potential (hence infinite at that point) and graph passing through that point.

We of course need just one minimizer u, but one that has infinite derivative at every point of the set E. Thus it is more natural to begin by defining u (via its derivative), because firstly this is very easy, and secondly this readily gives us a sequence of smooth admissible functions approximating u with which we shall see the Lavrentiev phenomenon does not occur. So we approach the construction of the Lagrangian with the derivative ψ of our intended minimizer already given. In the original construction of Csörnyei et al., the derivative of each intended minimizer was defined to be a function of the gradient of the potential. In order to be able to use this method, we construct the potential so that constant multiples of this same function of the gradient bound ψ from above and below (condition (3.18.4) of Lemma 3.18 below). The argument then proceeds similarly to that of Csörnyei et al., constructing a Lagrangian so that any primitive of ψ is a minimizer with respect to its own boundary conditions.

We first recall Lemma 10 from Csörnyei et al. [2008], stated and used almost as in this original paper, except that later we need also an upper bound of the function, for our (non-)Lavrentiev estimates. We do not repeat the (simple) proof of the other statements.

Lemma 3.17. There exists a C^∞ function $\gamma\colon \{(p,a,b) \in \mathbb{R}^3 : b > 0\} \to \mathbb{R}$ with the following properties:

(3.17.1) $p \mapsto \gamma(p,a,b)$ is convex;

(3.17.2) $\gamma(p,a,b) = 0$ for $p \leq a - 1$;

(3.17.3) $\gamma(p,a,b) = b(p-a)$ for $p \geq a+1$;

(3.17.4) $\gamma(p,a,b) \geq \max\{0, b(p-a)\}$; and

(3.17.5) $\gamma(p,a,b) \leq b|p-a+1|$.

Proof. Recalling the proof from Csörnyei et al. [2008], we see $\gamma(p,a,b) = b\int_\infty^{p-a}\eta$, where non-decreasing $\eta \in C^\infty(\mathbb{R})$ was chosen such that $\eta(x) = 0$ if $x \leq -1$, $\eta(x) = 1$ if $x \geq 1$, and $\int_{-1}^{1}\eta = 1$. The only new statement (3.17.5) is trivial: if $p \leq a-1$ or

$p \geq a+1$ then the result follows by (3.17.2) or (3.17.3) respectively. If $a-1 \leq p \leq a+1$, then

$$\gamma(p,a,b) = b\int_{\infty}^{p-a} \eta(x)\,dx \leq b\int_{-1}^{p-a} 1\,dx \leq b(p-a+1) \leq b|p-a+1|. \qquad \square$$

The next result is a version of Lemma 11 in Csörnyei et al. [2008]. The main difference, as discussed, is that ψ is given before the potential Φ. We recall that for a function $u\colon [a,b] \to \mathbb{R}$, the function $U\colon [a,b] \to \mathbb{R}^2$ is given by $U(x) = (x, u(x))$.

Lemma 3.18. Suppose $\psi \in C^\infty(\mathbb{R}\backslash E)$ is such that $\psi(x) \to \infty$ as $\operatorname{dist}(x,E) \to 0$ and $\omega(\psi(\cdot)) \in L^1(a,b)$, and $\Phi \in C^\infty(\mathbb{R}^2\backslash(E\times\mathbb{R})) \cap C(\mathbb{R}^2)$ satisfies the following conditions:

(3.18.1) Φ is decreasing in x and increasing in y on $\mathbb{R}^2$;

(3.18.2) $-\Phi_x(x,y) \geq 4\Phi_y(x,y) > 0$ for all $(x,y) \in \mathbb{R}^2\backslash(E\times\mathbb{R})$;

(3.18.3) $\Phi_y(x,y) > 320\omega'(\psi(x))$ for all $(x,y) \in \mathbb{R}^2\backslash(E\times\mathbb{R})$;

(3.18.4) $-2(\Phi_x/\Phi_y)(x,y) \leq \psi(x) \leq -160(\Phi_x/\Phi_y)(x,y)$ for all $(x,y) \in \mathbb{R}^2\backslash(E\times\mathbb{R})$; and

(3.18.5) for all non-decreasing $u \in \mathrm{AC}(a,b)$ such that $\omega(u'(\cdot)) \in L^1(a,b)$, the set $(\Phi\circ U)(E)$ is Lebesgue null.

Then there exists a Lagrangian $L \in C^\infty(\mathbb{R}^3)$, strictly convex in p and satisfying $L(x,y,p) \geq \omega(p)$ for all $(x,y,p) \in \mathbb{R}^3$, such that for all $u \in \mathrm{AC}(a,b)$

$$\mathscr{L}(u) = \int_a^b L(x,u(x),u'(x))\,dx \geq \Phi(U(b)) - \Phi(U(a)),$$

with equality if and only if $u' = \psi$ almost everywhere on $[a,b]$. In particular, any such u is the unique minimizer of (1.2) with respect to its boundary conditions.

Proof. This mimics the proof of Lemma 11 in Csörnyei et al. [2008]. Define $\theta, \xi \in C^\infty(\mathbb{R}^2\backslash(E\times\mathbb{R}))$ by

$$\theta(x,y) = \Phi_y(x,y) - \omega'(\psi(x)) \text{ and } \xi(x,y) = \frac{-\Phi_x(x,y) + \omega(\psi(x)) - \omega'(\psi(x))\psi(x)}{\theta(x,y)}.$$

Fix $(x,y) \in \mathbb{R}^2\backslash(E\times\mathbb{R})$. Then note by (3.18.4) and (3.18.2) that $\psi > 0$, so by properties of ω we have that $\omega'(\psi) > 0$. So using also (3.18.3) we have that

$$\theta > \Phi_y - \frac{1}{320}\Phi_y = \frac{319}{320}\Phi_y > 0, \tag{3.1}$$

so certainly ξ is well defined. By convexity of ω we have that $\omega(p)-\omega'(p)p \leq \omega(0) = 0$ for all $p \geq 0$. So, using this and properties (3.18.4) and (3.18.3), we see

$$-\Phi_x \geq -\Phi_x + \omega(\psi) - \omega'(\psi)\psi = \xi\theta \geq -\Phi_x - \omega'(\psi)\psi \geq -\Phi_x + \frac{\Phi_y}{320} \cdot \frac{160\Phi_x}{\Phi_y} = -\Phi_x/2. \tag{3.2}$$

So, since $\Phi_y = \theta + \omega'(\psi) > \theta > 0$, we see by (3.18.2) that

$$\xi \geq -\Phi_x/(2\theta) \geq -\Phi_x/(2\Phi_y) \geq 2, \tag{3.3}$$

and so, using (3.18.4), (3.1), and (3.2),

$$\psi \geq -2\Phi_x/\Phi_y \geq -2 \cdot 319\Phi_x/(320\theta) \geq 3\xi/2 \geq \xi + 1.$$

The point of these estimates, and the choice of constants in the assumptions which allows them to be derived, is that

$$0 \leq \xi - 1 \tag{3.4}$$

and

$$\psi \geq \xi + 1. \tag{3.5}$$

We use the corner-smoothing function γ from Lemma 3.17 to define

$$F(x, y, p) = \begin{cases} \gamma(p, \xi(x, y), \theta(x, y)) & (x, y) \in \mathbb{R}^2 \backslash (E \times \mathbb{R}) \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $F \in C^\infty((\mathbb{R}^2\backslash(E\times\mathbb{R}))\times\mathbb{R})$. Let $(x, y, p) \in E\times\mathbb{R}\times\mathbb{R}$. By (3.3) and (3.18.4) we see that $\xi(x, y) \geq -\Phi_x(x, y)/(2\Phi_y(x, y)) \geq \psi(x)/320$, and so $\xi(x, y) \to \infty$ as $\operatorname{dist}(x, E) \to 0$ by the assumption on ψ. Then we can find an open set $W \subseteq (a, b)$ containing x such that $\xi(t, z) \geq p + 2$ for any $(t, z) \in (W\backslash E) \times \mathbb{R}$, and hence that $F = 0$ on $(W \times \mathbb{R}) \times (-\infty, p+1)$, by property (3.17.2) of γ. So in fact $F \in C^\infty(\mathbb{R}^3)$. Clearly $F \geq 0$ by (3.17.4), and is convex in p by (3.17.1).

Defining $L(x, y, p) = F(x, y, p) + \omega(p)$ gives a Lagrangian $L \in C^\infty(\mathbb{R}^3)$ such that $L \geq \omega$ and L is strictly convex in p. The remainder of the construction is similar to that in Csörnyei et al. [2008]. Details to supplement the following can be found there. For $(x, y) \in \mathbb{R}^2\backslash(E \times \mathbb{R})$, we have, by strict convexity of ω and property (3.17.4) of γ, that

$$L(x,y,p) \geq \omega(\psi(x)) + \omega'(\psi(x))(p-\psi(x)) + \theta(x,y)(p-\xi(x,y))$$
$$= \Phi_x(x,y) + p\Phi_y(x,y).$$

Moreover, $p = \psi(x)$ implies equality by (3.5) and (3.17.3); and equality in this inequality implies $p = \psi(x)$ by strict convexity of ω. Thus equality holds in this inequality if and only if $p = \psi(x)$.

Let $u \in \mathrm{AC}(a,b)$. Then since $\Phi \in C^\infty(\mathbb{R}^2 \backslash (E \times \mathbb{R}))$ and $E \subseteq [a,b]$ is null, $(\Phi \circ U)\colon [a,b] \to \mathbb{R}$ is differentiable almost everywhere with $(\Phi \circ U)'(x) = \Phi_x(U(x)) + u'(x)(\Phi_y(U(x))$, and for almost every $x \in [a,b]$, the above inequality implies

$$L(x,u(x),u'(x)) \geq \Phi_x(x,u(x)) + u'(x)\Phi_y(x,u(x)) = (\Phi \circ U)'(x), \qquad (3.6)$$

with equality if and only if $u'(x) = \psi(x)$. Supposing further that u is non-decreasing and $\omega(u'(\cdot)) \in L^1(a,b)$, we note that $(\Phi \circ U)$ has the Lusin property, i.e. maps null sets to null sets: (3.18.5) implies that any subset of E is mapped to a null set, and on $[a,b]\backslash E$ the function $(\Phi \circ U)$ is locally absolutely continuous, since $\Phi \in C^\infty(\mathbb{R}^2 \backslash (E \times \mathbb{R}))$.

Given these observations, we now argue that it suffices to check the inequality in the conclusion of the Lemma when $u \in \mathrm{AC}(a,b)$ is non-decreasing and $\Phi(U(a) \leq \Phi(U(b))$. The result is trivial if $\Phi(U(a)) \geq \Phi(U(b))$, since $L \geq 0$ (in this situation the left-hand side of the inequality is non-negative while the right-hand side is non-positive). By (3.17.1), $\Phi(U(a)) \leq \Phi(U(b))$ only if $u(a) < u(b)$. So in this case, if u is not non-decreasing, we can construct non-decreasing $v \in \mathrm{AC}(a,b)$ such that $v(a) = u(a)$, $v(b) = u(b)$, and for almost every $x \in [a,b]$ either $v(x) = u(x)$ and $v'(x) = u'(x)$, or $v'(x) = 0$. We now observe that by (3.4) and (3.17.2) we have $\gamma(0,\xi,\theta) = 0$ on $\mathbb{R}^2 \backslash (E \times \mathbb{R})$. So for all $(x,y,p) \in \mathbb{R}^3$,

$$L(x,y,p) = \omega(p) + F(x,y,p) \geq \omega(p) \geq 0 = \omega(0) + F(x,y,0) = L(x,y,0),$$

where by properties of ω, the second inequality is strict whenever $p \neq 0$. Since $\{x \in [a,b] : v'(x) = 0\}$ must have positive measure, and on this set we have by the above note that $L(x,u,u') > 0 = L(x,v,v')$, we see that

$$\int_a^b L(x,u(x),u'(x))\,dx > \int_a^b L(x,v(x),v'(x))\,dx.$$

So we can indeed assume that $u \in \mathrm{AC}(a,b)$ is non-decreasing and such that $\Phi(U(a)) \leq \Phi(U(b))$. The result is again trivial if $\omega(u'(\cdot)) \notin L^1(a,b)$, since $L \geq \omega$ (in this situation the left-hand side is infinite). So we can also suppose that $\omega(u'(\cdot)) \in L^1(a,b)$. We let $\{(a_j,b_j)\}_{j\in J}$ be the (at most countable) sequence of components of $(a,b)\backslash E$ such that $\Phi(U(a_j)) < \Phi(U(b_j))$. Then using that $(\Phi \circ U)$ is locally absolutely continuous on $(a,b)\backslash E$ and the fact from (3.18.5) that $(\Phi \circ U)(E)$ is null, we see that, using (3.17.4) and (3.6),

$$\begin{aligned}\int_a^b L(x,u(x),u'(x))\,dx &\geq \sum_{j\in J}\int_{a_j}^{b_j} L(x,u(x),u'(x))\,dx\\ &\geq \sum_{j\in J}\int_{a_j}^{b_j} \max\{0,(\Phi\circ U)'\}\,dx\\ &\geq \sum_{j\in J}\Phi(U(b_j)) - \Phi(U(a_j))\\ &\geq \Phi(U(b)) - \Phi(U(a)).\end{aligned}$$

Equality in this relation implies that $L(x,u(x),u'(x)) = (\Phi\circ U)'(x)$ for almost every $x \in \bigcup_{j\in J}(a_j,b_j)$, but also that $\bigcup_{j\in J}(a_j,b_j) = (a,b)\backslash E$. Therefore in fact $L(x,u(x),u'(x)) = (\Phi\circ U)'(x)$ for almost every $x \in (a,b)\backslash E$. By (3.6) this implies that $u'(x) = \psi(x)$ for almost every $x \in [a,b]$, since E is null.

Conversely, $u'(x) = \psi(x)$ almost everywhere implies by (3.18.4) that

$$(\Phi\circ U)'(x) = (\Phi_x \circ U)(x) + \psi(x)(\Phi_y\circ U)(x) \geq (-\Phi_x\circ U)(x) \geq 0$$

almost everywhere. This, combined with the fact that $(\Phi\circ U)$ has the Lusin property (since $u' = \psi \geq 0$ almost everywhere implies that $u \in \mathrm{AC}(a,b)$ is non-decreasing, and by assumption $\omega(\psi(\cdot)) \in L^1(a,b)$) implies that $(\Phi\circ U)$ is absolutely continuous [see Saks, 1937, Chapter IX, Theorem 7.7]. Moreover, (3.6) gives that $L(x,u(x),u'(x)) = (\Phi\circ U)'(x)$ almost everywhere, hence

$$\int_a^b L(x,u(x),u'(x))\,dx = \int_a^b (\Phi\circ U)'(x)\,dx = \Phi(U(b)) - \Phi(U(a))$$

as required. □

3.2.2 Construction of the minimizer, u

We now begin the construction of our future minimizer u, by constructing first its derivative ψ. The essential property of ψ is that $\psi(x) \to \infty$ as $\mathrm{dist}(x,E) \to 0$.

We naturally define ψ as the limit of a sequence of non-negative $C^\infty(\mathbb{R})$ functions $\{\psi_k\}_{k=0}^\infty$, where each ψ_k is bounded above, and on an open set V_k covering E attains this bound (which tends to ∞ as $k \to \infty$). We construct ψ_k so that their primitives u_k will be admissible functions in problem (1.2) (i.e. have the same boundary conditions as u) and converge uniformly to u. In fact we shall guarantee that $u = u_k$ off V_k. So, since our Lagrangian will be constructed as in Lemma 3.18, our estimates showing that there is no Lavrentiev gap reduce just to estimates of the integral over V_k of a function involving the gradient of the potential Φ and ψ_k. This then requires a certain upper bound for the measure of V_k. We must also remember that our potential Φ must have a gradient which satisfies inequalities involving ψ and hence ψ_k. This Φ will—just as in Csörnyei et al. [2008]—be defined using a sequence of $C^\infty(\mathbb{R}^2)$ functions $\{\Phi^k\}_{k=0}^\infty$ which have appropriately steep gradients on open sets Ω_k around $E \times \mathbb{R}$. To guarantee that these Φ^k converge, these sets must be small in the directions of these gradients. We choose Ω_k so that this measure is controlled by that of V_k; this gives another upper bound for the measure of V_k. Other bounds are required for technical reasons in the proof; we impose just one inequality which suffices to give all the results.

For $k \geq 0$, let $\{h_k\}_{k=0}^\infty, \{t_k\}_{k=0}^\infty, \{A_k\}_{k=0}^\infty, \{B_k\}_{k=0}^\infty$ be strictly increasing sequences of real numbers tending to infinity, such that $h_0, t_0, A_0, B_0 \geq 1$. We will eventually need to define explicit values for these sequences to satisfy the exact inequalities required in Lemma 3.18, but until we make these definitions, the construction requires only these general assumptions.

By superlinearity of ω, for all $k \geq 0$ we can choose $l_k \geq k$ such that whenever $p \in \mathbb{R}$ satisfies $|p| \geq A_{l_k}$, we have

$$\omega(p) \geq 2(A_k + 1)|p|. \tag{3.7}$$

Define $V_0 = (a, b)$. For $k \geq 1$, we construct a decreasing sequence of open sets $V_k \subseteq (a, b)$ covering E, $V_k = \bigcup_{i=1}^{n_k} (a_k^i, b_k^i) \subseteq (a, b)$, where $V_k^i := (a_k^i, b_k^i)$ are indexed so that $a \leq a_k^1 < b_k^1 \leq a_k^2 < b_k^2 \leq \dots a_k^{n_k} < b_k^{n_k} \leq b$, such that

$$V_k \subseteq B_{2^{-k}}(E); \tag{3.8}$$

$$V_k \Subset V_{k-1}; \text{ and} \tag{3.9}$$

$$\operatorname{meas}(V_k) \leq \frac{\operatorname{dist}(E, \mathbb{R}\backslash V_{k-1})}{10 \cdot 2^{k+2}(b-a)(A_{l_k+2}^2 + 1)(h_{k+1} + 2)(\omega(h_{k+1} + 2) + 1)n_{k-1}}. \tag{3.10}$$

Suppose for $k \geq 1$ that V_{k-1} has been constructed. Since E is compact and null, we can choose $\rho_k \in (0, 2^{-k})$ such that $\operatorname{meas}(B_{2\rho_k}(E))$ is bounded above by the right-hand side of (3.10), and $B_{2\rho_k}(E) \subseteq V_{k-1}$. By compactness of E, we can choose $V_k \subseteq B_{\rho_k}(E)$ covering E which consists of a finite number of pairwise disjoint intervals. We discard any intervals not containing points of E. Conditions (3.8) and (3.10) are immediate. Since $V_k \subseteq B_{\rho_k}(E)$ where $B_{2\rho_k}(E) \subseteq V_{k-1}$ we have (3.9).

