

## A REGULARITY THEORY FOR VARIATIONAL PROBLEMS WITH HIGHER ORDER DERIVATIVES

F. H. CLARKE AND R. B. VINTER

**ABSTRACT.** We consider problems in the calculus of variations in one independent variable and where the Lagrangian involves derivatives up to order  $N$ ,  $N \geq 1$ . Existence theory supplies mild hypotheses under which there are minimizers for such problems, but they need to be strengthened for standard necessary conditions to apply.

For problems with  $N > 1$ , this paper initiates investigation of regularity properties, and associated necessary conditions, which obtain strictly under the hypotheses of existence theory. It is shown that the  $N$ th derivative of a minimizer is locally essentially bounded off a closed set of zero measure, the set of "points of bad behaviour". Additional hypotheses are shown to exclude occurrence of points of bad behaviour. Finally a counter example suggests respects in which problems with  $N > 1$  exhibit pathologies not present in the  $N = 1$  case.

### 1. INTRODUCTION

The basic problem in the calculus of variations is that of minimizing an integral functional

$$\int_a^b L(t, x(t), \dot{x}(t)) dt$$

over a suitable class of functions  $x$  with fixed endpoints. Two major issues in the theory are existence of minimizers, and necessary conditions to identify the minimizers. The foundations of a general existence theory were laid by Tonelli [11], who showed that existence of minimizers is guaranteed in the class of absolutely continuous functions, under weak and verifiable hypotheses, which for the moment we label (HE). A centrepiece of the theory of necessary conditions is the Pontryagin maximum principle of optimal control theory, a principle which subsumes the main classical necessary conditions of the calculus of variations.

Examination of the hypotheses (HE) of Tonelli's existence theory and those under which the maximum principle has been derived, or even makes sense, reveals a serious mismatch. We find there is a substantial class of problems

---

Received by the editors September 20, 1988.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 49B05.

The support of the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged. On remercie le FCAR (Québec) de son appui.

where existence theory predicts minimizers but where the hypotheses are not met under which we are permitted to identify them by means of such conditions as the maximum principle. Put another way, the difficulty is as follows. If we adopt hypotheses (HE), then knowledge merely that minimizers are absolutely continuous is inadequate for deriving the necessary conditions we should like to have. This suggests that we seek to overcome the difficulty by establishing that, while all absolutely continuous functions satisfying the end point conditions are considered in our search for minimizers, the minimizers are in fact confined to a subclass of arcs, which is more tractable from the point of view of deriving necessary conditions.

Tonelli was the first to prove significant results in this spirit [11, 12]. Tonelli regularity theory (as it is now called) establishes that, if we supplement (HE) by the conditions

- (i) the arcs are scalar valued functions,
- (ii)  $L$  is  $C^2$ , and
- (iii)  $L_{\dot{x}\dot{x}} > 0$  (strictly positive),

then the minimizers  $x$  have the property that  $\dot{x}$  is locally essentially bounded on an open subset  $\Omega \subset [a, b]$  of full measure. This extra information about minimizers suffices for derivation of necessary conditions, which have the character of the Euler equation, but are somewhat weaker than those implicit in the maximum principle. Tonelli regularity theory and its ramifications have been the subject of much research in recent years [4–8]. Notably, it was proved in [5] that there exists a set  $\Omega$  as described above, even when the supplementary hypotheses (i)–(iii) are dropped altogether.

In the present paper, our object is to develop a regularity theory, and associated necessary conditions, for problems involving higher order derivatives, for instance those where we seek minimizers for an integral functional of the form

$$\int_a^b L(t, x(t), \dot{x}(t), \ddot{x}(t)) dt.$$

Problems of this kind arise in the theory of beams and rods, for example, and have a long history (see [10] and references cited therein). Such problems are significant also because important cases of the optimal control problem, namely those with linear dynamics, can be reformulated as problems in the calculus of variations with higher derivatives, and regularity theory for the optimal control problem, outside very special situations where it can be reduced to the basic problem in the calculus of variations, is, to date, a completely undeveloped area of research.

At this stage we need to introduce some notation.  $W^{N,1}(I; \mathbf{R}^n)$  denotes the space of  $(N - 1)$  times continuously differentiable  $n$  vector valued functions on the closed interval  $I$  whose  $(N - 1)$ th derivatives are absolutely continuous in the case  $N \geq 2$ , and it denotes the space of absolutely continuous  $n$  vector valued functions on  $I$  in the case  $N = 1$ . (When we say a function is continuously differentiable on a closed interval, incidentally, we mean it is

continuously differentiable in the interior, and the derivatives have limits at the endpoints.) Given an arc  $y$ , we denote by  $D^0y, D^1y, \dots$  the derivatives of order  $0, 1, \dots$ .  $D^0y$  is of course just  $y$ . It is convenient often to write  $D^1y$  as  $Dy$ . For brevity we sometimes revert to standard notation  $Dy = \dot{y}, D^2y = \ddot{y}$  etc. Finally, the  $n \times (k - j + 1)$  matrix valued function  $D_j^k y$  is defined to be

$$D_j^k y(t) = \text{col}(D^j y(t), D^{j+1} y(t), \dots, D^k y(t)).$$

The following data is given: positive integers  $N$  and  $n$ , real numbers  $a$  and  $b, b > a$ , vectors  $A_0, \dots, A_{N-1}$  and  $B_0, \dots, B_{N-1}$  in  $\mathbf{R}^n$  and a function  $L: [a, b] \times \mathbf{R}^{n \times N} \times \mathbf{R}^n \rightarrow \mathbf{R}$ . We define

$$A := \text{col}(A_0, \dots, A_{N-1}) \quad \text{and} \quad B := \text{col}(B_0, \dots, B_{N-1}).$$

The following problem provides the framework for our study of problems in the calculus of variations with higher order derivatives. We label it

$$(P) \quad \begin{cases} \text{Minimize } J(x) := \int_a^b L(t, D_0^{N-1} y(t), D^N y(t)) dt \\ \text{over arcs } y \in W^{N,1}([a, b]; \mathbf{R}^n) \text{ which satisfy} \\ D_0^{N-1} y(a) = A \quad \text{and} \quad D_0^{N-1} y(b) = B. \end{cases}$$

Dependence of the Lagrangian on the highest order derivative will have special status in the statement of hypotheses and ensuing analysis and, cumbersome though it appears at present, we shall be grateful for our notation which groups the lower derivatives as a single argument.

