



Variations, approximation, and low regularity in one dimension

Richard Gratwick¹

Received: 12 March 2015 / Accepted: 19 February 2017 / Published online: 13 April 2017
© The Author(s) 2017. This article is an open access publication

Abstract We investigate the properties of minimizers of one-dimensional variational problems when the Lagrangian has no higher smoothness than continuity. An elementary approximation result is proved, but it is shown that this cannot be in general of the form of a standard Lipschitz “variation”. Part of this investigation, but of interest in its own right, is an example of a nowhere locally Lipschitz minimizer which serves as a counter-example to any putative Tonelli partial regularity statement. Under these low assumptions we find it nonetheless remains possible to derive necessary conditions for minimizers, in terms of approximate continuity and equality of the one-sided derivatives.

Mathematics Subject Classification 49N60

1 Introduction

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

$$\mathcal{L}(u) = \int_a^b L(t, u(t), u'(t)) dt$$

over some class of functions $u : [a, b] \rightarrow \mathbb{R}^n$ with fixed boundary conditions. The integrand $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is known as the *Lagrangian*. Tonelli [15, 16] presented rigorous existence results for minimizers of such a problem, demonstrating the need to work on

Communicated by J. Ball.

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement No. 291497.

✉ Richard Gratwick
R.Gratwick@ed.ac.uk

¹ School of Mathematics, James Clerk Maxwell Building, The King’s Buildings, Peter Guthrie Tait Road, Edinburgh EH9 3FD, UK

the function space of absolutely continuous functions, or what is now known also as the Sobolev space $W^{1,1}((a, b); \mathbb{R}^n)$. In particular such functions are only differentiable almost everywhere. Defining the functional \mathcal{L} on this space, Tonelli developed the *direct method* of the calculus of variations to deduce the existence of minimizers when certain conditions are imposed on the Lagrangian. The key assumptions are the conditions of convexity and superlinearity: i.e. that the function $p \mapsto L(t, y, p)$ is convex for each (t, y) , and that there exists some $\omega: \mathbb{R} \rightarrow \mathbb{R}$ satisfying $\omega(\|p\|)/\|p\| \rightarrow \infty$ as $\|p\| \rightarrow \infty$ such that $L(t, y, p) \geq \omega(\|p\|)$ for all (t, y, p) . Some minimal smoothness of the Lagrangian is also required, for example continuity suffices. The subject of this paper is what can happen at this level of regularity, i.e. when the Lagrangian is assumed only to be continuous.

The penalty paid for an abstract existence theorem is that one must work in a suitable function space, and therefore can only assert that the minimizer is $W^{1,1}$. A significant question is then whether it is possible to assert *a priori* any higher regularity of minimizers. Assuming appropriate growth conditions and C^k -regularity of the Lagrangian, one may prove C^k -regularity of the minimizers (see for example [3]). For scalar-valued functions u , Tonelli [15] provided a partial regularity theorem, asserting that C^∞ -regularity of the Lagrangian and strict convexity in p implies that any minimizer u is C^∞ on an open set of full measure. Clarke and Vinter [4] gave an analogous statement for vector-valued functions.

The assumption of strict convexity may not be weakened, but several authors have weakened the smoothness assumption on the Lagrangian. Clarke and Vinter imposed only a local Lipschitz condition in (y, p) . In the scalar case, Sych ev [14] imposed a local H older condition, and Cs ornyei et al. [5] imposed a local Lipschitz condition in y , locally uniformly in (t, p) . In the vectorial case, Ferriero [8,9] allowed this Lipschitz constant to vary as an integrable function of t . Recalling that no control of the modulus of continuity is required for the existence theorem, Gratwick and Preiss [11] gave a counter-example of a continuous Lagrangian which admits a minimizer non-differentiable on a dense set. So we are faced with the possibility of situations where minimizers over $W^{1,1}$ exist, but partial regularity results fail to hold. Section 2 presents a new counter-example illustrating this, with a minimizer having upper and lower derivatives $\pm\infty$ at a dense set of points.

A standard technique to prove necessary conditions of minimizers is to compute the first variation, i.e. to consider the limiting behaviour of the function $\gamma \mapsto \mathcal{L}(u + \gamma w)$ as $\gamma \rightarrow 0$. Following this path in the classical situation leads us to the Euler-Lagrange equation and other necessary conditions. In our low-level regularity situation, assuming only continuity of the Lagrangian, it is not immediately clear how such small perturbations behave. Ball and Mizel [2] gave examples of polynomial Lagrangians for which $\mathcal{L}(u + \gamma w) = \infty$ for a certain class of smooth functions w . In our case, when we do not have a partial regularity theorem, and must therefore admit the possibility of minimizers which are nowhere locally Lipschitz, it is not even immediately clear that it is possible to approximate the minimum value by any other trajectories at all.

The possibility of a complete failure of approximation is not absurd when one considers the possible presence of the *Lavrentiev phenomenon* [12], in which situation the energy of Lipschitz functions with the required boundary conditions is bounded away from the minimum value. That this can occur not only for polynomial integrands [13] but even for strictly convex and superlinear polynomial integrands [2] should warn us that we are wise to be wary of what might happen when we consider Lagrangians which satisfy only the bare continuity assumption. Ball and Mizel [2] gave another example of bad behaviour to keep us on guard: the *repulsion property* [1], whereby it can happen that $\mathcal{L}(u_n) \rightarrow \infty$ for any sequence of admissible Lipschitz functions u_n which converge uniformly to the minimizer.

Nevertheless, a general approximation result can be proved, indeed without great difficulty. This is the content of Theorem 15 in Sect. 3. In this section we go on to investigate how fruitful it may be to consider computing the variation as suggested above, and discover that in general it will not get us very far: there exist examples (Theorem 17), even superlinear and strictly convex examples (Theorem 18), of continuous Lagrangians where the addition of any Lipschitz variation to a minimizer results in an infinite value for the integral. We also investigate the relationship between approximation in this sense and the Lavrentiev phenomenon. We find in this section that we can make good use of the counter-example to partial regularity described in Sect. 2, using in an essential way the main new feature of this example, viz the fact that the minimizer is nowhere locally Lipschitz.

The technique used to construct the basic approximation in Theorem 15 is then put to repeated use in Sect. 4, where we pursue the question of whether any necessary conditions can be derived of minimizers in our setting. Having lost any hope of a general partial regularity statement, we are left wondering whether it might be the case that an arbitrary $W^{1,1}$ function can be a minimizer of a variational problem with a continuous Lagrangian. Under the assumption of *strict* convexity, we are able to show that, although the derivative of a minimizer need not exist at every point, at those points at which the derivative does exist, the derivative is approximately continuous. As a corollary of this, we can then show that when each one-sided derivative exists at a point, the two derivatives must in fact be equal. Such statements extend, suitably interpreted, to cases where the derivatives are infinite, and we may be more precise when infinite derivatives are confined to one component.

1.1 Notation and terminology

Throughout we fix $[a, b] \subseteq \mathbb{R}$, $n \geq 1$ and the euclidean norm $\| \cdot \|$ on \mathbb{R}^n . We shall consider $[a, b] \times \mathbb{R}^n \times \mathbb{R}^n$ to be equipped with the norm given by the maximum of the norms of the three components. The supremum norm of a real- or vector-valued function shall be denoted by $\| \cdot \|_\infty$, and the support of a such a function shall be denoted by spt . For a set $E \subseteq [a, b]$ we denote the Lebesgue measure of the set by $\lambda(E)$, and the characteristic function of the set by $\mathbb{1}_E$.

A *Lagrangian* shall be a function $L = L(t, y, p) : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$. Conditions on the Lagrangians shall be discussed at the relevant points, but in particular we demand that they are continuous, but never impose any stronger smoothness condition or prescribe any modulus of continuity. For a function $v \in W^{1,1}((a, b); \mathbb{R}^n)$ (if $n = 1$ we will usually suppress the notating of the target space), we let

$$\mathcal{L}(v) = \int_a^b L(t, v(t), v'(t)) dt.$$

Recall that *superlinearity* is the condition that for some $\omega : \mathbb{R} \rightarrow \mathbb{R}$ satisfying $\omega(\|p\|)/\|p\| \rightarrow \infty$ as $\|p\| \rightarrow \infty$, we have for all $(t, y, p) \in [a, b] \times \mathbb{R}^n \times \mathbb{R}^n$ that $L(t, y, p) \geq \omega(\|p\|)$. For $A, B \in \mathbb{R}^n$ we let

$$\mathcal{A}_{A,B} := \{v \in W^{1,1}((a, b); \mathbb{R}^n) : v(a) = A, v(b) = B\}.$$

2 Failure of partial regularity

In this section we present a counter-example to a putative partial regularity theorem in the manner of Tonelli for continuous Lagrangians. A first example of this kind was produced

by Gratwick and Preiss [11], exhibiting a Lipschitz minimizer which was non-differentiable on a dense set. The following example produces a minimizer with upper and lower Dini derivatives of $\pm\infty$ at a dense set of points, i.e. the derivative fails to exist at these points in as dramatic a way possible. That we have both Lipschitz and non-Lipschitz examples is worth emphasizing. The Lipschitz example serves to disillusion us should we be inclined to suspect, as can be the case, that *a priori* knowledge of boundedness of the derivative of a minimizer implies some higher regularity. The non-Lipschitz case is remarkable in that intuitively one does not expect superlinear Lagrangians to have minimizers with infinite derivatives at many points, far less minimizers with difference quotients oscillating arbitrarily largely.

This example was first presented by Gratwick [10, Example 2.35] as an application of a general construction scheme.

Definition 1 The upper and lower Dini derivatives, $\overline{D}v(t)$ and $\underline{D}v(t)$ respectively, of a function $v \in W^{1,1}(a, b)$ at a point $t \in (a, b)$ are given by

$$\overline{D}v(t) := \limsup_{s \rightarrow t} \frac{v(s) - v(t)}{s - t}, \text{ and } \underline{D}v(t) := \liminf_{s \rightarrow t} \frac{v(s) - v(t)}{s - t}.$$

Theorem 2 *There exist $T > 0$, $w \in W^{1,2}(-T, T)$, and a continuous $\phi: [-T, T] \times \mathbb{R} \rightarrow [0, \infty)$ such that*

$$\mathcal{L}(u) = \int_{-T}^T (\phi(t, u(t) - w(t)) + (u'(t))^2) dt$$

defines a functional on $W^{1,1}(-T, T)$ with a continuous Lagrangian such that w is a minimizer of \mathcal{L} over $\mathcal{A}_{w(-T), w(T)}$, but $\overline{D}w(t) = +\infty$ and $\underline{D}w(t) = -\infty$ for a dense set of points $t \in [-T, T]$.

The remainder of this section is devoted to a proof of this theorem.

2.1 Construction of the minimizer

Let $T \in (0, e^{-e^2}/2)$ be small enough such that for any $t \in [-T, T] \setminus \{0\}$,

$$(2|t|)^{1/3} \log \log 1/2|t| \leq \left(\frac{1}{\log 1/2|t|} \right)^{1/3} \log \log 1/2|t| \leq 1/125. \tag{1}$$

Given any sequence of points in $(-T, T)$, we can construct a Lagrangian 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 0, with difference quotients oscillating between arbitrarily large positive and negative values, but minimizes a problem with a continuous Lagrangian. This basic Lagrangian is of the form $(t, y, p) \mapsto \tilde{\phi}(t, y - \tilde{w}(t)) + p^2$ for a “weight function” $\tilde{\phi}: [-T, T] \times \mathbb{R} \rightarrow [0, \infty)$, i.e. such that $\tilde{\phi}(\cdot, 0) = 0$ and $|y| \mapsto \tilde{\phi}(t, y)$ is increasing, so $(t, y) \mapsto \tilde{\phi}(t, y - \tilde{w}(t))$ penalizes functions which stray from \tilde{w} . This summand of the Lagrangian then takes its minimum value along the graph of \tilde{w} , and assigns larger values to functions u the further their graph lies from that of \tilde{w} .

We sketch the main ideas behind the proof that \tilde{w} minimizes this “basic” problem

$$\mathcal{A}_{\tilde{w}(-T), \tilde{w}(T)} \ni u \mapsto \int_{-T}^T \left(\tilde{\phi}(t, u(t) - \tilde{w}(t)) + (u'(t))^2 \right) dt.$$

So suppose for now that $\tilde{u} \in \mathcal{A}_{\tilde{w}(-T), \tilde{w}(T)}$ is a minimizer for this problem.

If $\tilde{u}(0) = \tilde{w}(0)$, it suffices to argue separately on $[-T, 0]$ and $[0, T]$. We consider $[0, T]$. Note for any two functions $\tilde{u}, \tilde{w} : [-T, T] \rightarrow \mathbb{R}$, we have that

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

Assuming \tilde{w} is smooth enough that we can integrate by parts, our key argument is the following:

$$\begin{aligned} \int_0^T (\tilde{\phi}(t, \tilde{u} - \tilde{w}) + (\tilde{u}')^2) - \int_0^T (\tilde{w}')^2 &= \int_0^T ((\tilde{u}')^2 - (\tilde{w}')^2) + \tilde{\phi}(t, \tilde{u} - \tilde{w}) \\ &\geq \int_0^T (2(\tilde{u}' - \tilde{w}')\tilde{w}' + \tilde{\phi}(t, \tilde{u} - \tilde{w})) \\ &= [2(\tilde{u} - \tilde{w})\tilde{w}']_0^T \\ &\quad + \int_0^T (\tilde{\phi}(t, \tilde{u} - \tilde{w}) - 2(\tilde{u} - \tilde{w})\tilde{w}'') \\ &\geq \int_0^T (\tilde{\phi}(t, \tilde{u} - \tilde{w})) - 2|\tilde{u} - \tilde{w}||\tilde{w}''|, \end{aligned}$$

since the boundary terms vanish by assumption. Hence choosing $\tilde{\phi}(t, y) \geq 2|\tilde{w}''(t)||y|$, this final expression is non-negative, implying that \tilde{w} is indeed a minimizer with respect to its own boundary conditions. However, this inequality for $\tilde{\phi}$ cannot be enforced for all values of (t, y) , since $|\tilde{w}''(t)| \rightarrow \infty$ as $t \rightarrow 0$; this is the whole point of the example. Since we only need this inequality for values of $y = \tilde{u} - \tilde{w}$, we enforce the inequality for only values of (t, y) which lie in the (slightly expanded) convex hull of the graph of \tilde{w} , the shape of which is given by a function g . We choose the oscillations of \tilde{w} as we approach to 0 to be so slow that $|\tilde{w}''(t)g(t)| \rightarrow 0$ as $t \rightarrow 0$. So it is possible to construct a well-defined continuous function $\tilde{\phi}$ so that $\tilde{\phi}(t, y) \geq 2|g''(t)||y|$ for $|y| \leq c|g(t)|$, for some constant c . To exploit this definition, we then need to establish that $|\tilde{u}(t) - \tilde{w}(t)| \leq c|g(t)|$. It suffices to establish that $|\tilde{u}(t)| \leq c|g(t)|$. This is a consequence of the assumption that \tilde{u} is a minimizer and that $|g|$ is concave on $(0, T)$. Since \tilde{u} and \tilde{w} agree at 0 and T , any interval on which \tilde{u} lies outside the convex hull of \tilde{w} must be a proper subinterval of $(0, T)$. By the concavity of g on such an interval, we may find an affine function which lies strictly between \tilde{u} and g , and hence \tilde{u} and \tilde{w} . We then consider the competitor function in the minimization problem defined by replacing \tilde{u} with this affine function. Since affine functions minimize convex functionals, this strictly decreases the gradient term in the integrand. Since the affine function lies closer to \tilde{w} than \tilde{u} , this replacement cannot increase the $\tilde{\phi}(t, \cdot - \tilde{w}(t))$ term. Hence we get a contradiction: \tilde{u} must lie inside the (expanded) convex hull of \tilde{w} .

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 \tilde{l} on an interval around 0. This forces another requirement on the (slow) speed of the oscillations of \tilde{w} , since we incur two errors in making this replacement. We need to control the difference in the L^2 -norm of the gradients of \tilde{w} and \tilde{l} , and the difference in the gradients at the endpoints of this interval, since these latter terms appear as boundary terms when performing the integration by parts inside and outside the interval. The oscillations of \tilde{w} are carefully chosen so that these errors are controlled by a continuous function of the discrepancy $|\tilde{u}(0) - \tilde{w}(0)|$, which by a Lipschitz estimate on \tilde{u} in this situation, is comparable to the length of the interval on which we substitute \tilde{l} .

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 copies of \tilde{w} into the original \tilde{w} , and passing to the limit, w , say. The final Lagrangian is of the form $(t, y, p) \mapsto \phi(t, y - w(t)) + p^2$, 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 . Many of the technicalities of the following construction are related to guaranteeing the existence and appropriate properties of w and ϕ , and are in some sense secondary to the main points of the proof. As indicated by the sketch of the argument above, we need to understand the first and second derivatives of w , and the shape of its convex hull as seen from each point of singularity x_n . This demands a number of conditions in the inductive construction of w , ensuring that while the function w oscillates as required around x_n , elsewhere the derivatives do not interfere in a significant way with the basic argument performed around x_n .

Define $g, \tilde{w}: \mathbb{R} \rightarrow \mathbb{R}$ by

$$g(t) = \begin{cases} t \log \log 1/|t| & t \neq 0, \\ 0 & t = 0; \end{cases} \text{ and } \tilde{w}(t) = \begin{cases} g(t) \sin \log \log \log 1/|t| & t \neq 0, \\ 0 & t = 0. \end{cases}$$

Then

$$\tilde{w} \in C^\infty(\mathbb{R} \setminus \{0\}), \tag{3}$$

and in particular \tilde{w}'' is bounded away from 0 and \tilde{w}' satisfies the fundamental theorem of calculus on closed intervals not including 0. Note that for $t \neq 0$,

$$\tilde{w}'(t) = (\log \log 1/|t|)(\sin \log \log \log 1/|t|) - \left(\frac{\sin \log \log \log 1/|t| + \cos \log \log \log 1/|t|}{\log 1/|t|} \right), \tag{4}$$

which is an even function. We have chosen $T > 0$ small enough such that $1/\log 1/|t| \leq 1 \leq 2 \leq \log \log 1/|t|$ for all $t \in [-2T, 2T] \setminus \{0\}$, and so for such t we have that

$$|\tilde{w}'(t)| \leq \log \log 1/|t| + \frac{2}{\log 1/|t|} \leq 3 \log \log 1/|t|, \tag{5}$$

and that

$$|\tilde{w}''(t)| \leq \frac{2}{|t| \log 1/|t|} \left(\frac{1}{(\log 1/|t|)(\log \log 1/|t|)} + \frac{1}{\log 1/|t|} + 1 \right) \leq \frac{6}{|t| \log 1/|t|},$$

and hence it follows that

$$|g(t)\tilde{w}''(t)| \leq \frac{6 \log \log 1/|t|}{\log 1/|t|} \rightarrow 0 \text{ as } 0 < |t| \rightarrow 0, \tag{6}$$

which is a key fact discussed above which encapsulates one sense in which the oscillations of \tilde{w} are sufficiently slow.

The following functions give us for each $t \in [-T, T]$ the exact coefficients we shall eventually need in our weight function $\tilde{\phi}$. We define $\psi^1, \psi^2: \mathbb{R} \rightarrow [0, \infty)$ by

$$\psi^1(t) = \begin{cases} \frac{1812}{|t|(\log 1/|t|)^{1/3}} & t \neq 0, \\ 0 & t = 0; \end{cases} \text{ and } \psi^2(t) = \begin{cases} 3 + 2|\tilde{w}''(t)| & t \neq 0, \\ 0 & t = 0; \end{cases}$$

and $\psi: \mathbb{R} \rightarrow [0, \infty)$ by $\psi(t) = \psi^1(t) + \psi^2(t)$. Note that by (6),

$$t \mapsto g(t)\psi(t) \text{ defines a continuous function on } \mathbb{R} \text{ with value 0 at 0.} \tag{7}$$

We may therefore define a constant $C \in (1, \infty)$ by

$$C := 1 + \sup_{t \in [-T, T]} 5|g(t)|\psi(t).$$

Let $\{x_n\}_{n=0}^\infty$ be a sequence in $(-T, T)$, with $x_0 = 0$. For each $n \geq 0$ define the translated functions $\tilde{w}_n, g_n, \psi_n^1, \psi_n^2, \psi_n : [-T, T] \rightarrow \mathbb{R}$ by composing the respective function with the translation $t \mapsto (t - x_n)$, thus $\tilde{w}_n(t) = \tilde{w}(t - x_n)$, etc.

For each $n \geq 1$, we define $\sigma_n \in (0, 1)$ by

$$\sigma_n := \min_{0 \leq i \leq n-1} |x_i - x_n|/2.$$

Observe for future reference that

$$|t - x_n| \leq \sigma_n \text{ implies that } |t - x_i| \geq \sigma_n \text{ for all } 0 \leq i \leq n - 1, \tag{8}$$

for otherwise we should have for some $0 \leq i \leq n - 1$ that

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

which contradicts the definition of σ_n .

We want to construct a sequence of absolutely continuous functions w_n , where for each $0 \leq i \leq n$, up to the addition of a scalar, $w_n = \tilde{w}_i$ on a neighbourhood of x_i , thus w_n is singular at x_i . We first define a decreasing sequence $T_n \in (0, 1)$ and hence intervals $Y_n := [x_n - T_n, x_n + T_n]$. In the inductive construction of w_n we shall modify w_{n-1} only on Y_n . A requirement that these intervals be small and decreasing in measure is the first step towards guaranteeing that the w_n converge to some limit function.

Define a sequence $K_n \in [1, \infty)$ by setting $K_0 = 1$ and so that for $n \geq 1$, we have

$$\sum_{i=0}^{n-1} (|\tilde{w}_i''(t)| + |\tilde{w}_i'(t)| + 1) \leq K_n \text{ whenever } |t - x_i| \geq \sigma_n \text{ for all } 0 \leq i \leq n - 1; \tag{9}$$

and

$$K_n \geq 1 + K_{n-1}. \tag{10}$$

Also for $n \geq 0$ define a sequence $\theta_n \in [1, \infty)$ by setting $\theta_0 = 1$ and for $n \geq 1$ setting

$$\theta_n = 10K_n\sigma_n^{-1}. \tag{11}$$

The scaling constant θ_n is an unimportant technicality, which just permits some useful estimates, and is chosen so that the graph of w_n always lies inside the “multi-graph” of $w_n(x_i) \pm \theta_i |g_i|$, for all $0 \leq i \leq n$; see (3.3) below for a precise statement. Little conceptual understanding would be lost by regarding the θ_n as constant, and e.g. equal to 1.

For $n \geq 0$ we define $T_n \in (0, 1)$ by setting $T_0 = T$ and for $n \geq 1$ inductively defining T_n such that the following conditions hold:

- (T:1) $T_n \leq |x_n \pm T| \sigma_n T_{n-1} / 2$; and
- (T:2) $|g_n(t)\psi_n(t)| \leq 2^{-n} / 5\theta_n$ for $t \in Y_n$.

Note that (T:2) is possible by (7). Since we will only modify w_{n-1} on Y_n to construct w_n , we only need to add more weight to our Lagrangian for $t \in Y_n$. 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 these conditions on T_n . We observe that (T:1) guarantees in particular that

$$T_n \leq 2^{-n} \text{ for all } n \geq 0. \tag{12}$$

Condition (T:1) also guarantees that the points in Y_n are far from the previous x_i , in a certain sense. That $T_n \leq \sigma_n$ implies that (8) holds in particular on Y_n , i.e. that

$$t \in Y_n \text{ implies that } |x_i - t| \geq \sigma_n \text{ for all } 0 \leq i \leq n - 1. \tag{13}$$

This stops the subintervals we later consider from overlapping.

We emphasize that these values of T_n are constructed independently of the later constructed w_n ; the inductive construction of these functions will require us to pass further down the sequence of T_n than induction would otherwise allow, as we now see. For $n \geq 0$, find $m_n \geq n$ such that

$$2^{-m_n} \leq \frac{T_{n+1}^2}{32}. \tag{14}$$

Choose a small open cover $G_n \subseteq [-T, T]$ of the points $\{x_i\}_{i=0}^{m_n}$ such that

$$\lambda(G_n) \leq \frac{T_{n+1}^2}{16C}, \tag{15}$$

and choose $M_n \in (1, \infty)$ such that

$$\sum_{i=0}^{m_n} (\max\{\psi_i(t), \psi(T_i)\}) \leq M_n \text{ whenever } t \in [-T, T] \setminus G_n. \tag{16}$$

We note also that since g is strictly increasing and $g_i(x_i) = g(0) = 0$, for each $n \geq 0$ there exists $\eta_n \in (0, 1)$ such that for all $0 \leq i \leq n - 1$,

$$g_i(t) \geq \eta_n \text{ whenever } |x_i - t| \geq \sigma_n. \tag{17}$$

Let $R_0 = T$ and for $n \geq 1$ inductively construct decreasing numbers $R_n \in (0, T_n)$, progressively smaller fractions of the corresponding T_n , such that:

(R:1)

$$\int_{-R_n}^{R_n} |\tilde{w}'|^2 \leq \frac{T_n^4}{1024 (1 + \|\tilde{w}'\|_{L^2(-T, T)})^2};$$

and

(R:2)

$$g(R_n) \leq \frac{2^{-n} R_{n-1} T_n^5 \eta_n}{(344 \cdot 512) (1 + \|\tilde{w}'\|_{L^2(-T, T)})^2 K_n^2 M_{n-1}}.$$

In (R:2), for brevity we enforce one single inequality, relating the smallness of R_n to all the other construction constants with which we shall have cause to compare it. On no one application will we need the precise right-hand side as an upper bound. Rather at various points we need upper bounds which are most conveniently combined in this one expression. Now define progressively smaller 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 arbitrarily highly

on Z_n , so we have to make the measure of this set very small to have any control over the convergence of w_n in $W^{1,2}(-T, T)$.

The next lemma gives us the very delicate construction of the sequence of functions w_n by which we shall ultimately define our minimizer w . The basic key facts are that w_n oscillates precisely like $\tilde{w}_n = \tilde{w}(\cdot - x_n)$ in a neighbourhood of x_n no larger than Z_n ; that w_n equals w_{n-1} off Y_n ; that on $Y_n \setminus Z_n$ both the first and second derivatives of w_n are controlled in terms of those of w_{n-1} , in precise ways which are necessary for the inductive construction and the convergence of the w_n ; and that the graph of w_n lies within that of $w_n(x_i) \pm 2\theta_i |g_i|$ for each $0 \leq i \leq n$.

Lemma 3 *There exists a sequence of $w_n \in W^{1,2}(-T, T)$ satisfying, for $n \geq 0$:*

- (3.1) $w_n(t) = \tilde{w}_n(t) + \rho_n$ when $t \in [x_n - \tau_n, x_n + \tau_n]$, for some $\tau_n \in (0, R_n]$, and some $\rho_n \in \mathbb{R}$;
- (3.2) w'_n exists and is locally Lipschitz on $(-T, T) \setminus \{x_i\}_{i=0}^n$;
- (3.3) $|w_n(t) - w_n(x_i)| \leq (2 - 2^{-n})\theta_i |g_i(t)|$ for all $t \in [-T, T]$ and for all $0 \leq i \leq n$;
- (3.4) $|w'_n(t)| \leq K_{n+1}$ when $|t - x_{n+1}| \leq \sigma_{n+1}$, in particular on Y_{n+1} ;
- (3.5) w''_n exists almost everywhere and $|w''_n(t)| \leq K_{n+1}$ for almost every t such that $|t - x_{n+1}| \leq \sigma_{n+1}$, in particular on Y_{n+1} ;

and for $n \geq 1$:

- (3.6) $w_n = w_{n-1}$ off Y_n ;
- (3.7) $\|w_n - w_{n-1}\|_\infty \leq 5K_n g(R_n)$;
- (3.8) $w_n(x_i) = w_{n-1}(x_i)$ for all $0 \leq i \leq n$;
- (3.9) $\|w'_n - w'_{n-1}\|_{L^2(-T, T)} \leq \frac{T_n^2}{16(1 + \|\tilde{w}'\|_{L^2(-T, T)})}$;
- (3.10) $|w'_n(t)| \leq |w'_{n-1}(t)| + 2^{-n}$ for almost every $t \notin [x_n - \tau_n, x_n + \tau_n]$; and
- (3.11) $|w''_n(t)| \leq |w''_{n-1}(t)| + 2^{-n}$ for almost every $t \notin [x_n - \tau_n, x_n + \tau_n]$.