We also define

- $r_k^i = b_k^i - a_k^i > 0$, and $r_k = \min_{1\le i\le n_k} r_k^i > 0$; and
- $\delta_k = \mathrm{dist}(E, [a,b]\backslash V_k)$, where this is strictly positive by compactness of E, and satisfies $\delta_k < r_k/2$.

Note that for $x \in V_k$ and $y \notin V_{k-1}$, we in fact have, choosing $z \in E$ such that $|z-x| \le \mathrm{meas}(V_k)$, that

$$|x-y| \ge |y-z| - |z-x| \ge \delta_{k-1} - \mathrm{meas}(V_k) \ge \delta_{k-1}/2.$$

Thus

$$\mathrm{dist}(V_k, \mathbb{R}\backslash V_{k-1}) \ge \mathrm{dist}(E, \mathbb{R}\backslash V_{k-1})/2. \tag{3.11}$$

Lemma 3.19. There exist $\psi \in C^\infty(\mathbb{R}\backslash E)$ such that $\psi(x) \to \infty$ as $\mathrm{dist}(x,E) \to 0$ and $\omega(\psi(\cdot)) \in L^1(a,b)$, $u \in \mathrm{AC}(a,b)$ satisfying $u' = \psi$ almost everywhere on $[a,b]$, and sequence $\{u_k\}_{k=0}^\infty$ of functions $u_k \in C^\infty([a,b])$ such that, for all $k \ge 0$,

(3.19.1) $u(a) = u_k(a)$ and $u(b) = u_k(b)$;

(3.19.2) for $x \in [a,b]\backslash V_k$ we have $u(x) = u_k(x)$ (and consequently $u'(x) = u_k'(x)$ for $x \in [a,b]\backslash\overline{V_k}$);

(3.19.3) $u'(x) \ge h_k$ for all $x \in V_k\backslash E$;

(3.19.4) $u'(x) \le h_k + 2$ for all $x \in [a,b]\backslash V_k$; and

(3.19.5) $u_k \to u$ uniformly on $[a,b]$.

Proof. We first exhibit a sequence $\{\psi_k\}_{k=0}^\infty$ of functions $\psi_k \in C^\infty(\mathbb{R})$ such that for all $k \ge 0$

(3.19.a) $1 \le \psi_k(x) \le h_k + 2$ for all $x \in \mathbb{R}$;

(3.19.b) $h_k + 1 \le \psi_k(x)$ for all $x \in V_k$;

(3.19.c) $\psi_k(x) = \psi_l(x)$ for all $x \in \mathbb{R}\backslash V_l$ for all $0 \le l \le k$;

(3.19.d) $h_l \le \psi_k(x)$ for $x \in V_l$ for all $0 \le l \le k$; and

(3.19.e) $\int_{V_l^i} \psi_k = \int_{V_l^i} \psi_l$ for all $1 \le i \le n_l$ and all $0 \le l \le k$.

Define $\psi_0(x) = h_0 + 1$ for all $x \in \mathbb{R}$, which clearly satisfies (3.19.a)–(3.19.e). Let $k \ge 1$, and consider $1 \le j \le n_{k-1}$. Note that by (3.10)

$$\text{meas}(V_k) \le \frac{r_{k-1}}{2(h_k - h_{k-1}) + 1} \le \frac{\text{meas}(V_{k-1}^j)}{2(h_k - h_{k-1}) + 1}.$$

This implies that, for each $1 \le j \le n_{k-1}$, since $\text{meas}(V_{k-1}^j \cap V_k) \le \text{meas}(V_k)$,

$$\frac{\text{meas}(V_{k-1}^j)}{\text{meas}(V_k \cap V_{k-1}^j)} \ge 2(h_k - h_{k-1}) + 1 > h_k - h_{k-1} + 1.$$

Hence we can choose $\phi_k \in C^\infty(\mathbb{R})$ such that

$$\phi_k(x) = 0 \text{ for } x \in \mathbb{R} \backslash V_{k-1}; \tag{3.12}$$

$$-1 \le \phi_k(x) \le h_k - h_{k-1} \text{ for } x \in \mathbb{R}; \tag{3.13}$$

$$\phi_k(x) = h_k - h_{k-1} \text{ for } x \in V_k; \text{ and} \tag{3.14}$$

$$\int_{V_{k-1}^j} \phi_k = 0 \text{ for each } 1 \le j \le n_{k-1}. \tag{3.15}$$

For example, fix $1 \le j \le n_{k-1}$, and note that when considering open sets W_k^j, $\tilde{W}_k^j$ such that

$$V_k \cap V_{k-1}^j \Subset W_k^j \Subset \tilde{W}_k^j \Subset V_{k-1}^j,$$

the function $\text{meas}(\tilde{W}_k^j)/\text{meas}(W_k^j)$ depends continuously on the measures of the two sets $\tilde{W}_k^j$ and W_k^j, and takes values greater than but arbitrarily close to 1, and less than but arbitrarily close to $\text{meas}(V_{k-1}^j)/\text{meas}(V_{k-1}^j \cap V_k)$. Thus we may choose sets $\tilde{W}_k^j$ and W_k^j such that

$$\frac{\text{meas}(\tilde{W}_k^j)}{\text{meas}(W_k^j)} = h_k - h_{k-1} + 1,$$

that is

$$(h_k - h_{k-1})\text{meas}(W_k^j) = \text{meas}(\tilde{W}_k^j \backslash W_k^j).$$

Then defining $\phi_k^j : \mathbb{R} \to \mathbb{R}$ by

$$\phi_k^j(x) = \begin{cases} h_k - h_{k-1} & x \in W_k^j \\ -1 & x \in \tilde{W}_k^j \backslash W_k^j \\ 0 & \text{otherwise,} \end{cases}$$

we see that

$$\int_{-\infty}^{\infty} \phi_k^j = \int_{V_{k-1}^j} \phi_k^j = (h_k - h_{k-1})\text{meas}(W_k^j) - \text{meas}(\tilde{W}_k^j \backslash W_k^j) = 0.$$

Choosing an appropriate mollification, we can assume ϕ_k^j is of class C^∞, the same property holds, and ϕ_k^j satisfies (3.12)–(3.15), with $V_{k-1}^j \cap V_k$ replacing V_k in condition (3.14). Then defining $\phi_k = \sum_{j=1}^{n_{k-1}} \phi_k^j$ gives us ϕ_k as claimed.

Using this ϕ_k, we now suppose ψ_{k-1} to be defined, and set $\psi_k = \psi_{k-1} + \phi_k$. This defines our sequence $\{\psi_k\}_{k=0}^\infty$. We now show by induction on $k \geq 0$ that these functions satisfy the requirements (3.19.a)–(3.19.e). Let $k \geq 1$, and suppose ψ_{k-1} has been constructed in this way and satisfies all the conditions.

By (3.12) $\psi_k = \psi_{k-1}$ off V_{k-1}, which gives (3.19.c) by inductive hypothesis and since $\{V_k\}_{k=0}^\infty$ is a decreasing sequence. Then for points not in V_{k-1}, we see that the inequality in (3.19.a) holds by inductive hypothesis (3.19.a) and since $\{h_k\}_{k=0}^\infty$ is an increasing sequence.

For $x \in V_{k-1}$ we have, by (3.13) and inductive hypotheses (3.19.b) and (3.19.a), that

$$1 \leq h_{k-1} \leq \psi_{k-1}(x) - 1 \leq \psi_k(x) \leq \psi_{k-1}(x) + (h_k - h_{k-1}) \leq h_k + 2.$$

Hence the inequality in (3.19.a) holds everywhere, as required. Note that for $x \in V_k$ we have by (3.14) and inductive hypothesis (3.19.b), since $V_k \subseteq V_{k-1}$, that

$$\psi_k(x) = \psi_{k-1}(x) + h_k - h_{k-1} \geq h_k + 1,$$

as required for (3.19.b).

Let $x \in [a, b]$, and choose the greatest index $0 \leq l < k$ such that $x \in V_l$. If $l < k-1$, then $x \notin V_{k-1}$, so inequality (3.19.d) follows by inductive hypothesis. If $l = k-1$, then $x \in V_{k-1}$, and so by (3.13) and inductive hypothesis (3.19.b),

$$\psi_k(x) \geq \psi_{k-1}(x) - 1 \geq h_{k-1}.$$

In particular this gives (3.19.d) since $\{h_k\}_{k=0}^\infty$ is an increasing sequence.

For the claim (3.19.e), let $0 \leq l < k$ (there is nothing to prove for $l = k$), and fix $0 \leq i \leq n_l$. Then using (3.12), (3.15), and the inductive hypothesis we have

$$\int_{V_l^i} \psi_k = \int_{V_l^i} \psi_{k-1} + \int_{V_l^i} \phi_k = \int_{V_l^i} \psi_{k-1} + \int_{V_l^i \cap V_{k-1}} \phi_k = \int_{V_l^i} \psi_{k-1} = \int_{V_l^i} \psi_l,$$

since $\{V_k\}_{k=1}^{\infty}$ is decreasing, so $V_l^i \cap V_{k-1}$ consists of components V_{k-1}^j of V_{k-1}.

Using (3.8) we see that for all $x \notin E$ there is $k \geq 1$ such that $x \notin V_l$ for all $l \geq k$, thus by (3.19.c) letting $\psi(x) = \lim_{k\to\infty} \psi_k(x)$ gives a well-defined function $\psi \in C^{\infty}(\mathbb{R}\backslash E)$ such that

$$\psi(x) = \psi_k(x) \text{ for all } x \notin V_k. \tag{3.16}$$

We see that $\psi(x) \to \infty$ as $\operatorname{dist}(x, E) \to 0$ by (3.19.b), since $V_k \supseteq E$ for all $k \geq 0$ and $h_k \to \infty$ as $k \to \infty$. By (3.19.a) we have that $\psi(x) \geq 1$ for all $x \in \mathbb{R}\backslash E$.

Now, we have that $|\psi_k| \leq |\psi_0| + \sum_{l=1}^{\infty} |\phi_l|$ for all $k \geq 0$, and for any bounded set $M \subseteq \mathbb{R}$, we have, using (3.12), (3.13), and (3.10) that

$$\begin{aligned}
\int_M |\psi_0| + \int_M \sum_{l=1}^{\infty} |\phi_l| &\leq \operatorname{meas}(M)(h_0 + 1) + \sum_{l=1}^{\infty} \operatorname{meas}(M \cap V_{l-1})(h_l - h_{l-1} + 1) \\
&\leq \operatorname{meas}(M)(h_0 + h_1 + 2) + \sum_{l=1}^{\infty} (h_{l+1} + 1)\operatorname{meas}(V_l) \\
&\leq \operatorname{meas}(M)(2h_1 + 2) + \sum_{l=1}^{\infty} 2^{-l} \\
&< \infty.
\end{aligned}$$

So by the dominated convergence theorem $\psi \in L^1(a, b)$, and

$$\int_M \psi_k \to \int_M \psi \text{ as } k \to \infty \text{ for all } M \subseteq [a, b]. \tag{3.17}$$

We also note that, by (3.16), (3.19.a), and (3.10), we have

$$\begin{aligned}
\int_a^b \omega(\psi(x))\, dx &\leq \sum_{k=1}^{\infty} \int_{V_{k-1}\backslash V_k} \omega(\psi(x))\, dx \\
&\leq \sum_{k=1}^{\infty} \int_{V_{k-1}\backslash V_k} \omega(\psi_k(x))\, dx \\
&\leq \sum_{k=1}^{\infty} \int_{V_{k-1}} \omega(h_k + 2)\, dx \\
&\leq \sum_{k=1}^{\infty} \omega(h_k + 2)\operatorname{meas}(V_{k-1}) \\
&\leq \sum_{k=1}^{\infty} 2^{-k} + \omega(h_1 + 2)(b - a)
\end{aligned}$$

hence $\omega(\psi(\cdot)) \in L^1(a,b)$ as required.

We now define non-decreasing $u \in \mathrm{AC}(a,b)$ by

$$u(x) = \int_a^x \psi(t)\,dt,$$

and so $u' = \psi$ off E, in particular almost everywhere. Condition (3.19.3) follows immediately from (3.19.d). Condition (3.19.4) follows immediately from (3.19.c) and (3.19.a). For each $k \geq 0$ we define also non-decreasing $u_k \in C^\infty([a,b])$ by

$$u_k(x) = \int_a^x \psi_k(t)\,dt.$$

Property (3.19.1) follows since by (3.19.e) and (3.17),

$$u_k(b) = \int_a^b \psi_k = \int_{V_0} \psi_k = \int_{V_0} \psi_0 = \int_{V_0} \psi = \int_a^b \psi = u(b),$$

and since clearly $u_k(a) = 0 = u(a)$ by definition.

Suppose $x \in [a,b]\backslash V_k$. Then either we have $x \leq a_k^i$ for all $1 \leq i \leq n_k$, or we have for some $1 \leq i_x \leq n_k$ that $b_k^{i_x} \leq x$. In the first case we see immediately that, since $[a,x] \cap V_k = \emptyset$, (3.16) implies

$$u(x) = \int_a^x \psi(t)\,dt = \int_a^x \psi_k(t)\,dt = u_k(x)$$

by assumption. Otherwise we argue by (3.17), (3.19.e) and (3.19.c) that

$$\begin{aligned}
u(x) &= \int_a^x \psi \\
&= \sum_{i=1}^{i_x} \int_{V_k^i} \psi + \int_{[a,x]\backslash V_k} \psi \\
&= \sum_{i=1}^{i_x} \int_{V_k^i} \psi_k + \int_{[a,x]\backslash V_k} \psi_k \\
&= \int_a^x \psi_k \\
&= u_k(x)
\end{aligned}$$

as required for (3.19.2).

Fix $1 \leq i \leq n_k$, and let $x \in V_k^i$. Since u and u_k are non-decreasing, using (3.19.2), (3.19.a), and (3.10) we see that

$$
\begin{aligned}
|u_k(x) - u(x)| &\leq u_k(b_k^i) - u(a_k^i) \\
&= u_k(b_k^i) - u_k(a_k^i) \\
&= \int_{a_k^i}^{b_k^i} \psi_k \\
&\leq (h_k + 2)(b_k^i - a_k^i) \\
&\leq (h_k + 2)\text{meas}(V_k) \\
&\leq 2^{-k}.
\end{aligned}
$$

Since $u = u_k$ off V_k, we then have that $\sup_{x\in[a,b]} |u_k(x) - u(x)| \leq 2^{-k}$, hence u_k converges to u uniformly, as required for (3.19.5). □

3.2.3 Construction of the potential, Φ

The construction of our potential is based on that which constitutes the proof of Theorem 10 in Csörnyei et al. [2008]. We construct a sequence of $C^\infty(\mathbb{R}^2)$ functions $\{\Phi^k\}_{k=0}^\infty$ which have steep gradients on open sets Ω_k around $(E \times \mathbb{R})$. Because we have fixed our function $\psi \in C^\infty(\mathbb{R}\backslash E)$ with which we have to compare the derivatives of Φ, the sets Ω_k are now given before the construction. This contrasts with the situation of Csörnyei et al., where the sets could be chosen small enough at each stage of the construction of the sequence. We have of course carefully chosen Ω_k, or more precisely in fact V_k, so that all the properties required at this stage hold with these fixed sets. A final remark to make is that our sets Ω_k cannot shrink vertically, since the inequalities required of them are independent of the second variable y. This means that it takes a little effort to prove that the intersections with absolutely continuous curves are small for those curves we need to consider, since very steep curves can lie inside Ω_k and contribute a large linear measure. However, as we saw in Lemma 3.18, it suffices to consider those non-decreasing curves u with steepness controlled by the superlinearity, in the sense that $\omega(u'(\cdot)) \in L^1(a, b)$. This restricts the class of curves about which we need information to those whose intersections with Ω_k we can indeed control just by the measure of V_k, since our superlinearity ω is fixed.

Let $\Omega_0 = \mathbb{R}^2$, and for $k \geq 1$ define $\Omega_k = V_k \times \mathbb{R}$. Then by (3.11),

$$\text{dist}(\Omega_k, \mathbb{R}^2\backslash\Omega_{k-1}) \geq \delta_{k-1}/2. \tag{3.18}$$

We now state and prove appropriate versions of Lemmas 12 and 13 in Csörnyei et al. [2008]. For two vectors $x, y \in \mathbb{R}^2$, we write $[x, y]$ to denote the line segment in $\mathbb{R}^2$ connecting them.

Lemma 3.20. Let $\tau > 0$, $e \in \mathbb{R}^2 \backslash \{0\}$, and suppose $\Omega,\ \Omega' \subseteq \mathbb{R}^2$ are open sets such that $\Omega \Subset \Omega'$ and $\mathscr{H}^1(\Omega' \cap \Gamma) \leq \tau/2$ for any line Γ in the plane in direction e, i.e. $\Gamma \subseteq \mathbb{R}^2$ such that for distinct points $x, y \in \Gamma$, we have $\|e\|_2(y-x)/\|y-x\|_2 = \pm e$.

Then there exists $f \in C^\infty(\mathbb{R}^2)$ such that

- $0 \leq f(x) \leq \|e\|_2^{-1}\tau$ for all $x \in \mathbb{R}^2$;
- $\operatorname{dist}(\nabla f(x), [0, e]) < \tau$ for all $x \in \mathbb{R}^2$; and
- $\|\nabla f(x) - e\|_2 < \tau$ for $x \in \Omega$.