It is assumed throughout that the following hypotheses on  $L(t, z, w)$  are in force.

- (H1)  $L(t, z, w)$  is locally bounded, measurable in  $t$  and convex in  $w$ .
- (H2)  $L(t, z, w)$  is locally Lipschitz continuous in  $(z, w)$ , uniformly in  $t$ .

This means that for each bounded subset  $C$  of  $\mathbf{R}^{n \times N} \times \mathbf{R}^n$  there exists a constant  $K$  such that for all  $t \in [a, b]$  and  $(z_1, w_1), (z_2, w_2) \in C$  the following inequality holds

$$|L(t, z_1, w_1) - L(t, z_2, w_2)| \leq K|(z_1 - z_2, w_1 - w_2)|.$$

- (H3)  $L(t, z, w)$  is coercive in  $w$ , in the sense that there is a number  $\alpha > 0$  and a convex function  $\theta: [0, \infty) \rightarrow \mathbf{R}$  such that

$$(1.1) \quad L(t, z, w) \geq -\alpha|z| + \theta(|w|)$$

for all  $(t, z, w) \in [a, b] \times \mathbf{R}^{n \times N} \times \mathbf{R}^n$ , where  $\theta(r)/r \rightarrow \infty$  as  $r \rightarrow \infty$ .

These mild hypotheses are representative of hypotheses under which existence of solutions to problem (P) has been proved. They impose conditions on  $L$  regarding its dependence on the highest derivative variable which are precisely those regarding dependence on the velocity variable of the Lagrangian function in the earlier, one derivative, theory [5]. In fact in the case  $N = 1$ , the hypotheses are in all respects those of [5].

In view of the pathologies we might anticipate when we pass from problems in one derivative to problems in higher derivatives (see §4), it is unexpected that, under hypotheses (H1)–(H3) alone, bad behaviour of minimizers  $x$  can still be confined to the complement of an open set  $\Omega$  of full measure, now in the sense that  $D^n x$  is locally essentially bounded on  $\Omega$ . This is the main result of the paper.

Some comments on our proof techniques are now in order. These are based on construction of an auxiliary problem with Lagrangian  $\tilde{L}$ . The idea is to replace  $L$  in some local sense by  $\tilde{L}$ ; the special structure of  $\tilde{L}$  permits us to apply necessary conditions in [3] to minimizers for the auxiliary problem and thereby to establish the desired regularity properties of minimizers for (P). In general terms the techniques are similar to those for one derivative problems, previously treated in [5]. However the necessary conditions on minimizers for the auxiliary problem are of a more intricate nature for problems with higher derivatives, and this fact, together with the need to consider general polynomial interpolation, leads to serious technical difficulties not present in the  $N = 1$  case.

Nonsmooth analysis as developed in [3] enters into this paper in a fundamental way. Even when we restrict attention to problems where  $L$  is smooth, the auxiliary Lagrangian is not smooth, and we require the full power of nonsmooth optimization techniques to analyse it.

## 2. THE REGULARITY THEOREM

The regularity theorem involves a new notion of “regular point”  $\tau$  of an arc  $x \in W^{N,1}([a, b]; \mathbf{R}^n)$ . The definition of regular point is rather complicated, and we first look at special cases as an aid to understanding. In the case  $N = 1$ ,  $\tau$  is a regular point if there exist sequences  $\{t_i\}, \{s_i\}$  such that  $a \leq s_i \leq \tau \leq t_i \leq b$ ,  $s_i \neq t_i$ , for all  $i$ , and

$$\liminf_{i \rightarrow \infty} \frac{|x(t_i) - x(s_i)|}{|t_i - s_i|} < \infty.$$

In the case  $N = 2$ , the last inequality is replaced by

$$\liminf_{i \rightarrow \infty} \frac{|x(t_i) - x(s_i) - Dx(s_i)(t_i - s_i)|}{|t_i - s_i|^2} + \frac{|Dx(t_i) - Dx(s_i)|}{|t_i - s_i|} < \infty.$$

Definition in general requires introduction of the function

$$\eta: \mathbf{R}^{n \times N} \times \mathbf{R}^{n \times N} \times \mathbf{R} \rightarrow \mathbf{R},$$

$$\begin{aligned} &\eta((X_0, \dots, X_{N-1}), (Y_0, \dots, Y_{N-1}), \sigma) \\ &:= \sum_{i=0}^{N-1} \left| \left[ Y_i - \sum_{j=i}^{N-1} X_j \frac{\sigma^{j-i}}{(j-i)!} \right] / \sigma^{N-i} \right|. \end{aligned}$$

Let  $x \in W^{N,1}([a, b]; \mathbf{R}^n)$  and  $\tau \in [a, b]$ . We say that  $\tau$  is a *regular point* of  $x$  if

$$\liminf_{\substack{s, t \rightarrow \tau \\ a \leq s \leq \tau \leq t \leq b \\ s \neq t}} \eta(D_0^{N-1}x(s), D_0^{N-1}x(t), t - s) < \infty.$$

The stipulation that  $\tau$  be regular may seem rather restrictive. But, on the contrary, as is shown in the proof of Corollary 2.2, all Lebesgue points of  $D^N x$  are regular and so, in particular, the regular points have full measure.

**Theorem 2.1.** *A solution  $x$  to (P) exists. Let  $\tau$  be a regular point of  $x$ . Then*

- (i) *There is an interval  $I$  which is a neighbourhood of  $\tau$  in  $[a, b]$  in which  $D^N x$  is essentially bounded, and in which the arc  $x$  satisfies the higher order Euler inclusion, namely, there exist Lipschitz continuous functions  $p_0 \equiv 0, p_1, p_2, \dots, p_N$ , such that*

$$(2.1) \quad ((\dot{p}_1(t) + p_0(t), \dot{p}_2(t) + p_1(t), \dots, \dot{p}_N(t) + p_{N-1}(t)), p_N(t)) \in \partial L(t, D_0^{N-1}x(t), D^N x(t)) \quad \text{a.e. } t \in I.$$

- (ii) *If in addition, for each  $t$  in  $[a, b]$  and  $v$  in  $\mathbf{R}^n$  the function  $w \rightarrow L(t, D_0^{N-1}x(t), w)$  is strictly convex and the function  $s \rightarrow L(s, D_0^{N-1}x(s), v)$  is continuous at  $t$ , then  $x$  is  $C^N$  in  $I$ .*
- (iii) *If we add to the hypotheses of (ii) that, for each  $t \in [a, b]$  the function  $L$  is  $C^r$  in its arguments near  $(t, D_0^{N-1}x(t), D^N x(t))$  for some  $r \geq N+1$ , and  $L_{ww}(t, D_0^{N-1}x(t), D^N x(t)) > 0$ , then  $x$  is  $C^r$  in  $I$ .*