Proof For the case $n = 0$, we easily check that setting $w_0 = \tilde{w}_0$ satisfies all the required conditions. Condition (3.1) is trivial for $\tau_0 = T$ and $\rho_0 = 0$; and (3.2) follows from (3). (3.3) is evident from the definition of \tilde{w} . Condition (3.3) is evident from the definition of \tilde{w} , since $w_0(x_0) = \tilde{w}(0) = 0$ and $\theta_0 \geq 1$. Conditions (3.4) and (3.5) are given precisely by (9), since the inequality in (9) holds for $|t - x_1| \leq \sigma_1$, as observed in (8).

Suppose for $n \geq 1$ that we have constructed w_i as claimed for all $0 \leq i \leq n - 1$. We demonstrate how to insert a 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 $T_n \leq \sigma_n \leq |x_n - x_i|/2$ for all $0 \leq i \leq n - 1$, and so that $x_i \notin Y_n$. Thus w'_{n-1} exists and is Lipschitz on Y_n by inductive hypothesis (3.2). Let $m := w'_{n-1}(x_n)$, so $|m| \leq K_n$ by inductive hypothesis (3.4). 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 each side of x_n separately—so that on each side of x_n we approach x_n on an affine function of gradient m (a different function on each 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.

Now, $\limsup_{t \downarrow 0} \tilde{w}'(t) = +\infty$ and $\liminf_{t \downarrow 0} \tilde{w}'(t) = -\infty$, so, recalling that \tilde{w}' is a continuous and even function, we can find $\tau_n \in (0, R_n]$ such that $\tilde{w}'(\pm\tau_n) = m$.

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

Let $\delta_{\pm} := m - w'_{n-1}(x_n \pm R_n)$, so by inductive hypothesis (3.5) we see that

$$|\delta_{\pm}| = |w'_{n-1}(x_n) - w'_{n-1}(x_n \pm R_n)| \leq \|w''_{n-1}\|_{L^\infty(Y_n)} R_n \leq K_n R_n. \tag{18}$$

Define

$$c_{\pm} := w_{n-1}(x_n) - w_{n-1}(x_n \pm R_n) + \tilde{w}(\pm\tau_n) \pm m(R_n - \tau_n).$$

The point is that the functions $t \mapsto m(t - (x_n \pm R_n)) + w_{n-1}(x_n \pm R_n) + c_{\pm}$ are affine functions with gradient m which take the values $w_{n-1}(x_n \pm R_n) + c_{\pm}$ at $t = (x_n \pm R_n)$ and the values $w_{n-1}(x_n) + \tilde{w}(\pm\tau_n)$ at $t = (x_n \pm \tau_n)$. By an application of the mean value theorem, the definition of \tilde{w} , and the inductive hypothesis (3.4), we have, recalling that $|m| \leq K_n$, that

$$\begin{aligned} |c_{\pm}| &\leq |w_{n-1}(x_n) - w_{n-1}(x_n \pm R_n)| + |\tilde{w}(\pm\tau_n)| + |m||R_n - \tau_n| \\ &\leq \|w'_{n-1}\|_{L^\infty(Y_n)} R_n + g(\tau_n) + K_n R_n \\ &\leq K_n R_n + g(\tau_n) + K_n R_n \\ &\leq 3K_n g(R_n), \end{aligned} \tag{19}$$

using also that $g(R_n) \geq R_n$ and that $K_n \geq 1$. Now let

$$d_{\pm} := \frac{4}{T_n} \left(\pm \frac{\delta_{\pm}}{2} (R_n - T_n/2) - c_{\pm} \right),$$

and define the piecewise affine functions $q_{\pm} : [-T, T] \rightarrow \mathbb{R}$ by stipulating

$$q_{\pm}(x_n \pm T_n) = 0 = q_{\pm}(x_n \pm T_n/2), \quad q_{\pm}(x_n \pm 3T_n/4) = \pm d_{\pm},$$

and

$$q_-(t) = \begin{cases} 0 & t \leq x_n - T_n, \\ \delta_- & t \geq x_n - R_n, \\ \text{affine} & \text{otherwise;} \end{cases} \quad \text{and} \quad q_+(t) = \begin{cases} \delta_+ & t \leq x_n + R_n, \\ 0 & t \geq x_n + T_n, \\ \text{affine} & \text{otherwise.} \end{cases}$$

These q_{\pm} will be the derivatives of the cut-off functions we will use. So by definition of d_{\pm} ,

$$\int_{-T}^{x_n - R_n} q_-(t) dt = \int_{x_n - T_n}^{x_n - R_n} q_-(t) dt = \frac{1}{2} \left(-\frac{T_n d_-}{2} + (T_n/2 - R_n) \delta_- \right) = c_-, \tag{20}$$

and

$$\int_{x_n + R_n}^T q_+(t) dt = \int_{x_n + R_n}^{x_n + T_n} q_+(t) dt = \frac{1}{2} \left(\delta_+ (T_n/2 - R_n) + \frac{d_+ T_n}{2} \right) = -c_+. \tag{21}$$

We need bounds on the first and second derivatives of the cut-off functions we will use, so we establish appropriate bounds on q_{\pm} and q'_{\pm} . Now, $\|q_{\pm}\|_{\infty} = \max\{|\delta_{\pm}|, |d_{\pm}|\}$. Note that (18) and (19) imply, using again that $g(R_n) \geq R_n$, that

$$\begin{aligned}
 |d_{\pm}| &\leq \frac{4}{T_n} \left(\frac{|\delta_{\pm}|}{2} (T_n/2 - R_n) + |c_{\pm}| \right) \leq \frac{4}{T_n} \left(\frac{T_n K_n R_n}{4} + 3 K_n g(R_n) \right) \\
 &= K_n R_n + \frac{12 K_n g(R_n)}{T_n} \\
 &\leq \frac{13 K_n g(R_n)}{T_n}.
 \end{aligned} \tag{22}$$

So, comparing with (18), we have that

$$\|q_{\pm}\|_{\infty} \leq \frac{13 K_n g(R_n)}{T_n}. \tag{23}$$

Also, q'_{\pm} exists almost everywhere and satisfies $\|q'_{\pm}\|_{L^{\infty}(-T, T)} = \max\{\frac{4|d_{\pm}|}{T_n}, \frac{|\delta_{\pm}|}{T_n/2 - R_n}\}$. Note firstly by (22) and (R:2) that

$$\frac{4|d_{\pm}|}{T_n} \leq \frac{4}{T_n} \left(\frac{13 K_n g(R_n)}{T_n} \right) = \frac{52 K_n g(R_n)}{T_n^2} \leq 2^{-n},$$

and secondly that since (R:2) in particular implies that $R_n \leq T_n/4$, using (18) and (R:2) we see that

$$\frac{|\delta_{\pm}|}{(T_n/2) - R_n} \leq \frac{4 K_n R_n}{T_n} \leq 2^{-n}.$$

Hence

$$\|q'_{\pm}\|_{L^{\infty}(-T, T)} \leq 2^{-n}. \tag{24}$$

We can now define our cut-off functions $\chi_{\pm}: [-T, T] \rightarrow \mathbb{R}$ by

$$\chi_{-}(t) = \int_{-T}^t q_{-}(s) ds, \text{ and } \chi_{+}(t) = c_{+} - \delta_{+}((x_n + R_n) - (-T)) + \int_{-T}^t q_{+}(s) ds.$$

Then $\chi_{\pm} \in C^1(-T, T)$ are such that $\chi'_{\pm} = q_{\pm}$ everywhere, $\chi''_{\pm} = q'_{\pm}$ almost everywhere, and, by (20) and (21), and the definition of q_{\pm} , we have that

$$\chi_{\pm}(x_n \pm T_n) = 0, \chi_{\pm}(x_n \pm R_n) = c_{\pm}, \chi'_{\pm}(x_n \pm R_n) = q_{\pm}(x_n \pm R_n) = \delta_{\pm}.$$

We can now define $w_n: [-T, T] \rightarrow \mathbb{R}$ by

$$w_n(t) = \begin{cases} w_{n-1}(t) + \chi_{-}(t) & t \leq x_n - R_n, \\ m(t - (x_n - R_n)) + w_{n-1}(x_n - R_n) + c_{-} & x_n - R_n < t < x_n - \tau_n, \\ w_{n-1}(x_n) + \tilde{w}_n(t) & x_n - \tau_n \leq t \leq x_n + \tau_n, \\ m(t - (x_n + R_n)) + w_{n-1}(x_n + R_n) + c_{+} & x_n + \tau_n < t < x_n + R_n, \\ w_{n-1}(t) + \chi_{+}(t) & x_n + R_n \leq t. \end{cases}$$

We see that w_n is continuous by construction. Condition (3.1) is immediate, with τ_n as defined, and $\rho_n = w_{n-1}(x_n)$. We note that since, by the definitions of q_{\pm} , $\chi_{-}(t) = 0$ for $t \leq x_n - T_n$, and $\chi_{+}(t) = 0$ for $t \geq x_n + T_n$, we have that $w_n = w_{n-1}$ off Y_n , as required for (3.6). To check (3.8), we let $0 \leq i \leq n$. If $i \leq n - 1$, then $x_i \notin Y_n$ since $T_n \leq \sigma_n$, so $w_n(x_i) = w_{n-1}(x_i)$ by (3.6), which we have just checked for n . 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 see that w'_n exists off $\{x_i\}_{i=0}^n$ by inductive hypothesis (3.2) and by construction: the values of δ_{\pm} and τ_n were chosen precisely so that the derivatives agree across the joins in the definition of w_n . The derivative is given by

$$w'_n(t) = \begin{cases} w'_{n-1}(t) + q_-(t) & t \in [-T, x_n - R_n] \setminus \{x_i\}_{i=0}^{n-1}, \\ m & x_n - R_n < t < x_n - \tau_n, \\ \tilde{w}'_n(t) & x_n - \tau_n \leq t < x_n, \quad x_n < t \leq x_n + \tau_n, \\ m & x_n + \tau_n < t < x_n + R_n, \\ w'_{n-1}(t) + q_+(t) & t \in [x_n + R_n, T] \setminus \{x_i\}_{i=0}^{n-1}. \end{cases}$$

This is locally Lipschitz on $(-T, T) \setminus \bigcup_{i=0}^n \{x_i\}$ by inductive hypothesis (3.2) on w'_{n-1} and since q_{\pm} are Lipschitz, and by (3), as required for condition (3.2).

We have constructed w_n so that except on $[x_n - \tau_n, x_n + \tau_n]$, the derivative is comparable with that of w_{n-1} , wherever both exist, which is everywhere except for the points $\{x_i\}_{i=0}^n$. For $t \notin Z_n \cup \{x_i\}_{i=0}^{n-1}$, we see by (23) that

$$|w'_n(t) - w'_{n-1}(t)| \leq |q_{\pm}(t)| \leq \frac{13K_n g(R_n)}{T_n}.$$

For $t \in Z_n \setminus [x_n - \tau_n, x_n + \tau_n]$, we note that since $R_n \leq T_n \leq \sigma_n$, we have that $|x_n - t| \leq \sigma_n$, and so we may use inductive hypothesis (3.5) to see that

$$\begin{aligned} |w'_n(t) - w'_{n-1}(t)| &= |m - w'_{n-1}(t)| = |w'_{n-1}(x_n) - w'_{n-1}(t)| \leq \|w''_{n-1}\|_{L^\infty(Y_n)} |t - x_n| \\ &\leq K_n R_n. \end{aligned} \tag{25}$$

Since (R:2) implies that $K_n R_n \leq \frac{13K_n g(R_n)}{T_n} \leq 2^{-n}$, in particular we gain condition (3.10).

The pointwise comparison we have just established between w'_n and w'_{n-1} fails in general on $[x_n - \tau_n, x_n + \tau_n]$. On this set, in fact, the whole point of the construction is that w'_n now equals $\tilde{w}'_n = \tilde{w}'(\cdot - x_n)$, which oscillates between arbitrarily large positive and negative values. On the other hand, w'_{n-1} , since it is locally Lipschitz away from the points $\{x_i\}_{i=1}^{n-1}$, may be regarded as basically constant on $[x_n - \tau_n, x_n + \tau_n]$, at least compared to the behaviour of w'_n . We chose R_n to be so small that despite this (large!) discrepancy on $[x_n - \tau_n, x_n + \tau_n] \subseteq Z_n$, the two derivatives are close in $L^2(-T, T)$, as stated in (3.9), which we now check. First note that, using the definition of w_n and (3.6) (which we have checked for n),

$$\begin{aligned} \int_{-T}^T |w'_n(t) - w'_{n-1}(t)|^2 dt &= \int_{Y_n} |w'_n(t) - w'_{n-1}(t)|^2 dt \\ &= \int_{x_n - T_n}^{x_n - R_n} |q_-(t)|^2 dt + \int_{Z_n \setminus [x_n - \tau_n, x_n + \tau_n]} |m - w'_{n-1}(t)|^2 dt \\ &\quad + \int_{x_n - \tau_n}^{x_n + \tau_n} |\tilde{w}'_n(t) - w'_{n-1}(t)|^2 dt + \int_{x_n + R_n}^{x_n + T_n} |q_+(t)|^2 dt. \end{aligned}$$

Now, by (23),

$$\begin{aligned} &\int_{x_n - T_n}^{x_n - R_n} |q_-(t)|^2 dt + \int_{x_n + R_n}^{x_n + T_n} |q_+(t)|^2 dt \\ &\leq \int_{x_n - T_n}^{x_n - R_n} \frac{(13K_n g(R_n))^2}{T_n^2} dt + \int_{x_n + R_n}^{x_n + T_n} \frac{(13K_n g(R_n))^2}{T_n^2} dt \end{aligned}$$

$$\begin{aligned} &\leq \frac{2T_n 169g(R_n)^2 K_n^2}{T_n^2} \\ &\leq \frac{338g(R_n) K_n^2}{T_n}, \end{aligned}$$

using also that $g(R_n) \leq 1$. Further, by (25), we have that

$$\begin{aligned} \int_{Z_n \setminus [x_n - \tau_n, x_n + \tau_n]} |m - w'_{n-1}(t)|^2 dt &\leq \int_{Z_n \setminus [x_n - \tau_n, x_n + \tau_n]} (K_n R_n)^2 \leq 2R_n (K_n R_n)^2 \\ &\leq 2R_n K_n^2. \end{aligned}$$

Finally, by inductive hypothesis (3.4), we have that

$$\begin{aligned} \int_{x_n - \tau_n}^{x_n + \tau_n} |\tilde{w}'_n(t) - w'_{n-1}(t)|^2 dt &\leq \int_{x_n - \tau_n}^{x_n + \tau_n} 2(|\tilde{w}'_n(t)|^2 + |w'_{n-1}(t)|^2) dt \\ &\leq 2\left(\int_{-\tau_n}^{\tau_n} |\tilde{w}'(t)|^2 dt + \int_{-\tau_n}^{\tau_n} K_n^2 dt\right) \\ &\leq 2\int_{-R_n}^{R_n} |\tilde{w}'(t)|^2 dt + 4K_n^2 \tau_n \\ &\leq 2\int_{-R_n}^{R_n} |\tilde{w}'(t)|^2 dt + 4K_n^2 R_n. \end{aligned}$$

Combining these estimates, using (R:1), (R:2), and that $g(R_n) \geq R_n$, we see that

$$\begin{aligned} &\int_{-T}^T |w'_n(t) - w'_{n-1}(t)|^2 dt \\ &\leq \frac{338g(R_n) K_n^2}{T_n} + 2R_n K_n^2 + 2\int_{-R_n}^{R_n} |\tilde{w}'(t)|^2 dt + 4K_n^2 R_n \\ &\leq \frac{344g(R_n) K_n^2}{T_n} + \frac{T_n^4}{512(1 + \|\tilde{w}'\|_{L^2(-T, T)})^2} \\ &\leq \frac{T_n^4}{512(1 + \|\tilde{w}'\|_{L^2(-T, T)})^2} + \frac{T_n^4}{512(1 + \|\tilde{w}'\|_{L^2(-T, T)})^2} \\ &= \frac{T_n^4}{256(1 + \|\tilde{w}'\|_{L^2(-T, T)})^2}. \end{aligned}$$

Taking square roots gives condition (3.9). Now, w''_n exists almost everywhere by inductive hypothesis (3.5) and by construction, and, where it does, is given by

$$w''_n(t) = \begin{cases} w''_{n-1}(t) + q'_-(t) & t < x_n - R_n, \\ 0 & x_n - R_n < t < x_n - \tau_n, \\ \tilde{w}''_n(t) & x_n - \tau_n < t < x_n, \quad x_n < t < x_n + \tau_n, \\ 0 & x_n + \tau_n < t < x_n + R_n, \\ w''_{n-1}(t) + q'_+(t) & x_n + R_n < t. \end{cases}$$

Thus by (24), for almost every $t \in (-T, T) \setminus Z_n$, we have that

$$|w''_n(t)| \leq |w''_{n-1}(t)| + |q'_\pm(t)| \leq |w''_{n-1}(t)| + 2^{-n}.$$

Condition (3.11) follows, since $w''_n = 0$ on $Z_n \setminus [x_n - \tau_n, x_n + \tau_n]$.

We now check (3.4) and (3.5). Suppose $|t - x_{n+1}| \leq \sigma_{n+1}$. Then by (8) the inequality in (9) holds, in particular

$$\sum_{i=0}^n (|\tilde{w}''_i(t)| + |\tilde{w}'_i(t)|) \leq K_{n+1},$$

for all $0 \leq i \leq n$, precisely by choice of K_{n+1} .

Let $0 \leq k \leq n$ be such that $t \in Y_k \setminus \bigcup_{i=k+1}^n Y_i$. Then by inductive hypothesis (3.6) for $k + 1, \dots, n$ (note we have checked this for n), we have that $w_n = w_k$ on a neighbourhood of t , so $w'_n(t) = w'_k(t)$ and $w''_n(t) = w''_k(t)$ where both sides exist, i.e. almost everywhere. We have to distinguish the cases of when t lies in $[x_k - \tau_k, x_k + \tau_k]$ and when it does not.

If $t \notin [x_k - \tau_k, x_k + \tau_k]$, then by inductive hypotheses (3.10) (we have checked this for $k = n$) and (3.4) (since $t \in Y_k$), and by (10), we have that

$$|w'_n(t)| = |w'_k(t)| \leq |w'_{k-1}(t)| + 2^{-k} \leq K_k + 1 \leq K_{n+1},$$

as required for (3.4). Similarly by inductive hypotheses (3.11) (we have checked this for $k = n$) and (3.5), and by (10), we have almost everywhere that

$$|w''_n(t)| = |w''_k(t)| \leq |w''_{k-1}(t)| + 2^{-k} \leq K_k + 1 \leq K_{n+1},$$

as required for (3.5).

If $t \in [x_k - \tau_k, x_k + \tau_k]$, observe first that $t \neq x_k$ since $|t - x_{n+1}| \leq \sigma_{n+1}$. So by inductive hypothesis (3.1) (we have checked this for $k = n$), we have, using (9), that

$$|w'_n(t)| = |w'_k(t)| = |\tilde{w}'_k(t)| \leq \sum_{i=0}^k |\tilde{w}'_i(t)| \leq \sum_{i=0}^n |\tilde{w}'_i(t)| \leq K_{n+1},$$

and, if $t \in (x_k - \tau_k, x_k + \tau_k)$ (we claim nothing about w''_n at the endpoints $x_k \pm \tau_k$), almost everywhere we have that

$$|w''_n(t)| = |w''_k(t)| = |\tilde{w}''_k(t)| \leq \sum_{i=0}^k |\tilde{w}''_i(t)| \leq \sum_{i=0}^n |\tilde{w}''_i(t)| \leq K_{n+1},$$

as required. So (3.4) and (3.5) hold in all cases.

We now move towards checking (3.7). Now observe that (22) and (18) imply that on $[-T, x_n - R_n]$, we have, by definition, that

$$\begin{aligned} |\chi_-(t)| &\leq \int_{-T}^{x_n - R_n} |q_-(s)| ds \leq \frac{1}{2} \left(\frac{T_n}{2} |d_-| + (T_n/2 - R_n) |\delta_-| \right) \\ &\leq \frac{T_n}{4} \frac{13K_n g(R_n)}{T_n} + \frac{T_n K_n R_n}{2} \\ &\leq \frac{13K_n g(R_n)}{4} + \frac{K_n g(R_n)}{2} \\ &\leq 4K_n g(R_n). \end{aligned}$$

The same estimate holds for χ_+ on $[x_n + R_n, T]$: we note first by (21) that

$$\begin{aligned} \chi_+(t) &= c_+ - \delta_+((x_n + R_n) + T) + \int_{-T}^t q_+(s) ds \\ &= c_+ + \int_{x_n+R_n}^t q_+(s) ds \\ &= - \int_{x_n+R_n}^T q_+(s) ds + \int_{x_n+R_n}^t q_+(s) ds \\ &= - \int_t^T q_+(s) ds, \end{aligned}$$

and hence, since $|\chi_+| \leq \int_{x_n+R_n}^T |q_+|$ on $[x_n + R_n, T]$, we can estimate $|\chi_+|$ as we estimated $|\chi_-|$ on $[-T, x_n - R_n]$ above. So, we have for all $t \in [-T, T] \setminus Z_n$ that

$$|w_n(t) - w_{n-1}(t)| = |\chi_{\pm}(t)| \leq 4K_n g(R_n).$$

By inductive hypothesis (3.4) and (19), we have for $t \in Z_n \setminus [x_n - \tau_n, x_n + \tau_n]$ that

$$\begin{aligned} |w_n(t) - w_{n-1}(t)| &\leq |m(t - (x_n \pm R_n))| + |w_{n-1}(x_n \pm R_n) - w_{n-1}(t)| + |c_{\pm}| \\ &\leq |m|(t - (x_n \pm R_n))| + \|w'_{n-1}\|_{L^\infty(Y_n)} |t - (x_n \pm R_n)| + |c_{\pm}| \\ &\leq K_n R_n + K_n R_n + 3K_n g(R_n) \\ &\leq 5K_n g(R_n). \end{aligned}$$

Finally for $x_n - \tau_n \leq t \leq x_n + \tau_n$, by inductive hypothesis (3.4), the definition of \tilde{w} , and the monotonicity of g , we have that

$$\begin{aligned} |w_n(t) - w_{n-1}(t)| &= |(\tilde{w}_n(t) + w_{n-1}(x_n)) - w_{n-1}(t)| \leq |\tilde{w}_n(t)| + |w_{n-1}(x_n) - w_{n-1}(t)| \\ &\leq g(\tau_n) + \|w'_{n-1}\|_{L^\infty(Y_n)} |x_n - t| \\ &\leq g(R_n) + K_n R_n \\ &\leq 2K_n g(R_n). \end{aligned}$$

Hence we have that $\|w_n - w_{n-1}\|_\infty \leq 5K_n g(R_n)$, as required for (3.7).

We can now check (3.3). First consider $0 \leq i \leq n - 1$. The result is immediate by inductive hypothesis if $t \notin Y_n$, by (3.8) and (3.6), both of which we have checked for n . So suppose $t \in Y_n$. Then $|g_i(t)| \geq \eta_n$ by (13) and (17). Condition (3.7), which we have checked for n , and (R:2) imply that

$$|w_n(t) - w_{n-1}(t)| \leq 5K_n g(R_n) \leq 2^{-n} \eta_n \leq 2^{-n} |g_i(t)| \leq 2^{-n} \theta_i |g_i(t)|,$$

since $\theta_i \geq 1$. Condition (3.8) and inductive hypothesis (3.3) imply that

$$|w_{n-1}(t) - w_n(x_i)| = |w_{n-1}(t) - w_{n-1}(x_i)| \leq \left(2 - 2^{-(n-1)}\right) \theta_i |g_i(t)|.$$

So

$$\begin{aligned} |w_n(t) - w_n(x_i)| &\leq |w_n(t) - w_{n-1}(t)| + |w_{n-1}(t) - w_{n-1}(x_i)| \\ &\leq 2^{-n} \theta_i |g_i(t)| + \left(2 - 2^{-(n-1)}\right) \theta_i |g_i(t)| \\ &= (2 - 2^{-n}) \theta_i |g_i(t)|. \end{aligned}$$

It just remains to check (3.3) in the case $i = n$. We first show that for all $t \in [-T, T]$ we have chosen θ_n such that

$$|w_{n-1}(t) - w_{n-1}(x_n)| \leq \theta_n |g_n(t)|/2. \tag{26}$$

This is the motivating factor behind the choice of θ_n : blowing up the graph of $w_{n-1}(x_n) \pm |g_n| = w_n(x_n) \pm |g_n|$ so that it encloses that of w_{n-1} . Now, for $|t - x_n| \leq \sigma_n$, we have by inductive hypothesis (3.4) and (11), since $\log \log 1/|t - x_n| \geq 2$, that

$$\begin{aligned} |w_{n-1}(t) - w_{n-1}(x_n)| &\leq \|w'_{n-1}\|_{L^\infty(x_n - \sigma_n, x_n + \sigma_n)} |t - x_n| \leq K_n |t - x_n| \leq \theta_n |t - x_n| \\ &\leq \theta_n |g_n(t)|/2. \end{aligned}$$

If $|t - x_n| \geq \sigma_n$, then by inductive hypothesis (3.7), (R:2), and (11), and since T was chosen small enough such that $g(T) \leq 1 \leq 2 \leq \log \log 1/2T$, we have that

$$\begin{aligned} |w_{n-1}(t) - w_{n-1}(x_n)| &\leq |w_{n-1}(t) - w_0(t)| + |w_0(t) - w_0(x_n)| + |w_0(x_n) - w_{n-1}(x_n)| \\ &\leq 2\|w_0\|_\infty + 2\|w_{n-1} - w_0\|_\infty \\ &\leq 2 \left(g(T) + \sum_{i=1}^{n-1} \|w_i - w_{i-1}\|_\infty \right) \\ &\leq 2 \left(g(T) + \sum_{i=1}^{n-1} 5K_i g(R_i) \right) \\ &\leq 2 \left(g(T) + \sum_{i=1}^{n-1} 2^{-i} \right) \\ &\leq 4 \\ &\leq \theta_n \sigma_n \\ &\leq \frac{\theta_n \sigma_n (\log \log 1/2T)}{2} \\ &\leq \frac{\theta_n |t - x_n| (\log \log 1/2T)}{2} \\ &\leq \frac{\theta_n |t - x_n| \log \log 1/|t - x_n|}{2} \\ &= \theta_n |g_n(t)|/2, \end{aligned}$$

as claimed.

To check (3.3) in this final case, suppose first that $t \in [x_n - \tau_n, x_n + \tau_n]$. Then by (3.1), and the definition of \tilde{w} we have, since $\theta_n \geq 1$, that

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

To deal with the case $x_n - R_n \leq t < x_n - \tau_n$, we note first that the condition is satisfied at the endpoints of the interval. That it holds for $t = x_n - \tau_n$ has just been established. Using the definition of w_n , (3.8), inductive hypothesis (3.4), (19), and (11) we see that

$$\begin{aligned} |w_n(x_n - R_n) - w_n(x_n)| &= |w_{n-1}(x_n - R_n) + c_- - w_{n-1}(x_n)| \\ &\leq \|w'_{n-1}\|_{L^\infty(Y_n)} R_n + |c_-| \\ &\leq K_n R_n + 3K_n g(R_n) \\ &\leq 4K_n g(R_n) \end{aligned}$$

$$\begin{aligned} &\leq \theta_n g(R_n) \\ &\leq (2 - 2^{-n})\theta_n |g_n(x_n - R_n)|. \end{aligned}$$

So the condition holds at $x_n - R_n$ and $x_n - \tau_n$. Since w_n is defined to be affine between these points, and $|g_n|$ is concave on $[-T, x_n]$, the result holds for all $t \in [x_n - R_n, x_n - \tau_n]$. Similarly the result holds for all $t \in [x_n + \tau_n, x_n + R_n]$. Finally we have to consider $t \notin [x_n - R_n, x_n + R_n]$. In this case we have by monotonicity of g that $g_n(t) \geq g(R_n)$, and so we see using (3.8), (3.7), (26), and (11) that

$$\begin{aligned} |w_n(t) - w_n(x_n)| &\leq |w_n(t) - w_{n-1}(t)| + |w_{n-1}(t) - w_n(x_n)| \\ &\leq \|w_n - w_{n-1}\|_\infty + |w_{n-1}(t) - w_{n-1}(x_n)| \\ &\leq 5K_n g(R_n) + \theta_n |g_n(t)|/2 \\ &\leq 5K_n |g_n(t)| + \theta_n |g_n(t)|/2 \\ &\leq \theta_n |g_n(t)|/2 + \theta_n |g_n(t)|/2 \\ &\leq (2 - 2^{-n})\theta_n |g_n(t)|. \end{aligned}$$

Thus (3.3) holds for all $t \in [-T, T]$ as claimed. □

We now show that this sequence converges to some $w \in W^{1,2}(-T, T)$. This w will be our minimizer.