Proof. We first show it suffices to prove the result for $e = (1, 0)$. For arbitrary $e \in \mathbb{R}^2$, find rotation $R\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that $\|e\|_2^{-1} Re = (1, 0)$. Note of course that $\|e\|_2^{-1} R\Omega \Subset \|e\|_2^{-1} R\Omega'$. Also, if Γ is a horizontal line, then $\|e\|_2 R^{-1}\Gamma$ is a line in the direction of e, so by assumption, for horizontal lines Γ,

$$\mathscr{H}^1(\|e\|_2^{-1} R\Omega' \cap \Gamma) = \mathscr{H}^1(\|e\|_2^{-1} R(\Omega' \cap \|e\|_2 R^{-1}\Gamma)) \leq \|e\|_2^{-1}\tau/2,$$

so $\|e\|_2^{-1} R\Omega, \|e\|_2^{-1} R\Omega'$ satisfy the assumptions for $\tilde{\tau} := \|e\|_2^{-1}\tau$ and $\tilde{e} := (1, 0)$. So by assumption there exists $\tilde{f} \in C^\infty(\mathbb{R}^2)$ satisfying the three conclusions for $\tilde{e}$ and $\tilde{\tau}$. Define $f \in C^\infty(\mathbb{R}^2)$ by $f(x) = \tilde{f}(\|e\|_2 R^{-1}x)$. Fix $x \in \mathbb{R}^2$. Then firstly

$$0 \leq f(x) = \tilde{f}(\|e\|_2 R^{-1}x) \leq \|(1, 0)\|_2^{-1}\tilde{\tau} = \|e\|_2^{-1}\tau.$$

By assumption there exists $s \in [0, 1]$ such that

$$\left\|\nabla \tilde{f}(\|e\|_2 R^{-1}x) - s(1, 0)\right\|_2 < \tilde{\tau}.$$

Then for this $s \in [0, 1]$ we have

$$\begin{aligned}
\|\nabla f(x) - se\|_2 &= \left\|\|e\|_2 R^{-1}\nabla \tilde{f}(\|e\|_2 R^{-1}x) - s\|e\|_2 R^{-1}\|e\|_2^{-1} Re\right\|_2 \\
&= \left\|\|e\|_2 R^{-1}(\nabla \tilde{f}(\|e\|_2 R^{-1}x) - s\|e\|_2^{-1} Re)\right\|_2 \\
&= \|e\|_2 \left\|\nabla \tilde{f}(\|e\|_2 R^{-1}x) - s(1, 0)\right\|_2 \\
&< \|e\|_2 \tilde{\tau} \\
&= \tau.
\end{aligned}$$

Now let $x \in \Omega$. Then $\|e\|_2^{-1} Rx \in \|e\|_2^{-1} R\Omega$, so by assumption we have that

$$\left\| \nabla \tilde{f}(\|e\|_2^{-1} Rx) - (1,0) \right\|_2 < \tilde{\tau}.$$

Thus

$$\begin{aligned} \|\nabla f(x) - e\|_2 &= \left\| \|e\|_2 R^{-1} \nabla \tilde{f}(\|e\|_2 R^{-1} x) - \|e\|_2 R^{-1} \|e\|_2^{-1} Re \right\|_2 \\ &= \|e\|_2 \left\| \nabla \tilde{f}(\|e\|_2 R^{-1} x) - (1,0) \right\|_2 \\ &< \|e\|_2 \tilde{\tau} \\ &= \tau \end{aligned}$$

as required.

So we can assume without loss of generality that $e = (1,0)$. By using a suitable mollification, it suffices to construct a Lipschitz function $g\colon \mathbb{R}^2 \to \mathbb{R}$ such that

- $0 \leq g(x) \leq \tau/2$ for all $x \in \mathbb{R}^2$;
- $g_x(x) \in [0,1]$ and $g_y(x) = 0$ for every $x \in \mathbb{R}^2$; and
- $g_x(x) = 1$ for $x \in \Omega'$.

To do this we just define $\tilde{g}\colon \mathbb{R} \to \mathbb{R}$ by

$$g(x) = \sup_{y \in \mathbb{R}} \mathscr{H}^1(((-\infty, x] \times \{y\}) \cap \Omega'),$$

and define $g\colon \mathbb{R}^2 \to \mathbb{R}$ by $g(x,y) = \tilde{g}(x)$. This is clearly non-negative, non-decreasing in x, and independent of y. Further, $g(x,y) \leq \tau/2$ by the condition that Ω' meets all horizontal lines in a set of linear measure at most $\tau/2$.

For $x_1 \leq x_2$, we see that $g(x_1, y) \leq g(x_2, y) \leq g(x_1, y) + (x_2 - x_1)$, hence

$$0 \leq \frac{g(x_2, y) - g(x_1, y)}{x_2 - x_1} \leq 1.$$

Since Ω' is open, for $(x,y) \in \Omega'$, for sufficiently small $t > 0$ we have that the line segment $[(x,y),(x+t,y)]$ is contained in Ω', so $g(x+t,y) = g(x,y) + t$, and thus

$$\frac{g(x+t,y) - g(x,y)}{t} = 1$$

as required for the full result. □

Lemma 3.21. Let $\epsilon > 0$, $e^0, e^1 \in \mathbb{R}^2$ be distinct vectors, and $\Omega, V \subseteq \mathbb{R}^2$ be open sets such that $\Omega \Subset V$. Suppose $0 < \delta \le \operatorname{dist}(\Omega, \mathbb{R}^2 \backslash V)/2$ and there exists an open set $\Omega' \Supset \Omega$ such that for any line $\Gamma \subseteq \mathbb{R}^2$ in direction of $e^1 - e^0$, we have

$$\mathscr{H}^1(\Omega' \cap \Gamma) \le \frac{\epsilon}{2(1 + \delta^{-1}\|e^0 - e^1\|_2^{-1})}.$$

Let $g^0 \in C^\infty(\mathbb{R}^2)$.

Then there exists $g^1 \in C^\infty(\mathbb{R}^2)$ such that

- $\|g^1 - g^0\|_\infty < \epsilon\|e^0 - e^1\|_2^{-1}$;
- $g^1 = g^0$ off V;
- $\operatorname{dist}(\nabla g^1(x), [e^0, e^1]) < \epsilon + \|\nabla g^0(x) - e^0\|_2$ for $x \in \mathbb{R}^2$; and
- $\|\nabla g^1(x) - e^1\|_2 < \epsilon + \|\nabla g^0(x) - e^0\|_2$ for $x \in \Omega$.

Proof. Let $\tau = \frac{\epsilon}{1+\delta^{-1}\|e^0-e^1\|_2^{-1}}$ and apply Lemma 3.20 with this τ, the sets Ω and Ω' from the assumptions, and vector $e = e^1 - e^0$. Let f be the resulting function.

Choose $\chi \in C^\infty(\mathbb{R}^2)$ such that $0 \le \chi \le 1$, $\chi = 1$ on Ω, and $\chi = 0$ off V, and $\|\nabla\chi\|_2 \le \delta^{-1}$. Define $g^1 = g^0 + \chi f$. Clearly $g^1 \in C^\infty(\mathbb{R}^2)$. For $x \in \mathbb{R}^2$, we see immediately from Lemma 3.20 that

$$|g^1(x) - g^0(x)| \le |f(x)| \le \tau\|e^1 - e^0\|_2^{-1} < \epsilon\|e^1 - e^0\|_2^{-1}.$$

Also note that, by the properties of χ and Lemma 3.20, we have

$$\begin{aligned}
\operatorname{dist}(\nabla g^1(x), [e^0, e^1]) &\le \operatorname{dist}(\nabla g^0 - e^0 + (\chi\nabla f)(x) + (f\nabla\chi)(x), [0, e^1 - e^0]) \\
&\le \|\nabla g^0(x) - e^0\|_2 + \operatorname{dist}(\nabla f(x), [0, e]) + \|(f\nabla\chi)(x)\|_2 \\
&\le \|\nabla g^0(x) - e^0\|_2 + \tau + \tau\delta^{-1}\|e^0 - e^1\|_2^{-1} \\
&= \|\nabla g^0(x) - e^0\|_2 + \epsilon.
\end{aligned}$$

For $x \in \Omega$ we have, since $\chi(x) = 1$, that

$$\begin{aligned}
\|\nabla g^1(x) - e^1\|_2 &\le \|\nabla g^0(x) - e^0\|_2 + \|(\chi\nabla f)(x) - (e^1 - e^0)\|_2 + \|(f\nabla\chi)(x)\|_2 \\
&\le \|\nabla g^0(x) - e^0\|_2 + \|\nabla f(x) - e\|_2 + \tau\delta^{-1}\|e^0 - e^1\|_2^{-1} \\
&\le \|\nabla g^0(x) - e^0\|_2 + \tau(1 + \delta^{-1}\|e^0 - e^1\|_2^{-1}) \\
&= \|\nabla g^0(x) - e^0\|_2 + \epsilon.
\end{aligned}$$

□

We now construct a potential Φ which will satisfy the conditions of Lemma 3.18 with the function ψ given by Lemma 3.19.

Let our increasing sequences have the following values, for $k \geq 0$:

- $h_k = 10(3 + 2^{k+1})$;
- $t_k = 3 + 2^k$;
- $B_k = 4 + 320\omega'(h_k + 2)$, and $A_k = 3t_k B_k$.

Also for $k \geq 0$, define numbers $\eta_k = 1 - 2^{-k-1} > 0$ and $\epsilon_k = 2^{-k}(4n_k)^{-1} > 0$, and vector $e_k = (-A_k, B_k) \in \mathbb{R}^2$. We inductively construct a sequence $\{\Phi^k\}_{k=0}^{\infty}$ of functions $\Phi^k \in C^\infty(\mathbb{R}^2)$ satisfying, for $k \geq 0$,

$$\|\nabla\Phi^k(x,y) - e_k\|_2 < \eta_k \text{ for } (x,y) \in \overline{\Omega_k}; \tag{3.19}$$

and for $k \geq 1$,

$$\|\Phi^k - \Phi^{k-1}\|_\infty < \epsilon_{k-1}; \tag{3.20}$$

$$\Phi^k = \Phi^{k-1} \text{ off } \Omega_{k-1}; \text{ and} \tag{3.21}$$

$$\operatorname{dist}(\nabla\Phi^k(x,y), [e_{k-1}, e_k]) < \eta_k \text{ for } (x,y) \in \overline{\Omega_{k-1}}. \tag{3.22}$$

We define $\Phi^0(x,y) = -A_0 x + B_0 y$, which clearly satisfies (3.19). Suppose for $k \geq 1$ we have constructed Φ^{k-1} as claimed. To construct Φ^k we apply Lemma 3.21 with $\epsilon = \epsilon_{k-1}$, $e^0 = e_{k-1}$, $e^1 = e_k$, $\Omega = \Omega_k$, $V = \Omega_{k-1}$, $\Omega' = B_{r_k}(\Omega_k)$, $\delta = \delta_{k-1}/4$, $g^0 = \Phi^{k-1}$. We check this is possible by recalling (3.18) and observing that, regarding a line Γ in the direction of vector $e_k - e_{k-1}$ as the graph of a Lipschitz function $\gamma\colon \mathbb{R} \to \mathbb{R}$ with $\operatorname{Lip}(\gamma) = (B_k - B_{k-1})/(A_k - A_{k-1}) \leq B_k$, so $\Gamma(x) = (x, \gamma(x))$, we have by (3.10) that

$$\begin{aligned}
\mathscr{H}^1(B_{r_k}(\Omega_k) \cap \Gamma) &= \mathscr{H}^1(\Gamma(\Gamma^{-1}(B_{r_k}(\Omega_k)))) \\
&\leq (1 + (\operatorname{Lip}(\gamma))^2)^{1/2} \operatorname{meas}(\Gamma^{-1}(B_{r_k}(\Omega_k))) \\
&\leq 2B_k \operatorname{meas}(B_{r_k}(V_k)) \\
&\leq 2B_k \left(\sum_{i=1}^{n_k} (r_k^i + 2r_k) \right) \\
&\leq 6B_k \operatorname{meas}(V_k) \\
&\leq \frac{\epsilon_{k-1}}{2(1 + 4\delta_{k-1}^{-1})},
\end{aligned}$$

which suffices since $\|e_k - e_{k-1}\|_2 \geq 1$. We define Φ^k as the function g^1 given by the Lemma.

Then (3.21) is immediate, and again since $\|e_k - e_{k-1}\|_2 \geq 1$, we see that $\|\Phi^k - \Phi^{k-1}\|_\infty < \epsilon_{k-1}$ as required for (3.20). For (3.22) we let $(x, y) \in \overline{\Omega_{k-1}}$ and use inductive hypothesis (3.19) and the properties given by Lemma 3.21 to see that

$$\begin{aligned} \operatorname{dist}(\nabla\Phi^k(x, y), [e_{k-1}, e_k]) &< \epsilon_{k-1} + \|\nabla\Phi^{k-1}(x, y) - e_{k-1}\|_2 \\ &\leq \epsilon_{k-1} + \eta_{k-1} \\ &\leq 2^{-(k-1+2)} + 1 - 2^{-k} \\ &\leq \eta_k. \end{aligned}$$

Similarly for (3.19), we let $(x, y) \in \overline{\Omega_k}$ and use inductive hypothesis (3.19) again, noting that $\Omega_k \subseteq \Omega_{k-1}$, to see that

$$\|\nabla\Phi^k(x, y) - e_k\|_2 < \epsilon_{k-1} + \|\nabla\Phi^{k-1}(x, y) - e_{k-1}\|_2 \leq \eta_k.$$

Hence we can construct such a sequence $\{\Phi^k\}_{k=0}^\infty$ as claimed. We now check that this gives us the potential we require for Lemma 3.18. By (3.20) and since $\epsilon_k \leq 2^{-(k+2)}$, we see that Φ^k converge uniformly to some $\Phi \in C(\mathbb{R}^2)$.

Fix $(x, y) \in \mathbb{R}^2 \backslash (E \times \mathbb{R})$. By (3.8) and (3.9) there is $k \geq 1$ such that $(x, y) \in \overline{\Omega_{k-1}} \backslash \overline{\Omega_k}$, and hence $\Phi \in C^\infty(\mathbb{R}^2 \backslash (E \times \mathbb{R}))$ and $\nabla\Phi = \nabla\Phi^l$ on $\overline{\Omega_{k-1}} \backslash \overline{\Omega_k}$, for all $l \geq k$, by (3.21). Moreover, by (3.22),

$$\Phi_y(x, y) = \Phi^k_y(x, y) \geq B_{k-1} - \eta_{k-1} \geq B_0 - 1 \geq 3$$

and

$$\Phi_x(x, y) = \Phi^k_x(x, y) \leq -A_{k-1} + \eta_{k-1} \leq -A_0 + 1 \leq -3 \cdot 4t_0 + 1 \leq -47$$

as required for (3.18.1) and the second inequality of (3.18.2). By (3.22) there is $s \in [0, 1]$ such that $\|\nabla\Phi(x, y) - (se_{k-1} + (1 - s)e_k)\|_2 < 1$. Using this we see that

$$\begin{aligned} -\Phi_x(x, y) &\leq sA_{k-1} + (1 - s)A_k + 1 \\ &\leq 3t_k(sB_{k-1} + (1 - s)B_k) + 1 \\ &\leq 3t_k(\Phi_y(x, y) + 1) + 1 \\ &\leq 5t_k\Phi_y(x, y), \end{aligned}$$

thus $(-\Phi_x/\Phi_y)(x, y) \leq 5t_k$. Similarly

$$\begin{aligned} -\Phi_x(x, y) &\geq sA_{k-1} + (1 - s)A_k - 1 \\ &\geq 3t_{k-1}(sB_{k-1} + (1 - s)B_k) - 1 \end{aligned}$$

$$\geq 3t_{k-1}(\Phi_y(x,y) - 1) - 1$$
$$\geq t_{k-1}\Phi_y(x,y),$$

thus $(-\Phi_x/\Phi_y)(x,y) \geq t_{k-1}$. Condition (3.18.2) follows since $t_{k-1} \geq t_0 = 4$. Now, suppose further that $x \in (a,b)$. We know from (3.16), (3.19.d), and (3.19.a) that $h_{k-1} \leq \psi(x) \leq h_k + 2$. Thus, by properties of ω,

$$\Phi_y(x,y) > B_{k-1} - 1 = 3 + 320\omega'(h_k + 2) \geq 320\omega'(\psi(x))$$

as required for (3.18.3) in this case. If $x \notin (a,b)$, then $\psi(x) = \psi_0(x) = h_0 + 1$, so we see again that

$$\Phi_y(x,y) > B_0 - 1 = 3 + 320\omega'(h_0 + 2) \geq 320\omega'(\psi(x)),$$

so the full statement of (3.18.3) holds.

We note that, from the definitions,

$$10t_k = 10(3 + 2^k) = h_{k-1}$$

and

$$h_k + 2 = 10(3 + 2^{k+1}) + 2 \leq 10 \cdot 2^4(3 + 2^{k-1}) = 160t_{k-1}.$$

So again supposing first that $x \in (a,b)$, we see

$$-2\Phi_x(x,y)/\Phi_y(x,y) \leq 10t_k \leq \psi(x) \leq h_k + 2 \leq -160\Phi_x(x,y)/\Phi_y(x,y)$$

and hence get (3.18.4) in this case. For $x \notin (a,b)$, note then that $(x,y) \notin \overline{\Omega_1}$, so $\nabla\Phi(x,y) = \nabla\Phi_1(x,y)$. Hence again

$$-2\Phi_x(x,y)/\Phi_y(x,y) \leq 10t_1 \leq \psi(x) \leq h_1 + 2 \leq -160\Phi_x(x,y)/\Phi_y(x,y),$$

as required for the full statement of (3.18.4).