The generalized gradient  $\partial L(t, z, w)$  in (2.1) is taken with respect to the variables  $(z, w)$  (for fixed  $t$ ). If  $L$  is  $C^1$  in these variables, the condition implies

$$(2.2) \quad L_{z_0}(t, D_0^{N-1}x(t), D^N x(t)) + \sum_{i=1}^{N-1} (-1)^i \frac{d^i}{dt^i} L_{z_i}(t, D_0^{N-1}x(t), D^N x(t)) + (-1)^N \frac{d^N}{dt^N} L_w(t, D_0^{N-1}x(t), D^N x(t)) = 0, \quad \text{a.e.}$$

In this equation  $L_{z_0}, \dots, L_{z_{N-1}}, L_w$  denote gradients of

$$L(t, (z_0, \dots, z_{N-1}), w)$$

with respect to the variables  $z_0, \dots, z_{N-1}, w$ . This is a familiar higher order version of the Euler equation [2].

**Corollary 2.2.** *Let  $x$  solve (P) and let  $\Omega$  be the set of regular points of  $x$ . Then  $\Omega$  is an open set of full measure, and  $D^N x$  is locally essentially bounded on  $\Omega$ . For every point  $\tau$  in  $\Omega$  there is an interval  $I$ , which is a neighbourhood of  $\tau$  in  $[a, b]$ , and Lipschitz continuous functions  $p_1, \dots, p_N$  such that the Euler inclusion (2.1) holds a.e. in  $I$ .*

Note that in the case  $L(t, z, w)$  is  $C^1$  in  $(z, w)$  the necessary condition of optimality implicit in this corollary is that the higher order Euler equation (2.2) holds a.e. in  $[a, b]$ .

*Proof.* In view of Theorem 2.1, it suffices to show that  $\Omega$  is an open set of full measure. Let  $\tau$  be a point in  $(a, b)$  which is also a Lebesgue point of  $t \rightarrow |D^N x(t)|$ . Since  $D^N x$  is an integrable function, the following "exact" Taylor expansions are valid

$$(2.3) \quad D^i x(t) = \sum_{j=i}^{N-1} D^j x(t-\tau) \frac{(t-\tau)^{j-i}}{(j-i)!} + \int_{\tau}^t \int_{\tau}^{\sigma_{N-1}} \cdots \int_{\tau}^{\sigma_2} D^N x_1 \cdots d\sigma_{N-i}$$

for all points  $t \in [\tau, b]$  and for  $i = 0, \dots, N-1$ . But  $\tau$  is a Lebesgue point of  $|D^N(x)|$ . It follows that there exist a constant  $c$  and  $t_1$  in  $(\tau, b]$  such that

$$|t - \tau|^{-1} \left| \int_{\tau}^t D^N x(\sigma_1) d\sigma_1 \right| \leq |t - \tau|^{-1} \int_{\tau}^t |D^N x(\sigma_1)| d\sigma_1 \leq c$$

for all points  $t$  such that  $\tau < t < t_1$ .

A simple calculation now yields the information that, for each choice of  $t$  in  $[\tau, t_1]$  and of index value  $i$ , the multiple integral on the right side of (2.3) is bounded in norm by  $(t - \tau)^{N-i} c$ . It is evident from these observations that  $\tau$  is a regular point of  $x$ . We have shown that  $\Omega$  contains the Lebesgue points of  $|D^N x|$ . It follows that the set  $\Omega$  has full measure.

Take an arbitrary point  $\tau$  in  $\Omega$ . By Theorem 2.1 there is an interval  $I$  which is a neighbourhood of  $\tau$  in  $[a, b]$  in which  $D^N x$  is essentially bounded. But any point at which  $D^N x$  is essentially bounded is certainly a regular point. It follows that  $\tau$  lies in the interior of  $\Omega$  relative to  $[a, b]$ , and therefore that  $\Omega$  is open in  $[a, b]$ . The proof is complete.

### 3. PROOF OF THE THEOREM

**3.1 Existence.** For purposes of proving existence of solutions, it is convenient to reformulate (P) as a new problem (Q), where the number of derivatives involved is reduced to one, but where we are forced to consider an extended valued Lagrangian,  $L_0: \mathbf{R} \times \mathbf{R}^{n \times N} \times \mathbf{R}^{n \times N} \rightarrow \mathbf{R} \cup \{+\infty\}$ .

$$L_0(t, (x_1, \dots, x_N), (v_1, \dots, v_N)) := \begin{cases} L(t, (x_1, \dots, x_N), v_N), & \text{if } v_1 = x_2, \dots, v_{N-1} = x_N, \\ +\infty, & \text{otherwise.} \end{cases}$$

Problem (Q) is

$$\left\{ \begin{array}{l} \text{Minimize } J_0(x) := \int_a^b L_0(t, x(t), Dx(t)) dt \\ \text{over arcs } x \in W^{1,1}([a, b]: \mathbf{R}^{n \times N}) \\ \text{which satisfy} \\ x(a) = A, \quad x(b) = B. \end{array} \right.$$

It is known [9] that (P) and (Q) are equivalent in the following sense. On the one hand, if  $y \in W^{N,1}([a, b]; \mathbf{R}^n)$  then  $J(y) = J_0(D_0^{N-1}y)$ . On the other, if  $(y_1, \dots, y_N) \in W^{1,1}([a, b]; \mathbf{R}^{n \times N})$  and  $J_0((y_1, \dots, y_N)) < \infty$  then  $y_1 \in W^{N,1}([a, b]; \mathbf{R}^n)$  and  $J(y_1) = J_0((y_1, \dots, y_N))$ . However, under hypotheses (H1)–(H3), (Q) is known to have a solution (see [9, 1]). Existence of a solution to (P) follows.

**3.2 Hypothesis reduction.** We show at the outset that we suffer no loss of generality by augmenting the basic hypotheses in certain respects.