Lemma 4 *The sequence $\{w_n\}_{n=0}^\infty$ converges uniformly to some function $w \in W^{1,2}(-T, T)$ such that for all $n \geq 0$,*

- (4.1) $w(x_i) = w_n(x_i)$ for all $0 \leq i \leq n + 1$;
- (4.2) $\|w - w_n\|_\infty \leq 10K_{n+1}g(R_{n+1})$;
- (4.3) $\|w' - w'_n\|_{L^2(-T,T)} \leq \frac{T_{n+1}^2}{8(1 + \|\tilde{w}'\|_{L^2(-T,T)})}$; and
- (4.4) $|w(t) - w(x_n)| \leq 2\theta_n |g_n(t)|$ for all $t \in [-T, T]$.

Proof Let $m \geq n + 1 \geq 1$. (R:2) in particular implies that $K_m g(R_m) \leq g(R_{m-1})/2$, so combining with (3.7), we see that

$$\begin{aligned} \|w_m - w_n\|_\infty &\leq \|w_m - w_{m-1}\|_\infty + \dots + \|w_{n+1} - w_n\|_\infty \\ &\leq 5(K_m g(R_m) + \dots + K_{n+1}g(R_{n+1})) \\ &\leq 5\left(2^{-(m-(n+1))} + \dots + 1\right) K_{n+1}g(R_{n+1}) \\ &\leq 10K_{n+1}g(R_{n+1}). \end{aligned}$$

Hence, since (R:2) certainly implies that this tends to 0 as $n \rightarrow \infty$, the sequence $\{w_n\}_{n=0}^\infty$ is uniformly Cauchy, and so converges uniformly to some $w \in C(-T, T)$, which satisfies (4.2). Conditions (4.1) and (4.4) follow directly by taking limits in conditions (3.8) and (3.3) respectively.

Now, by (3.9) and (T:1), we have that

$$\begin{aligned} \|w'_m - w'_n\|_{L^2(-T,T)} &\leq \|w'_m - w'_{m-1}\|_{L^2(-T,T)} + \dots + \|w'_{n+1} - w'_n\|_{L^2(-T,T)} \\ &\leq \frac{T_m^2 + \dots + T_{n+1}^2}{16(1 + \|\tilde{w}'\|_{L^2(-T,T)})} \\ &\leq \frac{T_{n+1}^2}{8(1 + \|\tilde{w}'\|_{L^2(-T,T)})}, \end{aligned} \tag{27}$$

and, since likewise $T_n \rightarrow 0$ as $n \rightarrow \infty$, w'_n is Cauchy in $L^2(-T, T)$, and it follows that $w_n \rightarrow w$ in $W^{1,2}(-T, T)$, and that (4.3) holds. \square

2.2 Singularity

Having defined the function $w \in W^{1,2}(-T, T)$, we now check that it exhibits the required oscillating behaviour around each point x_n . The extra oscillations we added in to w_n are 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, the limiting behaviour of the difference quotients of w at x_n is the same as that of the difference quotients of \tilde{w} at 0, for each $n \geq 0$.

Lemma 5 *Let $n \geq 0$.*

Then $\overline{D}w(x_n) = +\infty$ and $\underline{D}w(x_n) = -\infty$.

Proof Let $m \geq n + 1$, and let $t \in [-T, T]$ be such that $|t - x_n| \leq T_m$. Note that if $t \in Y_i$ for $i \geq n + 1$, we have, since (T:1) implies that $T_i \leq \sigma_i$, that

$$|x_n - x_i| \leq |x_n - t| + |t - x_i| \leq |x_n - t| + T_i \leq |x_n - t| + |x_n - x_i|/2,$$

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

$$T_i \leq |x_n - x_i|/2 \leq |x_n - t| \leq T_m. \tag{28}$$

Since the T_i are decreasing, this implies that $i \geq m$.

If $t \notin Y_i$ for any $i \geq n + 1$ then $w(t) = w_n(t)$ by (3.6), and the following argument is trivial. Otherwise choose the least $i \geq n + 1$ such that $t \in Y_i$, so $w_n(t) = w_{i-1}(t)$. Then by (4.2), (R:2), and (28),

$$\begin{aligned} |w(t) - w_n(t)| &= |w(t) - w_{i-1}(t)| \leq \|w - w_{i-1}\|_\infty \leq 10K_i g(R_i) \\ &< 2^{-i} T_i \leq 2^{-i} |t - x_n|. \end{aligned}$$

Hence by (4.1) and since, by the above argument, $i \geq m$, we have that

$$\left| \frac{w(t) - w(x_n)}{t - x_n} - \frac{w_n(t) - w_n(x_n)}{t - x_n} \right| = \left| \frac{w(t) - w_n(t)}{t - x_n} \right| \leq \frac{2^{-i} |t - x_n|}{|t - x_n|} \leq 2^{-m}. \tag{29}$$

As $t \rightarrow x_n$, we may choose $m \rightarrow \infty$. Hence by (3.1) and the definition of \tilde{w}_n ,

$$\overline{D}w(x_n) = \overline{D}w_n(x_n) = \overline{D}\tilde{w}_n(x_n) = +\infty,$$

and

$$\underline{D}w(x_n) = \underline{D}w_n(x_n) = \underline{D}\tilde{w}_n(x_n) = -\infty.$$

\square

2.3 Construction of the Lagrangian

We now construct the Lagrangian which shall define the variational problem of which w will be the unique minimizer. Our basic weight function $\tilde{\phi}: [-T, T] \times \mathbb{R} \rightarrow [0, \infty)$ will be given by

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

We need some bound of the form $|\tilde{\phi}(t, y)| \leq c|g(t)|\psi(t)$ to ensure continuity of $\tilde{\phi}$; it turns out (see Lemma 7) that sensitive tracking of $|y|$, which shall represent the distance of a putative minimizer from our constructed function w , only for $|y| \leq 5|g(t)|$ suffices in the proof of minimality. Our function \tilde{w} was constructed precisely so that (6) and hence (7) 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 $\tilde{\phi}_n^1, \tilde{\phi}_n^2: [-T, T] \times \mathbb{R} \rightarrow [0, \infty)$ as follows. We recall that we need extra weight only on Y_n , so we define for $k = 1, 2$ and $(t, y) \in Y_n \times \mathbb{R}$,

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

and extend to a function on $[-T, T] \times \mathbb{R}$ by setting it to continue constantly at the value attained at the endpoints of Y_n , i.e. defining for $(t, y) \in ([-T, T] \setminus Y_n) \times \mathbb{R}$

$$\tilde{\phi}_n^k(t, y) = \begin{cases} 5\psi^k(T_n)\theta_n g(T_n) & |y| \geq 5\theta_n g(T_n), \\ \psi^k(T_n)|y| & |y| \leq 5\theta_n g(T_n). \end{cases}$$

Define $\tilde{\phi}_n: [-T, T] \times \mathbb{R} \rightarrow [0, \infty)$ by $\tilde{\phi}_n(t, y) = \tilde{\phi}_n^1(t, y) + \tilde{\phi}_n^2(t, y)$, which is continuous by (7).

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

$$\begin{aligned} \tilde{\phi}_n^k(t, y) &\leq \tilde{\phi}_n^k(t, z) \text{ whenever } |y| \leq |z|; \\ \text{Lip}(\tilde{\phi}_n^k(t, \cdot)) &\leq \max\{\psi_n^k(t), \psi^k(T_n)\}; \text{ and} \\ \tilde{\phi}_n^k(t, 0) &= 0. \end{aligned}$$

The last result is obvious, as are the other results for $t = x_n$. Suppose $t \in Y_n \setminus \{x_n\}$. First consider the case in which $|y| \leq |z| \leq 5\theta_n|g_n(t)|$. Then

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

and so

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

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

In the case when $5\theta_n|g(t)| \leq |y| \leq |z|$, we have that

$$\tilde{\phi}_n^k(t, y) = 5\theta_n|g_n(t)|\psi_n^k(t) = \tilde{\phi}_n^k(t, z),$$

and so both results are immediate. In the case in which $|y| \leq 5\theta_n|g_n(t)| \leq |z|$, we have that

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

and so

$$|\tilde{\phi}_n^k(t, z) - \tilde{\phi}_n^k(t, y)| = \psi_n^k(t)(5\theta_n|g_n(t)| - |y|) \leq \psi_n^k(t)(|z| - |y|) \leq \psi_n^k(t)|z - y|.$$

Thus in this case again $\text{Lip}(\tilde{\phi}_n^k(t, \cdot)) \leq \psi_n^k(t)$. Both results follow similarly for $t \notin Y_n$: we obtain instead that $\text{Lip}(\tilde{\phi}_n^k(t, \cdot)) \leq \psi^k(T_n)$, hence the full result claimed. Hence for all $t \in [-T, T]$, $|y| \mapsto \tilde{\phi}_n(t, y)$ is an increasing function with Lipschitz constant at most $\max\{\psi_n(t), \psi(T_n)\}$, and $\tilde{\phi}(t, 0) = 0$. This Lipschitz constant blows up as we approach x_n ,

since $\psi_n(t) \geq |\tilde{w}_n''(t)| \rightarrow \infty$, but we shall not need to use a Lipschitz estimate of this function arbitrarily close to x_n .

Defining $\phi_n : [-T, T] \times \mathbb{R} \rightarrow [0, \infty)$ by $\phi_n(t, y) = \sum_{i=0}^n \tilde{\phi}_i(t, y)$ gives a sequence of continuous functions such that for each $t \in [-T, T]$,

$$\phi_n(t, y) \leq \phi_n(t, z) \text{ whenever } |y| \leq |z|; \tag{30}$$

$$\text{Lip}(\phi_n(t, \cdot)) \leq \sum_{i=0}^n (\max\{\psi_i(t), \psi(T_i)\}); \text{ and} \tag{31}$$

$$\phi_n(t, 0) = 0. \tag{32}$$

For $n \geq 1$, by (T:2), we see that $0 \leq \tilde{\phi}_n(t, y) \leq \sup_{t \in Y_n} 5\psi_n(t)\theta_n|g_n(t)| \leq 2^{-n}$ for all $(t, y) \in [-T, T] \times \mathbb{R}$. Hence the sequence $\{\phi_n\}_{n=0}^\infty$ converges uniformly to a continuous function given by $\phi(t, y) = \sum_{i=0}^\infty \tilde{\phi}_i(t, y)$, satisfying

$$\|\phi\|_\infty \leq \|\tilde{\phi}_0\|_\infty + \sum_{i=1}^\infty \|\tilde{\phi}_i\|_\infty \leq \|\tilde{\phi}_0\|_\infty + \sum_{i=1}^\infty 2^{-i} = \|\tilde{\phi}_0\|_\infty + 1 = C; \text{ and} \tag{33}$$

$$\|\phi - \phi_n\|_\infty \leq \sum_{i=n+1}^\infty \|\tilde{\phi}_i\|_\infty \leq \sum_{i=n+1}^\infty 2^{-i} = 2^{-n}. \tag{34}$$

By passing to the limit in (30) and (32) we see that for each $t \in [-T, T]$,

$$\phi(t, y) \leq \phi(t, z) \text{ whenever } |y| \leq |z|; \text{ and} \tag{35}$$

$$\phi(t, 0) = 0. \tag{36}$$

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

We can now define a functional \mathcal{L} on $W^{1,1}(-T, T)$, with a continuous Lagrangian, superlinear and convex in p , by

$$\mathcal{L}(u) = \int_{-T}^T (\phi(t, u(t) - w(t)) + (u'(t))^2) dt,$$

and consider minimizing \mathcal{L} over $\mathcal{A}_{w(-T), w(T)}$. Since evidently \mathcal{L} is coercive on $W^{1,2}(-T, T)$, a minimizer over $\mathcal{A}_{w(-T), w(T)}$ exists in $W^{1,2}(-T, T)$, and we can regard the minimization problem as being defined on $W^{1,2}(-T, T)$.

2.4 Minimality

We shall find certain approximations to our functional \mathcal{L} useful, and so will define for all $n \geq 0$ the functional \mathcal{L}_n on $W^{1,2}(-T, T)$ by

$$\mathcal{L}_n(u) = \int_{-T}^T (\phi(t, u(t) - w_n(t)) + (u'(t))^2) dt.$$

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 6 *Let $u \in W^{1,2}(-T, T)$ and $n \geq 0$. Then*

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

Proof We first estimate $|\mathcal{L}(u) - \mathcal{L}_n(u)|$. Recall our definitions of $m_n \geq n$, $M_n \geq 0$, and $G_n \supseteq \bigcup_{i=0}^{m_n} \{x_i\}$ from the beginning of the construction. Let $t \in [-T, T] \setminus G_n$. We see by (31), and precisely by the choice of M_n in (16), that

$$\text{Lip}(\phi_{m_n}(t, \cdot)) \leq \sum_{i=0}^{m_n} (\max\{\psi_i(t), \psi(T_i)\}) \leq M_n.$$

This is the one occasion on which we shall use the (in principle very large) Lipschitz constant of $\phi_{m_n}(t, \cdot)$. The purpose of the open set G_n was to avoid using this number arbitrarily near $\{x_i\}_{i=0}^{m_n}$, at which points it blows up.

Using (4.2) and (R:2) we see that

$$\begin{aligned} |\phi_{m_n}(t, u - w) - \phi_{m_n}(t, u - w_n)| &\leq \text{Lip}(\phi_{m_n}(t, \cdot))|(u(t) - w(t)) - (u(t) - w_n(t))| \\ &\leq M_n \|w - w_n\|_\infty \\ &\leq 10M_n K_{n+1} g(R_{n+1}) \\ &\leq \frac{T_{n+1}^2}{16}. \end{aligned}$$

The choice of m_n in (14) and (34) imply that

$$\|\phi - \phi_{m_n}\|_\infty \leq 2^{-m_n} \leq \frac{T_{n+1}^2}{32}.$$

Hence

$$\begin{aligned} |\phi(t, u - w) - \phi(t, u - w_n)| &\leq |\phi(t, u - w) - \phi_{m_n}(t, u - w)| \\ &\quad + |\phi_{m_n}(t, u - w) - \phi_{m_n}(t, u - w_n)| \\ &\quad + |\phi_{m_n}(t, u - w_n) - \phi(t, u - w_n)| \\ &\leq 2\|\phi - \phi_{m_n}\|_\infty + \frac{T_{n+1}^2}{16} \\ &\leq \frac{2T_{n+1}^2}{32} + \frac{T_{n+1}^2}{16} \\ &= \frac{T_{n+1}^2}{8}. \end{aligned}$$

Now, using (33) and the choice of the measure of G_n in (15), we have that

$$\int_{G_n} |\phi(t, u - w) - \phi(t, u - w_n)| \leq 2 \int_{G_n} \|\phi\|_\infty \leq 2C\lambda(G_n) \leq 2C \frac{T_{n+1}^2}{16C} = \frac{T_{n+1}^2}{8}.$$

Combining these estimates, we see that

$$\begin{aligned} &|\mathcal{L}(u) - \mathcal{L}_n(u)| \\ &= \left| \int_{-T}^T (\phi(t, u - w) + (u')^2) - (\phi(t, u - w_n) + (u')^2) \right| \\ &\leq \int_{-T}^T |\phi(t, u - w) - \phi(t, u - w_n)| \\ &= \int_{G_n} |\phi(t, u - w) - \phi(t, u - w_n)| + \int_{[-T, T] \setminus G_n} |\phi(t, u - w) - \phi(t, u - w_n)| \end{aligned}$$

$$\begin{aligned} &\leq \frac{T_{n+1}^2}{8} + \int_{[-T, T] \setminus G_n} \frac{T_{n+1}^2}{8} \\ &\leq \frac{T_{n+1}^2}{8} + \frac{T_{n+1}^2}{8} \\ &= \frac{T_{n+1}^2}{4}. \end{aligned}$$

Now we estimate $|\mathcal{L}(w) - \mathcal{L}_n(w_n)|$. First we compare w' and w'_n with $w_0 = \tilde{w}$ in the L^2 -norm, noting that (4.3) and (27) in particular allow the estimates

$$\|w'\|_{L^2(-T, T)} \leq 1 + \|\tilde{w}'\|_{L^2(-T, T)}, \text{ and } \|w'_n\|_{L^2(-T, T)} \leq 1 + \|\tilde{w}'\|_{L^2(-T, T)}.$$

Hence it follows that

$$\|w' + w'_n\|_{L^2(-T, T)} \leq \|w'\|_{L^2(-T, T)} + \|w'_n\|_{L^2(-T, T)} \leq 2(1 + \|\tilde{w}'\|_{L^2(-T, T)}).$$

Thus using (36), Cauchy–Schwartz, and (4.3), we see that

$$\begin{aligned} |\mathcal{L}(w) - \mathcal{L}_n(w_n)| &\leq \int_{-T}^T |(w')^2 - (w'_n)^2| = \int_{-T}^T (|w' + w'_n||w' - w'_n|) \\ &\leq \|w' + w'_n\|_{L^2(-T, T)} \|w' - w'_n\|_{L^2(-T, T)} \\ &\leq \frac{2(1 + \|\tilde{w}'\|_{L^2(-T, T)}) T_{n+1}^2}{8(1 + \|\tilde{w}'\|_{L^2(-T, T)})} \\ &= \frac{T_{n+1}^2}{4}. \end{aligned}$$

Combining these two estimates we see that

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

□

We now show that w is the unique solution of our minimization problem. The basic idea on \tilde{w} which we sketched at the beginning of this section 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

$$W^{1,2}(-T, T) \ni u \mapsto \mathcal{L}_n(u)$$

over those u such that $u(\pm T) = w_n(\pm T) (= w(\pm T))$.

Let $u \in W^{1,2}(-T, T)$ be a minimizer of \mathcal{L} over $\mathcal{A}_{w(-T), w(T)}$, and suppose for a contradiction that $u \neq w$. Note that a minimizer certainly exists, since the Lagrangian is continuous, and superlinear and convex in p . We now make a number of estimates, with the eventual aim of showing that

$$\mathcal{L}(u) - \mathcal{L}(w) = \int_{-T}^T ((u')^2 + \phi(t, u - w) - (w')^2) > 0,$$

which contradicts the choice of u as a minimizer. If $u(x_n) = w(x_n)$ for all $n \geq 0$, then the proof is in principle 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 $[-T, T]$, we should immediately have $u = w$ by continuity, thus concluding the proof of minimality of w without using the assumption 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 such x_n .

Lemma 7 *Let $n \geq 0$ be such that $u(x_n) \neq w(x_n)$. Let $a_n, b_n > 0$ be such that $J_n := (x_n - a_n, x_n + b_n)$ is the connected component in $[-T, T]$ containing x_n of those points such that $|u(t) - w(x_n)| > 3\theta_n |g_n(t)|$, so $|u(x_n - a_n) - w(x_n)| = 3\theta_n g(a_n)$ and $|u(x_n + b_n) - w(x_n)| = 3\theta_n g(b_n)$. (Note that $J_n \subsetneq [-T, T]$ since u and w agree at $\pm T$ and so by (4.4)*

$$|u(\pm T) - w(x_n)| = |w(\pm T) - w(x_n)| \leq 2\theta_n |g_n(\pm T)|.$$

Then

$$\begin{cases} |(u - w_n)(t)| \geq \theta_n g(b_n) \text{ for } t \in [x_n, x_n + b_n] & b_n \geq a_n, \\ |(u - w_n)(t)| \geq \theta_n g(a_n) \text{ for } t \in [x_n - a_n, x_n] & a_n \geq b_n, \end{cases} \tag{37}$$

and

$$|u(t) - w(x_n)| \leq 3\theta_n |g_n(t)| \text{ for } t \notin J_n.$$

Proof We suppose that $u(x_n) > w(x_n)$. The argument for the case in which $u(x_n) < w(x_n)$ is very similar. We choose $\alpha_n, \beta_n > 0$ such that $(x_n - \alpha_n, x_n + \beta_n)$ is the connected component in $[-T, T]$ containing x_n of those points such that $|u(t) - w(x_n)| > 2\theta_n |g_n(t)|$. So $a_n \leq \alpha_n$ and $b_n \leq \beta_n$, and $[x_n - \alpha_n, x_n + \beta_n] \subseteq [-T, T]$. We prove that u is convex on $(x_n - \alpha_n, x_n + \beta_n)$. In the case in which $u(x_n) < w(x_n)$, we would prove that u is concave on $(x_n - \alpha_n, x_n + \beta_n)$. Suppose for a contradiction that on some non-trivial subinterval (t_1, t_2) of $(x_n - \alpha_n, x_n + \beta_n)$, u lies above its chord between the points t_1 and t_2 , i.e. that there exists some $\mu \in [0, 1]$ such that

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

The basic idea is that in this case we can redefine u to be affine on (t_1, t_2) or some subinterval of (t_1, t_2) , producing a function which does not increase the weight term $\phi(t, \cdot - w(t))$ of the integrand, since it only moves closer to w , and strictly decreases the gradient term, since it has constant gradient. Let $z: [-T, T] \rightarrow \mathbb{R}$ be the affine function with graph passing through $(t_1, u(t_1))$ and $(t_2, u(t_2))$, so

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

So we have by assumption on t_1, t_2 that

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

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

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

We then show that this contradicts the choice of u as a minimizer.

Since z is affine and $|g_n|$ is monotonic and concave on $[-T, x_n]$ and $[x_n, T_0]$, the equation $z = w(x_n) + 2\theta_n |g_n|$ can in principle have no or up to three distinct solutions on (t_1, t_2) . If

there is at most one solution, then since $z(t_i) = u(t_i) \geq w(x_n) + 2\theta_n|g_n(t_i)|$ for $i = 1, 2$, evidently $z \geq w(x_n) + 2\theta_n|g_n|$ on (t_1, t_2) . So 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 g_n and z , all three points cannot lie on one side of x_n . So suppose $s_1 < x_n \leq s_2 < s_3$. The principle here is that z must have a positive gradient if it intersects $w(x_n) + 2\theta_n|g_n|$ twice on the right of x_n . This then forces z to lie below $w(x_n) + 2\theta_n|g_n|$ for all points $t < s_1 < x_n$, which is a contradiction since it agrees with u at $t = t_1 < s_1$, and u lies above $w(x_n) + 2\theta_n|g_n|$ at this point. More precisely, for $t < x_n$, we have that

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

Since $t_1 < s_1 < x_n$, we have $|g_n(s_1)| = -g_n(s_1)$ and $|g_n(t_1)| = -g_n(t_1)$, so

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

This is a contradiction since $z(t_1) = u(t_1) \geq w(x_n) + 2\theta_n|g_n(t_1)|$. We can deal similarly with the case $s_1 < s_2 \leq x_n < s_3$.

So it remains to deal with the case in which we have precisely two distinct solutions (s_1, s_2) —this is the case in which we have in general to add a constant to z , since it is possible that $w(x_n) + 2\theta_n|g_n|$ lies above z on some subinterval of (s_1, s_2) . The same considerations as in the preceding paragraph show that we must have both solutions lying on one side of x_n . Suppose $x_n \leq s_1 < s_2$. Then $2|g_n| = 2g_n$ is C^∞ 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

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

Define \tilde{z} by

$$\tilde{z}(t) = z' \cdot (t - s_0) + w(x_n) + 2\theta_n g_n(s_0),$$

the tangent to $w(x_n) + 2\theta_n g_n$ at s_0 , so, since $s_0 \in (s_1, s_2) \subseteq (t_1, t_2) \subseteq (x_n - \alpha_n, x_n + \beta_n)$,

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

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

We claim that $\tilde{z} \geq w(x_n) + 2\theta_n|g_n|$ on $(\tilde{t}_1, \tilde{t}_2)$. Since $s_0 > s_1 \geq x_n$ and $\tilde{z}(s_0) = w(x_n) + 2\theta_n|g_n(s_0)|$, with $\tilde{z}' = z' = 2\theta_n g'_n(s_0)$, by concavity of g we have that $\tilde{z} \geq w(x_n) + 2\theta_n|g_n|$ on (x_n, T) . Suppose there existed $s \in (\tilde{t}_1, x_n]$ such that $\tilde{z}(s) < w(x_n) + 2\theta_n|g_n(s)| = w(x_n) - 2\theta_n g_n(s)$. Then we see as before, since $\tilde{z}' > 0 > -2\theta_n g'_n(t)$ for $t < x_n$, that

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

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

Thus, since $w \leq w(x_n) + 2\theta_n|g_n|$ by (4.4), we have on $(\tilde{t}_1, \tilde{t}_2)$ that

$$|u - w| = u - w \geq \tilde{z} - w = |\tilde{z} - w|. \tag{38}$$

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 that u is not affine on $(\tilde{t}_1, \tilde{t}_2)$, so we have strict inequality in the Cauchy–Schwartz inequality, thus

$$\begin{aligned} \int_{\tilde{t}_1}^{\tilde{t}_2} (u')^2 &= \frac{1}{\tilde{t}_2 - \tilde{t}_1} \left(\int_{\tilde{t}_1}^{\tilde{t}_2} 1^2 \right) \left(\int_{\tilde{t}_1}^{\tilde{t}_2} (u')^2 \right) > \frac{1}{\tilde{t}_2 - \tilde{t}_1} \left(\int_{\tilde{t}_1}^{\tilde{t}_2} u' \right)^2 \\ &= \frac{(u(\tilde{t}_2) - u(\tilde{t}_1))^2}{\tilde{t}_2 - \tilde{t}_1} \\ &= (\tilde{t}_2 - \tilde{t}_1) \left(\frac{z(\tilde{t}_2) - z(\tilde{t}_1)}{\tilde{t}_2 - \tilde{t}_1} \right)^2 \\ &= (\tilde{t}_2 - \tilde{t}_1)(\tilde{z}')^2 \\ &= \int_{\tilde{t}_1}^{\tilde{t}_2} (\tilde{z}')^2. \end{aligned} \tag{39}$$

Hence defining $\tilde{u}: [-T, T] \rightarrow \mathbb{R}$ by

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

we obtain a function $\tilde{u} \in W^{1,2}(-T, T)$ with $\tilde{u}(\pm T) = w(\pm T)$ and such that, using (35), (38), and (39),

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

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 $w(x_n) + 2\theta_n|g_n|$ at $(x_n - \alpha_n)$ and $(x_n + \beta_n)$:

$$u(t) \geq w(x_n) + 2\theta_n g(\beta_n) + 2\theta_n g'(\beta_n)(t - (x_n + \beta_n)),$$

and

$$u(t) \geq w(x_n) + 2\theta_n |g(-\alpha_n)| - 2\theta_n g'(-\alpha_n)(t - (x_n - \alpha_n)),$$

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

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

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 we have that

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

on some left neighbourhood of $x_n + \beta_n$. So for t in this neighbourhood, we have that

$$\begin{aligned} w(x_n) + 2\theta_n g_n(t) &= w(x_n) + 2\theta_n g_n(x_n + \beta_n) - \int_t^{x_n + \beta_n} 2\theta_n g'_n(s) ds \\ &> w(x_n) + 2\theta_n g(\beta_n) - \int_t^{x_n + \beta_n} \frac{w(x_n) + 2\theta_n g(\beta_n) - u(t_0)}{x_n + \beta_n - t_0} ds \\ &= w(x_n) + 2\theta_n g(\beta_n) - \frac{w(x_n) + 2\theta_n g(\beta_n) - u(t_0)}{x_n + \beta_n - t_0} (x_n + \beta_n - t) \\ &= u(x_n + \beta_n) - \frac{w(x_n) + 2\theta_n g(\beta_n) - u(t_0)}{x_n + \beta_n - t_0} (x_n + \beta_n - t) \\ &\geq u(t), \end{aligned}$$

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

We can now prove certain bounds on u' . Suppose that there exists $t_0 \in (x_n - \alpha_n, x_n + \beta_n)$ such that $u'(t_0) > 2\theta_n g'(\beta_n)$. Then we have that $u'(t) > 2\theta_n g'(\beta_n)$ for all $t \in (t_0, x_n + \beta_n)$ by convexity. Then we see, using the inequality proved in the previous paragraph, that

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

which is a contradiction since $u(x_n + \beta_n) = w(x_n) + 2\theta_n g(\beta_n)$ by the choice of β_n . So $u'(t) \leq 2\theta_n g'(\beta_n)$ for almost every $t \in (x_n - \alpha_n, x_n + \beta_n)$. Similarly we can prove that $u'(t) \geq -2\theta_n g'(-\alpha_n)$ for almost every $t \in (x_n - \alpha_n, x_n + \beta_n)$. In the case in which $u(x_n) < w(x_n)$ we would prove that $-2\theta_n g'(\beta_n) \leq u'(t) \leq 2\theta_n g'(-\alpha_n)$ for almost every $t \in (x_n - \alpha_n, x_n + \beta_n)$.