We finally check (3.18.5). Let $u \in \mathrm{AC}(a,b)$ be non-decreasing, and such that $\omega(u'(\cdot)) \in L^1(a,b)$. Fix $k \geq 1$, and note that by (3.20) we have that $\|\Phi - \Phi^k\|_\infty < 2\epsilon_k$. Fix $1 \leq i \leq n_k$. Now, the image of V_k^i under $\Phi \circ U$ is connected, thus

$$(\Phi \circ U)(V_k^i) \subseteq B_{2\epsilon_k}((\Phi_k \circ U)(V_k^i)),$$

and hence

$$\operatorname{meas}((\Phi \circ U)(V_k^i)) \le \operatorname{meas}(B_{2\epsilon_k}((\Phi_k \circ U)(V_k^i))) \le \operatorname{meas}((\Phi_k \circ U)(V_k^i)) + 4\epsilon_k.$$

Now,

$$\mathscr{H}^1(U(V_k^i)) \le \mathscr{H}^1(U(\{x \in V_k^i : |u'(x)| \le A_{l_k}\})) + \mathscr{H}^1(U(\{x \in V_k^i : |u'(x)| > A_{l_k}\})).$$

The first summand can be dealt with easily, since u is non-decreasing:

$$\mathscr{H}^1(U(\{x \in V_k^i : |u'(x)| \le 2A_{l_k}\})) \le \operatorname{meas}(V_k^i)(1 + 2A_{l_k}) \le 3A_{l_k}\operatorname{meas}(V_k^i).$$

On the other hand, again since u is non-decreasing, we have

$$\begin{aligned}
&\mathscr{H}^1(U(\{x \in V_k^i : |u'(x)| > A_{l_k}\})) \\
&\le \operatorname{meas}(\{x \in V_k^i : |u'(x)| > A_{l_k}\}) + \operatorname{meas}(u(\{x \in V_k^i : |u'(x)| > A_{l_k}\})) \\
&\le \operatorname{meas}(\{x \in V_k^i : |u'(x)| > A_{l_k}\}) + \int_{\{x \in V_k^i : |u'(x)| > A_{l_k}\}} u'(x)\,dx.
\end{aligned}$$

So, since $\{A_k\}_{k=0}^\infty$ are increasing, and by the choice of $l_k \ge k$ in (3.7), we see that

$$\begin{aligned}
&\operatorname{meas}((\Phi_k \circ U)(V_k^i)) \\
&\le \operatorname{Lip}(\Phi_k)(\mathscr{H}^1(U(V_k^i))) \\
&\le (A_k + B_k + 2)\Big(3A_{l_k}\operatorname{meas}(V_k^i) + \operatorname{meas}(\{x \in V_k^i : |u'(x)| > A_{l_k}\}) \\
&\quad + \int_{\{x \in V_k^i : |u'(x)| > A_{l_k}\}} u'(x)\,dx.\Big) \\
&\le 6A_{l_k}^2\operatorname{meas}(V_k^i) + 2A_{l_k}\operatorname{meas}(\{x \in V_k^i : |u'(x)| > A_{l_k}\}) + \int_{\{x \in V_k^i : |u'(x)| > A_{l_k}\}} 2A_k u'(x)\,dx \\
&\le 6A_{l_k}^2\operatorname{meas}(V_k^i) + \int_{\{x \in V_k^i : |u'(x)| > A_{l_k}\}} 2|u'(x)| + 2A_k|u'(x)|\,dx \\
&\le 6A_{l_k}^2\operatorname{meas}(V_k^i) + \int_{\{x \in V_k^i : |u'(x)| > A_{l_k}\}} \omega(u'(x))\,dx.
\end{aligned}$$

So, summing over $1 \le i \le n_k$ gives, using (3.10), and since the $\{V_k^i\}_{i=1}^{n_k}$ are disjoint,

$$\operatorname{meas}(\Phi \circ U)(V_k) \le \sum_{i=1}^{n_k} \operatorname{meas}(\Phi \circ U)(V_k^i)$$

$$
\begin{aligned}
&\leq \sum_{i=1}^{n_k} \big(\text{meas}((\Phi_k \circ U)(V_k^i)) + 4\epsilon_k\big) \\
&\leq \sum_{i=1}^{n_k} \left(6A_{l_k}^2 \text{meas}(V_k^i) + \int_{\{x \in V_k^i : |u'(x)| > A_{l_k}\}} \omega(u'(x))\,dx + 4\epsilon_k\right) \\
&\leq 6A_{l_k}^2 \text{meas}(V_k) + \int_{\{x \in V_k : |u'(x)| > A_{l_k}\}} \omega(u'(x))\,dx + 4\epsilon_k n_k \\
&\leq 2^{-k} + \int_{V_k} \omega(u'(x))\,dx + 2^{-k}.
\end{aligned}
$$

Since by assumption $\omega(u'(\cdot)) \in L^1(a,b)$, this tends to 0 as $k \to \infty$, since $\text{meas}(V_k) \to 0$. Therefore, since for all $k \geq 0$

$$(\Phi \circ U)(E) \subseteq (\Phi \circ U)(V_k),$$

we see that $(\Phi \circ U)(E)$ is indeed a null set.

3.2.4 Conclusion

Proof of Theorem 3.16. We let $L \in C^\infty(\mathbb{R}^3)$ be the Lagrangian given by Lemma 3.18, with ψ as given by Lemma 3.19, and this potential Φ. Since the function u given by Lemma 3.19 has by definition $u' = \psi$ almost everywhere on $[a,b]$, we know by Lemma 3.18 that the first statement of the theorem holds for this $u \in \text{AC}(a,b)$.

We check that the singular set is as claimed. Let $x \in E$. By the properties of ψ from Lemma 3.19, given $M > 0$ there is $\delta_0 > 0$ such that $0 < |x-y| < \delta_0$ implies $\psi(y) \geq M$. Suppose $y \in [a,b]$ is such that $x < y < x - \delta_0$. By definition we have we have

$$u(y) - u(x) = \int_x^y \psi(t)\,dt \geq (y-x)M$$

and hence

$$\frac{u(y)-u(x)}{y-x} \geq M,$$

therefore

$$\lim_{y \to x+} \frac{u(y)-u(x)}{y-x} \geq M.$$

Similarly we see that

$$\lim_{y \to x-} \frac{u(y)-u(x)}{y-x} \geq M,$$

thus $u'(x) \geq M$. $M > 0$ was arbitrary, so in fact $u'(x) = \infty$, thus E is contained in the singular set of u.

For $x \notin E$, since E is closed there is $\delta_0 > 0$ such that $[x-\delta_0, x+\delta_0] \cap E = \emptyset$. Since $\psi \in C^\infty(\mathbb{R}\backslash E)$, there exists $K > 0$ such that $0 \le \psi \le K$ on $[x-\delta_0, x+\delta_0]$. Hence for $y \in [a, b]$ such that $x < y < x+\delta_0$ we have, again by definition, that

$$0 \le u(y) - u(x) = \int_x^y \psi(t)\, dt \ge (y-x)K$$

hence

$$\frac{u(y)-u(x)}{y-x} \le K,$$

and so

$$0 \le \lim_{y \to x_+} \frac{u(y)-u(x)}{y-x} \le K.$$

Similarly

$$0 \le \lim_{y \to x_-} \frac{u(y)-u(x)}{y-x} \le K,$$

hence $|u'(x)| < \infty$. So x is not in the singular set of u, i.e. E contains the singular set of u. Thus $E = \{x \in [a, b] : |u'(x)| = \infty\}$ indeed.

We now prove the third statement of the theorem. Lemma 3.19 gives us a sequence of admissible functions $u_k \in C^\infty([a, b])$ which converge uniformly to u. We just need to prove that they also converge in energy. Let $\epsilon > 0$. By (3.19.2) we see that

$$\begin{aligned} 0 \le \mathscr{L}(u_k) - \mathscr{L}(u) &= \int_a^b L(x, u_k(x), u_k'(x)) - L(x, u(x), u'(x))\, dx \\ &= \int_{V_k} L(x, u_k(x), u_k'(x)) - L(x, u(x), u'(x))\, dx. \end{aligned}$$

We know from the precise conclusion of Lemma 3.18 that $x \mapsto L(x, u(x), u'(x))$ is integrable, so since $\text{meas}(V_k) \to 0$ as $k \to \infty$ by (3.10), we can choose $k_0 \ge 1$ such that $\int_{V_k} L(x, u(x), u'(x))\, dx < \epsilon/2$ whenever $k \ge k_0$.

Now, for each $k \ge 1$ and almost every $x \in [a, b]$, we have that

$$\begin{aligned} L(x, u_k(x), u_k'(x)) &= \omega(u_k'(x)) + F(x, u_k(x), u_k'(x)) \\ &= \omega(u_k'(x)) + \gamma(u_k'(x), \xi(x, u_k(x)), \theta(x, u_k(x))) \end{aligned}$$

by definition of the Lagrangian L in Lemma 3.18. Fix such an $x \in [a, b]$. Note by definition of u_k and (3.19.a) that $0 \le u_k'(x) \le h_k + 2$. We get the following upper bound for γ by using (3.17.5), (3.19.a), (3.2) (noting $\xi \ge 0$ by (3.3)), and (3.18.2):

$$\gamma(u_k'(x), \xi(x, u_k(x)), \theta(x, u_k(x))) \le \theta(x, u_k(x))|u_k'(x) - \xi(x, u_k(x)) + 1|$$

$$
\begin{aligned}
&\leq \Phi_y(x, u_k(x))(u_k'(x) + 1) + \theta(x, u_k(x))\xi(x, u_k(x)) \\
&\leq \Phi_y(x, u_k(x))((h_k + 2) + 1) - \Phi_x(x, u_k(x)) \\
&\leq -\Phi_x(x, u_k(x))(h_k + 7)/4 \\
&\leq -\Phi_x(x, u_k(x))(h_k/4 + h_k/4) \\
&\leq -\Phi_x(x, u_k(x))h_k/2.
\end{aligned}
$$

Now, $\omega(u_k'(x)) \leq \omega(h_k + 2)$ by properties of ω, and for sufficiently large $k \geq 0$, $h_k + 2 \leq \omega(h_k + 2)$, since ω is superlinear and $h_k \to \infty$ as $k \to \infty$. So, again using (3.19.a), and since certainly $-\Phi_x \geq 2$, we have, for large $k \geq 0$,

$$
L(x, u_k(x), u_k'(x)) \leq \omega(h_k + 2) - \Phi_x(x, u_k(x))h_k/2 \leq -\Phi_x(x, u_k(x))\omega(h_k + 2).
$$

Now, if $x \in \overline{V_{l-1}}\backslash\overline{V_l}$, then $(x, u_k(x)) \in \overline{\Omega_{l-1}}\backslash\overline{\Omega_l}$, so $\Phi_x(x, u_k(x)) = \Phi^l_x(x, u_k(x))$, and hence $-\Phi_x(x, u_k(x)) \leq A_l + 1$. Thus for large $l \geq 1$, almost everywhere on $\overline{V_{l-1}}\backslash\overline{V_l}$ we have

$$
L(x, u_k(x), u_k'(x)) \leq (A_l + 1)\omega(h_k + 2).
$$

So for sufficiently large $k \geq 1$, we have, since $\{h_k\}_{k=0}^{\infty}$ is increasing, by properties of ω, and (3.10), that

$$
\begin{aligned}
0 \leq \int_{V_k} L(x, u_k(x), u_k'(x))\, dx &\leq \sum_{l=k+1}^{\infty} \int_{\overline{V_{l-1}}\backslash\overline{V_l}} L(x, u_k(x), u_k'(x))\, dx \\
&\leq \sum_{l=k+1}^{\infty} \int_{\overline{V_{l-1}}\backslash\overline{V_l}} (A_l + 1)\omega(h_k + 2)\, dx \\
&\leq \sum_{l=k+1}^{\infty} \omega(h_l + 2)(A_l + 1)\mathrm{meas}(V_{l-1}) \\
&\leq \sum_{l=k}^{\infty} 2^{-l} \\
&\leq 2^{-k+1}.
\end{aligned}
$$

So choosing $k_1 \geq 1$ such that $2^{-k_1+1} \leq \epsilon/2$, we have for large $k \geq k_0, k_1$, that

$$
0 \leq \mathscr{L}(u_k) - \mathscr{L}(u) \leq \int_{V_k} L(x, u_k(x), u_k'(x))\, dx + \int_{V_k} L(x, u(x), u'(x))\, dx \leq \epsilon
$$

as required. $\square$

Chapter 4

Universal singular sets

4.1 Introduction

The superlinearity condition required for existence of minimizers should prevent infinite derivatives from occurring too often in solutions to minimization problems. To make this observation more precise, Ball and Nadirashvili [1993] introduced the *universal singular set* of a Lagrangian, which records for which points (x, y) in the plane there exists an interval $[a, b]$ and choice of boundary conditions $a = A$, $b = B$, such that the problem (1.2) has a minimizer with graph passing through (x, y) with infinite derivative.

Precisely, Ball and Nadirashvili [1993] give the following definition, for Lagrangians $L \in C^3(\mathbb{R}^3)$ satisfying $L_{pp} > 0$ and with superlinear growth in p for all points $(x, y) \in \mathbb{R}^2$.

Definition 4.1. The universal singular set of a Lagrangian $L \colon \mathbb{R}^3 \to \mathbb{R}$, which we shall write $\mathrm{uss}(L)$, is defined as those points $(x_0, y_0) \in \mathbb{R}^2$ where one can find an interval $[a, b]$ in $\mathbb{R}$ containing x_0 and a choice of boundary conditions $u(a) = A$, $u(b) = B$, such that there is a minimizer $u \in \mathrm{AC}(a, b)$ of the associated variational problem (1.2) with $u(x_0) = y_0$ and $|u'(x_0)| = \infty$.

Under these assumptions, in particular that $L \in C^3(\mathbb{R}^3)$, Ball and Mizel showed that this set is of the first Baire category. Sychëv [1994] lowered the smoothness assumption to $L \in C^1(\mathbb{R}^3)$, and showed in this situation that the universal singular set has zero two-dimensional measure.

4.1.1 Greater generality and geometric properties

Csörnyei et al. [2008] work in a more general setting, assuming no convexity of the Lagrangian, and hence no standard existence theory. They introduce a natural idea of generalized minimizers, and define the universal singular set with reference to these new objects. Throughout their paper, Lagrangians are assumed to be Borel measurable, bounded below and locally bounded above, and superlinear in p. For this discussion we assume this to be the setting. They introduce the notation, for a Lagrangian $L\colon \mathbb{R}^3 \to \mathbb{R}$, interval $[a,b] \subseteq \mathbb{R}$, and real numbers $A, B \in \mathbb{R}$,

$$\mathscr{L}(a, A; b, B) = \inf\{\mathscr{L}(u) : u \in \mathrm{AC}(a,b),\ u(a) = A,\ u(b) = B\}.$$

The *excess* of a function $u \in \mathrm{AC}(a,b)$ is the defined by

$$\mathscr{E}(u; a, b) = \int_a^b L(x, u(x), u'(x))\, dx - \mathscr{L}(a, u(a); b, u(b));$$

this measures in a natural way how far the function u is from being a minimizer with respect to its own boundary conditions.

Definition 4.2 (Generalized minimizer)**.** $u \in C(a,b)$ is a *generalized minimizer* for L on $[a,b]$ if the restriction of u to (a,b) is a locally uniform limit of some functions $u_n \in \mathrm{AC}(a_n, b_n)$ such that $\mathscr{E}(u_n; a_n, b_n) \to 0$.

Such functions are necessarily absolutely continuous. Different extensions of the concept of minimizer are required to ensure existence results in the non-convex theory, but it is shown that when the universal singular set is defined with reference to the existence of minimizers in these various senses, the resulting set is the same whatever precise notion was used. Moreover, in the convex case, the set is the same as that given by considering minimizers in the standard sense. So the distinction is not an important one as far as the results are concerned, but it should be noted that without assuming convexity of L in p, the notion of universal singular set is well-defined, and compatible with the standard definition in the convex case.

Csörnyei et al. [2008] investigate the geometric as well as topological properties of universal singular sets in this more general setting.

They showed that universal singular sets intersect most absolutely continuous curves in sets of zero linear measure, the exceptions being some curves with vertical tangents.

Theorem 4.3 (Csörnyei et al. [2008])**.** Let $L\colon \mathbb{R}^3 \to \mathbb{R}$ be a Lagrangian.

Then the graph of any absolutely continuous function, and any vertical line, meets the universal singular set of L in a set of linear measure zero.

By Fubini's theorem, this implies Sychëv's result on the two-dimensional measure of the universal singular set in a more general situation in terms of smoothness and convexity assumptions.

The result on category also then follows under low smoothness assumptions, although an extra condition on the modulus of continuity is imposed.

Theorem 4.4 (Csörnyei et al. [2008])**.** Let $L\colon \mathbb{R}^3 \to \mathbb{R}$ be a Lagrangian satisfying the Lipschitz condition (L) in the second variable (see Chapter 2).

Then the universal singular set is a countable union of closed sets. In particular, it is first category.

The final statement of this theorem follows since we know that the measure zero result holds in this more general case. The Lipschitz condition is in fact necessary:

Theorem 4.5 (Csörnyei et al. [2008])**.** Let $\omega\colon \mathbb{R} \to [0,\infty)$ be superlinear and such that $\omega(0) = 0$.

Then there is a continuous Lagrangian $L\colon \mathbb{R}^3 \to \mathbb{R}$ such that $L(x,y,p) \geq \omega(p)$ for all $(x,y,p) \in \mathbb{R}^3$ with universal singular set residual in $\mathbb{R}^2$.

Intersections of universal singular sets with absolutely continuous curves (i.e. not necessarily graphs of functions) are nearly always null, but some curves with vertical tangents cannot be outlawed.

Theorem 4.6 (Csörnyei et al. [2008])**.** Let ω be a given even, convex superlinearity. Suppose an absolutely continuous curve $\gamma\colon [a,b] \to \mathbb{R}^2$, $\gamma(t) = (x(t), y(t))$, is such that for almost all $t \in [a,b]$, one of the following holds:

$$\limsup_{s\to t} \left| \frac{y(s)-y(t)}{x(s)-x(t)} \right| < \infty$$

or

$$\liminf_{s\to t} |x(s)-x(t)|\omega\left(\frac{y(s)-y(t)}{x(s)-x(t)}\right) > 0.$$

(When $x(s) = x(t)$, we interpret $\frac{y(s)-y(t)}{x(s)-x(t)}$ as zero and $|x(s)-x(t)|\omega\left(\frac{y(s)-y(t)}{x(s)-x(t)}\right)$ as ∞.)

Then for any Lagrangian with superlinearity ω, the curve $\gamma(a,b) \subseteq \mathbb{R}^2$ meets the universal singular set of L in a set of linear measure zero.

Furthermore, there is no hope of outlawing those with vertical tangents, even under the classical assumptions.

Theorem 4.7 (Csörnyei et al. [2008])**.** Let $\omega \in C^\infty(\mathbb{R})$ be superlinear, strictly convex, and such that $\omega(p) \geq \omega(0) = 0$ for all $p \in \mathbb{R}$.

Then there exists a rectifiable compact set $S \subseteq \mathbb{R}^2$ of positive linear measure and a Lagrangian $L \in C^\infty(\mathbb{R}^3)$, strictly convex and with the prescribed superlinearity in p, such that $S \subseteq \mathrm{uss}(L)$.