The function  $\theta$  of hypothesis (H3) can be replaced by the nondecreasing function  $\tilde{\theta}(r): [0, \infty) \rightarrow \mathbf{R}$ :

$$\tilde{\theta}(r) := \inf\{\theta(r'): r' > 1\}$$

since  $\tilde{\theta}$  is superlinear, convex and is majorized by  $\theta$ . This observation coupled with the fact that adding a constant to  $L$  in problem (P) does not affect the minimizers, permits us to add to the basic hypotheses

(H4)  $\theta: [0, \infty) \rightarrow \mathbf{R}$  is positive-valued and nondecreasing.

Now let  $k$  be a constant such that the minimizer  $x$  (for problem (P)) under consideration satisfies  $\|D_0^{N-1}x\|_\infty < k$ . (Such a constant exists since  $x \in W^{N,1}$ .) Consider a new problem in which  $L$  in problem (P) is replaced by  $L_1(t, z, w) := \max[L(t, z, w), -\alpha k + \theta(|w|)] + \alpha k + 1$ . Hypotheses (H1)–(H4) continue to be satisfied (with  $\tilde{\theta} = \theta$  and  $\tilde{\alpha} = 0$ ). Note that (1.1) now holds with strict inequality. We have  $L_1 \leq L + \alpha k + 1$  everywhere and  $L_1(t, z, w) = L(t, z, w) + \alpha k + 1$  for all points in  $w$  in  $\mathbf{R}^n$  and  $(t, z)$  in some tube about the function  $D_0^{N-1}x$ . So  $x$  remains a minimizer and the assertions of the theorem for the new problem imply those for (P). Without loss of generality then we may add to the hypotheses:

(H5) In hypothesis (H3), inequality is strict and  $\alpha = 0$ .

**3.3 Polynomial interpolation.** Polynomial functions enter into the proof of the theorem at several points, both in the role of comparison arcs and when we come to interpret necessary conditions. In the present context a polynomial  $p$  of degree at most  $m$  is taken to be a function of the scalar variable  $t$ ,

$$p(t) = \sum_{i=0}^m g_i t^i$$

in which  $g_0, \dots, g_m$  are  $n$  vectors. Before proceeding, we gather together some useful properties.

**Lemma 3.1.** *Corresponding to any positive number  $k_1$  there exists a positive number  $k_2$  with the following properties. Let  $X (= \text{col}(X_0, \dots, X_{N-1}))$  and  $Y (= \text{col}(Y_0, \dots, Y_{N-1}))$  be arbitrary  $\mathbf{R}^{n \times N}$  vectors and  $s, t$  arbitrary real numbers,  $t > s$ , such that*

$$|t - s| \leq k_1$$

and

$$\eta((X_0, \dots, X_{N-1}), (Y_0, \dots, Y_{N-1}, t - s)) \leq k_1.$$

Then there is a unique polynomial  $p$  of degree at most  $(2N - 1)$  such that

$$D_0^{N-1} p(s) = X, \quad D_0^{N-1} p(t) = Y,$$

and  $p$  satisfies

$$\|D^N p\|_{\infty, s, t} \leq k_2$$

and

$$\|D_0^N p\|_{\infty, s, t} \leq k_2(\|X\| + \|Y\|).$$

In the lemma  $\|\cdot\|_{\infty, \alpha, \beta}$  denotes the  $L^\infty([\alpha, \beta])$  norm. We use this notation throughout the paper, when we wish to emphasize the domain of the functions concerned.

*Proof.* By translating the independent variable we can arrange that  $s = 0$ . An arbitrary polynomial  $p$  of degree at most  $2N - 1$  which satisfies the boundary conditions on the left,

$$D_0^{N-1} p(0) = X,$$

has representation

$$(3.1) \quad p(\tau) = g(\tau) \cdot X + h(\tau) \cdot Z$$

in which  $Z (= \text{col}(Z_0, \dots, Z_{N-1}))$  is some  $n \times N$  vector, and  $g, h$  are  $N$  vector valued functions of the scalar variable  $\tau$ ,

$$g(\tau) = [1, \tau, \tau^2/2!, \dots, \tau^{N-1}/(N-1)!]$$

and

$$h(\tau) = [\tau^N/N!, \dots, \tau^{2N-1}/(2N-1)!].$$

Observe that

$$D_0^{N-1} p(t) = PX + QZ$$

where

$$P := \begin{bmatrix} 1 & t & t^2/2! & \dots & t^{N-1}/(N-1)! \\ 0 & 1 & t & \dots & \cdot \\ & 0 & & & \cdot \\ & & & & 1 \end{bmatrix}$$

and

$$Q := \begin{bmatrix} t^N/N!, & t^{N+1}/(N+1)!, & \dots, & t^{2N-1}/(2N-1)! \\ t^{N-1}/(N-1)!, & t^N/N! & & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ t & \cdot & \cdot & t^N/N! \end{bmatrix}.$$

It follows that  $p$  also satisfies the boundary condition on the right

$$D_0^{N-1} p(t) = Y$$



if and only if  $Z$  satisfies

$$E \operatorname{col}(Z_0, tZ_1, \dots, t^{N-1}Z_{N-1}) = \xi$$

where

$$E := \begin{bmatrix} 1/N! & 1/(N+1)! & \dots & 1/(2N-1)! \\ 1/(N-1)! & 1/N! & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \dots & \dots & 1/N! \end{bmatrix}$$

and

$$\xi = \left\{ \left[ Y_i - \sum_{j=i}^{N-1} X_j t^{j-1} / (j-i)! \right] / t^{N-i} \right\}_{i=0}^{N-1}.$$

However  $E$  is invertible: an inductive argument yields the following formula for the determinant

$$\det E = \frac{(N-1)!(N-2)! \dots 1}{(2N-1)(2N-2)! \dots N!}.$$

We conclude that there is a unique polynomial  $p$  of degree at most  $2N-1$ , which satisfies all the boundary conditions. Notice that  $|\xi| \leq k_1$ , under the hypotheses. It follows that

$$|\operatorname{col}(Z_0, tZ_1, \dots, t^{N-1}Z_{N-1})| \leq k_1 |E^{-1}|$$

( $|E^{-1}|$  denotes the operator norm  $\max\{|E^{-1}x| : |x| = 1\}$ ). The assertions of the lemma all follow from this estimate, and equation (3.1), since  $|t| \leq k_1$ .

**Lemma 3.2.** *Let  $p$  be a polynomial of degree at most  $k$ , and let  $[t_0, t_1]$  be an arbitrary compact interval. Then there exists a vector  $v$ ,  $|v| = 1$ , and a subinterval  $J \subset [t_0, t_1]$  such that*

$$J \subset \{t \in [t_0, t_1] : p(t) \cdot v \geq \frac{1}{2} \|p\|_{\infty, t_0, t_1}\}$$

and

$$\operatorname{meas}\{J\} \geq d_k |t_1 - t_0|.$$

Here  $d_k = (16)^{-k}$ .