We now prove the important consequence (37) of these estimates. Suppose that $b_n \geq a_n$. Then using convexity of u , and, by monotonicity of g , the fact that $g(b_n) \geq g(-a_n) = -g(a_n)$, we see that for $t \in J_n$,

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

Fix $t \in [x_n, x_n + b_n]$. We then have by the estimates we have just proved that

$$\begin{aligned}
 u(t) &= u(x_n + b_n) - \int_t^{x_n + b_n} u'(s) ds \\
 &\geq w(x_n) + 3\theta_n g(b_n) - \int_t^{x_n + b_n} 2\theta_n g'(\beta_n) ds \\
 &= w(x_n) + 3\theta_n g(b_n) - 2((x_n + b_n) - t)\theta_n g'(\beta_n).
 \end{aligned}$$

Also, since $t \leq x_n + b_n$, we have, using (3.3), (4.1), and concavity of g ,

$$\begin{aligned}
 w_n(t) &\leq w(x_n) + 2\theta_n g_n(t) \\
 &\leq w(x_n) + 2\theta_n g'_n(x_n + b_n)(t - (x_n + b_n)) + 2\theta_n g_n(x_n + b_n) \\
 &\leq w(x_n) + 2\theta_n g'(\beta_n)(t - (x_n + b_n)) + 2\theta_n g(b_n).
 \end{aligned}$$

So we have that

$$\begin{aligned}
 u(t) - w_n(t) &\geq (w(x_n) + 3\theta_n g(b_n) - 2((x_n + b_n) - t)\theta_n g'(\beta_n)) \\
 &\quad - (w(x_n) + 2\theta_n g'(\beta_n)(t - (x_n + b_n)) + 2\theta_n g(b_n)) \\
 &= \theta_n g(b_n).
 \end{aligned}$$

Similarly we can prove that $u(t) - w_n(t) \geq \theta_n g(a_n)$ for $t \in [x_n - a_n, x_n]$ if $a_n \geq b_n$. In the case that $u(x_n) < w(x_n)$ we can prove in the same way that $u(t) - w_n(t) \leq -\theta_n g(b_n)$ on $[x_n, x_n + b_n]$ if $b_n \geq a_n$, or $u(t) - w_n(t) \leq -\theta_n g(a_n)$ on $[x_n - a_n, x_n]$ if $a_n \geq b_n$, 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 that there is a $t_0 \in (x_n + b_n, T)$ such that $u(t_0) > w(x_n) + 3\theta_n g_n(t_0)$. The argument on the left of x_n is the same. Defining affine $z: [-T, T] \rightarrow \mathbb{R}$ by

$$z(t) = w(x_n) + 3\theta_n g_n(t_0) + 3\theta_n g'_n(t_0)(t - t_0),$$

we see that $z(t_0) = w(x_n) + 3\theta_n g_n(t_0) < u(t_0)$, and, using the concavity of g_n , that $z \geq w(x_n) + 3\theta_n g_n$ on (x_n, T) . The connected component of $[-T, T]$ containing t_0 of those points for which $z < u$ is a subinterval of $(x_n + b_n, T)$, since

$$u(x_n + b_n) = w(x_n) + 3\theta_n g_n(b_n) \leq z(x_n + b_n),$$

and by (4.4),

$$u(T) = w(T) \leq w(x_n) + 2\theta_n g_n(T) < z(T).$$

So we have that $u(t) > z(t) \geq w(x_n) + 3\theta_n g_n(t)$ on some open subinterval of $(x_n + b_n, T)$. Hence we can perform the same trick as before, constructing a new function $\tilde{u} \in W^{1,2}(-T, T)$ by replacing u with z on this subinterval, such that $\mathcal{L}(\tilde{u}) < \mathcal{L}(u)$, which again contradicts the choice of u as a minimizer. □

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 region bounded by the graphs of $t \mapsto w(x_n) \pm 3\theta_n|g_n(t)|$ off this neighbourhood. We note that the final statement of the lemma holds by the same argument even when $u(x_n) = w(x_n)$, and thus when the set J_n introduced is empty.

For the remainder of the proof of minimality, 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 11 and Corollary 12 on the connected components of $[-T, T] \setminus \{x_n : u(x_n) = w(x_n)\}$. 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 $a_n, b_n > 0$ be such that $J_n := (x_n - a_n, x_n + b_n)$ is the connected component in $[-T, T]$ containing x_n of those points t such that $|u(t) - w(x_n)| > 3\theta_n|g_n(t)|$, as in Lemma 7. It will be easier to work on a symmetric interval around x_n , so let $c_n := \max\{a_n, b_n\}$, and $\tilde{J}_n := [x_n - c_n, x_n + c_n]$. We note the following immediate corollary of Lemma 7. Fix $n \geq 0$. For $t \notin J_n$, we have for any $i \geq n$, by (4.1) and (3.3), that

$$\begin{aligned} |(u - w_i)(t)| &\leq |u(t) - w(x_n)| + |w(x_n) - w_i(t)| = |u(t) - w(x_n)| + |w_i(x_n) - w_i(t)| \\ &\leq 3\theta_n|g_n(t)| + 2\theta_n|g_n(t)| \\ &= 5\theta_n|g_n(t)|. \end{aligned} \tag{40}$$

The inequalities (37) from Lemma 7 tell us that the graph of a putative minimizer u cannot get too close to that of w around x_n . In the next result, this lower bound on the distance between the two functions is shown to concentrate a certain amount of weight in the 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 8 *Let $n \geq 0$, and suppose $\tilde{J}_n \subseteq Y_n$.*

Then

$$\int_{\tilde{J}_n} \tilde{\phi}_n^1(t, u - w_n) \geq \frac{453\theta_n g(c_n)}{(\log 1/c_n)^{1/3}}.$$

Proof Choose $t_{c_n} \in (0, c_n)$ such that $g(t_{c_n}) = g(c_n)/5$. Noting that (1) in particular implies that $t^{1/2} \log \log 1/|t| \leq 1 \leq \log \log 1/|t|$, we see that if $0 < t^{1/2} \leq c_n/5$, we have that

$$\begin{aligned} g(t) &= t \log \log 1/|t| = t^{1/2} (t^{1/2} \log \log 1/|t|) \leq t^{1/2} \leq c_n/5 \leq (c_n \log \log 1/c_n)/5 \\ &= g(c_n)/5, \end{aligned}$$

hence we have the lower bound $t_{c_n}^{1/2} \geq c_n/5$, and thus the inequality

$$\log 1/c_n \geq \log 1/(5t_{c_n}^{1/2}) = (\log 1/25t_{c_n})/2.$$

Since (1) also in particular implies that $t_{c_n}^{1/2} \leq (g(c_n)/5)^{1/2} \leq (1/5 \cdot 125)^{1/2} = 1/25$, we have that $1/(25t_{c_n}) \geq (1/t_{c_n})^{1/2}$ and hence that

$$\log 1/c_n \geq (\log 1/25t_{c_n})/2 \geq (\log(1/t_{c_n})^{1/2})/2 = (\log 1/t_{c_n})/4,$$

the ultimate point being that

$$\frac{1}{(\log 1/c_n)^{1/3}} \leq \frac{4^{1/3}}{(\log 1/t_{c_n})^{1/3}} \leq \frac{2}{(\log 1/t_{c_n})^{1/3}}.$$

Suppose that $b_n \geq a_n$, so by definition $c_n = b_n$. The case in which $a_n > b_n$ differs only in trivial notation. For $t \in [x_n, x_n + t_{c_n}]$ we have by (37), the choice of t_{c_n} , and the monotonicity of g , that $|(u - w_n)(t)| \geq \theta_n g(c_n) = 5\theta_n g(t_{c_n}) \geq 5\theta_n g_n(t)$, hence by the definition of $\tilde{\phi}_n^1$ (noting our one assumption in the statement that $\tilde{J}_n \subseteq Y_n$), $\tilde{\phi}_n^1(t, u - w_n) = 5\theta_n g_n(t)\psi_n^1(t)$. On the interval $[x_n, x_n + t_{c_n}]$ this function is concave, so the integral admits an easy lower estimate by calculating the area of the triangle under the graph, using the definitions of t_{c_n} and ψ^1 :

$$\begin{aligned} \int_{\tilde{J}_n} \tilde{\phi}_n^1(t, u - w_n) &\geq \int_{x_n}^{x_n+t_{c_n}} \tilde{\phi}_n^1(t, u - w_n) = 5\theta_n \int_{x_n}^{x_n+t_{c_n}} g_n(t)\psi_n^1(t) \\ &\geq \frac{5}{2} \theta_n g(t_{c_n}) \psi^1(t_{c_n}) t_{c_n} \\ &= \frac{\theta_n g(c_n)}{2} \frac{1812}{t_{c_n} (\log 1/t_{c_n})^{1/3}} t_{c_n} \\ &\geq \frac{\theta_n g(c_n)}{4} \frac{1812}{(\log 1/c_n)^{1/3}} \\ &= \frac{453\theta_n g(c_n)}{(\log 1/c_n)^{1/3}}. \end{aligned}$$

□

We shall want to give special attention to that part of \tilde{J}_n on which $w_n = \tilde{w}_n$, so for $n \geq 0$ define $H_n \subseteq [-T, T]$ by setting $H_n := \tilde{J}_n \cap [x_n - \tau_n, x_n + \tau_n] = [x_n - d_n, x_n + d_n]$, say, so $d_n \leq c_n$. Note that by construction and (3.1),

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

We cannot immediately mimic the main principle of the proof and integrate by parts across x_n , since 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 w_n near x_n slow enough so that we can now 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 $\tilde{l}_n: [-T, T] \rightarrow \mathbb{R}$ be the affine function defined by

$$\tilde{l}_n(t) = \tilde{l}'_n \cdot (t - (x_n - d_n)) + \tilde{w}(-d_n),$$

where

$$\tilde{l}'_n := \frac{\tilde{w}(d_n) - \tilde{w}(-d_n)}{2d_n} = (\log \log 1/d_n)(\sin \log \log \log 1/d_n), \tag{41}$$

and define $l_n: [-T, T] \rightarrow \mathbb{R}$ by

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

Clearly $l_n \in W^{1,2}(-T, T)$.

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 $l'_n(u - w_n)'$, and secondly outside H_n , integrating $w'_n(u - w_n)'$:

$$\begin{aligned} I_{n,\pm} &= l'_n(u(x_n \pm d_n) - w_n(x_n \pm d_n)), \\ E_{n,\pm} &= w'_n(x_n \pm d_n)(u(x_n \pm d_n) - w_n(x_n \pm d_n)). \end{aligned}$$

Note that

$$|I_{n,\pm} - E_{n,\pm}| = |(l'_n - w'_n(x_n \pm d_n))(u(x_n \pm d_n) - w_n(x_n \pm d_n))|. \tag{42}$$

The next lemma describes the consequence for the derivative terms in the integrand of exchanging w_n with l_n on H_n . Integrating by parts gives us the boundary terms involving l'_n , and the second derivative term vanishes, since l_n is affine. The L^2 -norm of the difference between w'_n and l'_n gives us an error which we see, comparing with Lemma 8, will be absorbed into the weight term of the integrand.

Lemma 9 *Let $n \geq 0$.*

Then

$$\int_{H_n} ((u')^2 - (w'_n)^2) \geq 2(I_{n,+} - I_{n,-}) - \frac{432g(d_n)}{(\log 1/d_n)^{1/3}}.$$

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

$$\begin{aligned} \int_{H_n} ((u')^2 - (w'_n)^2) &= \int_{H_n} ((u')^2 - (l'_n)^2) + \int_{H_n} ((l'_n)^2 - (w'_n)^2) \\ &\geq \int_{H_n} ((u')^2 - (l'_n)^2) - \int_{H_n} |(l'_n)^2 - (w'_n)^2|. \end{aligned}$$

Since $w'_n = \tilde{w}'_n$ and $l'_n = \tilde{l}'_n$ on H_n , 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 for $t > 0$ that

$$\frac{d}{dt} ((\log \log 1/t)(\sin \log \log \log 1/t)) = -\frac{\sin \log \log \log 1/t + \cos \log \log \log 1/t}{t \log 1/t},$$

so

$$\left| \frac{d}{dt} ((\log \log 1/t)(\sin \log \log \log 1/t)) \right| \leq \frac{2}{t \log 1/t}.$$

Hence, recalling the expressions for the derivatives given in (41) and (4), and by applying the mean value theorem, we can see that for $0 < t \leq d$,

$$\begin{aligned} &|\tilde{l}' - \tilde{w}'(t)| \\ &= \left| (\log \log 1/d)(\sin \log \log \log 1/d) \right. \\ &\quad \left. - \left((\log \log 1/t)(\sin \log \log \log 1/t) - \frac{\sin \log \log \log 1/t + \cos \log \log \log 1/t}{\log 1/t} \right) \right| \end{aligned}$$

$$\begin{aligned}
 &\leq |(\log \log 1/d)(\sin \log \log \log 1/d) - (\log \log 1/t)(\sin \log \log \log 1/t)| + \frac{2}{\log 1/t} \\
 &\leq \frac{2(d-t)}{t \log 1/t} + \frac{2}{\log 1/t} \\
 &= \frac{2d}{t \log 1/t}.
 \end{aligned} \tag{43}$$

Now, let

$$\gamma(d) := \frac{d}{(\log 1/d)^{2/3}} \leq d, \tag{44}$$

so $\log 1/\gamma(d) = \log \left(\frac{(\log 1/d)^{2/3}}{d} \right) = \frac{2}{3} \log \log 1/d + \log 1/d \leq 2 \log 1/d$, and so we have that

$$\log \log 1/\gamma(d) \leq \log(2 \log 1/d) \leq \log(\log 1/d)^2 = 2 \log \log 1/d. \tag{45}$$

For $t \in [\gamma(d), d]$, we have by (43) and the definition of $\gamma(d)$ that

$$|\tilde{l}' - \tilde{w}'(t)| \leq \frac{2d}{t \log 1/t} \leq \frac{2d}{\gamma(d) \log 1/d} = \frac{2(\log 1/d)^{2/3}}{\log 1/d} = \frac{2}{(\log 1/d)^{1/3}}. \tag{46}$$

This is one sense in which \tilde{w} oscillates slowly enough: a good estimate for the discrepancy between the derivatives holds on an interval in the domain of integration large enough in measure. Noting that

$$\begin{aligned}
 &|\tilde{l}' \pm \tilde{w}'(t)| \\
 &= \left| (\log \log 1/d)(\sin \log \log \log 1/d) \right. \\
 &\quad \left. \pm \left((\log \log 1/t)(\sin \log \log \log 1/t) - \frac{\sin \log \log \log 1/t + \cos \log \log \log 1/t}{\log 1/t} \right) \right| \\
 &\leq |(\log \log 1/d)(\sin \log \log \log 1/d)| + |(\log \log 1/t)(\sin \log \log \log 1/t)| \\
 &\quad + \left| \frac{\sin \log \log \log 1/t + \cos \log \log \log 1/t}{\log 1/t} \right| \\
 &\leq \log \log 1/d + \log \log 1/t + \frac{2}{\log 1/t} \\
 &\leq 4 \log \log 1/t,
 \end{aligned}$$

we see, since the integrand is an even function, that

$$\begin{aligned}
 \int_H |(\tilde{l}')^2 - (\tilde{w}')^2| &= \int_{-d}^d |\tilde{l}' - \tilde{w}'| |\tilde{l}' + \tilde{w}'| \\
 &= 2 \int_0^d |\tilde{l}' - \tilde{w}'| |\tilde{l}' + \tilde{w}'| \\
 &= 2 \left(\int_0^{\gamma(d)} |\tilde{l}' - \tilde{w}'| |\tilde{l}' + \tilde{w}'| + \int_{\gamma(d)}^d |\tilde{l}' - \tilde{w}'| |\tilde{l}' + \tilde{w}'| \right) \\
 &\leq 2 \left(\int_0^{\gamma(d)} (4 \log \log 1/t)^2 + \int_{\gamma(d)}^d |\tilde{l}' - \tilde{w}'| |\tilde{l}' + \tilde{w}'| \right).
 \end{aligned} \tag{47}$$

We then use Cauchy–Schwartz and (46) to see that

$$\begin{aligned} \int_{\gamma(d)}^d |\tilde{l}' - \tilde{w}'| |\tilde{l}' + \tilde{w}'| &\leq \left(\int_{\gamma(d)}^d |\tilde{l}' - \tilde{w}'|^2 \right)^{1/2} \left(\int_{\gamma(d)}^d |\tilde{l}' + \tilde{w}'|^2 \right)^{1/2} \\ &\leq \left(\int_{\gamma(d)}^d \left(\frac{2}{(\log 1/d)^{1/3}} \right)^2 \right)^{1/2} \left(\int_{\gamma(d)}^d (4 \log \log 1/t)^2 \right)^{1/2} \\ &\leq \frac{8d^{1/2}}{(\log 1/d)^{1/3}} \left(\int_0^d (\log \log 1/t)^2 \right)^{1/2}. \end{aligned}$$

We now use repeated applications of integration by parts to derive the inequality

$$\int_0^d (\log \log 1/t)^2 dt \leq 3d(\log \log 1/d)^2.$$

First note, using the substitution $y = \log 1/t$, and integrating by parts, that

$$\begin{aligned} \int_0^d (\log \log 1/t)^2 dt &= \int_{\log 1/d}^\infty (\log y)^2 e^{-y} dy \\ &= \left([-e^{-y}(\log y)^2]_{\log 1/d}^\infty + \int_{\log 1/d}^\infty \frac{2(\log y)e^{-y}}{y} dy \right) \\ &= d(\log \log 1/d)^2 + \int_{\log 1/d}^\infty \frac{2(\log y)e^{-y}}{y} dy. \end{aligned}$$

Examining the second summand, we use Cauchy–Schwartz, and integration by parts twice more to see, using the simplifications that $\log 1/d \geq \log \log 1/d \geq 2 \geq 1$, that

$$\begin{aligned} &\int_{\log 1/d}^\infty \frac{(\log y)e^{-y}}{y} dy \\ &\leq \left(\int_{\log 1/d}^\infty e^{-2y} dy \right)^{1/2} \left(\int_{\log 1/d}^\infty \frac{(\log y)^2}{y^2} dy \right)^{1/2} \\ &= \left(\left[\frac{-e^{-2y}}{2} \right]_{\log 1/d}^\infty \right)^{1/2} \left(\left[\frac{-(\log y)^2}{y} \right]_{\log 1/d}^\infty - \int_{\log 1/d}^\infty \frac{-2 \log y}{y^2} dy \right)^{1/2} \\ &\leq 2^{-1/2} d \left(\frac{(\log \log 1/d)^2}{\log 1/d} - \left(\left[\frac{2 \log y}{y} \right]_{\log 1/d}^\infty - \int_{\log 1/d}^\infty \frac{2}{y^2} dy \right) \right)^{1/2} \\ &= 2^{-1/2} d \left(\frac{(\log \log 1/d)^2}{\log 1/d} - \left(\frac{-2 \log \log 1/d}{\log 1/d} - \left[\frac{-2}{y} \right]_{\log 1/d}^\infty \right) \right)^{1/2} \\ &= 2^{-1/2} d \left(\frac{(\log \log 1/d)^2}{\log 1/d} + \frac{2 \log \log 1/d}{\log 1/d} + \frac{2}{\log 1/d} \right)^{1/2} \\ &= \frac{2^{-1/2} d}{(\log 1/d)^{1/2}} ((\log \log 1/d + 1)^2 + 1)^{1/2} \end{aligned}$$

$$\begin{aligned} &\leq \frac{2^{-1/2}d}{(\log 1/d)^{1/2}} 2^{1/2}(\log \log 1/d + 1) \\ &\leq \frac{2^{-1/2}d}{(\log 1/d)^{1/2}} 2^{3/2} \log \log 1/d \\ &\leq 2 \log \log 1/d. \end{aligned}$$

Combining with the original expression, we have, using again that $\log \log 1/d \geq 2$, that

$$\begin{aligned} \int_0^d (\log \log 1/t)^2 dt &\leq d(\log \log 1/d) ((\log \log 1/d) + 4) \\ &\leq 3d(\log \log 1/d)^2, \end{aligned}$$

as claimed.

So we can conclude our estimates. Since (1) implies that $\log \log 1/d \leq (\log 1/d)^{1/3}$, applying this inequality to (47), using (44), and (45), we see that

$$\begin{aligned} \int_H |(\tilde{l}')^2 - (\tilde{w}')^2| &\leq 2 \left(48\gamma(d)(\log \log 1/\gamma(d))^2 + \frac{24d \log \log 1/d}{(\log 1/d)^{1/3}} \right) \\ &\leq \frac{384d(\log \log 1/d)^2}{(\log 1/d)^{2/3}} + \frac{48d \log \log 1/d}{(\log 1/d)^{1/3}} \\ &= \frac{g(d)}{(\log 1/d)^{1/3}} \left(\frac{384 \log \log 1/d}{(\log 1/d)^{1/3}} + 48 \right) \\ &\leq \frac{432g(d)}{(\log 1/d)^{1/3}}. \end{aligned} \tag{48}$$

By (2) we have, since $\tilde{l}(\pm d) = \tilde{w}(\pm d)$, that

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

Since

$$\begin{aligned} \int_H \left((u')^2 - (\tilde{w}')^2 \right) &= \int_H \left((u')^2 - (\tilde{l}')^2 \right) + \int_H \left((\tilde{l}')^2 - (\tilde{w}')^2 \right) \\ &\geq \int_H \left((u')^2 - (\tilde{l}')^2 \right) - \int_H |(\tilde{l}')^2 - (\tilde{w}')^2|, \end{aligned}$$

the result follows from (48). □

An estimate established in the preceding proof 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 .

Lemma 10 *Let $n \geq 0$.*

Then

$$|I_{n,+} - E_{n,+}| + |I_{n,-} - E_{n,-}| \leq \frac{20\theta_n g(c_n)}{\log 1/c_n}.$$

Proof We just have to estimate $|(u - w_n)(x_n \pm d_n)|$. Suppose that $u(x_n) > w(x_n)$; the argument for the case in which $u(x_n) < w(x_n)$ is similar. Suppose also that $b_n \geq a_n$, so by definition $c_n = b_n$. The case in which $a_n > b_n$ is similar. Then $u(t) \leq u(x_n + b_n)$ by the convexity of u established in Lemma 7, for all $t \in J_n$.

If $x_n - d_n \notin J_n$, then (40) implies that $|(u - w_n)(x_n - d_n)| \leq 5\theta_n g(d_n) \leq 5\theta_n g(b_n)$, by monotonicity of g , since $d_n \leq b_n$.

Certainly $x_n + d_n \in J_n$, since $d_n \leq b_n$ by definition of d_n , so if also $x_n - d_n \in J_n$, by definition of J_n we see that

$$w(x_n) \leq w(x_n) + 3\theta_n |g(d_n)| \leq u(x_n \pm d_n) \leq u(x_n + b_n) = w(x_n) + 3\theta_n g(b_n),$$

so $0 < u(x_n \pm d_n) - w(x_n) \leq 3\theta_n g(b_n)$. Hence, using (4.1) and (3.3), we have that, since $d_n \leq b_n$,

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

Hence in both cases $|(u - w_n)(x_n \pm d_n)| \leq 5\theta_n g(b_n)$. The result then follows by using (43) with $t = d$ in (42), and since $d_n \leq b_n$. □

The following is the key lemma, providing a positive lower bound for $\mathcal{L}_n(u) - \mathcal{L}_n(w_n)$. We combine our estimates for \mathcal{L}_n across the whole domain $[-T, T]$, integrating by parts off $\bigcup_{i=1}^n H_i$, and using the estimates from Lemmas 9 and 10 on each H_i . The argument is made more straightforward by assuming that the intervals \tilde{J}_i are small in a certain sense, which implies that the intervals on which we work do not overlap. Should this assumption fail for some n , then, as later lemmas will show, this means that the discrepancy $u - w$ around x_n is sufficiently large that we may ignore the fine detail of our construction at and beyond the stage n , and we can conclude the proof using just \mathcal{L}_{n-1} .

Lemma 11 *Suppose $n \geq 0$ is such that for all $0 \leq j \leq n$,*

$$\tilde{J}_k \cap Y_j = \emptyset \text{ for all } 0 \leq k \leq j - 1; \text{ and} \tag{49}$$

$$\tilde{J}_j \subseteq Y_j. \tag{50}$$

Then

$$\mathcal{L}_n(u) - \mathcal{L}_n(w_n) \geq \sum_{i=0}^n \left(\frac{\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} \right) + \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} |u - w_n|.$$

Proof By (3.6) and assumption (49) we have for all $0 \leq k \leq j - 1 \leq j \leq n$ that $w_j = w_k$ on \tilde{J}_k , in particular that

$$w_n = w_k, \quad w'_n = w'_k, \quad \text{and } w''_n = w''_k \text{ on } \tilde{J}_k, \text{ whenever both sides exist.} \tag{51}$$

Also, assumptions (49) and (50) together imply that $\{\tilde{J}_i\}_{i=0}^n$ is pairwise disjoint.