They also showed that any set covered by universal singular sets of Lagrangians with arbitrary superlinearity is purely unrectifiable.

Theorem 4.8 (Csörnyei et al. [2008])**.** Let $E \subseteq \mathbb{R}^2$ be such that for any superlinearity ω there is a Lagrangian $L\colon \mathbb{R}^3 \to \mathbb{R}$ with this prescribed superlinearity such that the universal singular set of the Lagrangian contains E.

Then E is purely unrectifiable.

This result is nearly optimal, in the following sense.

Theorem 4.9 (Csörnyei et al. [2008])**.** Let $\omega \in C^\infty(\mathbb{R})$ be superlinear, strictly convex, and such that $\omega(p) \geq \omega(0) = 0$ for all $p \in \mathbb{R}$, and let $S \subseteq \mathbb{R}^2$ be a compact purely unrectifiable set.

Then there exists a smooth Lagrangian $L = L(x, y, p)$, strictly convex and with the prescribed superlinearity ω in p, such that the universal singular set of L contains S.

In this chapter we show that this final result is also true of F_σ purely unrectifiable sets. Thus we are near to a complete characterization of such sets. A natural converse to this new result would be that any set E which can be covered by universal singular sets of smooth Lagrangians with arbitrary superlinearity must admit a purely unrectifiable F_σ cover. That this might be true seems plausible: by Theorem 4.8 E is purely unrectifiable, and moreover, since the Lagrangians are smooth, each universal singular set is F_σ, by Theorem 4.4. However, it is not true in general that these universal singular sets are purely unrectifiable, see Theorem 4.7 for a counterexample. It is not currently known whether E must in fact admit an F_σ purely unrectifiable cover.

4.2 Towards a characterization

4.2.1 Preliminaries

Given $(a, A), (b, B) \in \mathbb{R}^2$, we let $Q(a, A; b, B)$ denote the smallest closed rectangle in $\mathbb{R}^2$ with two vertices at (a, A) and (b, B) and sides parallel to the coordinate axes (we admit the possibility that this contains zero area).

We recall that a set $S \subseteq \mathbb{R}^2$ is *purely unrectifiable* if it meets every Lipschitz curve $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ in a set of linear measure zero.

We shall call a function $\omega \in C^\infty(\mathbb{R})$ a *superlinearity* if

- $\omega(p) \geq \omega(0) = 0$ for all $p \in \mathbb{R}$;
- ω is strictly convex; and
- (superlinearity) $\omega(p)/|p| \to \infty$ as $|p| \to \infty$.

For this chapter, a *Lagrangian* shall be a function $L = L(x, y, p) \colon \mathbb{R}^3 \to \mathbb{R}$, of class C^∞, superlinear and strictly convex in p, where here superlinear means that for some superlinearity ω, $L(x, y, p) \geq \omega(p)$ for all $(x, y, p) \in \mathbb{R}^3$. By Theorem 1.1, these assumptions suffice to guarantee existence and partial regularity of a solution to the minimization problem (1.2) over those $u \in \mathrm{AC}(a, b)$ satisfying $u(a) = A$ and $u(b) = B$.

All of our Lagrangians will be of the form $L(x, y, p) = F(x, y, p) + \omega(p)$, for functions $F \colon \mathbb{R}^3 \to \mathbb{R}$ satisfying the following conditions, which we shall refer to as $(\star_F)$:

$(\star_1)$ $F \in C^\infty(\mathbb{R}^3)$;

$(\star_2)$ $F \geq 0$ on $\mathbb{R}^3$ and $F(x, y, 0) = 0$ for all $(x, y) \in \mathbb{R}^2$; and

$(\star_3)$ $p \mapsto F(x, y, p)$ is convex for each fixed $(x, y) \in \mathbb{R}^2$.

We shall say a Lagrangian L of this form is of form $(\star)$ (so this terminology agrees with that of Csörnyei et al. [2008]).

In this chapter we prove the following theorem.

Theorem 4.10. Let ω be a given superlinearity, and let $S \subseteq \mathbb{R}^2$ be an F_σ purely unrectifiable set.

Then there exists a Lagrangian L of form $(\star)$ with the prescribed superlinearity ω such that the universal singular set of L contains S.

We first record a straightforward construction of a cut-off function for sets which do not quite satisfy the usual compact containment requirement.

Lemma 4.11. Let $S \subseteq \mathbb{R}^2$ be closed, $V \subseteq \mathbb{R}^2$ be open, $W \subseteq V$ be bounded and such that $\overline{W} \backslash V \subseteq S$.

Then there exists $\phi \colon \mathbb{R}^2 \to [0, \infty)$ such that

- $\phi \in C^\infty(\mathbb{R}^2 \backslash ((\partial V) \cap S))$;
- $\phi = 1$ on W; and

- $\phi = 0$ off V.

Proof. For each $n \geq 1$, define

$$W_n = \{x \in W : \text{dist}(x, \mathbb{R}^2 \backslash V) \in [1/n, 1/(n-1))\}.$$

Then $W = \bigcup_{n=1}^{\infty} W_n$ and each $\overline{W_n}$ is a compact set contained in V. Thus for each $n \geq 1$ we can choose open Y_n such that $\overline{W_n} \subseteq Y_n \subseteq B_{1/n}(W_n) \cap V$. Then for any $y \in Y_n$, choosing $z \in W_n$ such that $\|y - z\|_2 < 1/n$, we have

$$\text{dist}(y, \mathbb{R}^2 \backslash V) \leq \|y - z\|_2 + \text{dist}(z, \mathbb{R}^2 \backslash V) < 1/n + 1/(n-1) < 2/(n-1). \quad (4.1)$$

We now show that $\{Y_n\}_{n=1}^{\infty}$ is a locally finite collection on $\mathbb{R}^2 \backslash ((\partial V) \cap S)$. So let $x \in \mathbb{R}^2 \backslash ((\partial V) \cap S)$.

Case i) $x \in V$. Choose $m \geq 1$ such that $B_{1/m}(x) \subseteq V$ and let $n \geq 4m + 1$. Then for $y \in Y_n$, $\text{dist}(y, \mathbb{R}^2 \backslash V) < 1/2m$ by (4.1), so

$$\|y - x\|_2 \geq \text{dist}(x, \mathbb{R}^2 \backslash V) - \text{dist}(y, \mathbb{R}^2 \backslash V) > 1/m - 1/2m = 1/2m.$$

Hence $Y_n \cap B_{1/2m}(x) = \emptyset$ for all $n \geq 4m + 1$.

Case ii) $x \notin \overline{V}$. Then since $Y_n \subseteq V$ for all $n \geq 1$, $\mathbb{R}^2 \backslash \overline{V}$ is an open set containing x which meets no Y_n.

Case iii) $x \in (\partial V) \backslash S$. Then by the assumption on W, $x \notin \overline{W}$. Choose $m \geq 1$ such that $B_{1/m}(x) \cap W = \emptyset$ and let $n \geq 2m$. Then for $y \in Y_n$, we choose $w_n \in W_n \subseteq W$ such that $\|y - w_n\|_2 < 1/n$ and argue that

$$\|x - y\|_2 \geq \|x - w_n\|_2 - \|w_n - y\|_2 > 1/m - 1/n \geq 1/m - 1/2m = 1/2m.$$

Hence $Y_n \cap B_{1/2m}(x) = \emptyset$ for $n \geq 2m$.

Thus $\{Y_n\}_{n=1}^{\infty}$ is indeed a locally finite collection on $\mathbb{R}^2 \backslash ((\partial V) \cap S)$. We now choose a sequence $\{\phi_n\}_{n=1}^{\infty}$ of functions $\phi_n \in C^{\infty}(\mathbb{R}^2)$ such that

- $0 \leq \phi_n \leq 1$;
- $\text{spt}(\phi_n) \subseteq Y_n$; and
- for $x \in W_n$, $\sum_{i=1}^{m} \phi_i(x) = 1$ for all $m \geq n$.

This can be done by, for example, choosing for each $n \geq 1$ a function $\psi_n \in C^{\infty}(\mathbb{R}^2)$

such that $0 \leq \psi_n \leq 1$, $\psi_n = 1$ on W_n, and $\operatorname{spt}(\psi_n) \subseteq Y_n$, and then defining

$$\phi_n = \psi_n \prod_{i=1}^{n-1}(1 - \psi_i).$$

Clearly, $0 \leq \phi_n \leq 1$. Since $\operatorname{spt}(\psi_n) \subseteq Y_n$, we see that $\operatorname{spt}(\phi_n) \subseteq Y_n$ as required. One easily sees that

$$\sum_{i=1}^{n} \phi_i = 1 - \prod_{i=1}^{n}(1 - \psi_i),$$

so for $x \in W_n$ and $m \geq n$, we see that since $\psi_n(x) = 1$ by choice,

$$\sum_{i=1}^{m} \phi_i(x) = 1 - 0 = 1$$

as required. (See the partition of unity construction in e.g. Rudin [1966, Theorem 2.3] for essentially this construction.)

Then define $\phi \colon \mathbb{R}^2 \to \mathbb{R}$ by

$$\phi(x) = \begin{cases} \sum_{i=1}^{\infty} \phi_i(x) & x \in \mathbb{R}^2 \backslash ((\partial V) \cap S) \\ 0 & x \in (\partial V) \cap S. \end{cases}$$

Then since $\operatorname{spt}(\phi_n) \subseteq Y_n$, and $\{Y_n\}_{n=1}^{\infty}$ is a locally finite collection on $\mathbb{R}^2 \backslash ((\partial V) \cap S)$, ϕ is well-defined and $\phi \in C^{\infty}(\mathbb{R}^2 \backslash ((\partial V) \cap S))$. Clearly, ϕ is non-negative. Let $x \in W$. Then $x \in W_n$ for some $n \geq 1$, and since $W \subseteq V$, we know $x \notin \partial V$, so

$$\phi(x) = \sum_{i=1}^{\infty} \phi_i(x) = \lim_{m \to \infty} \sum_{i=1}^{m} \phi_i(x) = 1.$$

If $\phi(y) \neq 0$, then $\phi_n(y) \neq 0$ for some $n \geq 1$, thus $y \in \operatorname{spt}(\phi_n) \subseteq Y_n \subseteq V$. Taking the contrapositive, we see that if $y \notin V$, then $\phi(y) = 0$. Hence ϕ is as required. $\square$

4.2.2 The construction: general discussion

Suppose $S = \bigcup_{n=1}^{\infty} S_n$, where each S_n is compact and purely unrectifiable. We construct by induction a sequence of Lagrangians L_n such that for each $n \geq 1$ we have $\operatorname{uss}(L_n) \supseteq \bigcup_{m=1}^{n} S_m$. We discuss how to do this so that the L_n converge to a Lagrangian L with $\operatorname{uss}(L) \supseteq S$.

Fix a point $(x_0, y_0) \in S_n \backslash \bigcup_{m=1}^{n-1} S_m$. We construct Lagrangian L_n and function $\Phi_n \in C(\mathbb{R}^2) \cap C^{\infty}(\mathbb{R}^2 \backslash S_n)$ such that there is a rectangular neighbourhood $Q(a_0, A_0; b_0, B_0)$ of (x_0, y_0) such that for any $u \in \operatorname{AC}(a_0, b_0)$ with graph lying in

$Q(a_0, A_0; b_0, B_0)$, we have

$$\int_{a_0}^{b_0} L_n(x, u(x), u'(x))\,dx \geq \Phi_n(U(b_0)) - \Phi_n(U(a_0)), \tag{4.2}$$

with equality if and only if $u'(x) = \psi_n(x, u(x))$ almost everywhere, where we ensure $\psi_n := -2(\Phi_n)_x/(\Phi_n)_y$ is well-defined on $\mathbb{R}^2 \backslash S_n$. We then solve the ordinary differential equation $u_0'(x) = \psi_n(x, u_0(x))$ for a locally absolutely continuous $u_0\colon \mathbb{R} \to \mathbb{R}$ with $u_0(x_0) = y_0$. If Φ_n was constructed so that $\psi_n(x, y) \to \infty$ as $\operatorname{dist}((x, y), S_n) \to 0$, we then have $u_0'(x) \to \infty$ as $x \to x_0$. Moreover, by a trick from Csörnyei et al. [2008] which uses properties of Φ_n and S_n, and which we have already seen in Chapter 3, inequality (4.2) suffices to prove that u_0 is a minimizer with respect to its own boundary conditions on $[a_0, b_0]$. This shows that $(x_0, y_0) \in \operatorname{uss}(L_n)$.

Let $m \geq n$. If we have constructed our Lagrangians so that firstly $L_m \geq L_n$, we have

$$\int_{a_0}^{b_0} L_m(x, u(x), u'(x))\,dx \geq \int_{a_0}^{b_0} L_n(x, u(x), u'(x))\,dx \geq \Phi_n(U(b_0)) - \Phi_n(U(a_0))$$

for all $u \in \mathrm{AC}(a_0, b_0)$. If secondly we can guarantee that $L_m(x, u_0, u_0') = L_n(x, u_0, u_0')$ for almost every $x \in (a_0, b_0)$, where u_0 is the solution of the ODE mentioned above, then we have that

$$\int_{a_0}^{b_0} L_m(x, u_0(x), u_0'(x))\,dx = \int_{a_0}^{b_0} L_n(x, u_0(x), u_0'(x))\,dx = \Phi_n(U_0(b_0)) - \Phi_n(U_0(a_0)).$$

Thus u_0 is a minimizer of the functional given via Lagrangian L_m over $\mathrm{AC}(a_0, b_0)$ with respect to its own boundary conditions. Assuming the Lagrangians L_n converge pointwise to a Lagrangian L, we let $m \to \infty$ in these two relations to see that u_0 is a minimizer of the functional given via Lagrangian L over $\mathrm{AC}(a_0, b_0)$ with respect to its own boundary conditions.

This outline of our strategy gives us two requirements at the inductive step of constructing L_n. The details of this inductive step mimic those of the original proof in Csörnyei et al. [2008]. For a given superlinearity ω, they construct a function $F\colon \mathbb{R}^3 \to \mathbb{R}$ of form $(\star_F)$ and define $L(x, y, p) = F(x, y, p) + \omega(p)$. The key observation to make about this proof when seeking to generalize it for our purposes is that ω may be regarded just as a Lagrangian depending only on p. Or rather, the role of ω may be taken by any Lagrangian strictly convex and superlinear in p, with partial derivatives with respect to p replacing any ω' terms. In particular, the argument may be applied to a previously constructed L_{n-1}. The arguments of Csörnyei et al. then tell us how to construct an F_n of form $(\star_F)$ to add to this L_{n-1}, via the construction of a potential $\Phi_n \in C(\mathbb{R}^2) \cap C^\infty(\mathbb{R}^2 \backslash S_n)$. The considerations of the preceding paragraphs mean the argument is rather more intricate, but the general strategy is the same.

Ensuring that $L_m \geq L_n$ for all $m \geq n$ is easy; this just requires the stipulation that each F_n is non-negative, which is already given by the methods of Csörnyei et al.. Harder is ensuring that for each point $(x_0, y_0) \in \bigcup_{m=1}^{n-1} S_m$, we have $L_n = L_{n-1}$ along the trajectory u_0 (except perhaps on a null set) on some fixed neighbourhood of x_0. The key fact here is that precisely by construction we know that $u_0' = \psi_m(x, u_0)$ for some $1 \leq m < n$, where $\psi_m \in C^\infty(\mathbb{R}^2 \backslash S_m)$, and therefore ψ_m is bounded on sets positively separated from S_m.

At this point it is easiest to first suppose that the $\{S_n\}_{n=1}^\infty$ are pairwise disjoint. Thus S_n is positively separated from $\bigcup_{m=1}^{n-1} S_m$, so we can choose a neighbourhood H_n of S_n on which ψ_m is bounded above for all $1 \leq m < n$, by M_n say. So to construct an appropriate L_n, the condition is now just that F_n is only non-zero on $H_n \times (M_n, \infty)$. A straightforward use of a cut-off function on $\mathbb{R}^2$ ensures $F_n(x, y, p) = 0$ for $(x, y) \notin H_n$. The demand that $F_n(x, y, p) = 0$ if $(x, y, p) \in H_n \times (M_n, \infty)$ reduces to certain inequalities involving the derivatives of the potential Φ_n on the set H_n. These can be satisfied using a construction of the potential similar to the construction used by Csörnyei et al..

The existence of a pointwise limit $L(x, y, p) := \lim_{n\to\infty} L_n(x, y, p)$ is trivial if the lower bounds M_n tend to infinity: then for each fixed $(x, y, p) \in \mathbb{R}^3$, for large enough n, L_n does not change on a neighbourhood of (x, y, p), so the limit L exists and is smooth. The arguments sketched above show that $\operatorname{uss}(L) \supseteq \bigcup_{m=1}^\infty S_m$.

This discussion applies directly only to the disjoint case, but the spirit of the proof is retained in the full version. The issue in the general case is that we of course no longer have positive separation of our compact sets S_n, and hence cannot in general find an upper bound on S_n for the derivative of a minimizer u_0 witnessing $\bigcup_{m=1}^{n-1} S_m \subseteq \operatorname{uss}(L_{n-1})$. This problem can be overcome by covering $S_n \backslash \bigcup_{m=1}^{n-1} S_m$ with an increasing sequence of open sets $\{V_n^i\}_{i=1}^\infty$, each positively separated from $\bigcup_{m=1}^{n-1} S_m$. For each $i \geq 1$, on V_n^i there is an upper bound, M_n^i say, of ψ_m for all $1 \leq m < n$. Our requirement now is that F_n is non-zero only on $\bigcup_{i=1}^\infty (V_n^i \times (M_n^i, \infty))$. As $i \to \infty$, we shall have $\operatorname{dist}(V_n^i, \bigcup_{m=1}^{n-1} S_m) \to 0$, and thus $\sup_{V_n^i} \psi_m \to \infty$. So we have $M_n^i \to \infty$ as $i \to \infty$, i.e. the closer in the plane we get to $\bigcup_{m=1}^{n-1} S_m$, the higher in the third coordinate of $\mathbb{R}^3$ we must go before we may alter L_{n-1}. Thus we can think of the region of permitted change to L_{n-1} as being a "cylinder" in $\mathbb{R}^3$ with "base sloping up to infinity" as we approach $\bigcup_{m=1}^{n-1} S_m$.