*Proof.* We may assume that  $t_1 > t_0$  and that  $p$  is not the zero function since, in these cases, the assertions of the lemma are trivial. Besides which, we may add to the hypotheses

$$(3.2) \quad t_0 = 0, \quad t_1 = 1 \quad \text{and} \quad \|p\|_{\infty, t_0, t_1} = 1,$$

for, if these conditions were not satisfied, we could replace  $p$  by  $\tilde{p}$  and  $[t_0, t_1]$  by  $[0, 1]$  where  $\tilde{p}(t) = (\|p\|_{\infty, t_0, t_1}^{-1})^{-1} p([t_1 - t_0]t + t_0)$ . The extra hypotheses are now satisfied; we deduce the desired properties of  $p$  from those of  $\tilde{p}$ . In summary, we may assume (3.2).

From this point, proof proceeds by induction on  $k$ . The case  $k = 0$  is trivial. Suppose now the assertions of the lemma are true for  $k$ . Let  $p$  be

any polynomial of degree at most  $k + 1$  and such that  $\|p\|_{\infty,0,1} = 1$ . Two possibilities need to be considered.

(i) Suppose  $\|Dp\|_{\infty,0,1} \leq 8/d_k$ . Let  $s$  be a point such that  $|p(s)| = 1$  and set  $v = p(s)$ . Then  $p(t) \cdot v \leq \frac{1}{2}$  for all  $t$  in the interval  $[t - d_k/16, t + d_k/16] \cap [0, 1]$ . But this interval contains a subinterval of length  $d_{k+1}$  ( $= d_k/16$ ), as required.

(ii) Suppose  $\|Dp\|_{\infty,0,1} \leq 8/d_k$ . The polynomial  $Dp$  has degree at most  $k$ . By the induction hypothesis then, there exists a subinterval  $S \subset [0, 1]$  and a vector  $v$ ,  $|v| = 1$ , such that

$$Dp(t) \cdot v > \frac{1}{2}(8/d_k), \quad \text{all } t \in S,$$

and  $\text{meas}\{S\} \geq d_k$ . Now if  $p(t) \cdot v > \frac{1}{2}$  for all  $t \in S$ ,  $p$  clearly has the desired properties. The remaining case to be considered is that where  $p(s) \cdot v \leq \frac{1}{2}$  for some  $s \in S$ . Observe that there is a subinterval  $S'$  of  $S$ , of length  $d_k/4$ , separated from  $s$  by at least  $d_k/4$ . For  $t \in S'$  we have

$$\left| \int_s^t Dp(\sigma) \cdot v \, d\sigma \right| > (4/d_k) \cdot d_k/4 = 1$$

whence

$$|p(t) \cdot v| > 1 - |p(s) \cdot v| \geq \frac{1}{2}.$$

In this case too then,  $p(t) \cdot v \geq \frac{1}{2}$  on a subinterval of length  $d_k/4$ . The proof is complete.

Let  $x$  be a minimizer for (P), and let  $\tau$  be a regular point. By definition of regular point there are sequences  $\{s_i\}$  and  $\{t_i\}$ , converging to  $\tau$ , with  $a \leq s_i \leq \tau \leq t_i \leq b$ ,  $s_i \neq t_i$ , and a constant  $k_1$  such that

$$(3.3) \quad \eta(D_0^{N-1}x(s_i), D_0^{N-1}x(t_i), t_i - s_i) < k_1$$

for  $i = 1, 2, \dots$ . Note that, in view of the continuity of  $D_0^{N-1}x$  and  $\eta$ , we can adjust the  $s_i$ 's and  $t_i$ 's a small amount if necessary and thereby arrange that, for each  $i$ ,  $[s_i, t_i]$  is a neighbourhood of  $\tau$  in  $[a, b]$ , while at the same time preserving the inequality (3.3).

We denote by  $J_i$  the integral functional

$$J_i(y) = \int_{s_i}^{t_i} L(t, D_0^{N-1}y(t), D^N y(t)) \, dt,$$

for  $i = 1, 2, \dots$ . Evidently the minimizer  $x$  for (P) (strictly speaking its restriction to  $[s_i, t_i]$ ) is also a minimizer for  $(P_i)$  for  $i = 1, 2, \dots$ , where  $(P_i)$  is the problem

$$(3.4) \quad \begin{cases} \text{Minimize } J_i(y) \\ \text{over } y \in W^{N,1}([s_i, t_i]; \mathbf{R}^n) \text{ which satisfy} \\ D_0^{N-1}y(s_i) = D_0^{N-1}x(s_i), \quad D_0^{N-1}y(t_i) = D_0^{N-1}x(t_i). \end{cases}$$

For  $i = 1, 2, \dots$ , we take  $y_i$  to be the polynomial of degree at most  $2N - 1$  which satisfies the boundary conditions (3.4) in problem  $(P_i)$ . (See Lemma 3.1.) We shall refer to  $y_i$  as the "interpolating polynomial" (for  $(P_i)$ ).

**Lemma 3.3.** For  $i = 1, 2, \dots$  let  $z_i \in W^{N,1}([s_i, t_i], \mathbf{R}^n)$  be an arc which satisfies the boundary conditions (3.4) in problem  $(P_i)$  and also

$$\frac{1}{2} \int_{s_i}^{t_i} \theta(|D^N z_i(t)|) dt \leq \int_{s_i}^{t_i} L(t, D_0^{N-1} y(t), D^N y(t)) dt.$$

Here  $y_i$  is the interpolating polynomial for  $(P_i)$ . Then there exist numbers  $M > 0$  and  $R_0 > 0$  with the following properties

- (i) The real numbers  $\|D_0^{N-1} y_i\|_{\infty, s_i, t_i}$ ,  $\|D_0^{N-1} z_i\|_{\infty, s_i, t_i}$  for  $i = 1, 2, \dots$ , and also  $\|D_0^{N-1} x\|_{\infty, a, b}$  are all bounded by  $M$ , and
- (ii)  $\|D^N y\|_{\infty, s_i, t_i} < R_0$ , for  $i = 1, 2, \dots$ .