Now, let $0 \leq i \leq n$. We split up the integral into summands which we shall tackle separately:

$$\begin{aligned} & \int_{\tilde{J}_i} ((u')^2 + \phi(t, u - w_i) - (w'_i)^2) \\ &= \int_{\tilde{J}_i} (\phi^1(t, u - w_i) + \phi^2(t, u - w_i)) + \int_{H_i} ((u')^2 - (w'_i)^2) + \int_{\tilde{J}_i \setminus H_i} ((u')^2 - (w'_i)^2) \\ &\geq \int_{\tilde{J}_i} \phi^1(t, u - w_i) + \int_{H_i} ((u')^2 - (w'_i)^2) + \int_{\tilde{J}_i \setminus H_i} (\phi^2(t, u - w_i) + (u')^2 - (w'_i)^2). \end{aligned}$$

Now, by Lemma 8 [note that this applies by assumption (50)] and Lemma 9, and since $c_i \geq d_i$ and $\theta_i \geq 1$,

$$\begin{aligned} \int_{\tilde{J}_i} \phi^1(t, u - w_i) + \int_{H_i} ((u')^2 - (w'_i)^2) &\geq \int_{\tilde{J}_i} \tilde{\phi}_i^1(t, u - w_i) + \int_{H_i} ((u')^2 - (w'_i)^2) \\ &\geq \frac{453\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} + 2(I_{i,+} - I_{i,-}) - \frac{432g(d_i)}{(\log 1/d_i)^{1/3}} \\ &\geq \frac{21\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} + 2(I_{i,+} - I_{i,-}). \end{aligned}$$

So, combining, we have that

$$\begin{aligned} \int_{\tilde{J}_i} ((u')^2 + \phi(t, u - w_i) - (w'_i)^2) &\geq \frac{21\theta g(c_i)}{(\log 1/c_i)^{1/3}} + 2(I_{i,+} - I_{i,-}) \\ &\quad + \int_{\tilde{J}_i \setminus H_i} (\phi^2(t, u - w_i) + (u')^2 - (w'_i)^2). \end{aligned} \tag{52}$$

Now, for any $t \in [-T, T]$, let $\mathcal{S}_n(t) := \{j = 0, \dots, n : t \in Y_j\}$. We now show by an easy induction on n that

$$\sum_{j \in \mathcal{S}_n(t)} \psi_j^2(t) \geq 2|w''_n(t)| + 1 + 2^{-(n-1)}, \tag{53}$$

for almost every $t \in [-T, T]$. For $n = 0$, we have by definition of ψ^2 that for all $t \neq x_0$, $\psi_0^2(t) = 3 + 2|w''_0(t)|$, as required. Suppose the result holds for all $0 \leq i \leq n - 1$, where $n \geq 1$. Let $i_n(t) \leq n$ denote the greatest index in $\mathcal{S}_n(t)$, i.e. the greatest index $j \leq n$ such that $t \in Y_j$. By (3.6) we have that $w''_n(t) = w''_{i_n(t)}(t)$ almost everywhere. If $t \in (x_{i_n(t)} - \tau_{i_n(t)}, x_{i_n(t)} + \tau_{i_n(t)})$, then $w''_{i_n(t)}(t) = \tilde{w}''_{i_n(t)}(t)$ by (3.1), and by definition of ψ^2 , for $t \neq x_{i_n(t)}$,

$$\sum_{j \in \mathcal{S}_n(t)} \psi_j^2(t) \geq \psi_{i_n(t)}^2(t) = 3 + 2|\tilde{w}''_{i_n(t)}(t)| \geq 1 + 2^{-(n-1)} + 2|w''_{i_n(t)}(t)|,$$

as required. If $t \notin [x_{i_n(t)} - \tau_{i_n(t)}, x_{i_n(t)} + \tau_{i_n(t)}]$ (note then necessarily $i_n(t) \geq 1$ since $\tau_0 = T$), then $|w''_{i_n(t)}(t)| \leq |w''_{i_n(t)-1}(t)| + 2^{-i_n(t)}$ almost everywhere by (3.11). So by inductive hypothesis

$$\begin{aligned} \sum_{j \in \mathcal{S}_n(t)} \psi_j^2(t) &\geq \sum_{j \in \mathcal{S}_{i_n(t)-1}(t)} \psi_j^2(t) \\ &\geq 2|w''_{i_n(t)-1}(t)| + 1 + 2^{-((i_n(t)-1)-1)} \\ &\geq 2|w''_{i_n(t)}(t)| - 2 \cdot 2^{-i_n(t)} + 1 + 2^{-((i_n(t)-1)-1)} \end{aligned}$$

$$\begin{aligned}
 &= 2|w''_{i_n}(t)| + 1 + 2^{-(i_n(t)-1)} \\
 &\geq 2|w''_n(t)| + 1 + 2^{-(n-1)},
 \end{aligned}$$

as required for (53).

Given this, now consider $t \notin \bigcup_{i=0}^n \tilde{J}_i$. Then since by definition $\tilde{J}_j \supseteq J_j$ for all $j \geq 0$, (40) implies that $|(u - w_n)(t)| \leq 5\theta_j |g_j(t)|$ for all $0 \leq j \leq n$. Therefore $\tilde{\phi}_j^2(t, u - w_n) = \psi_j^2(t)|u - w_n|$ by definition of $\tilde{\phi}^2$, for $j \in \mathcal{J}_n(t)$. Thus almost everywhere, we have by (53) that

$$\begin{aligned}
 \phi^2(t, u - w_n) - 2(u - w_n)w''_n &\geq \sum_{j \in \mathcal{J}_n(t)} \left(\tilde{\phi}_j^2(t, u - w_n) \right) - 2|u - w_n||w''_n| \\
 &= \sum_{j \in \mathcal{J}_n(t)} \left(\psi_j^2(t)|u - w_n| \right) - 2|u - w_n||w''_n| \\
 &= |u - w_n| \left(\sum_{j \in \mathcal{J}_n(t)} (\psi_j^2(t)) - 2|w''_n(t)| \right) \\
 &\geq |u - w_n|.
 \end{aligned}$$

Now, let $t \in \tilde{J}_i \setminus H_i$ for some $0 \leq i \leq n$. Then note that we must have $i \geq 1$, since $\tau_0 = T$, so $H_0 = \tilde{J}_0$. Since $\{\tilde{J}_j\}_{j=0}^n$ is pairwise disjoint, we have that $t \notin \tilde{J}_j$ for $j \leq i - 1$. Hence, again by (40), $|(u - w_i)(t)| \leq 5\theta_j |g_j(t)|$, so by definition of $\tilde{\phi}^2$, $\tilde{\phi}_j^2(t, u - w_i) = \psi_j^2(t)|u - w_i|$ for all $j \leq i - 1$, recalling assumption (50). Since $t \notin H_i$, we have $t \notin [x_i - \tau_i, x_i + \tau_i]$, and hence that $|w''_i(t)| \leq |w''_{i-1}(t)| + 2^{-i}$ almost everywhere by (3.11). Hence by (53) we have almost everywhere that

$$\begin{aligned}
 \sum_{j \in \mathcal{J}_{i-1}(t)} \psi_j^2(t) &\geq 1 + 2|w''_{i-1}(t)| + 2^{-(i-2)} \geq 1 + 2|w''_i(t)| - 2^{-(i-1)} + 2^{-(i-2)} \\
 &\geq 1 + 2|w''_i(t)|,
 \end{aligned}$$

and so

$$\begin{aligned}
 \phi^2(t, u - w_i) - 2(u - w_i)w''_i &\geq \sum_{j \in \mathcal{J}_{i-1}(t)} \left(\tilde{\phi}_j^2(t, u - w_i) \right) - 2|u - w_i||w''_i| \\
 &= \sum_{j \in \mathcal{J}_{i-1}(t)} \left(\psi_j^2(t)|u - w_i| \right) - 2|u - w_i||w''_i| \\
 &\geq |u - w_i|.
 \end{aligned}$$

Thus we have for almost every $t \notin \bigcup_{i=0}^n H_i$, noting the argument on $\tilde{J}_i \setminus H_i$ above applies by (51), that

$$\phi^2(t, u - w_n) - 2(u - w_n)w''_n \geq |u - w_n|,$$

and hence that

$$\int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} (\phi^2(t, u - w_n) - 2(u - w_n)w''_n) \geq \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} |u - w_n|. \tag{54}$$

The reason for making this estimate is that we want to integrate $(u' - w'_n)w'_n$ by parts on $[-T, T] \setminus \bigcup_{i=0}^n H_i$. 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 $(u - w_n)$ and w'_n are bounded and absolutely

continuous on $[-T, T] \setminus \bigcup_{i=0}^n H_i$ by (3.2), and thus $(u - w_n)w'_n$ is absolutely continuous on $[-T, T] \setminus \bigcup_{i=0}^n H_i$. However, in the general case in which $w(x_j) = u(x_j)$ for some $0 \leq j \leq n$, and thus that $w_n(x_j) = u(x_j)$, we have to argue a little more carefully.

We claim that even in this general case the parts formula is still valid on $[-T, T] \setminus \bigcup_{i=0}^n H_i$; this is the assertion that $(u - w_n)w'_n$ can be written as an indefinite integral on $[-T, T] \setminus \bigcup_{i=0}^n H_i$. The argument of the preceding paragraph gives us that $(u - w_n)w'_n$ 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 (13), and since $\{\sigma_n\}_{n=1}^\infty$ is decreasing, we know that $[x_j - \sigma_n, x_j + \sigma_n] \cap Y_m = \emptyset$ for all $j + 1 \leq m \leq n$. So by (3.6) and (3.1), $w_n = \tilde{w}_j + \rho_j$ on $[x_j - t_j, x_j + t_j]$. It suffices to check that $(u - w_n)w'_n$ can be written as an indefinite integral on $(x_j - t_j, x_j + t_j)$. We check that

$$\int_{x_j - t_j}^{x_j} ((u - w_n)w'_n)'(s) ds = -((u - w_n)(x_j - t_j))w'_n(x_j - t_j);$$

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

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

Since $J_j = \emptyset$, we see by (40) that $|(u - w_n)(t)| \leq 5\theta_j |g_j(t)|$ on $(x_j - t_j, x_j + t_j)$. Thus we see by (5) that

$$\begin{aligned} |(u - w_n)(t)w'_n(t)| &= |(u - w_n)(t)\tilde{w}'_j(t)| \leq 5\theta_j |g_j(t)| 3 \log \log 1/|t - x_j| \\ &\leq 15\theta_j |t - x_j| (\log \log 1/|t - x_j|)^2 \\ &\rightarrow 0 \text{ as } t \rightarrow x_j. \end{aligned}$$

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

$$\begin{aligned} -((u - w_n)(x_j - t_j))w'_n(x_j - t_j) &= \lim_{t \uparrow x_j} (((u - w_n)(t))w'_n(t) \\ &\quad - ((u - w_n)(x_j - t_j))w'_n(x_j - t_j)) \\ &= \lim_{t \uparrow x_j} \int_{x_j - t_j}^t ((u - w_n)w'_n)'(s) ds \\ &= \int_{x_j - t_j}^{x_j} ((u - w_n)w'_n)'(s) ds, \end{aligned}$$

as required. It just remains to justify our use of the dominated convergence theorem, i.e. to show that $((u - w_n)w'_n)' \in L^1(x_j - t_j, x_j + t_j)$. Again, noting that (40) still holds, we have, using (3.1) and Cauchy–Schwartz, that

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

$$\begin{aligned} &\leq \int_{x_j-t_j}^{x_j+t_j} |5\theta_j g_j \tilde{w}''_j| + \int_{x_j-t_j}^{x_j+t_j} |u' \tilde{w}'_j| + \int_{x_j-t_j}^{x_j+t_j} |\tilde{w}'_j|^2 \\ &\leq 5\theta_j \int_{-t_j}^{t_j} |g \tilde{w}''| + \left(\int_{x_j-t_j}^{x_j+t_j} |u'|^2 \right)^{1/2} \left(\int_{-t_j}^{t_j} |\tilde{w}'|^2 \right)^{1/2} + \int_{-t_j}^{t_j} |\tilde{w}'|^2. \end{aligned}$$

This right hand side is finite by (6), and since $u, \tilde{w} \in W^{1,2}(-T, T)$.

So, using (2), and recalling that $u(\pm T) = w(\pm T)$, and using (54) (recalling that $H_i \subseteq \tilde{J}_i$), we have, integrating by parts as we now know we can do, that

$$\begin{aligned} &\int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} (\phi^2(t, (u - w_n)) + (u')^2 - (w'_n)^2) \\ &\geq \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} (\phi^2(t, u - w_n) + 2(u' - w'_n)w'_n) \\ &= 2[(u - w_n)w'_n]_{[-T, T] \setminus \bigcup_{i=0}^n H_i} + \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} (\phi^2(t, u - w_n) - 2(u - w_n)w''_n) \\ &= -2 \sum_{i=0}^n [(u - w_i)w'_i]_{x_i-d_i}^{x_i+d_i} + \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} (\phi^2(t, u - w_n) - 2(u - w_n)w''_n) \\ &\geq -2 \sum_{i=0}^n (E_{i,+} - E_{i,-}) + \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} |u - w_n|. \tag{55} \end{aligned}$$

So, since $\{\tilde{J}_i\}_{i=0}^n$ is pairwise disjoint, we can argue as follows, using (51), (52), (55), and Lemma 10 to see that

$$\begin{aligned} &\mathcal{L}_n(u) - \mathcal{L}_n(w_n) \\ &= \int_{-T}^T ((u')^2 + \phi(t, u - w_n) - (w'_n)^2) \\ &= \int_{\bigcup_{i=0}^n \tilde{J}_i} ((u')^2 + \phi(t, u - w_n) - (w'_n)^2) \\ &\quad + \int_{[-T, T] \setminus \bigcup_{i=0}^n \tilde{J}_i} ((u')^2 + \phi(t, u - w_n) - (w'_n)^2) \\ &= \sum_{i=0}^n \int_{\tilde{J}_i} ((u')^2 + \phi(t, u - w_i) - (w'_i)^2) \\ &\quad + \int_{[-T, T] \setminus \bigcup_{i=0}^n \tilde{J}_i} ((u')^2 + \phi(t, u - w_n) - (w'_n)^2) \\ &\geq \sum_{i=0}^n \left(\frac{21\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} + 2(I_{i,+} - I_{i,-}) \right) \\ &\quad + \int_{\tilde{J}_i \setminus H_i} (\phi^2(t, u - w_i) + (u')^2 - (w'_i)^2) \\ &\quad + \int_{[-T, T] \setminus \bigcup_{i=0}^n \tilde{J}_i} (\phi^2(t, u - w_n) + (u')^2 - (w'_n)^2) \\ &\geq \sum_{i=0}^n \left(\frac{21\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} + 2(I_{i,+} - I_{i,-}) \right) \end{aligned}$$

$$\begin{aligned}
 & + \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} (\phi^2(t, u - w_n) + (u')^2 - (w'_n)^2) \\
 \geq & \sum_{i=0}^n \left(-2(E_{i,+} - E_{i,-}) + \frac{21\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} + 2(I_{i,+} - I_{i,-}) \right) \\
 & + \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} |u - w_n| \\
 = & \sum_{i=0}^n \left(2((I_{i,+} - E_{i,+}) - (I_{i,-} - E_{i,-})) + \frac{21\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} \right) \\
 & + \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} |u - w_n| \\
 \geq & \sum_{i=0}^n \left(\frac{21\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} - 2(|I_{i,+} - E_{i,+}| + |I_{i,-} - E_{i,-}|) \right) \\
 & + \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} |u - w_n| \\
 \geq & \sum_{i=0}^n \left(\frac{\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} \right) \\
 & + \int_{[-T, T] \setminus \bigcup_{i=0}^n H_i} |u - w_n|.
 \end{aligned}$$

□

Corollary 12 *Suppose for all $n \geq 0$ that the assumptions (49) and (50) hold.*

Then

$$\mathcal{L}(u) - \mathcal{L}(w) \geq \sum_{i=0}^{\infty} \left(\frac{\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} \right) + \int_{[-T, T] \setminus \bigcup_{i=0}^{\infty} H_i} |u - w| > 0.$$

Proof This follows by the preceding lemma and the dominated convergence theorem as follows. It is straightforward to see that

$$\lim_{n \rightarrow \infty} \left(|u - w_n| \mathbb{1}_{[-T, T] \setminus \bigcup_{i=0}^n H_i} \right) (t) = \left(|u - w| \mathbb{1}_{[-T, T] \setminus \bigcup_{i=0}^{\infty} H_i} \right) (t)$$

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

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

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

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

So the dominated convergence theorem implies that

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{[-T, T] \setminus \cup_{i=0}^n H_i} |u - w_n| &= \lim_{n \rightarrow \infty} \int_{-T}^T (|u - w_n| \mathbb{1}_{[-T, T] \setminus \cup_{i=0}^n H_i}) \\ &= \int_{-T}^T \lim_{n \rightarrow \infty} (|u - w_n| \mathbb{1}_{[-T, T] \setminus \cup_{i=0}^n H_i}) \\ &= \int_{-T}^T (|u - w| \mathbb{1}_{[-T, T] \setminus \cup_{i=0}^\infty H_i}) \\ &= \int_{[-T, T] \setminus \cup_{i=0}^\infty H_i} |u - w|. \end{aligned}$$

Lemma 6 and (12) give that

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

So since, by assumption, Lemma 11 applies 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 $[-T, T] \setminus \cup_{i=0}^\infty H_i = [-T, T]$, so on the assumption that $u \neq w$, where both are continuous functions, the integral term must be strictly positive. □

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 \mathcal{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 apparently 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 13 *Let $n \geq 1$ be such that assumptions (49) and (50) hold for $n - 1$, but for some $0 \leq k \leq n - 1$ we have that $\tilde{J}_k \cap Y_n \neq \emptyset$, i.e. (49) fails for n .*

Then

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

Proof That (49) fails for n implies that $c_k \geq T_n$, otherwise choosing $t \in \tilde{J}_k \cap Y_n$ we would have, since (T:1) implies that $T_n \leq |x_n - x_k|/2$, that

$$|x_n - x_k| \leq |x_n - t| + |t - x_k| \leq T_n + c_k < 2T_n \leq |x_n - x_k|,$$

which is a contradiction. So, applying Lemma 11 to $n - 1$ we see, using this fact, that $\theta_k \geq 1$, and since (1) implies that $c_k \leq c_k^{1/3} \leq (1/\log 1/c_k)^{1/3}$, that

$$\begin{aligned}
 \mathcal{L}_{n-1}(u) - \mathcal{L}_{n-1}(w_{n-1}) &\geq \sum_{i=0}^{n-1} \left(\frac{\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} \right) + \int_{[-T, T] \setminus \bigcup_{i=0}^{n-1} H_i} |u - w_{n-1}| \\
 &\geq \frac{\theta_k g(c_k)}{(\log 1/c_k)^{1/3}} \\
 &\geq \frac{c_k \log \log 1/c_k}{(\log 1/c_k)^{1/3}} \\
 &\geq c_k^2 \log \log 1/c_k \\
 &\geq c_k^2 \\
 &\geq T_n^2.
 \end{aligned}$$

□

Lemma 14 *Let $n \geq 1$ be such that assumption (49) holds for n , assumption (50) holds for $n - 1$, but $\tilde{J}_n \not\subseteq Y_n$, i.e. (50) fails for n .*

Then

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

Proof We suppose that $c_n = b_n$. The case in which $a_n > b_n$ differs only in trivial notation. That (50) fails for n implies that $b_n \geq T_n$. That (49) holds for n implies in particular that $Y_n \cap \bigcup_{i=0}^{n-1} \tilde{J}_i = \emptyset$. Thus by Lemma 11 for $n - 1$, since by definition $H_i \subseteq \tilde{J}_i$ for all $0 \leq i \leq n - 1$,

$$\begin{aligned}
 \mathcal{L}_{n-1}(u) - \mathcal{L}_{n-1}(w_{n-1}) &\geq \sum_{i=0}^{n-1} \left(\frac{\theta_i g(c_i)}{(\log 1/c_i)^{1/3}} \right) + \int_{[-T, T] \setminus \bigcup_{i=0}^{n-1} H_i} |u - w_{n-1}| \\
 &\geq \int_{[-T, T] \setminus \bigcup_{i=0}^{n-1} \tilde{J}_i} |u - w_{n-1}| \\
 &\geq \int_{Y_n} |u - w_{n-1}| \\
 &\geq \int_{x_n}^{x_n + T_n} |u - w_{n-1}|.
 \end{aligned}$$

But we know by (37), also using (3.7), monotonicity of g , (R:2), and that $\theta_n \geq 1$, that for $t \in [x_n, x_n + b_n]$ we have

$$\begin{aligned}
 |(u - w_{n-1})(t)| &\geq |(u - w_n)(t)| - |w_n(t) - w_{n-1}(t)| \geq \theta_n g(b_n) - \|w_n - w_{n-1}\|_\infty \\
 &\geq g(T_n) - 5K_n g(R_n) \\
 &\geq g(T_n)/2.
 \end{aligned}$$

Hence we see, since $[x_n, x_n + T_n] \subseteq [x_n, x_n + b_n]$, and since $\log \log 1/T_n \geq 1$, that

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

□

We are now in a position to conclude our argument. If our crucial assumptions (49) and (50) hold for all $n \geq 0$, then we are in the case of Corollary 12 and we are done. Otherwise, choose

the least $n \geq 0$ such that one of (49) or (50) fails. We observe that then $n \geq 1$ necessarily, since $\tilde{J}_0 \subseteq [-T, T]$.

Suppose $n \geq 1$ is such that (49) fails for n . Then we are in the case of Lemma 13 and we see by Lemma 6 that

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

Suppose $n \geq 1$ is such that (49) holds for n but (50) fails. Then we are in the case of Lemma 14 and we see again by Lemma 6 that

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

This contradicts the choice of u as a minimizer, so we know that no minimizer $u \neq w$ exists. Letting $\{x_n\}_{n=0}^\infty$ be an enumeration of $\mathbb{Q} \cap (-T, T)$ concludes the proof.

3 Approximation and variations

In this section we investigate different ways of approximating the minimum value of a variational problem. Throughout we continue to assume only that the Lagrangian L is continuous.

Question 1 *Let $v \in W^{1,1}(a, b)$ be a minimizer of \mathcal{L} over $\mathcal{A}_{v(a),v(b)}$. Does there exist a sequence $u_j \in W^{1,\infty}(a, b) \cap \mathcal{A}_{v(a),v(b)}$ such that $\mathcal{L}(u_j) \rightarrow \mathcal{L}(v)$?*

The answer to this in general is a well-known “no”, and in situations where the answer is negative, the *Lavrentiev phenomenon* is said to occur. Lavrentiev [12] gave the first example, and Manià [13] gave an example with a polynomial Lagrangian. Both examples have Lagrangians which vanish along the minimizing trajectory. Ball and Mizel [2] gave the first superlinear examples, with polynomial L for with $L_{pp} \geq \varepsilon > 0$ for some $\varepsilon > 0$.

This settles the question of whether in general the minimum value can be approximated by Lipschitz trajectories: no. A related question is whether the minimum value can be approximated by adding Lipschitz functions to the minimizing trajectory. One way of motivating this question is to consider that classically one finds minimizers by taking the first variations in the direction of functions $u \in C_0^\infty(a, b)$, i.e. computing $\frac{d}{d\gamma} \mathcal{L}(v + \gamma u)|_{\gamma=0}$, and looking for functions v for which this value is 0 for all such u . Under appropriate assumptions on L one can thereby derive the Euler-Lagrange equation, and look for minimizers among solutions to that. But how does the function $\gamma \mapsto \mathcal{L}(v + \gamma u)$ behave in general?

First we investigate this question forgetting for the moment that u is taken to be Lipschitz.

Question 2 *Let $v \in W^{1,1}(a, b)$. Does there exist a sequence $u_j \in \mathcal{A}_{v(a),v(b)}$, $u_j \neq v$, such that $\mathcal{L}(u_j) \rightarrow \mathcal{L}(v)$?*

The answer is an easy but apparently unrecorded “yes”, assuming only continuity of L , and holds for vector-valued trajectories v without too much extra work.

Theorem 15 *Let $v \in W^{1,1}((a, b); \mathbb{R}^n)$ be such that $t \mapsto L(t, v(t), v'(t))$ is integrable, $U \subseteq (a, b)$ be open and non-empty, and $\varepsilon > 0$.*

Then there exists $u \in \mathcal{A}_{v(a),v(b)}$ such that $\emptyset \neq \{t \in [a, b] : u(t) \neq v(t)\} \subseteq U$ and $|\mathcal{L}(u) - \mathcal{L}(v)| \leq \varepsilon$.

Remark Our method of proof gives the immediate further information that u is locally Lipschitz on $\{t \in [a, b] : u(t) \neq v(t)\}$.

If the function v is somewhere locally Lipschitz in U , then the approximation is obvious and can be done by adding to v a non-zero function of small norm in $W_0^{1,\infty}((a, b); \mathbb{R}^n)$ which is zero where v is not locally Lipschitz. If v is nowhere locally Lipschitz in U —which if v is a minimizer implies that L does not admit a partial regularity theorem—then the approximation is only slightly less obvious, and is done by replacing v with an affine function on appropriately small intervals. Notice however that the difference between v and the approximating function is non-Lipschitz.

The proof requires an easy lemma. For $v \in W^{1,1}((a, b); \mathbb{R}^n)$, $m > 0$, and $t \in (a, b)$, define

$$E_t := \{s \in [a, b] : \|v(s) - v(t)\| > m|s - t|\}; \text{ and}$$

$$M_t := \{s \in [a, b] : \|v(s) - v(t)\| = m|s - t|\}.$$

Lemma 16 *Let $v \in W^{1,1}((a, b); \mathbb{R}^n)$ and $m > 0$, and suppose that $t \in (a, b)$ is such that $\lambda(M_t) = 0$.*

Then $\lambda(E_t) \geq \limsup_{s \rightarrow t} \lambda(E_s)$.

Proof Let $t_k \in (a, b)$ be such that $t_k \rightarrow t$, and suppose that $s \in \bigcap_{k=1}^{\infty} \bigcup_{l=k}^{\infty} E_{t_l}$. So for all $k \geq 1$ there exists an $l \geq k$ such that $s \in E_{t_k}$, which by definition implies that

$$\|v(s) - v(t_k)\| > m|s - t_k|.$$

Letting $k \rightarrow \infty$, continuity of v implies that then

$$\|v(s) - v(t)\| \geq m|s - t|,$$

showing that $s \in E_t \cup M_t$. Thus $\bigcap_{k=1}^{\infty} \bigcup_{l=k}^{\infty} E_{t_l} \subseteq E_t \cup M_t$. Therefore, since by assumption $\lambda(M_t) = 0$, we have that

$$\begin{aligned} \lim_{k \rightarrow \infty} \lambda(E_{t_k}) &\leq \lim_{k \rightarrow \infty} \lambda\left(\bigcup_{l=k}^{\infty} E_{t_l}\right) = \lambda\left(\bigcap_{k=1}^{\infty} \bigcup_{l=k}^{\infty} E_{t_l}\right) \leq \lambda(E_t \cup M_t) \leq \lambda(E_t) + \lambda(M_t) \\ &= \lambda(E_t), \end{aligned}$$

as required. □

Proof (of Theorem 15) Choose $t_0 \in U$ such that $v'(t_0)$ exists and $\|v'(t_0)\| < \infty$. Then there exists $\rho > 0$ such that $|t_0 - t| \leq \rho$ implies that $\|v(t) - v(t_0)\| \leq (\|v'(t_0)\| + 1)|t - t_0|$. For $t \in [a, b]$ such that $|t - t_0| \geq \rho$, we have $\|v(t) - v(t_0)\| \leq 2 \sup_{s \in [a, b]} \|v(s)\| \leq 2 \sup_{s \in [a, b]} \|v(s)\| \rho^{-1} |t - t_0|$. So, for all $m \geq \max\{\|v'(t_0)\| + 1, 2 \sup_{s \in [a, b]} \|v(s)\| \rho^{-1}\}$, we have that $\|v(t) - v(t_0)\| \leq m|t - t_0|$ for all $t \in [a, b]$. Choose such an m , moreover such that $\lambda(\{s \in [a, b] : \|v(s) - v(t_0)\| = m|s - t_0|\}) = 0$; this is possible since this condition fails for at most countably many values of m . Then Lemma 16 implies that

$$0 \leq \limsup_{t \rightarrow t_0} \lambda(E_t) \leq \lambda(E_{t_0}) = \lambda(\emptyset) = 0. \tag{56}$$

By continuity of L we can choose $\delta \in (0, 1)$ such that $|L(t, y, p) - L(s, z, q)| \leq \varepsilon/2(b-a)$ whenever $\max\{|t|, \|y\|, \|p\|\} \leq |a| + |b| + \|v\|_{\infty} + m$ and $\max\{|s - t|, \|y - z\|, \|p - q\|\} \leq \delta$. Choose $\tau \in (0, \text{dist}(t_0, [a, b] \setminus U)/2)$ such that

- (15:1) $\tau \leq \delta/2m$;
- (15:2) $\tau \leq \varepsilon/(4 \sup_{t \in [a,b], \|p\| \leq m} |L(t, v(t), p)|)$;
- (15:3) $\int_E \|v'(t)\| dt \leq \delta/2$ whenever $\lambda(E) \leq \tau$; and
- (15:4) $\int_E |L(t, v(t), v'(t))| dt \leq \varepsilon/4$ whenever $\lambda(E) \leq \tau$.

By (56) we can choose $\eta \in (0, \tau)$ such that $|t - t_0| \leq \eta$ implies that $0 \leq \lambda(E_t) \leq \tau$.

Now, if $\|v'(t)\| \leq m$ for almost every $t \in (t_0 - \eta, t_0 + \eta)$, then we can construct a trivial variation in the usual way, by taking some non-zero $\psi \in C^\infty((a, b); \mathbb{R}^n)$ with $\text{spt } \psi \subseteq (t_0 - \eta, t_0 + \eta)$, and considering the sequence of functions $(v + j^{-1}\psi)$ as $j \rightarrow \infty$.