Throughout the proof we adhere to the indexing suggested above. Subscripts such as m, n refer to the inductive step. Superscripts such as i, j, k, l are used to index sequences of objects discussed within the argument at a fixed inductive step. This superscript notation is retained even in arguments presented independently of the induction (e.g. in Lemma 4.12) to avoid confusion.

4.2.3 The construction: details

Our first result is a modified version of Lemma 11 from Csörnyei et al. [2008]. This tells us how to modify a given Lagrangian so as to include new points in its universal singular set, but without changing it on certain "cylinders" in $\mathbb{R}^3$. We try to motivate the exact assumptions made in the next lemma by sketching its role in the inductive construction of L_n. First we note that the set G does not appear in the conclusions, only in the assumptions regarding Φ. G_n will be chosen to be a bounded open cover of S_n, but there is no loss of understanding in assuming $G = \mathbb{R}^2$ for this first lemma. We choose a sequence $\{V_n^i\}_{i=1}^{\infty}$ covering $S_n \setminus \bigcup_{m=1}^{n-1} S_m$, but such that each V_n^i is positively separated from $\bigcup_{m=1}^{n-1} S_m$, and also a sequence of upper bounds $\{M_n^i\}_{i=1}^{\infty}$ of ψ_m $(1 \leq m < n)$ on V_n^i, and thus a sequence of "cylinders" $\{V_n^i \times (M_n^i, \infty)\}_{i=1}^{\infty}$ in $\mathbb{R}^3$. Our goal, as discussed above, is to construct a function F_n of form $(\star_F)$ which is zero off these sets. We show, just as in Csörnyei et al. [2008], that such an F_n is given by a potential $\Phi_n \in C(\mathbb{R}) \cap C^{\infty}(\mathbb{R}^2 \setminus S_n)$, where the derivatives of Φ_n satisfy certain inequalities. The inequalities we require are similar to but more complicated than those from Csörnyei et al. [2008], since we demand also some information about our resulting function F_n on the sets $V_n^i \times (M_n^i, \infty)$. We also need to fix a neighbourhood W_n of $S_n \setminus \bigcup_{m=1}^{n-1} S_m$ which will contain the graphs of $u_0 \in \mathrm{AC}(a_0, b_0)$ which witness that $(x_0, y_0) \in \mathrm{uss}(L_n)$ for each $(x_0, y_0) \in S_n \setminus \bigcup_{m=1}^{n-1} S_m$. Since L_n is already determined on $\bigcup_{m=1}^{n-1} S_m$, we keep this neighbourhood W_n in some sense as far from $\bigcup_{m=1}^{n-1} S_m$ as possible. Ideally (viz in the disjoint case) we would have that W_n is compactly contained in $V_n := \bigcup_{i=1}^{\infty} V_n^i$, but since W_n must cover $S_n \setminus \bigcup_{m=1}^{n-1} S_m$ and V_n must not intersect $\bigcup_{m=1}^{n-1} S_m$, this is not in general possible. The best we can ask for is that W_n does not approach the boundary of V_n unless it is required to do so to cover all the points of S_n, hence the condition on $\overline{W}$ below.

Since both this result and Lemma 3.18 in Chapter 3 are versions of the same result, Lemma 11 from Csörnyei et al. [2008], there are some similarities between the two results. In particular we shall again use the corner-smoothing result of Lemma 3.17.

Lemma 4.12. Let $F\colon \mathbb{R}^3 \to \mathbb{R}$ be of form $(\star_F)$, $S \subseteq \mathbb{R}^2$ be compact, and $G \supseteq S$ be open. Let $L(x,y,p) = \omega(p) + F(x,y,p)$, where ω is a given superlinearity.

Suppose further that the function $\Phi \in C^\infty(\mathbb{R}^2 \backslash S) \cap C(\mathbb{R}^2)$, sequence $\{V^i\}_{i=1}^\infty$ of sets $V^i \subseteq \mathbb{R}^2$, and sequence of non-negative constants $\{M^i\}_{i=1}^\infty$ are such that $V := \bigcup_{i=1}^\infty V^i$ is open and bounded, $\overline{V} \subseteq G$, and the following conditions hold:

(4.12.1) Φ is decreasing in x and increasing in y on $\mathbb{R}^2$;

(4.12.2) $-\Phi_x(x,y) \geq (2M^i + 4)\Phi_y(x,y)$ for $(x,y) \in V^i \backslash S$ for all $i \geq 1$, and $-\Phi_x(x,y) \geq 4\Phi_y(x,y) > 0$ for $(x,y) \in \mathbb{R}^2 \backslash S$;

(4.12.3) $\Phi_y(x,y) \geq 4L_p(x,y,(-2\Phi_x/\Phi_y)(x,y))$ for $(x,y) \in G \backslash S$;

(4.12.4) $\lim_{0 < \operatorname{dist}((x,y),S) \to 0} (\Phi_x/\Phi_y)(x,y) = -\infty$;

(4.12.5) for all $a < b$ and non-decreasing $u \in \mathrm{AC}(a,b)$, the sets $\{x : U(x) \in S\}$ and $\{(\Phi \circ U)(x) : U(x) \in S\}$ are Lebesgue null.

Then for any $W \subseteq V$ such that $\overline{W} \backslash V \subseteq S$, there exists $\hat{F}\colon \mathbb{R}^3 \to \mathbb{R}$ of form $(\star_F)$ with the following properties:

(4.12.6) $\hat{F} \geq F$ on $\mathbb{R}^3$;

(4.12.7) $\hat{F} = F$ on $\mathbb{R}^3 \backslash \bigcup_{i=1}^\infty (V^i \times (M^i, \infty))$; and

(4.12.8) $\hat{L}\colon \mathbb{R}^3 \to \mathbb{R}$ defined by

$$\hat{L}(x,y,p) = \omega(p) + \hat{F}(x,y,p)$$

has the property that for all $a < b$ and all $u \in \mathrm{AC}(a,b)$ such that $Q(a,u(a);b,u(b)) \subseteq W$, we have

$$\int_a^b \hat{L}(x,u(x),u'(x))\,dx \geq \Phi(U(b)) - \Phi(U(a)),$$

with equality if and only if $u'(x) = (-2\Phi_x/\Phi_y)(x,u(x))$ for almost every $x \in [a,b]$.

Proof. We mimic the proof of Lemma 11 in Csörnyei et al. [2008], but now working with $L = F + \omega$, in place of just ω. The main difference in our assumptions from those in the original lemma of Csörnyei et al. is the dependence of the inequality in (4.12.2) on the sets V_i. This is exactly the stronger information we need to guarantee the conclusion (4.12.7) which we now require.

Define $\psi \in C^\infty(\mathbb{R}^2 \backslash S)$ and $\theta, \xi \in C^\infty(G \backslash S)$ by

$$\psi = -2\Phi_x/\Phi_y, \ \theta = \Phi_y - L_p(x,y,\psi), \ \xi = \frac{-\Phi_x + L(x,y,\psi) - \psi L_p(x,y,\psi)}{\theta}.$$

Condition (4.12.2) ensures ψ is well-defined and strictly positive everywhere on $\mathbb{R}^2 \backslash S$. By properties of ω and F, we have for all $(x,y) \in \mathbb{R}^2 \backslash S$ that $L(x,y,0) = 0$ and L is strictly convex in p. Since $\psi > 0$, we know also by properties of ω that $\omega(\psi) > 0$. So using the mean value theorem, and property $(\star_2)$ of F, we have for all $(x,y) \in \mathbb{R}^2 \backslash S$ that

$$L_p(x,y,\psi) > \frac{L(x,y,\psi) - L(x,y,0)}{\psi - 0} \geq \frac{\omega(\psi)}{\psi} > 0. \tag{4.3}$$

So for $(x,y) \in G \backslash S$, we have by (4.12.3) that $\theta \geq 3L_p(x,y,\psi) > 0$ and hence that ξ is well-defined.

Fix $(x,y) \in G \backslash S$. Note that by our definitions of θ and ξ,

$$L(x,y,\psi) + (p - \psi)L_p(x,y,\psi) + \theta(p - \xi) = \Phi_x + p\Phi_y. \tag{4.4}$$

Also note that the strict convexity of L in p and the mean value theorem give us the relation

$$L(x,y,p) \geq L(x,y,\psi) + (p - \psi)L_p(x,y,\psi) \tag{4.5}$$

with equality if and only if $p = \psi$. By (4.3) and (4.12.3) we have

$$\Phi_y > \theta \geq \Phi_y - \Phi_y/4 = 3\Phi_y/4. \tag{4.6}$$

By (4.5) for case $p = 0$ and the fact that $L(x,y,0) = 0$, we have that

$$\xi\theta = -\Phi_x + L(x,y,\psi) - \psi L_p(x,y,\psi) < -\Phi_x. \tag{4.7}$$

Further, by (4.12.3), the definition of ψ, and the facts that $L \geq 0$ and $-\Phi_x > 0$, we also have

$$\begin{aligned}
\xi\theta &= -\Phi_x + L(x,y,\psi) - (-2\Phi_x/\Phi_y)L_p(x,y,\psi) \\
&\geq -\Phi_x + L(x,y,\psi) - (-2\Phi_x)/4 \\
&> -\Phi_x/2.
\end{aligned}$$

Hence, using (4.6) and the fact that $\Phi_y > 0$,

$$\xi > -\Phi_x/2\theta \geq -\Phi_x/2\Phi_y. \tag{4.8}$$

This implies, using (4.12.2), that

$$\xi \geq M^i + 2 > M^i + 1 \text{ if } (x,y) \in V^i \backslash S; \text{ and } \xi \geq 2 > 1 \text{ on } G \backslash S. \tag{4.9}$$

The latter gives, using the definition of ψ, (4.7), (4.6), and that $\xi\theta > 0$, that

$$\psi = -2\Phi_x/\Phi_y > 2\xi\theta/\Phi_y \geq 3\xi\Phi_y/2\Phi_y = \xi + \xi/2 \geq \xi + 1 \tag{4.10}$$

on $G\backslash S$. Since G is open and covers S, for (x,y) sufficiently close to S we have $(x,y) \in G$, so by (4.8) and (4.12.4) we have that

$$\lim_{0<\operatorname{dist}((x,y),S)\to 0} \xi \geq -\frac{1}{2}\left(\lim_{0<\operatorname{dist}((x,y),S)\to 0} \frac{\Phi_x}{\Phi_y}\right) = \infty. \tag{4.11}$$

We now use the corner-smoothing γ constructed in Lemma 3.17 to define $f\colon G \times \mathbb{R} \to \mathbb{R}$ by

$$f(x,y,p) = \begin{cases} \gamma(p, \xi(x,y), \theta(x,y)) & (x,y) \in G\backslash S \\ 0 & (x,y) \in S. \end{cases}$$

Evidently $f \geq 0$ on $G \times \mathbb{R}$ by (3.17.4). Since $\xi \geq 1$ on $G\backslash S$ from (4.9), property (3.17.2) of γ implies that $f(x,y,0) = 0$ for all $(x,y) \in G$. For fixed $(x,y) \in G$, that $p \mapsto f(x,y,p)$ is convex follows from (3.17.1). Clearly $f \in C^\infty((G\backslash S) \times \mathbb{R})$. But for given $p \in \mathbb{R}$, by (4.11) there is an open set Ω with $S \subseteq \Omega \subseteq G$ such that $\xi \geq p + 2$ on $\Omega\backslash S$. Hence $f = 0$ on $\Omega \times (-\infty, p+1)$ by (3.17.2). That is, for given $(x,y,p) \in S \times \mathbb{R}$, there is an open set $\Omega \times (-\infty, p+1)$ containing (x,y,p) on which $f = 0$. Hence $f \in C^\infty(G \times \mathbb{R})$.

Let $i \geq 1$ and suppose $(x,y,p) \in (V^i\backslash S) \times (-\infty, M^i]$. Then by (4.9) we see that $\xi > M^i + 1 \geq p + 1$, so $f(x,y,p) = 0$ by (3.17.2). Hence

$$f(x,y,p) = 0 \text{ for all } (x,y,p) \in \bigcup_{i=1}^{\infty} (V^i \times (-\infty, M^i]). \tag{4.12}$$

Let $W \subseteq V$ be such that $\overline{W}\backslash V \subseteq S$. By Lemma 4.11 we can find a function $\phi\colon \mathbb{R}^2 \to [0,\infty)$ such that $\phi \in C^\infty(\mathbb{R}^2\backslash((\partial V) \cap S))$, $\phi = 1$ on W and $\phi = 0$ off V; i.e. a cut-off function which necessarily fails to be smooth on $\overline{W}\backslash V$.

Choose an open set $G' \supseteq S \cup \overline{V}$ such that $\overline{G'} \subseteq G$, and define $\tilde{F} : \mathbb{R}^3 \to \mathbb{R}$ by

$$\tilde{F}(x,y,p) = \begin{cases} \phi(x,y)f(x,y,p) & (x,y) \in G' \\ 0 & (x,y) \notin G'. \end{cases}$$

We claim $\tilde{F} \in C^\infty(\mathbb{R}^3)$. Clearly $\tilde{F} \in C^\infty(((\mathbb{R}^2 \backslash \overline{G'}) \times \mathbb{R}) \cup ((G' \backslash ((\partial V) \cap S)) \times \mathbb{R}))$.

So consider first $(x,y,p) \in \partial G' \times \mathbb{R}$. Since $(x,y) \notin G' \supseteq \overline{V}$, we can find an open set Ω such that $(x,y) \in \Omega \subseteq \mathbb{R}^2 \backslash V$. So $\phi = 0$ on Ω, hence $\tilde{F} = 0$ on $\Omega \times \mathbb{R}$, hence $\tilde{F} \in C^\infty(\Omega \times \mathbb{R})$.

Consider now the case $(x,y,p) \in (G' \cap ((\partial V) \cap S)) \times \mathbb{R}$. By (4.11), there exists an open set Ω with $S \subseteq \Omega \subseteq G'$ such that $\xi \geq p+2$ on $\Omega \backslash S$. Since $(x,y) \in S$, we have $(x,y,p) \in \Omega \times (-\infty, p+1)$. By (3.17.2), $f = 0$ and hence $\tilde{F} = 0$ on $\Omega \times (-\infty, p+1)$. So $\tilde{F} \in C^\infty(\Omega \times (-\infty, p+1))$.

So indeed $\tilde{F} \in C^\infty(\mathbb{R}^3)$. That $\tilde{F}$ satisfies the remaining properties of $(\star_F)$ follows by the analogous properties proved above of f. Now define $\hat{F} = F + \tilde{F}$. Thus $\hat{F}$ is also of form $(\star_F)$. Property (4.12.6) follows since $\tilde{F}$ satisfies $(\star_2)$.

Let $(x,y,p) \in \mathbb{R}^3 \backslash \bigcup_{i=1}^\infty (V^i \times (M^i, \infty))$. If $(x,y) \notin V$, then $\tilde{F}(x,y,p) = 0$, by choice of ϕ. If $(x,y) \in V^i \subseteq G'$ for some $i \geq 1$, then $p \leq M^i$, and so $\tilde{F}(x,y,p) = 0$ by (4.12). Thus $\hat{F}$ satisfies (4.12.7).

We define $\hat{L}(x,y,p) = \omega(p) + \hat{F}(x,y,p)$ and are just required to check (4.12.8). So let $(x,y) \in W \backslash S$. Since $W \subseteq V \subseteq G'$ and $\phi = 1$ on W, we have by definition and (3.17.4) that

$$\tilde{F}(x,y,p) = f(x,y,p) = \gamma(p, \xi(x,y), \theta(x,y)) \geq \theta(x,y)(p - \xi(x,y)).$$

Hence by (4.5) and (4.4)

$$\begin{aligned} \hat{L}(x,y,p) = \omega(p) + \hat{F}(x,y,p) &= \omega(p) + F(x,y,p) + \tilde{F}(x,y,p) \\ &= L(x,y,p) + \tilde{F}(x,y,p) \\ &\geq L(x,y,p) + \theta(p - \xi) \\ &\geq L(x,y,\psi) + (p-\psi)L_p(x,y,\psi) + \theta(p-\xi) \\ &= \Phi_x + p\Phi_y. \end{aligned}$$

For the case $p = \psi(x,y)$, (4.10) and (3.17.3) imply that the first inequality above is an equality, as clearly the second is, thus $\hat{L}(x,y,\psi) = \Phi_x + \psi\Phi_y$. Moreover, should the equality $\hat{L}(x,y,p) = \Phi_x + p\Phi_y$ hold, then in particular the second inequality in the above calculation must be an equality, which by strict convexity of L forces

$p = \psi$. That is, we have equality in this inequality if and only if $p = \psi(x, y)$.

The remainder of the proof is very similar to that of Lemma 3.18 in Chapter 3. We sketch the structure and refer to this previous result for the details.

First suppose $a < b$ and $u \in \mathrm{AC}(a, b)$ is non-decreasing and such that $U([a, b]) \subseteq W$. Then since (4.12.5) states that $U(x) \notin S$ for almost every $x \in [a, b]$, $(\Phi \circ U)\colon [a, b] \to \mathbb{R}$ is differentiable almost everywhere, and with the above observations about $\hat{L}$ we see that for almost every $x \in [a, b]$,

$$\hat{L}(x, u(x), u'(x)) \geq (\Phi \circ U)'(x), \tag{4.13}$$

with equality if and only if $u'(x) = \psi(x, u(x))$. Note again that by (4.12.5), $(\Phi \circ U)$ has the Lusin property, i.e. maps null sets to null sets.

Given these observations, we now check (4.12.8). Let $a < b$ and $u \in \mathrm{AC}(a, b)$ be such that $Q(a, u(a); b, u(b)) \subseteq W$. Exactly as in Lemma 3.18, we can assume when checking (4.12.8) that u is non-decreasing and such that $\Phi(U(a)) \leq \Phi(U(b))$. That u is non-decreasing implies, since $Q(a, u(a); b, u(b)) \subseteq W$, that in fact $U([a, b]) \subseteq W$. So the relation (4.13) holds for almost every $x \in [a, b]$. Arguing again with $\{(a_j, b_j)\}_{j \in J}$, the (at most countable) sequence of components of $(a, b) \backslash U^{-1}(S)$ such that $\Phi(U(a_j)) < \Phi(U(b_j))$, we see that

$$\begin{aligned} \int_a^b \hat{L}(x, u(x), u'(x))\, dx &\geq \sum_{j \in J} \int_{a_j}^{b_j} \hat{L}(x, u(x), u'(x))\, dx \\ &\geq (\Phi \circ U)(b) - (\Phi \circ U)(a). \end{aligned}$$

Equality in this relation implies that $\hat{L}(x, u(x), u'(x)) = (\Phi \circ U)'(x)$ for almost every $x \in (a, b) \backslash U^{-1}(S)$. By (4.13) and (4.12.5) this implies that $u'(x) = \psi(x, u(x))$ for almost every $x \in [a, b]$.