*Proof.* The assertions concerning the arcs  $x$  and  $y_i$  follow from Lemma 3.1 and the fact that  $x$  is a  $W^{N,1}$  function. It remains to bound  $D_0^{N-1} z_i$ .

Since  $\theta$  is superlinear, there exists a number  $a$  such that  $\frac{1}{2}\theta(a') > a'$  whenever  $a' \geq a$ . For each  $i$  we have

$$\begin{aligned} \|D^{N-1} z_i(t)\|_{\infty, s_i, t_i} &\leq |D^{N-1} z_i(s_i)| + \int_{s_i}^{t_i} |D^N z_i(s)| ds \\ &\leq |D^{N-1} z_i(s_i)| + a|t_i - s_i| + \frac{1}{2} \int_{s_i}^{t_i} \theta(|D^N z_i(s)|) ds \\ &\leq |D^{N-1} z_i(s_i)| + a|t_i - s_i| + \frac{1}{2} \int_{s_i}^{t_i} L(t, D_0^{N-1} y_i(t), D^N y_i(t)) dt. \end{aligned}$$

All the terms on the right are bounded by some constant, independent of  $i$ . This, together with similarly derived bounds on the lower derivatives, yields the desired uniform bound on  $\|D_0^{N-1} z_i\|_{\infty, s_i, t_i}$ .

**3.4 The auxiliary Lagrangian.** In what follows,  $M$  and  $R_0$  are as in Lemma 3.3. Define  $c_0$ ,

$$(3.5) \quad c_0 := \max\{|L(t, z, w)| : t \in [a, b], |z| \leq M, |w| \leq R_0\}$$

and the function  $d: \mathbf{R} \rightarrow \mathbf{R} \cup \{+\infty\}$

$$(3.6) \quad d(P) := \inf\{|w| : p \in \partial_w L(t, z, w), t \in [a, b], |z| \leq M, |p| \geq P\}$$

(the infimum here is interpreted as “ $+\infty$ ” when no elements exist which satisfy the constraints).

Since  $L(t, z, w)$  is locally Lipschitz continuous in  $(z, w)$ , uniformly in  $t$ , we have

**Lemma 3.4.** The function  $d$  is monotone increasing and

$$\lim_{P \rightarrow +\infty} d(P) = +\infty.$$

Choose a number  $\varepsilon > 0$ . This will remain fixed hereafter. In view of Lemma 3.4 and the superlinear growth of  $\theta$ , we may choose the constant  $R_1$  to satisfy

$R_1 > R_0$  and

$$(3.7) \quad \theta \circ d \left( \frac{\theta(r')}{4r'} - \frac{L(t, z, 0)}{2r'} - \frac{3\varepsilon}{2} \right) > 2 \cdot (16)^{N-1} c_0$$

whenever  $r' \geq R_1$ ,  $t \in [a, b]$  and  $|z| \leq M$ .

Set

$$(3.8) \quad c_1 := \max\{|L(t, z, w)| : t \in [a, b], |z| \leq M, |w| \leq R_1\}$$

and

$$(3.9) \quad \sigma_1 := \max\{|\xi| : \xi \in \partial_w L(t, z, w), t \in [a, b], |z| \leq M, |w| \leq R_1\}.$$

Appealing once again to the convexity and superlinear growth of  $\theta$ , we may select a constant  $R_2 > R_1$  such that  $\theta$  is strictly increasing on  $[R_2, \infty)$  and

$$(3.10) \quad \frac{1}{2}\theta(r) \geq \sigma_1(R_1 + r) + c_1$$

for all  $r \geq R_2$ .

The function  $\varphi : \mathbf{R}^n \rightarrow \mathbf{R}$  is defined as follows:  $\varphi(w) := \frac{1}{2} \max[\theta(|w|), \theta(R_2)]$ .

We are now ready to construct the auxiliary Lagrangian  $\tilde{L} : \mathbf{R} \times \mathbf{R}^{n \times N} \times \mathbf{R}^n \rightarrow \mathbf{R}$ . For each  $(t, z)$ ,  $\tilde{L}(t, z, \cdot)$  is taken to be the convex hull of the functions  $w \rightarrow L(t, z, w)$ , restricted to  $|w| \leq R_2$  and  $w \rightarrow \varphi(w)$ . Precisely stated,

$$\begin{aligned} \tilde{L}(t, z, w) &:= \inf\{\lambda L(t, z, u) + (1 - \lambda)\varphi(v) : \\ &\quad 0 \leq \lambda \leq 1, |u| \leq R_2 \text{ and } \lambda u + (1 - \lambda)v = w\}. \end{aligned}$$

We list important properties of the auxiliary Lagrangian.

**Proposition 3.5.** (a)  $\tilde{L}(t, z, w)$  is locally bounded, measurable in  $t$  and convex in  $w$ .

(b)  $\tilde{L}(t, z, w)$  is locally Lipschitz continuous in  $(z, w)$ , uniformly in  $t \in [a, b]$ .

(c)  $\tilde{L}(t, z, w) \geq \theta(|w|)/2$  for all  $(t, z, w)$ .

(d) For  $t \in [a, b]$  and  $|z| \leq M$  we have  $\tilde{L}(t, z, w) = L(t, z, w)$  if  $|w| \leq R_1$ ,  $\tilde{L}(t, z, w) \leq L(t, z, w)$  if  $|w| \leq R_2$  and  $\tilde{L}(t, z, w) < L(t, z, w)$  if  $|w| > R_2$ .

(e) For  $(t, z) \in [a, b] \times \mathbf{R}^{n \times N}$  we have  $\tilde{L}(t, z, w) = \theta(|w|)/2$  if  $|w| \geq R_2$ .