So suppose otherwise, i.e. that there exists $s_0 \in (t_0 - \eta, t_0 + \eta)$ such that $v'(s_0)$ exists and $\|v'(s_0)\| > m$. Then s_0 is an endpoint of some connected component (s_0, s_1) of the set E_{s_0} , by choice of s_0 . Notice, since $|s_0 - t_0| < \eta$, by the choice of η we have that $0 < (s_1 - s_0) \leq \lambda(E_{s_0}) \leq \tau$. Since $\eta < \tau$, we see that

$$|s_1 - t_0| \leq |s_1 - s_0| + |s_0 - t_0| \leq \tau + \eta < 2\tau < \text{dist}(t_0, [a, b] \setminus U).$$

So $s_1 \in U \subseteq (a, b)$, and we must have that $\|v(s_0) - v(s_1)\| = m|s_0 - s_1|$, since the only other way in which s_1 could be an endpoint of a component of E_{s_0} would be for it to be an endpoint of $[a, b]$, which we have now excluded.

So we can define $u \in \mathcal{A}_{v(a), v(b)}$ by

$$u(t) := \begin{cases} v(t) & t \notin (s_0, s_1), \\ \text{affine} & t \in (s_0, s_1); \end{cases}$$

so $u \neq v$, but $u = v$ off the set $(s_0, s_1) \subseteq U$, where $0 \leq s_1 - s_0 \leq \tau$. Moreover, on (s_0, s_1) we have that $\|u'\| = m$, and by (15:3) and (15:1) that

$$\begin{aligned} \|v(t) - u(t)\| &\leq \|v(t) - v(s_0)\| + \|u(s_0) - u(t)\| \leq \int_{s_0}^t \|v'(s)\| ds + m(t - s_0) \\ &\leq \int_{s_0}^{s_1} \|v'(s)\| ds + m|s_1 - s_0| \\ &\leq \delta/2 + \delta/2. \end{aligned}$$

So $\|v(t) - u(t)\| \leq \delta$ for all $t \in [a, b]$. So by the choice of δ as witnessing the continuity of L , (15:4), and (15:2), we have that

$$\begin{aligned} |\mathcal{L}(u) - \mathcal{L}(v)| &\leq \int_{s_0}^{s_1} |L(t, u(t), u') - L(t, v(t), u')| dt \\ &\quad + \int_{s_0}^{s_1} |L(t, v(t), u') - L(t, v(t), v'(t))| dt \\ &\leq \int_{s_0}^{s_1} \varepsilon/2(b - a) dt + \int_{s_0}^{s_1} |L(t, v(t), u')| dt + \int_{s_0}^{s_1} |L(t, v(t), v'(t))| dt \\ &\leq \varepsilon/2 + |(s_1 - s_0)| \left(\sup_{t \in [a,b], \|p\| \leq m} |L(t, v(t), p)| \right) + \varepsilon/4 \\ &\leq 3\varepsilon/4 + \tau \left(\sup_{t \in [a,b], \|p\| \leq m} |L(t, v(t), p)| \right) \\ &\leq \varepsilon, \end{aligned}$$

as required. □

Question 3 Let $v \in W^{1,1}(a, b)$. Does there exist a sequence of non-zero $u_j \in W_0^{1,\infty}(a, b)$ such that $\mathcal{L}(v + u_j) \rightarrow \mathcal{L}(v)$?

Ball and Mizel [2] gave examples exhibiting the Lavrentiev phenomenon for which they made the incidental observation that $\mathcal{L}(v + tu) = \infty$ for all $t \neq 0$, for a large class of $u \in C_0^\infty(a, b)$, viz those u which are non-zero at a certain point in the domain (at which the minimizer v is singular). The Lagrangians are polynomial, superlinear, and satisfy $L_{pp} \geq \varepsilon > 0$. This would seem to suggest that the same could happen for all $u \in C_0^\infty(a, b)$ if a minimizer was singular on a dense set. Indeed this is the case, as we shall shortly show, so the answer to our question, even if v is a minimizer, is “no”. The construction is straightforward if we do not concern ourselves with superlinearity and strict convexity; we have to try rather harder to get $L_{pp} > 0$, since in this case partial regularity statements follow given only the mildest assumptions on the modulus of continuity of the Lagrangian [4, 5, 8, 14].

The following example is not at all difficult but I am not aware of it being presented elsewhere.

Theorem 17 *There exists $v \in W^{1,1}(0, 1)$ and a continuous Lagrangian $L : [0, 1] \times \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$, convex in p , such that v is a minimizer of \mathcal{L} over $\mathcal{A}_{v(0),v(1)}$, $0 \leq \mathcal{L}(v) < \infty$, but $\mathcal{L}(v + u) = \infty$ for all non-zero $u \in W_0^{1,\infty}(0, 1)$.*

Proof Let $\{x_n\}_{n=0}^\infty$ be an enumeration of $\mathbb{Q} \cap [0, 1]$. For $n, k \geq 0$ define $U_{n,k} := (x_n - 2^{-n-3k}, x_n + 2^{-n-3k}) \cap [0, 1]$. For $n \geq 0$ define a non-negative function $\rho_n \in L^1(0, 1)$ by

$$\rho_n(t) := \sum_{k=0}^\infty 2^k \mathbb{1}_{U_{n,k}}(t).$$

So

$$\int_0^1 \rho_n(t) dt \leq \sum_{k=0}^\infty 2^{k+1} 2^{-3k-n} \leq 2^{-n+2},$$

and we can define $\rho \in L^1(0, 1)$ by $\rho := \sum_{n=0}^\infty \rho_n$ and $v \in W^{1,1}(0, 1)$ by

$$v(t) := \int_0^t (1 + \rho(s)) ds.$$

So for all $n, k \geq 0$, for almost every $t \in U_{n,k}$ we have that

$$v'(t) = 1 + \rho(t) \geq \rho_n(t) \geq 2^k. \tag{57}$$

Let $L : [0, 1] \times \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$ be given by

$$L(t, y, p) := (y - v(t))^2 p^8.$$

Then L is continuous, and convex in p , and v is clearly a minimizer of \mathcal{L} over $\mathcal{A}_{v(0),v(1)}$.

For any non-zero $u \in W_0^{1,\infty}(0, 1)$, there exist $n \geq 0, K \geq 0$, and $\varepsilon > 0$ such that $|u| \geq \varepsilon$ on $U_{n,k}$ for all $k \geq K$. Without loss of generality, we may assume that $2^K \geq 2\|u'\|_{L^\infty(0,1)}$. Let $k \geq K$. So we have that $2^k - \|u'\|_{L^\infty(0,1)} \geq 2^k - 2^{k-1} = 2^{k-1}$, which, with (57), implies that

$$\int_{U_{n,k}} (v'(t) + u'(t)) dt \geq \int_{U_{n,k}} (2^k - \|u'\|_{L^\infty(0,1)}) dt \geq 2^{k-1} \lambda(U_{n,k}).$$

So we have, using Jensen’s inequality, that

$$\begin{aligned} \mathcal{L}(v + u) &\geq \int_{U_{n,k}} (u(t))^2(v'(t) + u'(t))^8 dt \\ &\geq \varepsilon^2 \lambda(U_{n,k}) \left(\frac{1}{\lambda(U_{n,k})} \int_{U_{n,k}} (v'(t) + u'(t)) dt \right)^8 \\ &\geq \varepsilon^2 2^{-n-3k+1} \cdot 2^{8(k-1)} \\ &= \varepsilon^2 2^{5k-n-7}. \end{aligned}$$

Since this holds for all $k \geq K$, we have that $\mathcal{L}(v + u) = \infty$, as required. □

This example will also serve to demonstrate that the general approximation provided in Theorem 15 is not of the form $v + u$ for some $u \in W_0^{1,1}(a, b)$ for which $\mathcal{L}(v + \gamma u)$ is finite for the range of values $\gamma \in (0, 1]$. So question 2, while admitting a positive answer, cannot in general be answered in the manner provided by Theorem 15 by passing far enough down a sequence of the form $v + j^{-1}u$ for some $u \in W_0^{1,1}(a, b)$.

Let L and v be as constructed in Theorem 17, $U \subseteq (0, 1)$ be open and non-empty, and let $u \in \mathcal{A}_{v(0),v(1)}$ be as constructed in Theorem 15 for some $\varepsilon > 0$. Then the key point of the construction of u is that there exists a subinterval (s_0, s_1) of U and a fixed gradient m , say, such that $u(t) \neq v(t)$ implies that $t \in (s_0, s_1)$ and $u'(t) = m$, for almost every $t \in (0, 1)$. Define $w \in W_0^{1,1}(0, 1)$ by $w(t) := u(t) - v(t)$, so $v'(t) + w'(t) = m$ for almost every $t \in \{s \in (0, 1) : w(s) \neq 0\} \neq \emptyset$. Let $\gamma \in (0, 1)$.

Then there exist $n, K \geq 0$ and $\delta > 0$ such that $|\gamma w(t)| \geq \delta$ for $t \in U_{n,k} \subseteq (s_0, s_1)$ for all $k \geq K$, where $U_{n,k}$ are as in Theorem 17. Without loss of generality we may choose $K \geq 0$ such that $2^K \geq 2|m|/(1 - \gamma)$. Let $k \geq K$. So $m \leq (1 - \gamma)2^{k-1}$ and so

$$(1 - \gamma)2^k - m \geq (1 - \gamma)(2^k - 2^{k-1}) = (1 - \gamma)2^{k-1}.$$

Notice that $v'(t) + \gamma w'(t) = v'(t) + \gamma(m - v'(t)) = (1 - \gamma)v'(t) + \gamma m$ for almost every $t \in U_{n,k}$, so, by (57), we have that

$$\begin{aligned} \int_{U_{n,k}} (v'(t) + \gamma w'(t)) dt &= \int_{U_{n,k}} ((1 - \gamma)v'(t) + \gamma m) dt \geq \int_{U_{n,k}} ((1 - \gamma)2^k - \gamma|m|) dt \\ &\geq ((1 - \gamma)2^k - |m|)\lambda(U_{n,k}) \\ &\geq (1 - \gamma)2^{k-1}\lambda(U_{n,k}), \end{aligned}$$

since $\gamma \leq 1$. Hence we have, using Jensen’s inequality, that

$$\begin{aligned} \mathcal{L}(v + \gamma w) &\geq \int_{U_{n,k}} (\gamma w(t))^2(v'(t) + \gamma w'(t))^8 dt \\ &\geq \delta^2 \lambda(U_{n,k}) \left(\frac{1}{\lambda(U_{n,k})} \int_{U_{n,k}} (v'(t) + \gamma w'(t)) dt \right)^8 \\ &\geq \delta^2 2^{-n-3k+1} \cdot (1 - \gamma)^8 2^{8(k-1)} \\ &= \delta^2 (1 - \gamma)^8 2^{5k-n-7}. \end{aligned}$$

Since this holds for all $k \geq K$, we have that $\mathcal{L}(v + \gamma w) = \infty$, for all $\gamma \in (0, 1)$.

The Lagrangian we have constructed in Theorem 17, however, vanishes along the minimizer, and so is not superlinear. Gratwick and Preiss [11] show that it is possible to have a

continuous, superlinear Lagrangian with $L_{pp} \geq 2 > 0$ for which the minimizer is nowhere locally differentiable. That minimizer is, however, Lipschitz. The example of Sect. 2 is a non-Lipschitz version of this construction, which gives a minimizer which has upper and lower Dini derivatives of $\pm\infty$ at every point of a dense set.

Theorem 18 *There exist $T > 0$, $w \in W^{1,2}(-T, T)$, and a continuous Lagrangian $L: [-T, T] \times \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$, superlinear and satisfying $L_{pp} \geq 2 > 0$, such that w is a minimizer of \mathcal{L} over $\mathcal{A}_{w(-T), w(T)}$, $0 \leq \mathcal{L}(w) < \infty$, but $\mathcal{L}(w + u) = \infty$ for all non-zero $u \in W_0^{1,\infty}(-T, T)$.*

Proof We let T , w , and ϕ be as from Theorem 2, the notation of which we retain. We have to add another term to that Lagrangian. For $k \geq 0$ we choose a decreasing sequence of numbers $t_k \in (0, T)$ such that

$$\frac{\tilde{w}(t_k)}{t_k} \geq 2k + 1,$$

and recall that (29) implies that

$$\frac{w(x_n + t_k) - w(x_n)}{t_k} \geq \frac{\tilde{w}(t_k)}{t_k} - 1 \geq 2k, \tag{58}$$

for all $n \geq 0$ and $k \geq 0$ large enough, depending on n . Define a convex and superlinear function $\omega: \mathbb{R} \rightarrow [0, \infty)$ as follows. Set $\omega(0) := 0$ and $\omega(1) := t_1^{-1}$. Suppose $\omega(l)$ has been defined such that $\omega(l) \geq 2\omega(l - 1)$ for each $1 \leq l \leq k - 1$ for $k \geq 2$. Define

$$\omega(k) := \max\{2\omega(k - 1), kt_k^{-1}\}.$$

This defines $\omega(k)$ for all $k \geq 0$, and we then define ω to be affine between the specified endpoints on each interval $[k, k + 1]$. Define $\omega(p) := \omega(-p)$ for $p \in (-\infty, 0)$. Define $L: [-T, T] \times \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$ by

$$L(t, y, p) := \phi(t, y - w(t)) + p^2 + (y - w(t))^2\omega(p),$$

which is continuous, superlinear, with $L_{pp} \geq 2 > 0$, for which by Theorem 2, w is a minimizer of \mathcal{L} over $\mathcal{A}_{w(-T), w(T)}$.

Let $u \in W_0^{1,\infty}(-T, T)$ be non-zero. So there exist $n \geq 0$ and $\varepsilon > 0$ such that $|u| \geq \varepsilon$ on a neighbourhood of x_n . Choosing $K \geq \|u'\|_{L^\infty(-T, T)}$, we have for all large $k \geq K$, by (58), that

$$\begin{aligned} \int_{x_n}^{x_n+t_k} (w'(t) + u'(t)) dt &\geq w(x_n + t_k) - w(x_n) - \|u'\|_{L^\infty(-T, T)}t_k \\ &\geq (2k - \|u'\|_{L^\infty(-T, T)})t_k \\ &\geq kt_k. \end{aligned}$$

So, using Jensen's inequality and that ω is non-decreasing on $[0, \infty)$, we have that

$$\begin{aligned} \int_{x_n}^{x_n+t_k} L(t, w + u, w' + u') dt &\geq \int_{x_n}^{x_n+t_k} (u(t))^2\omega(w'(t) + u'(t)) dt \\ &\geq \varepsilon^2 t_k \omega\left(\frac{1}{t_k} \int_{x_n}^{x_n+t_k} (w'(t) + u'(t)) dt\right) \\ &\geq \varepsilon^2 t_k \omega(k) \\ &\geq \varepsilon^2 k, \end{aligned}$$

where the final inequality follows by the definition of ω . Since this holds for all large $k \geq K$, $\mathcal{L}(w + u) = \infty$, as required. \square

We might now speculate whether in any one given problem, it is always possible to approximate the minimum value *either* by Lipschitz trajectories *or* by adding Lipschitz trajectories to the minimizer. We know that neither approach alone succeeds in general, but is it possible that both fail simultaneously?

Question 4 *Let $v \in W^{1,1}(a, b)$ be a minimizer of \mathcal{L} over $\mathcal{A}_{v(a),v(b)}$. Suppose that the Lavrentiev phenomenon occurs. Does there exist a sequence of non-zero $u_j \in W_0^{1,\infty}(a, b)$ such that $\mathcal{L}(v + u_j) \rightarrow \mathcal{L}(v)$?*

There seems to be very little reason to think this might be true: inferring a positive approximation result from the presence of the Lavrentiev phenomenon seems eccentric. (In contrast, the principle of gaining information about the minimizer assuming *non*-occurrence of the Lavrentiev phenomenon is used, for example, by Esposito, Leonetti, and Mingione [6].) To show it to be false, we just need to show that the Lavrentiev phenomenon occurs in (a modified version of) one of the examples from Theorems 17 or 18.

Theorem 19 *There exists $v \in W^{1,1}(0, 1)$ and a continuous Lagrangian $L : [0, 1] \times \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$, convex in p , such that v is a minimizer of \mathcal{L} over $\mathcal{A}_{v(0),v(1)}$, $0 \leq \mathcal{L}(v) < \infty$, but $\mathcal{L}(v + u) = \infty$ for all non-zero $u \in W_0^{1,\infty}(0, 1)$. Moreover, the Lavrentiev phenomenon occurs.*

Proof We show that the example from Theorem 17 exhibits the Lavrentiev phenomenon. The argument is based on the presentation given in [3] of the example given by Manià [13]. We borrow our notation from the proof of Theorem 17. Without loss of generality $x_0 = 0$.

Let $u \in W^{1,\infty}(0, 1) \cap \mathcal{A}_{v(0),v(1)}$. The definition of v implies that for $t \in U_{0,k}$, $v(t) \geq t\rho_0(t) \geq 2^k$ and so, since as $t \rightarrow 0$ we may choose $k \rightarrow \infty$, and $v(0) = 0$, we see that $v'(0) = \infty$. Since $u(0) = v(0) = 0$ and u is Lipschitz, we must have that $u < v/4$ on a right neighbourhood of 0. Since also $u(1) = v(1)$, by the intermediate value theorem, $\{t \in (0, 1) : u(t) = v(t)/4\} \neq \emptyset$. Define

$$\tau_1 := \sup\{t \in (0, 1) : u(t) = v(t)/4\} < 1.$$

Similarly $\{t \in (\tau_1, 1) : u(t) = v(t)/2\} \neq \emptyset$, so we may define

$$\tau_2 := \min\{2\tau_1, \inf\{t \in (\tau_1, 1) : u(t) = v(t)/2\}\}.$$

Choose $k \geq 0$ such that $2^{-3k-3} \leq \tau_1 \leq 2^{-3k}$, so $\tau_1 \in U_{0,k}$. Then for $t \in (\tau_1, \tau_2)$, we have by definition of τ_2 , monotonicity of v , and (57) that

$$v(t) - u(t) \geq v(t) - v(t)/2 = v(t)/2 \geq v(\tau_1)/2 \geq 2^{k-1}\tau_1 \geq 2^{k-1} \cdot 2^{-3k-3} = 2^{-2k-4},$$

so $(v(t) - u(t))^2 \geq 2^{-4k-8}$.

If $\tau_2 = 2\tau_1$, then $\tau_1, \tau_2 \in U_{0,k-1}$ if $k \geq 1$, and so $v' \geq 2^{k-1}$ almost everywhere on (τ_1, τ_2) . If $k = 0$, then all we may say is that $\tau_1, \tau_2 \in U_{0,0} = [0, 1)$, on which $v' \geq 2$ by definition of v , since $\rho \geq 1$ everywhere. So in general we may say that $v' \geq 2^{k-1}$ almost everywhere on (τ_1, τ_2) . Hence, by definition of τ_1 ,

$$\begin{aligned} u(\tau_2) - u(\tau_1) &\geq v(\tau_2)/4 - v(\tau_1)/4 = 2^{-2}(v(2\tau_1) - v(\tau_1)) \geq 2^{-2} \cdot 2^{k-1}\tau_1 \\ &\geq 2^{k-3} \cdot 2^{-3k-3} \\ &= 2^{-2k-6}. \end{aligned}$$

Otherwise, we have by the definitions of τ_1 and τ_2 , the monotonicity of v , and (57) that

$$u(\tau_2) - u(\tau_1) = v(\tau_2)/2 - v(\tau_1)/4 \geq v(\tau_1)/4 \geq 2^{-2} \cdot 2^k \tau_1 \geq 2^{-2+k} \cdot 2^{-3k-3} = 2^{-2k-5}.$$

Hence in both cases we have that $(u(\tau_2) - u(\tau_1))^8 \geq 2^{-16k-48}$. So, using Jensen’s inequality, we see, since $(\tau_2 - \tau_1)^{-1} \geq \tau_1^{-1} \geq 2^{3k}$ by definition of τ_2 , that

$$\begin{aligned} \mathcal{L}(u) &\geq \int_{\tau_1}^{\tau_2} (u(t) - v(t))^2 (u'(t))^8 dt \geq 2^{-4k-8} (\tau_2 - \tau_1) \left((\tau_2 - \tau_1)^{-1} \int_{\tau_1}^{\tau_2} u'(t) dt \right)^8 \\ &\geq 2^{-4k-8} (\tau_2 - \tau_1)^{-7} 2^{-16k-48} \\ &\geq 2^{-20k-56} \cdot 2^{21k} \\ &\geq 2^{-56}. \end{aligned}$$

Since this number is independent of k and therefore of u , we see that the Lavrentiev phenomenon occurs, as claimed. □

The Lagrangian in Theorem 19 can be adapted to have $L_{pp} \geq \varepsilon > 0$ and be superlinear, while still exhibiting the Lavrentiev phenomenon. Notice that

$$\int_0^1 \rho_n^2 \leq \sum_{k=0}^{\infty} 2^{2k} \lambda(U_{n,k}) \leq \sum_{k=0}^{\infty} 2^{2k} \cdot 2^{-3k-n+1} = 2^{-n+2},$$

and therefore that $\sum_{n=0}^{\infty} \rho_n$ converges in $L^2(0, 1)$, thus $v \in W^{1,2}(0, 1)$. Following [3] (see p. 148), we set

$$\tilde{L}(t, y, p) = (y - v(t))^2 p^8 + \varepsilon p^2,$$

for some $0 < \varepsilon < 2^{-56} \|v'\|_{L^2(0,1)}^{-2}$. Then $\tilde{L}_{pp} \geq 2\varepsilon > 0$ and \tilde{L} is superlinear, and moreover, letting $\tilde{\mathcal{L}}$ denote the corresponding functional,

$$\inf_{w \in \mathcal{A}_{v(0),v(1)}} \tilde{\mathcal{L}}(w) \leq \tilde{\mathcal{L}}(v) = \varepsilon \|v'\|_{L^2(0,1)}^2 < 2^{-56} \leq \tilde{\mathcal{L}}(u),$$

for all $u \in W^{1,\infty}(0, 1) \cap \mathcal{A}_{v(0),v(1)}$, so the Lavrentiev phenomenon persists. However, we lose the easy observation that v is a minimizer, and the result about Lipschitz variations is no longer clear.

The example of Theorem 18 sadly rather readily fails to exhibit the Lavrentiev phenomenon: consider following the near-minimizer w_n everywhere except on small intervals around its singularities, on which one just remains constant until one can pick up the minimizer on the other side of the singularity (this argument is made precise by Gratwick [10]). However, it can be modified, as suggested by the standard computations involved in Manià’s example, into an example which does exhibit the Lavrentiev phenomenon, by adding a non-decreasing trajectory with a vertical tangent at 0.

Theorem 20 *There exist $T > 0$, $w \in W^{1,2}(0, T)$, and a continuous Lagrangian $L : [0, T] \times \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$, superlinear in p and with $L_{pp} \geq 2 > 0$, such that w is a minimizer of \mathcal{L} over $\mathcal{A}_{w(0),w(T)}$, $0 \leq \mathcal{L}(w) < \infty$, but $\mathcal{L}(w + u) = \infty$ for all non-zero $u \in W_0^{1,\infty}(0, T)$. Moreover, the Lavrentiev phenomenon occurs.*

Proof We adapt the example from Theorem 18, and borrow the notation from that proof. We consider only the interval $[0, T]$, and observe that the function w from this example is a

minimizer of \mathcal{L} on $[0, T]$ over $\mathcal{A}_{w(0),w(T)}$, and note that $w(0) = w(x_0) = w_0(x_0) = \tilde{w}(0) = 0$ by (4.1) and the definition of \tilde{w} .

From (4.4) and (4.1) we know that

$$|w(t)| = |w(t) - w(0)| \leq 2g(t), \tag{59}$$

for all $t \in [0, T]$. So $w(t) + 3g(t) \geq g(t)$. This “3g-centred” version of w will be our new minimizer with respect to its own boundary conditions $w(0) + 3g(0) = 0$ and $w(T) + 3g(T)$. We modify our Lagrangian from Theorem 18 to construct a problem which this new function minimizes; to do this we need to add a new weight function containing a term in g'' .

Let $\Phi : [0, T] \times \mathbb{R} \rightarrow [0, \infty)$ be given by

$$\Phi(t, y) := \begin{cases} 0 & t = 0, \\ 7|g''(t)||y| & t \neq 0, |y| \leq 6g(t), \\ 42|g''(t)|g(t) & t \neq 0, |y| > 6g(t). \end{cases}$$

Now,

$$g''(t) = \frac{-1}{t \log 1/t} \left(\frac{1}{\log 1/t} + 1 \right),$$

so recalling that $T > 0$ was chosen small enough that $\log 1/t \geq 1$ on $(0, T)$, we see that

$$|g''(t)g(t)| \leq \frac{2t \log \log 1/t}{t \log 1/t} = \frac{2 \log \log 1/t}{\log 1/t} \rightarrow 0 \text{ as } t \rightarrow 0, \tag{60}$$

so Φ is continuous. Now define $F : [0, T] \times \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$ by

$$F(t, y, p) := \phi(t, y - (w + 3g)) + \Phi(t, y - (w + 3g)) + p^2,$$

and consider the corresponding functional $\mathcal{F}(u) := \int_0^T F(t, u, u')$ defined for $u \in W^{1,1}(0, T)$. We claim that $w + 3g$ is a minimizer over $\mathcal{A}_{0,(w+3g)(T)}$ of \mathcal{F} . Let $u \in \mathcal{A}_{0,w(T)}$ be such that $u + 3g$ is a minimizer of \mathcal{F} over $\mathcal{A}_{0,(w+3g)(T)}$.

First we claim that $|u(t)| \leq 4g(t)$ on $[0, T]$. This is the same strategy of proof as found in Lemma 7, so we give no more than a sketch of the argument. Suppose for a contradiction that $u(t) > 4g(t)$ on some interval I . Then since $|w(t) + 3g(t)| \leq 5g(t)$ by (59), we see that $u(t) + 3g(t) > 7g(t) \geq w(t) + 3g(t)$ on I , where $7g$ is a concave function, and in particular that $I \subsetneq (0, T)$. Therefore we can find an affine function l such that $u(t) + 3g(t) > l(t) \geq w(t) + 3g(t)$ on some subinterval of I . Defining a new trajectory $u_l \in \mathcal{A}_{0,(w+3g)(T)}$ by replacing $u + 3g$ with l on this subinterval of I , we see that u_l does not increase the “weight terms” ϕ and Φ in F , and strictly decreases the gradient term, since affine functions are the unique minimizers of quadratic functionals, so

$$\mathcal{F}(u_l) < \mathcal{F}(u + 3g),$$

which contradicts the choice of u as being such that $u + 3g$ is a minimizer. Supposing in the other case for a contradiction that $u(t) < -4g(t)$ on some interval I , we see that $u(t) + 3g(t) < -g(t) \leq 0 \leq w(t) + 3g(t)$, where $-g$ is a convex function, and we use a similar argument to gain a contradiction.

So indeed $|u(t)| \leq 4g(t)$ on $[0, T]$, hence $|u(t) - w(t)| \leq 6g(t)$ by (59) and thus $\Phi(t, u - w) = 7|g''||u - w|$ on $(0, T]$, by definition.

We now claim that, extended to have value 0 at $t = 0$, the function $g'(u - w)$ is absolutely continuous on $[0, T]$, i.e. can be written as an indefinite integral on $[0, T]$. That this definition makes it continuous follows since

$$|g'(t)(u(t) - w(t))| \leq 6|g'(t)g(t)| \leq 6 \left(\frac{1}{\log 1/t} + \log \log 1/t \right) t \log \log 1/t \rightarrow 0$$

as $t \rightarrow 0$. Clearly $g'(u - w)$ is absolutely continuous on subintervals of $(0, T)$ bounded away from 0, so by the dominated convergence theorem it suffices to show that $(g'(u - w))' \in L^1(0, T)$. Now,

$$|(g'(u - w))'| \leq |g''||u - w| + |g'|(|u' + 3g'| + |w' + 3g'|),$$

and by the above we have that $|g''||u - w| \leq 6|g''||g|$, which is bounded on $[0, T]$ by (60), so certainly integrable. Using Cauchy–Schwartz, we see further that

$$\int_0^T |g'|(|u' + 3g'| + |w' + 3g'|) \leq \left(\int_0^T |g'|^2 \right)^{1/2} \left(\int_0^T (|u' + 3g'| + |w' + 3g'|)^2 \right)^{1/2},$$

which is finite since $g' \in L^2(0, T)$, $w \in W^{1,2}(0, T)$, and since $u + 3g$ is a minimizer of \mathcal{F} by assumption. So indeed $g'(u - w)$ is absolutely continuous on $[0, T]$.