Conversely, $u'(x) = \psi(x, u(x))$ almost everywhere implies

$$(\Phi \circ U)'(x) = (\Phi_x \circ U)(x) + \psi(x, u(x))(\Phi_y \circ U)(x) = (-\Phi_x \circ U)(x) \geq 0$$

almost everywhere. This, combined with the fact that $(\Phi \circ U)$ has the Lusin property, implies that $(\Phi \circ U)$ is absolutely continuous, and moreover, (4.13) gives that $\hat{L}(x, u(x), u'(x)) = (\Phi \circ U)'(x)$ almost everywhere. Hence

$$\int_a^b \hat{L}(x, u(x), u'(x))\, dx = \int_a^b (\Phi \circ U)'(x)\, dx = (\Phi \circ U)(b) - (\Phi \circ U)(a)$$

as required. □

We now give the construction of the potential required for an application of this lemma. This is a version of the proof of Theorem 10 in Csörnyei et al. [2008], i.e. the construction of a potential satisfying the conditions of their Lemma 11. This is done entirely independently of the sequence of constants $\{M^i\}_{i=1}^{\infty}$, which are therefore taken to be arbitrary. We then simply define subsets $\{V^i\}_{i=1}^{\infty}$ of $\mathbb{R}^2$ so that the required inequalities hold. The final statement (4.15.3) falls naturally out of the proof in Csörnyei et al. [2008]; it is only now in our case that it is relevant to emphasize it.

As part of the proof of Lemma 4.15 we recall Lemmas 12 and 13 stated and proved in Csörnyei et al. [2008], which are used to prove our statement in exactly the same way as they are used by Csörnyei et al.. The only difference in our presentation here is that we use Euclidean norms rather than supremum norms on $\mathbb{R}^2$, but this involves no change in either the proofs or the applications of the results. We do not give the proofs. The second lemma follows easily from the first, in exactly the same way that Lemma 3.21 follows from Lemma 3.20 in Chapter 3. The first relies on using the pure unrectifiability of S to find, given $\epsilon > 0$ and $C > 0$, an open set Ω around S such that the graph of any Lipschitz function from $\mathbb{R}$ to $\mathbb{R}$ with Lipschitz constant less than C intersects Ω in a set of length at most ϵ.

For two vectors $x, y \in \mathbb{R}^2$, we let $[x, y]$ denote the line segment in $\mathbb{R}^2$ with these points as endpoints.

Lemma 4.13. Let $S \subseteq \mathbb{R}^2$ be a compact purely unrectifiable set, $e \in \mathbb{R}^2$, and $\tau > 0$.

Then there is $g \in C^\infty(\mathbb{R}^2)$ such that

- $0 \leq g(x) \leq \tau$ for all $x \in \mathbb{R}^2$;
- $\operatorname{dist}(\nabla g(x), [0, e]) < \tau$ for all $x \in \mathbb{R}^2$; and
- $\sup_{x \in S} \|\nabla g(x) - e\|_2 < \tau$.

Lemma 4.14. Let $S \subseteq \mathbb{R}^2$ be a compact purely unrectifiable set, $\Omega \supseteq S$ be open, $h^0 \in C^\infty(\mathbb{R}^2)$, $e^0, e^1 \in \mathbb{R}^2$, and $\epsilon > 0$.

Then there is $h^1 \in C^\infty(\mathbb{R}^2)$ such that

- $\|h^1 - h^0\|_\infty < \epsilon$;
- $h^1 = h^0$ outside Ω;
- $\operatorname{dist}(\nabla h^1(x), [e^0, e^1]) < \epsilon + \|\nabla h^0(x) - e^0\|_2$ for $x \in \mathbb{R}^2$; and
- $\|\nabla h^1(x) - e^1\|_2 < \epsilon + \|\nabla h^0(x) - e^0\|_2$ for $x \in S$.

Lemma 4.15. Let $S \subseteq \mathbb{R}^2$ be compact and purely unrectifiable, $G, H \subseteq \mathbb{R}^2$ be bounded such that H is open and $\overline{H} \subseteq G$, and $\{M^i\}_{i=1}^{\infty}$ be a sequence of constants. Let $F\colon \mathbb{R}^3 \to \mathbb{R}$ be such that F_p exists and is bounded above on $G \times [8, n]$ for each $n \geq 9$. Let $L(x, y, p) = \omega(p) + F(x, y, p)$, where ω is a given superlinearity.

Then there is $\Phi \in C^\infty(\mathbb{R}^2 \backslash S) \cap C(\mathbb{R}^2)$ and a sequence $\{V^i\}_{i=1}^{\infty}$ of open sets $V^i \subseteq \mathbb{R}^2$ such that the conditions (4.12.1)–(4.12.5) of Lemma 4.12 hold, and

(4.15.1) $H \cap S \subseteq V := \bigcup_{i=1}^{\infty} V^i \subseteq \overline{V} \subseteq G$;

(4.15.2) $V^i \subseteq \{(x, y) \in H : \operatorname{dist}((x, y), \mathbb{R}^2 \backslash H) > 1/i\}$ for all $i \geq 1$; and

(4.15.3) $\psi \in C^\infty(\mathbb{R}^2 \backslash S)$ defined by $\psi := -2\Phi_x/\Phi_y$ is bounded above on any subset of $\mathbb{R}^2$ positively separated from S.

Proof. We use a slight variant of the construction which comprises the proof of Theorem 10 in Csörnyei et al. [2008].

We define an increasing sequence $\{c^k\}_{k=0}^{\infty}$ by, for each $k \geq 0$, choosing $c^k \geq 0$ such that $L_p(x, y, p) \leq c^k$ for all $(x, y, p) \in G \times [8, 5 \cdot 2^{k+4}]$. We now define

$$B^k = 4 + 4c^k \text{ and } A^k = 3 \cdot 2^{k+2} B^k.$$

The construction of Φ is then similar to that in Csörnyei et al. [2008], with these new definitions of A^k and B^k. We sketch the proof; and refer to Csörnyei et al. for more details. The construction we give here is closer to this original proof than the version given in Chapter 3, since we are now free to choose Ω^k at each stage, whereas in the previous chapter they were fixed. The construction relies on the exhibiting of a sequence, for $k \geq 0$, of functions $\Phi^k \in C^\infty(\mathbb{R}^2)$, open sets Ω^k, and $\epsilon^k > 0$ such that, where $\eta^k = 1 - 2^{k-1}$,

$$\Phi^0(x) = -A^0 x + B^0 y, \quad \Omega^0 = \mathbb{R}^2, \quad \epsilon^0 = 1/4; \tag{4.14}$$

$$\|\nabla\Phi^k(x) - e^k\|_2 < \eta^k \text{ for } x \in \overline{\Omega^k}; \tag{4.15}$$

if $a < b$, $u \in C([a, b])$ is non-decreasing and $\Phi \in C(\mathbb{R}^2)$ satisfies $\|\Phi - \Phi^k\|_\infty < 2\epsilon^k$, then

$$\operatorname{meas}(\{(\Phi \circ U)(x) : U(x) \in \Omega^k\}) \leq 1/k; \tag{4.16}$$

and for $k \geq 1$

$$\|\Phi^k - \Phi^{k-1}\|_\infty < \epsilon^{k-1}; \tag{4.17}$$

$$\Phi^1 = \Phi^0 \text{ off } B_1(S), \quad \text{and for } k \geq 2, \quad \Phi^k = \Phi^{k-1} \text{ outside } \overline{\Omega^{k-1}}; \tag{4.18}$$

$$\operatorname{dist}(\nabla\Phi^k(x), [e^{k-1}, e^k]) < \eta^k \text{ for } x \in \overline{\Omega^{k-1}}; \tag{4.19}$$

$$S \subseteq \Omega^k, \quad \overline{\Omega^k} \subseteq B_{2^{-k}}(S) \cap \Omega^{k-1}, \quad \epsilon^k \leq \epsilon^{k-1}/2. \tag{4.20}$$

(Interpret $1/0$ as ∞ in (4.16).)

This can be done inductively, using (4.14) to define Φ^0, Ω^0, and ϵ^0, and for $k \geq 1$ applying Lemma 4.14 with $\Omega = \Omega^{k-1}$ (except for $k = 1$ when we put $\Omega = B_1(S)$), $h^0 = \Phi^{k-1}$, $e^0 = e^{k-1}$, $e^1 = e^k$, $\epsilon = \epsilon^{k-1}$ and defining $\Phi^k = h^1$. Properties (4.17)–(4.18) are immediate from the definition of Φ^k, and (4.19) follows by induction, as in Chapter 3. Defining $\Omega^k = B_\delta(S)$ for sufficiently small $\delta > 0$, chosen using the pure unrectifiability of S, and defining $\epsilon^k = \min\{\epsilon^{k-1}/2, \delta\}$ gives the remaining properties (4.15), (4.16), and (4.20).

By (4.17), the sequence Φ^k converges uniformly to some $\Phi \in C(\mathbb{R}^2)$. By (4.18) and the nesting of $\{\Omega^k\}_{k=1}^\infty$, $\Phi^l = \Phi^k$ on $\mathbb{R}^2 \backslash \overline{\Omega^k}$ for all $l \geq k$. Hence, by (4.20), $\Phi \in C^\infty(\mathbb{R}^2 \backslash S)$ and $\nabla\Phi = \nabla\Phi^k$ on $\mathbb{R}^2 \backslash \overline{\Omega^k}$. For $(x, y) \in \mathbb{R}^2 \backslash S$, by (4.20) there is a smallest $k \geq 1$ such that $(x, y) \in \overline{\Omega^{k-1}} \backslash \overline{\Omega^k}$, and so $\nabla\Phi(x, y) = \nabla\Phi^k(x, y)$. Hence by (4.19)

$$\Phi_y \geq B^{k-1} - 1 \geq B^0 - 1 = 3 + 4c^0 \geq 3$$

and

$$\Phi_x \leq -A^{k-1} + 1 = -3 \cdot 2^{k+1}(4 + 4c^{k-1}) + 1 \leq -3 \cdot 4 \cdot 4 + 1 = -47.$$

Thus we have (4.12.1) and the very last inequality of (4.12.2). More precisely, by (4.19) there is $s \in [0, 1]$ such that

$$\begin{aligned} -\Phi_x &\geq sA^{k-1} + (1 - s)A^k - 1 \\ &= s3 \cdot 2^{k+1}B^{k-1} + (1 - s)3 \cdot 2^{k+2}B^k - 1 \\ &\geq 3 \cdot 2^{k+1}(sB^{k-1} + (1 - s)B^k) - 1 \\ &\geq 3 \cdot 2^{k+1}(\Phi_y - 1) - 1 \\ &\geq 2^{k+1}\Phi_y. \end{aligned} \tag{4.21}$$

This gives the penultimate inequality of (4.12.2), since $k \geq 1$, and also (4.12.4), since as $0 < \operatorname{dist}((x, y), S) \to 0$, we have $k \to \infty$, by (4.20). We now check (4.12.3). Again, there is $s \in [0, 1]$ such that

$$\begin{aligned} -\Phi_x &< sA^{k-1} + (1 - s)A^k + 1 \\ &= s3 \cdot 2^{k+1}B^{k-1} + (1 - s)3 \cdot 2^{k+2}B^k + 1 \\ &\leq 3 \cdot 2^{k+2}(sB^{k-1} + (1 - s)B^k) + 1 \end{aligned}$$

$$\leq 3 \cdot 2^{k+2}(\Phi_y + 1) + 1$$
$$\leq 5 \cdot 2^{k+2}\Phi_y$$

whence

$$\frac{-2\Phi_x}{\Phi_y} \leq 5 \cdot 2^{k+3} \text{ on } \mathbb{R}^2 \backslash \overline{\Omega^k}. \tag{4.22}$$

In particular for given $(x, y) \in G\backslash S$ there is a $k \geq 1$ such that $(x, y) \in G\backslash \overline{\Omega^k}$; thus

$$\Phi_y \geq B^{k-1} - 1 \geq 4c^{k-1} \geq 4L_p(x, y, \psi),$$

as required for (4.12.3) since $(x, y, \psi) \in G \times [8, 5 \cdot 2^{k+3}]$ from (4.21) and (4.22).

Condition (4.22) also gives (4.15.3), since by (4.20) for any set $X \subseteq \mathbb{R}^2$ positively separated from S there is $k \geq 1$ such that $X \subseteq \mathbb{R}^2 \backslash \overline{\Omega^k}$, and hence $5 \cdot 2^{k+3}$ is an upper bound for ψ on X.

We are now obliged to construct $\{V^i\}_{i=1}^{\infty}$. For each $i \geq 1$, choose $k_i \geq 1$ such that $2^{k_i+2} \geq 2M^i + 4$ and define open $V^i \subseteq H$ by

$$V^i = \Omega^{k_i} \cap \{(x, y) \in H : \text{dist}((x, y), \mathbb{R}^2 \backslash H) > 1/i\}.$$

Evidently the V^i satisfy (4.15.2). For $(x, y) \in H \cap S$, we see $(x, y) \in V^i$ for $i \geq 2$ such that $B_{1/(i-1)}((x, y)) \subseteq H$, since $S \subseteq \Omega^{k_i}$ for all $i \geq 1$. Hence $V := \bigcup_{i=1}^{\infty} V^i$ is an open set such that $H \cap S \subseteq V \subseteq H$. Since $\overline{H} \subseteq G$, we then have that $\overline{V} \subseteq G$, as required for (4.15.1).

All that remains to check of $\{V^i\}_{i=1}^{\infty}$ is (4.12.2). Let $(x, y) \in V^i \backslash S$. Then $(x, y) \in \overline{\Omega^{k-1}} \backslash \overline{\Omega^k}$ for some $k > k_i$. So by (4.21), and recalling $\Phi_y > 0$, we see that, by choice of $k_i \geq 1$,

$$-\Phi_x \geq 2^{k+1}\Phi_y \geq 2^{k_i+2}\Phi_y \geq (2M^i + 4)\Phi_y$$

as required.

We easily check condition (4.12.5). Let $a < b$ and $u \in \text{AC}(a, b)$ be non-decreasing. The $\{x \in (a, b) : U(x) \in S\}$ is null since S is purely unrectifiable. For all $k \geq 0$, properties (4.17) and (4.20) imply that $\|\Phi - \Phi^k\|_\infty < 2\epsilon^k$ for all $k \geq 0$, and hence by property (4.16) that $\text{meas}(\{(\Phi \circ U)(x) : U(x) \in S\}) \leq 1/k$. Hence this set is also null. □

We now give the exact details of the inductive construction of our Lagrangians L_n. Let $S \subseteq \mathbb{R}^2$ be a purely unrectifiable set such that $S = \bigcup_{n=1}^{\infty} S_n$

for some compact S_n, and let ω be a fixed superlinearity. For each $n \geq 1$ define

$$G_n = B_1(S_n) \text{ and } H_n = B_{1/2}(S_n) \setminus \bigcup_{m=1}^{n-1} S_m.$$

The set H_n is a neighbourhood of the set $S_n \setminus \bigcup_{m=1}^{n-1} S_m$ which we want to cover by $\operatorname{uss}(L_n)$, but contains no points of $\bigcup_{m=1}^{n-1} S_m$, which we assume to be covered by $\operatorname{uss}(L_{n-1})$. Thus $H_n \times \mathbb{R} \subseteq \mathbb{R}^3$ is the domain on which we modify a given L_{n-1}, building L_n to deal with the points in S_n, without interfering with the structure of L_{n-1} on $\bigcup_{m=1}^{n-1} S_m$. In the case that the S_n are pairwise disjoint, H_n could be chosen to be any open neighbourhood of S_n positively separated from $\bigcup_{m=1}^{n-1} S_m$.

Lemma 4.16. For each $n \geq 1$ there exist $F_n \colon \mathbb{R}^3 \to \mathbb{R}$ of form $(\star_F)$, $\Phi_n \in C^\infty(\mathbb{R}^2 \setminus S_n) \cap C(\mathbb{R}^2)$, sequence $\{V_n^i\}_{i=1}^\infty$ of open sets $V_n^i \subseteq H_n$, sequence of constants $\{M_n^i\}_{i=1}^\infty$, and an open set $W_n \subseteq \mathbb{R}^2$ such that the following relations hold:

(4.16.1) $H_n \cap S_n \subseteq W_n \subseteq V_n := \bigcup_{i=1}^\infty V_n^i \subseteq \overline{V_n} \subseteq G_n$;

(4.16.2) $\overline{W_n} \setminus V_n \subseteq S_n$;

(4.16.3) $\{M_n^i\}_{i=1}^\infty$ is a non-decreasing sequence and $M_n^1 \geq n$;

(4.16.4) $\lim_{0<\operatorname{dist}((x,y),S_n)\to 0}((\Phi_n)_x/(\Phi_n)_y)(x,y) = -\infty$;

(4.16.5) $L_n \colon \mathbb{R}^3 \to \mathbb{R}$ defined by

$$L_n(x,y,p) = \omega(p) + F_n(x,y,p)$$

has the property that for all $a < b$ and all functions $u \in \mathrm{AC}(a,b)$ such that $Q(a,u(a);b,u(b)) \subseteq W_n$, we have

$$\int_a^b L_n(x,u(x),u'(x))\,dx \geq \Phi_n(U(b)) - \Phi_n(U(a))$$

with equality if and only if $u'(x) = (-2(\Phi_n)_x/(\Phi_n)_y)(x,u(x))$ for almost every $x \in [a,b]$;

and for $n \geq 2$,

(4.16.6) $F_n \geq F_{n-1}$ on $\mathbb{R}^3$;

(4.16.7) $F_n = F_{n-1}$ on $\mathbb{R}^3 \setminus \bigcup_{i=1}^\infty (V_n^i \times (M_n^i, \infty))$; and

(4.16.8) $\psi_m \in C^\infty(\mathbb{R}^2 \setminus S_m)$ defined by $\psi_m := -2(\Phi_m)_x/(\Phi_m)_y$ satisfies $\psi_m \leq M_n^i$ on V_n^i for all $i \geq 1$, for each $1 \leq m < n$.