*Proof.*  $\tilde{L}$  is convex in  $w$  by construction, and locally bounded since  $0 \leq \tilde{L} \leq \varphi$  and  $\varphi$  is locally bounded. To see that  $\tilde{L}$  is measurable in the  $t$  variable, we use the fact that, for fixed  $z, w$ , the inf defining  $\tilde{L}$  can equivalently be taken over countably many measurable functions of  $t$  (obtained by taking  $(\lambda, u, v)$  in a suitable dense set). (a) has been proved. (c) is true because  $L$  and  $\varphi$  satisfy the desired inequality and  $w \rightarrow \frac{1}{2}\theta(|w|)$  is convex. Consider (e). Take arbitrary points  $w' \in \mathbf{R}^n$ ,  $|w'| > R_2$ , and  $(t, z) \in [a, b] \times \mathbf{R}^{n \times N}$ . Choose  $\zeta \in \partial\varphi(w')$ . By the subgradient inequality

$$(3.11) \quad \varphi(v) - \varphi(w') - \zeta \cdot (v - w') \geq 0, \quad \text{for } v \in \mathbf{R}^n.$$

Since  $\theta$  is continuous and strictly increasing on  $[R_2, \infty)$ ,  $\varphi(w)$  and  $\frac{1}{2}\theta(|w|)$  coincide on a neighbourhood of  $w = w'$ . It follows that  $\zeta$  is also a subgradient of  $w \rightarrow \frac{1}{2}\theta(|w|)$  and, for all  $u \in \mathbf{R}^n$  and  $(t, z) \in [a, b] \times \mathbf{R}^{n \times N}$ ,

$$(3.12) \quad L(t, z, u) - \varphi(w') - \zeta \cdot (u - w') \geq \frac{1}{2}\theta(|u|) - \frac{1}{2}\theta(|w'|) - \zeta \cdot (u - w') \geq 0.$$

From (3.11) and (3.12) we deduce

$$\tilde{L}(t, z, w) = \inf\{\lambda L(t, z, u) + (1 - \lambda)\varphi(v)\}$$

(the infimum is over  $0 \leq \lambda \leq 1$ ,  $|u| \leq R_2$ ,  $\lambda u + (1 - \lambda)v = w$ )

$$\geq \varphi(w') - \zeta \cdot (w - w').$$

Setting  $w = w'$  we obtain  $\tilde{L}(t, z, w) \geq \varphi(w')$ . But then  $\tilde{L}(t, z, w) = \frac{1}{2}\theta(|w|)$  in the region  $|w| > R_2$  since  $\varphi$  and  $w \rightarrow \frac{1}{2}\theta(|w|)$  coincide here and  $\varphi$  majorizes  $\tilde{L}$ . This remains true in the region  $|w| \geq R_2$ , by the continuity properties of convex functions.

We turn next to (b). Take  $k_1 > 0$ , and let  $K$  be a Lipschitz constant for  $z \rightarrow L(t, z, w)$ , uniformly valid for  $t$  in  $[a, b]$ ,  $|w| \leq R_2$  and  $|z| \leq k_1$ . Take  $z_1, z_2 \in \mathbf{R}^{n \times N}$  such that  $|z_1|, |z_2| \leq k_1$ . Then for any  $\delta > 0$ ,  $t$  in  $[a, b]$  and  $w$ , we may choose  $u, v$  and  $\lambda$  as in the definition of  $\tilde{L}$  such that

$$\tilde{L}(t, z_1, w) + \delta \geq \lambda L(t, z_1, u) + (1 - \lambda)\varphi(v).$$

It follows that

$$\begin{aligned} \tilde{L}(t, z_2, w) &\leq \lambda L(t, z_2, u) + (1 - \lambda)\varphi(v) \\ &\leq \lambda L(t, z_1, u) + K|z_1 - z_2| + (1 - \lambda)\varphi(v) \\ &\leq \tilde{L}(t, z_1, w) + K|z_1 - z_2| + \delta. \end{aligned}$$

Since  $z_1$  and  $z_2$  are interchangeable and  $\delta > 0$  is arbitrary, it follows that  $z \rightarrow \tilde{L}(t, z, w)$  has Lipschitz rank at most  $K$  on  $|z| \leq k_1$  for all  $t \in [a, b]$  and  $w \in \mathbf{R}^n$ .

Now take  $k_2 \geq R_2$ . We shall show that  $w \rightarrow \tilde{L}(t, z, w)$  is Lipschitz continuous in the region  $|w| \leq k_2$  uniformly in  $(t, z) \in [a, b] \times \mathbf{R}^{n \times N}$ . It will follow that  $\tilde{L}$  is locally Lipschitz continuous jointly in the variables  $z, w$ , uniformly in  $t \in [a, b]$ , since it has this property in the individual variables.

Choose  $(t, z) \in [a, b] \times \mathbf{R}^{n \times N}$ , and let  $w \rightarrow p \cdot w + q$  be an arbitrary, nonconstant, affine function which is majorized by  $w \rightarrow \tilde{L}(t, z, w)$ . By (e), we must have

$$p \cdot w + q \leq \frac{1}{2}\theta(|w|)$$

for  $|w| \geq k_2$ . Setting  $w = (k_2 + 1)p/|p|$ , we obtain

$$(3.13) \quad (k_2 + 1)|p| + q \leq \frac{1}{2}\theta(k_2 + 1).$$

However, since  $\tilde{L} \geq 0$ , we also have

$$(3.14) \quad \tilde{L}(t, z, w) - p \cdot w - q \geq -|p|k_2 - q$$

for  $|w| \leq k_2$ . (3.13) and (3.14) yield

$$\tilde{L}(t, z, w) - p \cdot w - q \geq |p| - \frac{1}{2}\theta(k_2 + 1)$$

for  $|w| \leq k_2$ .

Set  $K_1 = \frac{1}{2}\theta(k_2 + 1) + 1$ . It follows that

$$(3.15) \quad \tilde{L}(t, z, w) - p \cdot w - q \geq 1$$

for  $|w| \leq k_2$ ,  $|p| \geq K_1$ .

Now the function  $w \rightarrow \tilde{L}(t, z, w)$  is expressible as the pointwise supremum of affine functionals majorized by  $\tilde{L}$ . However inequality (3.15) tells us that, to evaluate this function in the region  $|w| \leq k_2$ , we may limit attention to affine functionals with Lipschitz rank at most  $K_1$ . It follows that  $w \rightarrow \tilde{L}(t, z, w)$  has Lipschitz rank at most  $K_1$  here for arbitrary  $(t, z) \in [a, b] \times \mathbf{R}^{n \times N}$ .

Consider finally (d). The cases  $|w| \geq R_2$  and  $R_2 > |w| > R_1$  follow from (e) and from the facts that  $L$  majorizes  $\tilde{L}$  and  $L$  strictly majorizes  $u \rightarrow \frac{1}{2}\theta(|u|)$ . It remains to show then that the values  $L(t, z, w)$  and  $\tilde{L}(t, z, w)$  coincide for  $(t, z) \in [a, b] \times \mathbf{R}^{n \times N}$  and  $|w| \leq R_1$ .