The minimality of w established in Theorem 2 implies that

$$\int_0^T (\phi(t, u - w) + (u')^2 - (w')^2) \geq 0.$$

So, recalling also the simple pointwise inequality (2), and integrating $g'(u' - w')$ by parts, we see that

$$\begin{aligned} & \mathcal{F}(u + 3g) - \mathcal{F}(w + 3g) \\ &= \int_0^T (\phi(t, u - w) + \Phi(t, u - w) + (u' + 3g')^2 - (w' + 3g')^2) dt \\ &= \int_0^T (\phi(t, u - w) + (u')^2 - (w')^2) dt \\ & \quad + \int_0^T (\Phi(t, u - w) + 6g'(u' - w')) dt \\ & \geq 0 + 6[g'(u - w)]_0^T + \int_0^T (\Phi(t, u - w) - 6g''(u - w)) dt \\ & \geq \int_0^T (7|g''||u - w| - 6|g''||u - w|) dt \\ & \geq 0. \end{aligned}$$

So $w + 3g$ is indeed a minimizer of \mathcal{F} .

Since g' increases to ∞ as we approach 0, we can find a sequence $r_k > 0$ such that $r_k \downarrow 0$ and $g' \geq k + 1$ on $(0, r_k)$. $T > 0$ was chosen small enough that we may consistently set $r_0 = 2T$. Define a convex and superlinear function $\Theta: \mathbb{R} \rightarrow [0, \infty)$ such that

$$\Theta(p) \geq 2^8 \|w' + 3g'\|_{L^2(0,T)}^2 p r_k^{-3} \text{ for } p \geq k/4, \tag{61}$$

for all $k \geq 0$, as follows. Set $\Theta(0) := 0$ and $\Theta(1/4) := 2^6 \|w' + 3g'\|_{L^2(0,T)}^2 r_1^{-3}$. Suppose that $\Theta(l/4)$ has been defined such that $\Theta(l/4) \geq 2\Theta((l - 1)/4)$ for all $1 \leq l \leq k - 1$, for some $k \geq 2$. Define

$$\Theta(k/4) := \max \left\{ 2\Theta((k-1)/4), 2^6 \|w' + 3g'\|_{L^2(0,T)}^2 kr_k^{-3} \right\}.$$

This defines Θ inductively at the points $k/4$, and we extend it to be affine on each interval $[k/4, (k+1)/4]$. Define $\Theta(p) = \Theta(-p)$ for $p \in (-\infty, 0)$.

Define $L(t, y, p) : [0, T] \times \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$ by

$$\begin{aligned} L(t, y, p) &:= \phi(t, y - (w + 3g)) + \Phi(t, y - (w + 3g)) + p^2 \\ &\quad + (y - (w + 3g))^2(\omega(p) + \Theta(p)) \\ &= F(t, y, p) + (y - (w + 3g))^2(\omega(p) + \Theta(p)). \end{aligned}$$

So L is continuous, superlinear in p and has $L_{pp} \geq 2 > 0$, and, since $w + 3g$ is a minimizer of \mathcal{F} over $\mathcal{A}_{0,(w+3g)}(T)$, clearly $w + 3g$ is a minimizer of the associated functional \mathcal{L} over $\mathcal{A}_{0,(w+3g)}(T)$.

By monotonicity of g and (58) we have that

$$\frac{(w + 3g)(x_n + t_k) - (w + 3g)(x_n)}{t_k} \geq \frac{w(x_n + t_k) - w(x_n)}{t_k} \geq 2k,$$

for all $n \geq 0$ and sufficiently large $k \geq 0$, depending on n . Given this, the argument that $\mathcal{L}(w + 3g + u) = \infty$ for all $u \in W_0^{1,\infty}(0, T)$ follows exactly as in Theorem 18.

It just remains to show that the Lavrentiev phenomenon occurs. The Manià-style estimates follow exactly the same pattern as before. Let $u \in W^{1,\infty}(0, T) \cap \mathcal{A}_{0,(w+3g)}(T)$. Since $g'(0) = \infty$, and that $u(0) = w(0) = g(0) = 0$ and u is Lipschitz, we have that $u < g/4$ on some right neighbourhood of 0. Since $u(T) = (w + 3g)(T) \geq g(T)$, the intermediate value theorem implies that $\{t \in (0, T) : u(t) = g(t)/4\} \neq \emptyset$. Define

$$\tau_1 := \sup\{t \in (0, T) : u(t) = g(t)/4\} < T.$$

Similarly $\{t \in (\tau_1, T) : u(t) = g(t)/2\} \neq \emptyset$, so we may define

$$\tau_2 = \min\{2\tau_1, \inf\{t \in (\tau_1, T) : u(t) = g(t)/2\}\}.$$

So $0 < \tau_2 - \tau_1 \leq \tau_1$. Choose $k \geq 0$ such that $r_{k+1} \leq 2\tau_1 \leq r_k$. By (59), the definition of τ_2 , the monotonicity of g , and the choice of r_k , for all $t \in (\tau_1, \tau_2)$ we have that

$$(w + 3g)(t) - u(t) \geq g(t) - g(t)/2 \geq g(\tau_1)/2 \geq (k + 1)\tau_1/2.$$

Thus $((w + 3g)(t) - u(t))^2 \geq 2^{-2}(k + 1)^2\tau_1^2$ on (τ_1, τ_2) .

Now, if $\tau_2 = 2\tau_1$, then by the definition of τ_1 and choice of r_k , we have that

$$u(\tau_2) - u(\tau_1) \geq g(\tau_2)/4 - g(\tau_1)/4 = 2^{-2}(g(2\tau_1) - g(\tau_1)) \geq 2^{-2}(k + 1)\tau_1.$$

Otherwise, we have by the definitions of τ_1 and τ_2 , the monotonicity of g , and the choice of r_k that

$$u(\tau_2) - u(\tau_1) = g(\tau_2)/2 - g(\tau_1)/4 \geq g(\tau_1)/2 - g(\tau_1)/4 = g(\tau_1)/4 \geq 2^{-2}(k + 1)\tau_1.$$

So in either case we have that

$$\frac{u(\tau_2) - u(\tau_1)}{\tau_2 - \tau_1} \geq \frac{u(\tau_2) - u(\tau_1)}{\tau_1} \geq 2^{-2}(k + 1).$$

Therefore by Jensen’s inequality we have, by the choice of Θ as satisfying (61), that

$$\begin{aligned} \int_{\tau_1}^{\tau_2} \Theta(u'(t)) dt &\geq (\tau_2 - \tau_1) \Theta\left(\frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} u'(t) dt\right) \\ &= (\tau_2 - \tau_1) \Theta\left(\frac{u(\tau_2) - u(\tau_1)}{\tau_2 - \tau_1}\right) \\ &\geq (\tau_2 - \tau_1) 2^8 \|w' + 3g'\|_{L^2(0,T)}^2 \frac{u(\tau_2) - u(\tau_1)}{\tau_2 - \tau_1} r_{k+1}^{-3} \\ &\geq 2^6 \|w' + 3g'\|_{L^2(0,T)}^2 (k + 1) \tau_1 r_{k+1}^{-3}. \end{aligned}$$

So, since $\tau_1 \geq r_{k+1}/2$, we have that

$$\begin{aligned} \mathcal{L}(u) &\geq \int_{\tau_1}^{\tau_2} (u(t) - (w + 3g)(t))^2 \Theta(u'(t)) dt \\ &\geq 2^{-2} (k + 1)^2 \tau_1^2 \int_{\tau_1}^{\tau_2} \Theta(u'(t)) dt \\ &\geq 2^{-2} (k + 1)^2 \tau_1^2 2^6 \|w' + 3g'\|_{L^2(0,T)}^2 (k + 1) \tau_1 r_{k+1}^{-3} \\ &= 2^4 (k + 1)^3 \tau_1^3 \|w' + 3g'\|_{L^2(0,T)}^2 r_{k+1}^{-3} \\ &\geq 2 \|w' + 3g'\|_{L^2(0,T)}^2. \end{aligned}$$

That is, for all $u \in W^{1,\infty}(0, T) \cap \mathcal{A}_{0,(w+3g)}(T)$, we have that

$$\mathcal{L}(u) \geq 2 \|w' + 3g'\|_{L^2(0,T)}^2 > \|w' + 3g'\|_{L^2(0,T)}^2 = \mathcal{L}(w + 3g),$$

which is precisely to say that the Lavrentiev phenomenon occurs, as required. □

4 Minimal regularity

We know that under our standing assumptions, minimizers need not be everywhere differentiable. In this section, we deduce properties of the derivatives of minimizers at points at which they do exist. We show that derivatives must be approximately continuous at points where they exist (Theorems 24 and 25), and that “kinks” may not appear, i.e. if both one-sided derivatives exist at a point, they must be equal (Corollaries 26 and 27).

Our results in this section apply to vector-valued trajectories $v : [a, b] \rightarrow \mathbb{R}^n$. Our proofs proceed by contradiction, assuming a minimizer v has a derivative which fails to be well-behaved in a certain way, and thereby constructing a competitor trajectory with strictly lower energy, by replacing v with affine pieces on open subintervals of the domain.

Definition 21 The left and right Dini derivatives, $D^-v(t)$ and $D^+v(t)$ respectively, of a function $v \in W^{1,1}((a, b); \mathbb{R})$ at a point $t \in [a, b]$ are given by

$$D^-v(t) := \lim_{s \uparrow t} \frac{v(s) - v(t)}{s - t}, \text{ and } D^+v(t) := \lim_{s \downarrow t} \frac{v(s) - v(t)}{s - t},$$

whenever these limits make sense and exist as finite or infinite values. The left and right derivatives of vector-valued functions $v \in W^{1,1}((a, b); \mathbb{R}^n)$ are formed by taking the vectors of the corresponding left and right derivatives of the components; thus these exist at a point if and only if the corresponding derivatives of each component function exist at that point. In principle, then, such vectors of derivatives might contain components with infinite values.

We shall clearly distinguish the cases when all the components are finite, and when one or more may be infinite.

Definition 22 (See for example [7]) We recall the usual definition of approximate continuity. Let $f : [a, b] \rightarrow \mathbb{R}^n$ be measurable. We say that f is approximately continuous on the left at $t \in (a, b]$ if, for all $c > 0$,

$$\lim_{s \uparrow t} (t - s)^{-1} \lambda(\{r \in (s, t) : \|f(r) - f(t)\| \geq c\}) = 0;$$

similarly, we say that f is approximately continuous on the right at $t \in [a, b)$ if, for all $c > 0$,

$$\lim_{s \downarrow t} (s - t)^{-1} \lambda(\{r \in (t, s) : \|f(r) - f(t)\| \geq c\}) = 0.$$

We retain our standing assumption of continuity of the Lagrangian. Some further assumption of strict convexity is required to deduce any regularity results. We impose the following condition on L : that for all $R \in [1, \infty)$, there exists $\tau_R > 0$ such that for all (t, y, p) with $\max\{|t|, \|y\|, \|p\|\} \leq R + 1$, there exists a subdifferential $\xi \in \mathbb{R}^n$ of $L(t, y, \cdot)$ at p such that

$$L(t, y, q) \geq L(t, y, p) + \xi \cdot (q - p) + 2\tau_R, \tag{62}$$

whenever $\|q - p\| \geq R^{-1}$. This holds in particular if the (partial) Hessian L_{pp} exists and is continuous and strictly positive for all (t, y, p) .

The following lemma is our key tool, which we use repeatedly in the remainder of the section.

Lemma 23 Let $v \in W^{1,1}((a, b); \mathbb{R}^n)$, $R \geq |a| + |b| + \|v\|_\infty + 1$, and $\varepsilon > 0$.

Then there exists $\delta > 0$ such that if $(t_1, t_2) \subseteq [a, b]$ satisfies $t_2 - t_1 \leq \delta$ and $\|v(t_2) - v(t_1)\| \leq R(t_2 - t_1)$, then $u \in W^{1,1}((a, b); \mathbb{R}^n)$ defined by

$$u(r) := \begin{cases} v(r) & r \notin (t_1, t_2), \\ \text{affine} & \text{otherwise;} \end{cases}$$

satisfies

$$\begin{aligned} & \int_{t_1}^{t_2} L(s, v(s), v'(s)) \, ds \\ & \geq \int_{t_1}^{t_2} L(s, u(s), u'(s)) \, ds + \tau_R \lambda(\{s \in (t_1, t_2) : \|v'(s) - u'(s)\| \geq R^{-1}\}) - \varepsilon(t_2 - t_1). \end{aligned}$$

Proof First we show that the strict convexity and continuity of L conspire to allow us to use a subdifferential of $L(t, y, \cdot)$ in the convexity inequality involving the function $L(s, z, \cdot)$, when (s, z) is near to (t, y) .

We choose $\delta_1 \in (0, 1)$ witnessing the uniform continuity for $\min\{\tau_R/2, \varepsilon/3\}$ of L for (t, y, p) such that $\max\{|t|, \|y\|, \|p\|\} \leq R + 1$. Let $(t, y), (s, z) \in [a, b] \times \mathbb{R}^n$ be such that $\max\{|t|, \|y\|\} \leq R$ and $\max\{|s - t|, \|y - z\|\} \leq \delta_1$, and let $p, q \in \mathbb{R}^n$ be such that $\|p\| \leq R$ and $\|q - p\| \geq R^{-1}$. Define $\tilde{q} := p + R^{-1}\|q - p\|^{-1}(q - p)$, so $\|\tilde{q} - p\| = R^{-1}$ and $\|\tilde{q}\| \leq R + R^{-1}$. Letting $\xi \in \mathbb{R}^n$ be a subdifferential of $L(t, y, \cdot)$ at p which satisfies (62), using continuity twice we see that

$$\begin{aligned} L(s, z, \tilde{q}) & \geq L(t, y, \tilde{q}) - \tau_R/2 \geq L(t, y, p) + \xi \cdot (\tilde{q} - p) + 3\tau_R/2 \\ & \geq L(s, z, p) + \xi \cdot (\tilde{q} - p) + \tau_R. \end{aligned}$$

Convexity of the one-dimensional function $\mu \mapsto L(s, z, p + \mu(q - p))$ allows us to infer, since $R^{-1}\|q - p\|^{-1} \leq 1$, that

$$\begin{aligned} L(s, z, q) - L(s, z, p) &\geq \frac{L(s, z, \tilde{q}) - L(s, z, p)}{R^{-1}\|q - p\|^{-1}} \geq R\|q - p\|(\xi \cdot (\tilde{q} - p) + \tau_R) \\ &\geq \xi \cdot (q - p) + \tau_R. \end{aligned}$$

That is, $\delta_1 > 0$ is such that for all (t, y, p) with $\max\{|t|, \|y\|, \|p\|\} \leq R$, there exists a subdifferential $\xi \in \mathbb{R}^n$ of $L(t, y, \cdot)$ at p such that

$$L(s, z, q) \geq L(s, z, p) + \xi \cdot (q - p) + \tau_R, \tag{63}$$

whenever $\max\{|s - t|, \|y - z\|\} \leq \delta_1$ and $\|q - p\| \geq R^{-1}$.

Since v is absolutely continuous, when seeking to apply this inequality along the graph of the trajectory of v we can reduce the condition on proximity of the y variable (i.e. the $v(t)$ term) to a condition only on proximity of the time variable. Choose $\delta \in (0, \delta_1/2R)$ such that $\int_E \|v'\| \leq \delta_1/2$ whenever $E \subseteq [a, b]$ satisfies $\lambda(E) \leq \delta$.

Now let $(t_1, t_2) \subseteq [a, b]$ be such that $0 < t_2 - t_1 \leq \delta$, and $\|v(t_2) - v(t_1)\| \leq R(t_2 - t_1)$, and let $s \in (t_1, t_2)$. Then for $u \in W^{1,1}((a, b); \mathbb{R}^n)$ as defined in the statement, $u'(s) = l := (t_2 - t_1)^{-1}(v(t_2) - v(t_1))$, and so $\|l\| \leq R$ by assumption. Then

$$\begin{aligned} \|u(s) - v(t_1)\| &\leq R(s - t_1) \leq R\delta \leq \delta_1/2; \\ \|v(s) - v(t_1)\| &\leq \int_{t_1}^s \|v'\| \leq \delta_1/2; \end{aligned}$$

and so

$$\|u(s) - v(s)\| \leq \|u(s) - v(t_1)\| + \|v(t_1) - v(s)\| \leq \delta_1.$$

So $\max\{|s - t_1|, \|u(s) - v(t_1)\|, \|v(s) - v(t_1)\|, \|u(s) - v(s)\|\} \leq \delta_1$. Moreover, $\max\{|t|, \|v\|_\infty\} \leq R$. So there exists a subdifferential $\xi \in \mathbb{R}^n$ of $L(t_1, v(t_1), \cdot)$ at l for which (63) holds. Suppose first that $s \in (t_1, t_2)$ is such that $v'(s)$ exists and $\|v'(s) - l\| \geq R^{-1}$. Then using (63) and the continuity of L , we have

$$\begin{aligned} L(s, v(s), v'(s)) &\geq L(s, v(s), l) + \xi \cdot (v'(s) - l) + \tau_R \\ &\geq L(s, u(s), l) + \xi \cdot (v'(s) - l) + \tau_R - \varepsilon/3. \end{aligned}$$

Suppose now that $s \in (t_1, t_2)$ is such that $v'(s)$ exists and $\|v'(s) - l\| < R^{-1}$. Then $\|v'(s)\| \leq \|l\| + R^{-1} \leq R + 1$, and we may use continuity and a (non-strict) application of our convexity assumption (62) to see that

$$\begin{aligned} L(s, v(s), v'(s)) &\geq L(t_1, v(t_1), v'(s)) - \varepsilon/3 \\ &\geq L(t_1, v(t_1), l) + \xi \cdot (v'(s) - l) - \varepsilon/3 \\ &\geq L(s, u(s), l) + \xi \cdot (v'(s) - l) - 2\varepsilon/3. \end{aligned}$$

Since almost every $s \in (t_1, t_2)$ falls into one of these two cases, we can now integrate and see that

$$\begin{aligned} &\int_{t_1}^{t_2} L(s, v(s), v'(s)) \, ds \\ &\geq \int_{\{s \in (t_1, t_2) : \|v'(s) - l\| \geq R^{-1}\}} (L(s, u(s), l) + \xi \cdot (v'(s) - l) + \tau_R - \varepsilon/3) \, ds \end{aligned}$$

$$\begin{aligned}
 & + \int_{\{s \in (t_1, t_2) : \|v'(s) - l\| < R^{-1}\}} (L(s, u(s), l) + \xi \cdot (v'(s) - l) - 2\varepsilon/3) ds \\
 & \geq \int_{t_1}^{t_2} L(s, u(s), l) ds \\
 & \quad + \int_{t_1}^{t_2} \xi \cdot (v'(s) - l) ds + \tau_R \lambda(\{s \in (t_1, t_2) : \|v'(s) - l\| \geq R^{-1}\}) - \varepsilon(t_2 - t_1) \\
 & = \int_{t_1}^{t_2} L(s, u(s), u'(s)) ds + \tau_R \lambda(\{s \in (t_1, t_2) : \|v'(s) - l\| \geq R^{-1}\}) - \varepsilon(t_2 - t_1),
 \end{aligned}$$

recalling that $l = (t_2 - t_1)^{-1}(v(t_2) - v(t_1))$, and therefore that $\int_{t_1}^{t_2} \xi \cdot (v'(s) - l) ds = 0$. □

Armed with this tool, we may swiftly deduce some facts about the behaviour of the derivatives of minimizers. Assuming for a contradiction some bad behaviour of the derivative of a minimizer, each proof comes down to the ability to insert small affine segments into the trajectory, with slopes which differ significantly from the derivative of the minimizer. The construction of these affine segments is slightly easier when the range is one-dimensional, but the general case is not particularly difficult, so we content ourselves with attacking immediately the vector-valued case. The reader who believes that our proofs need not be quite as fussy in the one-dimensional case is quite right.

Theorem 24 *Let $v \in W^{1,1}((a, b); \mathbb{R}^n)$ be a minimizer of \mathcal{L} over $\mathcal{A}_{v(a), v(b)}$, and suppose for some $t \in [a, b]$ that, respectively, $t \in (a, b]$ and $D^-v(t)$ exists and each component is finite; or $t \in [a, b)$ and $D^+v(t)$ exists and each component is finite.*

Then v' is approximately continuous on the left, respectively right, at t .

Proof We consider the case in which $t \in [a, b)$; the other case is similar.

By translating our domain $[a, b]$, subtracting an affine function from v , and making the corresponding corrections to L , without loss of generality we may assume that $t = 0 \in [a, b)$, $v(t) = 0$, and $D^+v(t) = 0$.

Suppose for a contradiction that the result is false, so there exist $c, \alpha \in (0, 1)$ and arbitrarily small $s > 0$ such that

$$\lambda(\{r \in (0, s) : \|v'(r)\| \geq c\}) > \alpha s. \tag{64}$$

Let $\delta \in (0, b)$ be as given by Lemma 23 for $R \geq 2c^{-1}$ and $\varepsilon \leq \alpha \tau_R/8$. Let $s_0 \in (0, \delta/2)$ be such that (64) holds and $s \in (0, s_0)$ implies that

$$\frac{\|v(2s)\|}{2s} \leq c/8. \tag{65}$$

Consider $s \in (0, s_0)$ such that $\|v'(s)\| \geq c$. Then there exist s^\pm such that (s^-, s) and (s, s^+) are connected components of the set $\{r \in (0, 2s_0) : \|v(r) - v(s)\| > c|r - s|/2\}$. Note that (65) implies that $s^- > 0$, and that $s^+ < 2s_0$, since

$$\frac{\|v(2s_0) - v(s)\|}{|2s_0 - s|} \leq \frac{\|v(2s_0)\|}{s_0} + \frac{\|v(s)\|}{s_0} \leq 2 \frac{\|v(2s_0)\|}{2s_0} + \frac{\|v(s)\|}{s} \leq 3c/8.$$

By the Besicovitch covering theorem we may extract from the collection $\{(s^-, s^+) : s \in (0, s_0), \|v'(s)\| \geq c\}$ a pairwise disjoint subcollection $\mathcal{S} = \{(s_i^-, s_i^+)\}_{i=1}^\infty$, say, such that

$$\lambda\left(\bigcup \mathcal{S} \cap \{r \in (0, s_0) : \|v'(r)\| \geq c\}\right) \geq \lambda(\{r \in (0, s_0) : \|v'(r)\| \geq c\})/2 > \alpha s_0/2.$$

Define $u \in \mathcal{A}_{v(a),v(b)}$ by

$$u(r) := \begin{cases} v(r) & r \notin \bigcup_{i=1}^\infty (s_i^-, s_i) \cup (s_i, s_i^+), \\ \text{affine} & \text{otherwise.} \end{cases}$$

Let $i \geq 1$, and let $I_i^- := (s_i^-, s_i)$ and $I_i^+ := (s_i, s_i^+)$. By choice of s_i^\pm , on I_i^\pm we have that $\|u'\| = c/2$, and, furthermore, that $\|v'(r)\| \geq c$ implies that $\|v'(r) - u'\| \geq c/2$. Hence

$$\lambda(\{r \in I_i^\pm : \|v'(r) - u'\| \geq c/2\}) \geq \lambda(\{r \in I_i^\pm : \|v'(r)\| \geq c\}).$$

So Lemma 23 implies that

$$\int_{I_i^\pm} L(r, v, v') \, dr \geq \int_{I_i^\pm} L(r, u, u') \, dr + \tau_R \lambda(\{r \in I_i^\pm : \|v'(r)\| \geq c\}) - \alpha \tau_R \lambda(I_i^\pm) / 8,$$

and so summing, since \mathcal{J} is pairwise disjoint and $\bigcup \mathcal{J} \subseteq (0, 2s_0)$, gives that

$$\begin{aligned} & \int_{\bigcup \mathcal{J}} L(r, v, v') \, dr \\ & \geq \int_{\bigcup \mathcal{J}} L(r, u, u') \, dr + \tau_R \lambda\left(\bigcup \mathcal{J} \cap \{r \in (0, s_0) : \|v'(r)\| \geq c\}\right) - \alpha \tau_R \lambda\left(\bigcup \mathcal{J}\right) / 8 \\ & \geq \int_{\bigcup \mathcal{J}} L(r, u, u') \, dr + \alpha \tau_R s_0 / 2 - \alpha \tau_R s_0 / 4 \\ & = \int_{\bigcup \mathcal{J}} L(r, u, u') + \tau_R \alpha s_0 / 4, \end{aligned}$$

which is a contradiction. □

Theorem 25 *Let $v \in W^{1,1}((a, b); \mathbb{R}^n)$ be a minimizer of \mathcal{L} over $\mathcal{A}_{v(a),v(b)}$, and suppose for some $t \in [a, b]$ that, respectively, $t \in (a, b)$ and $D^-v(t)$ exists and at least one component is infinite; or $t \in [a, b)$ and $D^+v(t)$ exists and at least one component is infinite.*

Then v' is approximately continuous on the left, respectively right, at t , in the sense that for all $m > 0$,

$$\begin{aligned} \lim_{s \uparrow t} (t - s)^{-1} \lambda(\{r \in (s, t) : \|v'(r)\| \leq m\}) &= 0, \text{ respectively} \\ \lim_{s \downarrow t} (s - t)^{-1} \lambda(\{r \in (t, s) : \|v'(r)\| \leq m\}) &= 0. \end{aligned}$$

Proof We consider that case in which $t \in [a, b)$; the other case is similar.