Proof. For each $n \geq 1$, we want to apply Lemma 4.15 to get a potential with which we can apply Lemma 4.12. To begin, we define $M_1^i = 1$ for all $i \geq 1$, and $F_0 \colon \mathbb{R}^3 \to \mathbb{R}$ to be the zero function.

For $n \geq 2$ we suppose $\Phi_m \in C^\infty(\mathbb{R}^2 \backslash S_m) \cap C(\mathbb{R}^2)$ to have been constructed as claimed, and moreover such that ψ_m and S_m satisfy (4.15.3) for each $1 \leq m < n$. For each $i \geq 1$ define

$$\tilde{V}_n^i = \{(x,y) \in H_n : \operatorname{dist}((x,y), \mathbb{R}^2 \backslash H_n) > 1/i\}.$$

So for all $i \geq 1$ we have $\operatorname{dist}(\tilde{V}_n^i, \mathbb{R}^2 \backslash H_n) > 0$, and also therefore $\operatorname{dist}(\tilde{V}_n^i, S_m) > 0$ for each $1 \leq m < n$, since $\bigcup_{m=1}^{n-1} S_m \subseteq \mathbb{R}^2 \backslash H_n$. So by the assumption (4.15.3) on each ψ_m, we can choose $M_n^1 \geq n$ such that $\psi^m \leq M_n^1$ on $\tilde{V}_n^1$ for all $1 \leq m < n$, and inductively $M_n^i \geq M_n^{i-1}$ such that $\psi^m \leq M_n^i$ on $\tilde{V}_n^i$ for all $1 \leq m < n$. This gives us a sequence $\{M_n^i\}_{i=1}^\infty$ satisfying (4.16.3).

We can now apply Lemma 4.15 inductively for each $n \geq 1$, using data $S = S_n$, $G = G_n$, $H = H_n$, $\{M^i\}_{i=1}^\infty = \{M_n^i\}_{i=1}^\infty$, $F = F_{n-1}$. This gives us a function $\Phi = \Phi_n$ of the required form, and a sequence of open sets $\{V^i\}_{i=1}^\infty = \{V_n^i\}_{i=1}^\infty$ such that by (4.15.1),

$$H_n \cap S_n \subseteq V_n := \bigcup_{i=1}^\infty V_n^i \subseteq \overline{V_n} \subseteq G_n. \tag{4.23}$$

For $n \geq 2$, we have by (4.15.2) that $V_n^i \subseteq \tilde{V}_n^i$ for each $i \geq 1$, so (4.16.8) holds, from the above discussion on $\tilde{V}_n^i$.

Lemma 4.15 also asserts that all the conditions of Lemma 4.12 hold, using this data, which gives us in particular (4.16.4). To apply Lemma 4.12, we need a suitable W_n.

Since Lemma 4.15 tells us V_n is open, for all $x \in H_n \cap S_n$, there is $\delta_x > 0$ such that $B_{\delta_x}(x) \subseteq V_n$. Then defining

$$W_n = \bigcup_{x \in H_n \cap S_n} B_{\delta_x/2}(x)$$

gives an open set W_n which, in conjunction with (4.23), gives (4.16.1). We can now easily check that (4.16.2) holds.

Suppose $x \in \overline{W_n} \backslash S_n$. Choose $\epsilon > 0$ such that $B_\epsilon(x) \cap S_n = \emptyset$, and find $w \in W_n \cap B_{\epsilon/2}(x)$. Then by definition of W_n there exists $y \in H_n \cap S_n \subseteq S_n$ such that $w \in B_{\delta_y/2}(y)$. If $\delta_y \leq \epsilon$, then

$$\|y - x\|_2 \leq \|y - w\|_2 + \|w - x\|_2 < \delta_y/2 + \epsilon/2 \leq \epsilon,$$

which contradicts the choice of ϵ, since $y \in S_n$. So $\delta_y > \epsilon$, hence $\|y - x\|_2 < \delta_y$, and thus $x \in B_{\delta_y}(y) \subseteq V$ by choice of δ_y. Thus $\overline{W_n} \backslash S_n \subseteq V_n$, and hence $\overline{W_n} \backslash V_n \subseteq S_n$.

Then we are in the situation of Lemma 4.12. Set $F_n = \widehat{F_{n-1}}$ as given in Lemma 4.12 for this W_n. The remaining conclusions then follow directly from those of the lemma. Since Lemma 4.15 asserts that ψ_n and S_n also satisfy (4.15.3), we are able to iterate the construction to produce the required sequence.

□

Proof of Theorem 4.10. By Lemma 4.16 we have a sequence $\{F_n\}_{n=1}^{\infty}$ of functions $F_n \colon \mathbb{R}^3 \to \mathbb{R}$ of form $(\star_F)$. Note that for $n_0 \geq 1$, we have by (4.16.3) that $p \notin (M_n^i, \infty)$ for all $i \geq 1$ and all $n \geq n_0$ whenever $p \in (-\infty, n_0)$. Hence by (4.16.7), $F_n = F_{n_0}$ on $\mathbb{R}^2 \times (-\infty, n_0)$ for all $n \geq n_0$. Then for $(x, y, p) \in \mathbb{R}^3$, choosing $n_0 > p$, we have that $F_n = F_{n_0}$ for all $n \geq n_0$ on an open set around (x, y, p). We then define $F \colon \mathbb{R}^3 \to \mathbb{R}$ by $F(x, y, p) = \lim_{n\to\infty} F_n(x, y, p)$, and it is clear that F is of form $(\star_F)$.

So we can define Lagrangian $L \colon \mathbb{R}^3 \to \mathbb{R}$ of form $(\star)$ by defining

$$L(x, y, p) = \omega(p) + F(x, y, p).$$

We claim S lies in the universal singular set of L.

Let $(x_0, y_0) \in S$. Choose $n_0 \geq 1$ such that $(x_0, y_0) \in S_{n_0} \backslash \bigcup_{m=1}^{n_0-1} S_m$. As in Csörnyei et al. [2008], we can construct a locally absolutely continuous $u_0 \colon \mathbb{R} \to \mathbb{R}$ such that $u_0'(x) = \psi_{n_0}(x, u_0(x))$ for almost every $x \in \mathbb{R}$ and $u(x_0) = y_0$. For each $k \geq 0$ we find $u^k \in C^1(\mathbb{R})$ such that $(u^k)'(x) = \psi_{n_0}^k(x, u^k(x))$ for all $k \geq 0$, and show that $\{u^k\}_{k=0}^{\infty}$ is an equicontinuous family. Some subsequence therefore converges locally uniformly to a non-decreasing function $u_0 \in C(\mathbb{R})$ which solves $u_0'(x) = \psi_{n_0}(x, u_0(x))$ whenever $(x, u_0(x)) \notin S_{n_0}$, i.e. almost everywhere. Thus u_0 is locally absolutely continuous. We observe that $(x_0, y_0) \in S_{n_0} \cap H_{n_0} \subseteq W_{n_0}$, using (4.16.1). Since W_{n_0} is open we can choose real numbers $a_0 < b_0$ such that $(x_0, y_0) \in Q(a_0, u(a_0); b_0, u(b_0)) \subseteq W_{n_0}$.

We claim we have constructed $\{F_n\}_{n=1}^{\infty}$ in such a way that

$$L(x, u_0(x), u_0'(x)) = L_{n_0}(x, u_0(x), u_0'(x)) \text{ for almost every } x \in [a_0, b_0].$$

We show in fact that for all $n \geq n_0$,

$$L_n(x, u_0(x), u_0'(x)) = L_{n_0}(x, u_0(x), u_0'(x)) \text{ for almost every } x \in [a_0, b_0].$$

This suffices since a countable union of null sets is null.

We proceed by induction. The claim is obvious for $n = n_0$, so let $n > n_0$ and assume that the statement is true for $n-1$. At points where the graph of the trajectory u_0 lies outside V_n, we see the result immediately since we know L_{n-1} was not changed there: for $x \in [a_0, b_0] \backslash U_0^{-1}(V_n)$, by (4.16.7) we have

$$L_n(x, u_0(x), u_0'(x)) = L_{n-1}(x, u_0(x), u_0'(x)).$$

When the graph of the trajectory u_0 lies inside V_n, we have to use some information about the derivative of u_0. Let $i \geq 1$. By choice of u_0, (4.16.8), and (4.16.3), for almost every $x \in [a_0, b_0] \cap U_0^{-1}(V_n^i)$ we have that

$$u_0'(x) = \psi_{n_0}(x, u_0(x)) \leq M_n^i \leq M_n^j \text{ for all } j \geq i.$$

So for almost every $x \in [a_0, b_0] \cap U_0^{-1}(V_n^i)$, we have $u_0'(x) \notin (M_n^j, \infty)$ for all $j \geq i$. For each $x \in [a_0, b_0] \cap U_0^{-1}(V_n)$, there is a least $i \geq 1$ such that $(x, u_0(x)) \in V_n^i$; so $(x, u_0(x)) \notin V_n^j$ for all $1 \leq j < i$. Then, since $U_0^{-1}(V_n) = \bigcup_{i=1}^{\infty} U_0^{-1}(V_n^i)$ and a countable union of null sets is null, for almost every $x \in [a_0, b_0] \cap U_0^{-1}(V_n)$ we have that

$$(x, u_0(x), u_0'(x)) \in \mathbb{R}^3 \backslash \bigcup_{j=1}^{\infty} (V_n^j \times (M_n^j, \infty)).$$

But then, by (4.16.7), we see that indeed

$$L_n(x, u_0(x), u_0'(x)) = L_{n-1}(x, u_0(x), u_0'(x))$$

for almost every $x \in [a_0, b_0] \cap U_0^{-1}(V_n)$. The result then follows by the inductive hypothesis.

So applying (4.16.5) to L_{n_0}, we see, since $u_0'(x) = \psi_{n_0}(x, u_0(x))$ for almost every $x \in [a_0, b_0]$,

$$\begin{aligned} \int_{a_0}^{b_0} L(x, u_0(x), u_0'(x))\, dx &= \int_{a_0}^{b_0} L_{n_0}(x, u_0(x), u_0'(x))\, dx \\ &= \Phi_{n_0}(U(b_0)) - \Phi_{n_0}(U(a_0)). \end{aligned}$$

By (4.16.6) and (4.16.5), we see

$$\begin{aligned} \int_{a_0}^{b_0} L(x, u(x), u'(x))\, dx &\geq \int_{a_0}^{b_0} L_{n_0}(x, u(x), u'(x))\, dx \\ &\geq \Phi_{n_0}(U(b_0)) - \Phi_{n_0}(U(a_0)) \end{aligned}$$

for any $u \in \mathrm{AC}(a_0, b_0)$ such that $Q(a_0, u(a_0); b, u(b_0)) \subseteq W_{n_0}$. Thus u_0 is a minimizer for (1.2) over those functions $u \in \mathrm{AC}(a_0, b_0)$ such that $u(a_0) = u_0(a_0)$ and $u(b_0) = u_0(b_0)$. Tonelli's partial regularity result and (4.16.4) then imply that $u_0'(x_0) = \infty$. Hence (x_0, y_0) lies in the universal singular set of L, as required. □

Bibliography

E. Acerbi and N. Fusco. A regularity theorem for minimizers of quasiconvex integrals. *Arch. Ration. Mech. Anal.*, 99(3):261–281, 1987.

J. M. Ball and V. J. Mizel. Singular minimizers for regular one-dimensional problems in the calculus of variations. *Bull. Amer. Math. Soc. (N.S.)*, 11(1):143–146, 1984.

J. M. Ball and V. J. Mizel. One-dimensional variational problems whose minimizers do not satisfy the Euler-Lagrange equation. *Arch. Ration. Mech. Anal.*, 90(4): 325–388, 1985.

J. M. Ball and N. S. Nadirashvili. Universal singular sets for the one-dimensional variational problems. *Calc. Var. Partial Differential Equations*, 1(4):429–438, 1993.

S. Bernstein. Sur les équations du calcul des variations. *Ann. Sci. École Norm. Sup. (3)*, 29:431–485, 1912.

G. Buttazzo, M. Giaquinta, and S. Hildebrandt. *One-dimensional variational problems. An Introduction*, volume 15 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1998.

L. Cesari. *Optimization—theory and applications. Problems with ordinary differential equations*, volume 17 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1983.

F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

F. H. Clarke and R. B. Vinter. On the conditions under which the Euler equation or the maximum principle hold. *Appl. Math. Optim.*, 12(1):73–79, 1984.

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(1):73–98, 1985a.

F. H. Clarke and R. B. Vinter. Existence and regularity in the small in the calculus of variations. *J. Differential Equations*, 59(3):336–354, 1985b.

F. H. Clarke and R. B. Vinter. Regularity of solutions to variational problems with polynomial Lagrangians. *Bull. Polish Acad. Sci. Math.*, 34(1-2):73–81, 1986.

M. Csörnyei, B. Kirchheim, T. C. O'Neil, D. Preiss, and S. Winter. Universal singular sets in the calculus of variations. *Arch. Ration. Mech. Anal.*, 190(3): 371–424, 2008.

B. Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2008.

A. M. Davie. Singular minimisers in the calculus of variations in one dimension. *Arch. Ration. Mech. Anal.*, 101(2):161–177, 1988.

E. De Giorgi. Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)*, 3:25–43, 1957.

E. De Giorgi. Un esempio di estremali discontinue per un problema variazionale di tipo ellittico. *Boll. Un. Mat. Ital. (4)*, 1:135–137, 1968.

L. C. Evans. Quasiconvexity and partial regularity in the calculus of variations. *Arch. Ration. Mech. Anal.*, 95(3):227–252, 1986.

H. Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

A. Ferriero. Relaxation and regularity in the calculus of variations. *J. Differential Equations*, 249(10):2548–2560, 2010.

M. Foss and G. Mingione. Partial continuity for elliptic problems. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 25(3):471–503, 2008.

M. Giaquinta. *Multiple integrals in the calculus of variations and nonlinear elliptic systems*, volume 105 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1983.

M. Giaquinta and E. Giusti. Nonlinear elliptic systems with quadratic growth. *Manuscripta Math.*, 24(3):323–349, 1978.

M. Giaquinta and G. Modica. Partial regularity of minimizers of quasiconvex integrals. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 3(3):185–208, 1986.

E. Giusti. *Direct methods in the calculus of variations.* World Scientific Publishing Co. Inc., River Edge, NJ, 2003.

E. Giusti and M. Miranda. Un esempio di soluzioni discontinue per un problema di minimo relativo ad un integrale regolare del calcolo delle variazioni. *Boll. Un. Mat. Ital. (4)*, 1:219–226, 1968.

E. Giusti and M. Miranda. Sulla regolarità delle soluzioni deboli di una classe di sistemi ellittici quasi-lineari. *Arch. Ration. Mech. Anal.*, 31:173–184, 1968/1969.

J. Kristensen and G. Mingione. The singular set of minima of integral functionals. *Arch. Ration. Mech. Anal.*, 180(3):331–398, 2006.

J. Kristensen and G. Mingione. The singular set of Lipschitzian minima of multiple integrals. *Arch. Ration. Mech. Anal.*, 184(2):341–369, 2007.

M. Lavrentiev. Sur quelques problèmes du calcul des variations. *Ann. Mat. Pura Appl.*, 4:7–28, 1926.

B. Manià. Sopra un essempio di Lavrentieff. *Boll. Un. Mat Ital.*, 13:147–153, 1934.

G. Mingione. Regularity of minima: an invitation to the dark side of the calculus of variations. *Appl. Math.*, 51(4):355–426, 2006.

C. B. Morrey, Jr. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952.

C. B. Morrey, Jr. Partial regularity results for non-linear elliptic systems. *J. Math. Mech.*, 17:649–670, 1967/1968.

C. B. Morrey, Jr. *Multiple integrals in the calculus of variations.* Classics in Mathematics. Springer-Verlag, Berlin, 2008. Reprint of the 1966 edition.

S. Müller and V. Šverák. Convex integration for Lipschitz mappings and counterexamples to regularity. *Ann. of Math. (2)*, 157(3):715–742, 2003.

J. Nečas. Example of an irregular solution to a nonlinear elliptic system with analytic coefficients and conditions for regularity. In *Theory of nonlinear operators (Proc. Fourth Internat. Summer School, Acad. Sci., Berlin, 1975)*, pages 197–206. Abh. Akad. Wiss. DDR Abt. Math.–Natur.–Tech., Jahrgang 1977, 1. Akademie-Verlag, Berlin, 1977.

W. Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, 1966.

S. Saks. *Theory of the integral*, volume VII of *Monografie Matematyczne.* Stechert, New York, 1937.

V. Šverák and X. Yan. A singular minimizer of a smooth strongly convex functional in three dimensions. *Calc. Var. Partial Differential Equations*, 10(3):213–221, 2000.

V. Sverák and X. Yan. Non-Lipschitz minimizers of smooth uniformly convex functionals. *Proc. Natl. Acad. Sci. USA*, 99(24):15269–15276, 2002.

M. A. Sychëv. Regularity of solutions of some variational problems. *Soviet Math. Dokl.*, 43(1):292–296, 1991.

M. A. Sychëv. A classical problem of the calculus of variations. *Soviet Math. Dokl.*, 44(1):116–120, 1992.

M. A. Sychëv. On the regularity of solutions of variational problems. *Russian Acad. Sci. Sb. Math*, 75(2):535–556, 1993.

M. A. Sychëv. The Lebesgue measure of a universal singular set in the simplest problems of the calculus of variations. *Sibirsk. Mat. Zh.*, 35(6):1373–1389, iv, 1994.

M. A. Sychev and V. J. Mizel. A condition on the value function both necessary and sufficient for full regularity of minimizers of one-dimensional variational problems. *Trans. Amer. Math. Soc.*, 350(1):119–133, 1998.

L. Tonelli. Sur un méthode directe du calcul des variations. *Rend. Circ. Mat. Palermo*, 39:233–264, 1915.

L. Tonelli. *Fondamentica di Calcolo delle variazioni*, volume 2. Zanichelli, Bologna, 1923.

L. Tonelli. Su gli integrali del calcolo delle variazioni in forma ordinaria. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (2)*, 3(3-4):401–450, 1934.

W. P. Ziemer. *Weakly differentiable functions. Sobolev spaces and functions of bounded variation*, volume 120 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1989.