Without loss of generality we may assume that $t = 0 \in [a, b)$ and $v(t) = 0$. We suppose for a contradiction that there exist $m \in (1, \infty)$, $\alpha \in (0, 1)$, and arbitrarily small $s > 0$ such that

$$\lambda(\{r \in (0, s) : \|v'(r)\| \leq m\}) > \alpha s. \tag{66}$$

Let $\delta \in (0, b)$ be as given by Lemma 23 for $R \geq 2m$, and $\varepsilon \leq \tau_R \alpha / 16$. Choose $s_0 \in (0, \delta/2)$ such that (66) holds and such that $s \in (0, s_0)$ satisfies

$$\frac{\|v(s)\|}{s} \geq 3m. \tag{67}$$

Consider $s \in (0, s_0)$ such that $\|v'(s)\| \leq m$. Then there exist s^\pm such that (s^-, s) and (s, s^+) are connected components of the set $\{r \in (0, s_0) : \|v(r) - v(s)\| < 2m|r - s|\}$. Note that $s^- > 0$ by (67). Define $\sigma_s := s - s^- > 0$. By the Besicovitch covering theorem we may

extract from the collection $\{(s - \sigma_s, s + \alpha\sigma_s/8) : s \in (0, s_0), \|v'(s)\| \leq m\}$ a pairwise disjoint subcollection $\mathcal{J} = \{(s_i - \sigma_i, s_i + \alpha\sigma_i/8)\}_{i=1}^\infty$, say, such that

$$\lambda\left(\bigcup \mathcal{J} \cap \{r \in (0, s_0) : \|v'(r)\| \leq m\}\right) > \alpha s_0/2.$$

For each $i \geq 1$, since $\alpha/8 < 1$ and $0 < \sigma_i < s_i$, we see that $s_i + \alpha\sigma_i/8 < 2s_i < 2s_0$, i.e. $\bigcup \mathcal{J} \subseteq (0, 2s_0)$. Since \mathcal{J} is pairwise disjoint, we see that

$$\sum_{i=1}^\infty \sigma_i = \lambda\left(\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i)\right) \leq 2s_0,$$

so

$$\begin{aligned} \lambda\left(\bigcup_{i=1}^\infty (s_i, s_i + \alpha\sigma_i/8) \cap \{r \in (0, s_0) : \|v'(r)\| \leq m\}\right) &\leq \lambda\left(\bigcup_{i=1}^\infty (s_i, s_i + \alpha\sigma_i/8)\right) \\ &= \alpha\left(\sum_{i=1}^\infty \sigma_i\right)/8 \\ &\leq \alpha s_0/4, \end{aligned}$$

and so

$$\lambda\left(\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i) \cap \{r \in (0, s_0) : \|v'(r)\| \leq m\}\right) \geq \alpha s_0/4. \tag{68}$$

Define $u \in \mathcal{A}_{v(a),v(b)}$ by

$$u(r) := \begin{cases} v(r) & r \notin \bigcup_{i=1}^\infty (s_i - \sigma_i, s_i), \\ \text{affine} & \text{otherwise.} \end{cases}$$

Fix $i \geq 1$. On $(s_i - \sigma_i, s_i)$, we have that $\|u'\| = 2m$, and, furthermore, that $\|v'(r)\| \leq m$ implies $\|v'(r) - u'\| \geq m$, so

$$\lambda(\{r \in (s_i - \sigma_i, s_i) : \|v'(r) - u'\| \geq m\}) \geq \lambda(\{r \in (s_i - \sigma_i, s_i) : \|v'(r)\| \leq m\}).$$

So Lemma 23 implies that

$$\begin{aligned} &\int_{s_i - \sigma_i}^{s_i} L(r, v, v') dr \\ &\geq \int_{s_i - \sigma_i}^{s_i} L(r, u, u') dr + \tau_R \lambda(\{r \in (s_i - \sigma_i, s_i) : \|v'(r)\| \leq m\}) - \alpha \tau_R \sigma_i/16, \end{aligned}$$

and so, summing, since \mathcal{J} is pairwise disjoint and $\bigcup \mathcal{J} \subseteq (0, 2s_0)$, gives, by (68), that

$$\begin{aligned} &\int_{\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i)} L(r, v, v') dr \\ &\geq \int_{\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i)} L(r, u, u') dr \\ &\quad + \tau_R \lambda\left(\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i) \cap \{r \in (0, s_0) : \|v'(r)\| \leq m\}\right) - \alpha \tau_R \lambda\left(\bigcup \mathcal{J}\right)/16 \end{aligned}$$

$$\begin{aligned} &\geq \int_{\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i)} L(r, u, u') \, dr + \alpha \tau_R s_0 / 4 - \alpha \tau_R s_0 / 8 \\ &= \int_{\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i)} L(r, u, u') \, dr + \alpha \tau_R s_0 / 8, \end{aligned}$$

which is a contradiction. □

Corollary 26 *Let $v \in W^{1,1}((a, b); \mathbb{R}^n)$ be a minimizer of \mathcal{L} over $\mathcal{A}_{v(a),v(b)}$, and $t \in (a, b)$ be such that $D^\pm v(t)$ both exist, and each component of both is finite.*

Then $D^-v(t) = D^+v(t)$.

Proof Without loss of generality we may assume that $t = 0 \in (a, b)$ and $v(t) = 0$. We assume for a contradiction that $D^-v(0) \neq D^+v(0)$, so we may further suppose without loss of generality that $-D^-v_1(0) = D^+v_1(0) = m > 0$, say.

Let $\delta \in (0, \min\{|a|, |b|\})$ be as given by Lemma 23 for $R \geq \|D^+v(0)\| + \|D^-v(0)\| + m + 2m^{-1} + 1$, and $\varepsilon \leq \tau_R/4$.

By Theorem 24 we can choose $s_0 \in (0, \delta/3)$ such that $s \in (0, s_0)$ satisfies

$$\|(3s)^{-1}v(3s) - D^+v(0)\| < m/2 \text{ and } \|(-3s)^{-1}v(-3s) - D^-v(0)\| < m/2; \text{ and} \tag{69}$$

$$\begin{cases} \lambda(\{r \in (0, s) : \|v'(r) - D^+v(0)\| \geq m/2\}) < s/2, \\ \lambda(\{r \in (-s, 0) : \|v'(r) - D^-v(0)\| \geq m/2\}) < s/2. \end{cases} \tag{70}$$

Fix $s_1 \in (0, s_0)$. Now, (69) implies that $\left| \frac{v_1(-s_1)}{-s_1} + m \right| < m/2$, and hence that $ms_1/2 < v_1(-s_1) < 3ms_1/2$, and, furthermore, that $\left| \frac{v_1(3s_1)}{3s_1} - m \right| < m/2$, and hence that $v_1(3s_1) > 3ms_1/2$. So $v_1(0) = 0 < ms_1/2 < v_1(-s_1) < 3ms_1/2 < v_1(3s_1)$, and therefore there exists $s_2 \in (0, 3s_1)$ such that $v_1(s_2) = v_1(-s_1)$. Define $u \in \mathcal{A}_{v(a),v(b)}$ by

$$u(r) := \begin{cases} v(r) & r \notin (-s_1, s_2), \\ \text{affine} & \text{otherwise.} \end{cases}$$

Then on $(-s_2, s_1)$,

$$\begin{aligned} \|u'\| &\leq \frac{\|v(s_2)\|}{s_2 + s_1} + \frac{\|v(-s_1)\|}{s_2 + s_1} \leq \frac{\|v(s_2)\|}{s_2} + \frac{\|v(-s_1)\|}{s_1} \\ &\leq \|D^+v(0)\| + m/2 + \|D^-v(0)\| + m/2 \\ &\leq R. \end{aligned}$$

For $r \in (0, s_2)$, we see that $\|v'(r) - D^+v(0)\| \leq m/2$ implies that $\|v'(r) - u'\| \geq |v'_1(r) - u'_1(r)| = |v'_1(r)| > m/2$, hence, by (70), that

$$\begin{aligned} \lambda(\{r \in (0, s_2) : \|v'(r) - u'\| \geq m/2\}) &\geq \lambda(\{r \in (0, s_2) : \|v'(r) - D^+v(0)\| \leq m/2\}) \\ &\geq s_2/2. \end{aligned}$$

Similarly,

$$\begin{aligned} \lambda(\{r \in (-s_1, 0) : \|v'(r) - u'\| \geq m/2\}) &\geq \lambda(\{r \in (-s_1, 0) : \|v'(r) - D^-v(0)\| \leq m/2\}) \\ &\geq s_1/2. \end{aligned}$$

So Lemma 23 implies that

$$\begin{aligned} & \int_{-s_1}^{s_2} L(r, v, v') \, dr \\ & \geq \int_{-s_1}^{s_2} L(r, u, u') \, dr + \tau_R \lambda (\{r \in (-s_1, s_2) : \|v'(r) - u'\| \geq m/2\}) - \tau_R(s_1 + s_2)/4 \\ & \geq \int_{-s_1}^{s_2} L(r, u, u') \, dr + \tau_R(s_1 + s_2)/2 - \tau_R(s_1 + s_2)/4 \\ & = \int_{-s_1}^{s_2} L(r, u, u') \, dr + \tau_R(s_1 + s_2)/4, \end{aligned}$$

which is a contradiction. □

Corollary 27 *Let $v \in W^{1,1}((a, b); \mathbb{R}^n)$ be a minimizer of \mathcal{L} over $\mathcal{A}_{v(a),v(b)}$, and suppose that $t \in (a, b)$ is such that $D^\pm v(t)$ both exist.*

Then if at least one component of one one-sided derivative is infinite, then at least one component of the other one-sided derivative is infinite. Note that this statement does not assert that there is one coordinate function with infinite left and right derivatives, and in fact I do not know whether such an assertion can be made in general.

Proof We suppose for a contradiction that the result is false. We consider the case in which at least one component of $D^+ v(t)$ is infinite, but all components of $D^- v(t)$ are finite; the other case is similar.

Without loss of generality we may assume that $t = 0 \in (a, b)$, $v(t) = 0$, and $D^- v(0) = 0$.

Let $\delta \in (0, \min\{|a|, |b|\})$ be as given by Lemma 23 for $R \geq 2$ and $\varepsilon \leq \tau_R/4$. By Theorems 24 and 25 we can choose $s_0 \in (0, \delta)$ such that $s \in (0, s_0)$ satisfies

$$\frac{\|v(-s)\|}{s} < 1/2, \text{ and } \frac{\|v(s)\|}{s} > 3; \text{ and} \tag{71}$$

$$\begin{cases} \lambda (\{r \in (-s, 0) : \|v'(r)\| \geq 1/2\}) < s/2, \\ \lambda (\{r \in (0, s) : \|v'(r)\| \leq 2\}) < s/2. \end{cases} \tag{72}$$

Fix $s_1 \in (0, s_0)$, and consider the set $\{r \in (-s_1, s_0) : \|v(r) - v(-s_1)\| < |r + s_1|\}$. By (71) we know that 0 lies in this set, but since

$$\|v(s_1) - v(-s_1)\| \geq \|v(s_1)\| - \|v(-s_1)\| \geq 3s_1 - s_1/2 > 2s_1,$$

we see that s_1 does not. Therefore there exists $s_2 \in (0, s_1)$ such that $\|v(s_2) - v(-s_1)\| = |s_2 + s_1|$. Define $u \in \mathcal{A}_{v(a),v(b)}$ by

$$u(r) := \begin{cases} v(r) & r \notin (-s_1, s_2), \\ \text{affine} & \text{otherwise.} \end{cases}$$

Then on $(-s_1, s_2)$, we have that $\|u'\| = 1$, and, furthermore, that $\|v'(r)\| \leq 1/2$ implies that $\|v'(r) - u'\| \geq 1/2$, and $\|v'(r)\| \geq 2$ implies that $\|v'(r) - u'\| \geq 1/2$, so by (72),

$$\lambda (\{r \in (-s_1, 0) : \|v'(r) - u'\| \geq 1/2\}) \geq \lambda (\{r \in (-s_1, 0) : \|v'(r)\| \leq 1/2\}) \geq s_1/2,$$

and

$$\lambda (\{r \in (0, s_2) : \|v'(r) - u'\| \geq 1/2\}) \geq \lambda (\{r \in (0, s_2) : \|v'(r)\| \geq 2\}) \geq s_2/2.$$

So Lemma 23 implies that

$$\begin{aligned} & \int_{-s_1}^{s_2} L(r, v, v') \, dr \\ & \geq \int_{-s_1}^{s_2} L(r, u, u') \, dr + \tau_R \lambda (\{r \in (-s_1, s_2) : \|v'(r) - u'\| \geq 1/2\}) + \tau_R (s_1 + s_2)/4 \\ & \geq \int_{-s_1}^{s_2} L(r, u, u') \, dr + \tau_R (s_1 + s_2)/2 - \tau_R (s_1 + s_2)/4 \\ & = \int_{-s_1}^{s_2} L(r, u, u') \, dr + \tau_R (s_1 + s_2)/4, \end{aligned}$$

which is a contradiction. □

More information is available about the behaviour of infinite derivatives if we can locate them only in one coordinate function.

Theorem 28 *Let $v \in W^{1,1}((a, b); \mathbb{R}^n)$ be a minimizer of \mathcal{L} over $\mathcal{A}_{v(a),v(b)}$, and suppose for some $t \in [a, b]$ that, respectively, $t \in (a, b]$, $D^-v_1(t)$ exists as an infinite value, and v_j are Lipschitz in a left-neighbourhood of t for $2 \leq j \leq n$; or that $t \in [a, b)$, $D^+v_1(t)$ exists as an infinite value, and v_j are Lipschitz in a right-neighbourhood of t for $2 \leq j \leq n$.*

Then v'_1 is approximately continuous on the left, respectively right, at t , in the sense that for all $m > 0$,

$$\begin{aligned} & \lim_{s \uparrow t} (t - s)^{-1} \lambda (\{r \in (s, t) : |v'_1(r)| \leq m\}) = 0, \text{ respectively} \\ & \lim_{s \downarrow t} (s - t)^{-1} \lambda (\{r \in (t, s) : |v'_1(r)| \leq m\}) = 0. \end{aligned}$$

Note that we do not assume that the derivatives of the components v_j for $2 \leq j \leq n$ exist at t .

Proof We consider the case in which $t \in [a, b)$; the other case is similar.

Without loss of generality we may assume that $t = 0 \in [a, b)$ and $v(t) = 0$. Suppose for a contradiction that there exist $m \in (1, \infty)$, $\alpha \in (0, 1)$, and arbitrarily small $s > 0$ such that

$$\lambda (\{r \in (0, s) : |v'_1(r)| \leq m\}) > \alpha s. \tag{73}$$

Choose $\eta \in (0, b)$ such that v_j are Lipschitz on $[0, \eta]$ for $2 \leq j \leq n$, and let $\delta \in (0, b)$ be as given by Lemma 23 for $R \geq 2m + \sum_{j=2}^n \text{Lip}(v_j|_{[0,\eta]})$, and $\varepsilon \leq \alpha \tau_R/16$.

Choose $s_0 \in (0, \min\{\delta, \eta\}/2)$ such that (73) holds and $s \in (0, s_0)$ satisfies

$$|v_1(2s)|/2s > 2R. \tag{74}$$

Consider $s \in (0, s_0)$ such that $|v'_1(s)| \leq m$. Then $|v_1(r) - v_1(s)|/|r - s| < 2m$ for r in some neighbourhood of s contained in $(0, s_0)$, and so

$$\frac{\|v(r) - v(s)\|}{|r - s|} < 2m + \sum_{j=2}^n \text{Lip}(v_j|_{[0,\eta]}) \leq R,$$

for r in some neighbourhood of s contained in $(0, s_0)$. So there exist s^\pm such that (s^-, s) and (s, s^+) are connected components of the set $\{r \in (0, s_0) : \|v(r) - v(s)\| < R|r - s|\}$. Note that $s^- > 0$ by (74). We now proceed similarly to the proof of Theorem 25. Let $\sigma_s := s - s^- > 0$. By the Besicovitch covering theorem we may extract from the collection $\{(s - \sigma_s, s + \alpha \sigma_s/8) :$

$s \in (0, s_0)$, $|v'_1(s)| \leq m$ a pairwise disjoint collection $\mathcal{S} = \{(s_i - \sigma_i, s_i + \alpha\sigma_i/8)\}_{i=1}^\infty$, say, such that $\bigcup \mathcal{S} \subseteq (0, 2s_0)$, and

$$\lambda \left(\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i) \cap \{r \in (0, s_0) : |v'_1(r)| \leq m\} \right) \geq \alpha s_0/4. \tag{75}$$

Define $u \in \mathcal{A}_{v(a),v(b)}$ by

$$u(r) := \begin{cases} v(r) & r \notin \bigcup_{i=1}^\infty (s_i - \sigma_i, s_i), \\ \text{affine} & \text{otherwise.} \end{cases}$$

Fix $1 \geq \epsilon$. On $(s_i - \sigma_i, s_i)$, we have that $\|u'\| = R$, and, furthermore, that $|v'_1(r)| \leq m$ implies that $\|v'(r)\| \leq m + \sum_{j=2}^n \text{Lip}(v_j|_{[0,\eta]})$, which implies that $\|v'(r) - u'\| \geq \epsilon$. So

$$\lambda(\{r \in (s_i - \sigma_i, s_i) : \|v'(r) - u'\| \geq \epsilon\}) \geq \lambda(\{r \in (s_i - \sigma_i, s_i) : |v'_1(r)| \leq m\}).$$

So Lemma 23 implies that

$$\begin{aligned} & \int_{s_i - \sigma_i}^{s_i} L(r, v, v') \, dr \\ & \geq \int_{s_i - \sigma_i}^{s_i} L(r, u, u') \, dr + \tau_R \lambda(\{r \in (s_i - \sigma_i, s_i) : |v'_1(r)| \leq m\}) - \alpha \tau_R \sigma_i/16, \end{aligned}$$

and so, summing, since \mathcal{S} is pairwise disjoint and $\bigcup \mathcal{S} \subseteq (0, 2s_0)$, gives, by (75), that

$$\begin{aligned} & \int_{\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i)} L(r, v, v') \, dr \\ & \geq \int_{\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i)} L(r, u, u') \, dr \\ & \quad + \tau_R \lambda \left(\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i) \cap \{r \in (0, s_0) : |v'_1(r)| \leq m\} \right) - \alpha \tau_R \lambda(\bigcup \mathcal{S})/16 \\ & \geq \int_{\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i)} L(r, u, u') \, dr + \alpha \tau_R s_0/4 - \alpha \tau_R s_0/8 \\ & = \int_{\bigcup_{i=1}^\infty (s_i - \sigma_i, s_i)} L(r, u, u') \, dr + \alpha \tau_R s_0/8, \end{aligned}$$

which is a contradiction. □

Corollary 29 *Let $v \in W^{1,1}((a, b); \mathbb{R}^n)$ be a minimizer of \mathcal{L} over $\mathcal{A}_{v(a),v(b)}$, and suppose that $t \in (a, b)$ is such that $D^\pm v(t)$ both exist, and v_j is Lipschitz in a neighbourhood of t for $2 \leq j \leq n$.*

Then if one one-sided derivative of v_1 is infinite at t , then the two one-sided derivatives of v_1 are equal to the same infinite value.

Proof We suppose that $D^+ v_1(t) = +\infty$; the other cases are similar.

Without loss of generality we may assume that $t = 0 \in (a, b)$ and $v(t) = 0$. By Corollary 27, we are just required to prove that $D^- v_1(0) > -\infty$. Suppose for a contradiction that $D^- v_1(0) = -\infty$. Choose $\eta \in (0, \min\{|a|, |b|\})$ such that v_j are Lipschitz on $(-\eta, \eta)$ for $2 \leq j \leq n$. Let $\delta \in (0, \min\{|a|, |b|\})$ be as given by Lemma 23 for $R \geq 1 + \sum_{j=2}^n \text{Lip}(v_j|_{(-\eta,\eta)})$

and $\varepsilon \leq \tau_R/4$. By Theorem 25 we may choose $s_0 \in (0, \min\{\eta, \delta\})$ such that $s \in (0, s_0)$ satisfies

$$\frac{v_1(s)}{s} > 1 \text{ and } \frac{v_1(-s)}{-s} < -1; \text{ and} \tag{76}$$

$$\begin{cases} \lambda(\{r \in (-s, 0) : |v'_1(r)| \leq 3R\}) < s/2, \\ \lambda(\{r \in (0, s) : |v'_1(r)| \leq 3R\}) < s/2. \end{cases} \tag{77}$$

By (76), $v_1(s_0) \geq s_0 > 0 = v_1(0)$, so we can choose $s_1 \in (0, s_0)$ such that $v_1(-s_1) < v_1(s_0)$. Since, by (76), $v_1(0) = 0 < s_1 < v_1(-s_1) < v_1(s_0)$, there exists $s_2 \in (0, s_0)$ such that $v_1(-s_1) = v_1(s_2)$. Then $\|v(-s_1) - v(s_2)\| \leq \sum_{j=2}^n \text{Lip}(v_j|_{(-\eta, \eta)})|s_2 + s_1| \leq R|s_2 + s_1|$. Define $u \in \mathcal{A}_{v(a), v(b)}$ by

$$u(r) := \begin{cases} v(r) & r \notin (-s_1, s_2), \\ \text{affine} & \text{otherwise.} \end{cases}$$

So on $(-s_1, s_2)$ we have that $\|u'\| \leq R$, and, furthermore, that $|v'_1(r)| \geq 3R$ implies that $\|v'(r)\| \geq 2R$ and hence that $\|v'(r) - u'\| \geq R$, and so, by (77),

$$\begin{aligned} \lambda(\{r \in (-s_1, s_2) : \|v'(r) - u'\| \geq R\}) &\geq \lambda(\{r \in (-s_1, s_2) : |v'_1(r)| \geq 3R\}) \\ &\geq (s_1 + s_2)/2. \end{aligned}$$

So Lemma 23 implies that

$$\begin{aligned} &\int_{-s_1}^{s_2} L(r, v, v') \, dr \\ &\geq \int_{-s_1}^{s_2} L(r, u, u') \, dr + \tau_R \lambda(\{r \in (-s_1, s_2) : \|v'(r) - u'\| \geq R\}) - \tau_R(s_1 + s_2)/4 \\ &\geq \int_{-s_1}^{s_2} L(r, u, u') \, dr + \tau_R(s_1 + s_2)/2 - \tau_R(s_1 + s_2)/4 \\ &= \int_{-s_1}^{s_2} L(r, u, u') \, dr + \tau_R(s_1 + s_2)/4, \end{aligned}$$

which is a contradiction. □

It is a rather delicate matter to investigate the behaviour of derivatives of individual coordinate functions once we admit infinite derivatives. It seems that the presence of arbitrarily steep tangents in one variable can mask a multitude of sins in the others. The following example demonstrates one such case: a vertical tangent in the second variable allows a cusp point in the first variable.

Example 30 Let $v \in W^{1,1}((-1, 1), \mathbb{R}^2)$ be defined by $v(t) = (|t|, \text{sign}(t)|t|^{1/3})$, and let $\tilde{L} : [-1, 1] \times \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow [0, \infty)$ be given by

$$\tilde{L}(t, y, p) = ((y_1 - v_1(t))^2 + (y_2^3 - t)^2)p_2^6.$$

Let $\eta > 0$ be the constant from the usual Manià estimates (see for example [3]), i.e. such that

$$\int_0^1 (u(t)^3 - t)^2 (u'(t))^6 \, dt \geq \eta,$$

for any $u \in W^{1,1}(0, 1)$ such that $u(1) = 1$ and $u(t) < t^{1/3}/4$ for some $t > 0$.

Suppose $u \in \mathcal{A}_{v(-1),v(1)}$. If $u_2(t) < t^{1/3}/4$ for some $t \in (0, 1)$, then $\int_{-1}^1 \tilde{L}(s, u, u') ds \geq \eta$; similarly if $u_2(t) > -t^{1/3}/4$ for some $t \in (-1, 0)$, then $\int_{-1}^1 \tilde{L}(s, u, u') ds \geq \eta$.

Suppose now that $u_2(t) \geq t^{1/3}/4$ on $(0, 1]$ and $u_2(t) \leq -t^{1/3}/4$ on $[-1, 0)$. Suppose further that $u_1(0) \neq v_1(0) = 0$. So there exists some $t \in (0, 1)$ such that $|u_1(s) - v_1(s)| > t$ on $(-t, t)$. Then by Jensen’s inequality,

$$\begin{aligned} \int_{-1}^1 \tilde{L}(s, u, u') ds &\geq \int_{-t}^t (u_1(s) - v_1(s))^2 (u_2'(s))^6 ds \geq t^2 \int_{-t}^t (u_2'(s))^6 ds \\ &\geq 2t^3 \left(\frac{u_2(t) - u_2(-t)}{2t} \right)^6 \\ &\geq 2t^3 \left(\frac{t^{1/3} + t^{1/3}}{8t} \right)^6 \\ &\geq 2^{-11}. \end{aligned}$$

So if u is such that $\int_{-1}^1 \tilde{L}(s, u, u') ds < \eta$, then $u_1(0) \neq v_1(0)$ implies that $\int_{-1}^1 \tilde{L}(s, u, u') ds \geq 2^{-11}$. So if u is such that $u_1(0) \neq v_1(0)$, then

$$\int_{-1}^1 \tilde{L}(s, u, u') ds \geq \min\{\eta, 2^{-11}\}.$$

Choose $\sigma \in (1, 3/2)$ and $\varepsilon < \min\{\eta, 2^{-11}\} \left(\int_{-1}^1 (|v_1'|^2 + |v_2'|^\sigma) \right)^{-1}$, and define $L : [-1, 1] \times \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow [0, \infty)$ by

$$L(t, y, p) = ((y_1 - v_1(t))^2 + (y_2^3 - t)^2) p_2^6 + \varepsilon(p_1^2 + p_2^\sigma).$$

Let $u \in \mathcal{A}_{v(-1),v(1)}$ be a minimizer of \mathcal{L} , so by the above argument and the choice of ε , $\mathcal{L}(u) \leq \mathcal{L}(v) < \min\{\eta, 2^{-11}\}$, so $u_1(0) = v_1(0)$. Suppose I is a non-trivial component of $\{s \in (-1, 1) : u_1(s) \neq v_1(s)\}$, then $0 \notin I$, and so v_1 is linear on I , and therefore is the unique minimizer of the functional $v_1 \mapsto \int_I (v_1')^2$. Define $\hat{u} \in \mathcal{A}_{v(-1),v(1)}$ by

$$\hat{u}(r) := \begin{cases} u(r) & r \notin I, \\ (v_1(r), u_2(r)) & r \in I. \end{cases}$$

Then

$$\begin{aligned} \int_I L(r, u, u') dr &= \int_I ((u_1 - v_1)^2 + (u_2^3 - t)^2) (u_2')^6 + \varepsilon((u_1')^2 + (u_2')^\sigma) dr \\ &> \int_I (u_2^3 - t)^2 (u_2')^6 + \varepsilon((v_1')^2 + (u_2')^\sigma) dr \\ &= \int_I L(r, \hat{u}, \hat{u}') dr, \end{aligned}$$

which is a contradiction. Hence $u_1 = v_1$, and in particular $D^\pm u_1(0) = \pm 1$.

Acknowledgements I wish to thank John Ball for first posing and discussing with me the original question of approximation (question 2 in Sect. 3), which proved surprisingly diverting and served as the starting point for the current paper. I am also grateful to David Preiss for encouragement, advice, and conversation on these topics. Finally, I am indebted to an anonymous referee for comments on an earlier version of this article, in particular for suggesting and motivating significant improvements to the presentation.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Ball, J.M.: Singularities and computation of minimizers for variational problems. In: Foundations of Computational Mathematics (Oxford, 1999), London Mathematical Society Lecture Note Series, vol. 284, pp. 1–20. Cambridge University Press, Cambridge (2001)
2. Ball, J.M., Mizel, V.J.: One-dimensional variational problems whose minimizers do not satisfy the Euler–Lagrange equation. *Arch. Ration. Mech. Anal.* **90**(4), 325–388 (1985)
3. Buttazzo, G., Giaquinta, M., Hildebrandt, S.: One-dimensional variational problems. An introduction, Oxford Lecture Series in Mathematics and its Applications, vol. 15. The Clarendon Press Oxford University Press, New York (1998)
4. Clarke, F.H., Vinter, R.B.: Regularity properties of solutions to the basic problem in the calculus of variations. *Trans. Am. Math. Soc.* **289**(1), 73–98 (1985)
5. Csörnyei, M., Kirchheim, B., O’Neil, T.C., Preiss, D., Winter, S.: Universal singular sets in the calculus of variations. *Arch. Ration. Mech. Anal.* **190**(3), 371–424 (2008)
6. Esposito, L., Leonetti, F., Mingione, G.: Sharp regularity for functionals with (p, q) growth. *J. Differ. Equ.* **204**(1), 5–55 (2004)
7. Evans, L., Gariepy, R.: *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics. CRC Press, Boca Raton (1992)
8. Ferriero, A.: A direct proof of the Tonelli’s partial regularity result. *Discrete Contin. Dyn. Syst.* **32**(6), 2089–2099 (2012)
9. Ferriero, A.: On the Tonelli’s partial regularity. *Differ. Integral Equ.* **26**(1–2), 1–9 (2013)
10. Gratwick, R.: Singular minimizers in the calculus of variations. Ph.D. thesis, University of Warwick (2011)
11. Gratwick, R., Preiss, D.: A one-dimensional variational problem with continuous Lagrangian and singular minimizer. *Arch. Ration. Mech. Anal.* **202**(1), 177–211 (2011)
12. Lavrentiev, M.: Sur quelques problèmes du calcul des variations. *Ann. Mat. Pura Appl.* **4**, 7–28 (1926)
13. Manià, B.: Sopra un esempio di Lavrentieff. *Bull. Un. Mat. Ital.* **13**, 147–153 (1934)
14. Sychëv, M.A.: On the regularity of solutions of variational problems. *Russ. Acad. Sci. Sb. Math.* **75**(2), 535–556 (1993)
15. Tonelli, L.: *Fondamentica di Calcolo delle variazioni*, vol. 2. Zanichelli, Bologna (1923)
16. Tonelli, L.: Su gli integrali del calcolo delle variazioni in forma ordinaria. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (2)* **3**(3–4), 401–450 (1934)