Take  $(t, z)$  and  $w$  as above and choose  $\zeta \in \partial_w L(t, z, w)$ . By (3.10) and definition of the constants  $c_1, \sigma_1$  (see (3.8) and (3.9)) we have that

$$\varphi(v) \geq \frac{1}{2}\theta(|v|) \geq \sigma_1(R_1 + |v|) + c_1 \geq L(t, z, w) + \zeta \cdot (v - w)$$

for points  $v$  which satisfy  $|v| \geq R_2$ . On the other hand

$$\varphi(v) \geq \frac{1}{2}\theta(|R_2|) \geq \sigma_1(R_1 + R_2) + c_1 \geq L(t, z, w) + \zeta \cdot (v - w)$$

also for points  $v$  which satisfy  $|v| \leq R_2$ . By the subgradient inequality however,

$$L(t, z, u) - L(t, z, w) - \zeta \cdot (u - w) \geq 0$$

for all  $u \in \mathbf{R}^n$ . Scaling and adding these inequalities, we arrive at

$$\begin{aligned} \tilde{L}(t, z, w') &= \inf\{\lambda L(t, z, u) + (1 - \lambda)\varphi(v) : \\ &\quad 0 \leq \lambda \leq 1, |u| \leq R_2 \text{ and } \lambda u + (1 - \lambda)v = w'\} \\ &\leq L(t, z, w) + \zeta \cdot (w' - w), \end{aligned}$$

for all points  $w' \in \mathbf{R}^n$ . Setting  $w' = w$  yields  $\tilde{L}(t, z, w) \geq L(t, z, w)$ . Since however  $L$  majorizes  $\tilde{L}$  we may replace inequality here by equality. The proof is complete.

**3.5 End of proof.** Consider now the auxiliary problems  $(\tilde{P}_i)$ ,  $i = 1, 2, \dots$ :

$$\left\{ \begin{array}{l} \text{Minimize } \int_{s_i}^{t_i} \tilde{L}(t, D_0^{N-1}y(t), D^N y(t)) dt \\ \text{over arcs } y \in W^{N,1}(s_i, t_i) \text{ which satisfy} \\ D_0^{N-1}y(s_i) = D_0^{N-1}x(s_i), \quad D_0^{N-1}y(t_i) = D_0^{N-1}x(t_i). \end{array} \right.$$

Parts (a)–(c) of Proposition 3.5 ensure that  $(\tilde{P}_i)$  has a solution, we write it  $x_i$ , for each  $i$ .

In order to take advantage of the remaining properties of  $\tilde{L}$  listed in Proposition 3.5, we need to show that  $D_0^{N-1}x_i$  is suitably confined.

**Lemma 3.6.**  $\|D_0^{N-1}x_i\|_{\infty, s_i, t_i} \leq M$ , for all  $i$ , where  $M$  is the constant of Lemma 3.3.

*Proof.* We have

$$\frac{1}{2} \int_{s_i}^{t_i} \theta(|D^N x_i(s)|) ds \leq \int_{s_i}^{t_i} \tilde{L}(t, D_0^{N-1}x_i(t), D^N x_i(t)) dt,$$

by Proposition 3.5, part (c),

$$\leq \int_{s_i}^{t_i} \tilde{L}(t, D_0^{N-1}y_i(t), D^N y_i(t)) dt,$$

since  $x_i$  is a minimizer (here  $y_i$  is the interpolating polynomial),

$$= \int_{s_i}^{t_i} L(t, D_0^{N-1}y_i(t), D^N y_i(t)) dt,$$

by Proposition 3.5, part (d) and Lemma 3.3, part (ii). The result now follows from Lemma 3.3, part (i).

The next step is to apply necessary conditions from optimal control theory to the minimizer  $x_i$ . To this end we reformulate  $(\tilde{P}_i)$  as an optimal control problem  $(CP_i)$ :

$$\left\{ \begin{array}{l} \text{Minimize } \int_{s_i}^{t_i} \tilde{L}(t, \xi(t), u(t)) dt \text{ over functions} \\ \xi \in W^{1,1}([s_i, t_i]; \mathbf{R}^{n \times N}) \text{ and } u \in L^1([s_i, t_i]; \mathbf{R}^n) \\ \text{which satisfy} \\ \dot{\xi}(t) = F\xi(t) + Gu(t) \text{ a.e. } t \in [s_i, t_i] \text{ and} \\ \xi(s_i) = D_0^{N-1}x(s_i), \quad \xi(t_i) = D_0^{N-1}x(t_i). \end{array} \right.$$

Here, once again,  $\tilde{L}$  is the auxiliary Lagrangian. The matrices  $F$  ( $nN$  by  $nN$ ) and  $G$  ( $nN$  by  $n$ ) are

$$(3.16) \quad F = \begin{bmatrix} 0 & M & 0 & \dots & 0 \\ 0 & 0 & M & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & M \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \quad \text{and} \quad G = \text{col}\{0, \dots, 0, M\}$$

in which  $M$  denotes the  $n \times n$  identity matrix, and  $0$  the  $n \times n$  zero matrix.

Our choice of differential equation in problem  $(CP_i)$  ensures that functions  $\xi, u$  satisfy the constraints of problem  $(CP_i)$  only if  $\xi = D_0^{N-1}\xi_1$ , and  $u = D^N \xi_1$ , where  $\xi_1$  is the first block component of  $\xi$ . Evidently  $(\xi_i = D_0^{N-1}x_i, u_i = D^N x_i)$  solves problem  $(CP_i)$ .

Now  $(CP_i)$  is a problem to which the maximum principle [3, Theorem 5.2.3] applies, with reference to the minimizer  $(D_0^{N-1}x_i, D^N x_i)$ . In checking the hypotheses involved, we notice in particular that there is a tube  $T$  about  $D_0^{N-1}x_i$  and a constant  $c$  such that

$$|\tilde{L}(t, z', u) - \tilde{L}(t, z'', u)| \leq c|z' - z''|$$

whenever  $(t, z'), (t, z'') \in T$  and  $u \in \mathbf{R}^n$ . (The fact that  $\partial_z \tilde{L}(t, z, w) (= \partial_z \varphi(|w|)) = 0$  when  $t \in [a, b]$ , and  $|w| > R_2$  is crucial to this assertion.) Thus  $\tilde{L}$  has the required uniform Lipschitz continuity properties for application of the maximum principle. A further important observation is that the underlying controlled differential equation in problem  $(CP_i)$  is controllable, and therefore the maximum principle applies in normal form (i.e., the cost multiplier may be set to one).

For each  $i$ , we deduce existence of an absolutely continuous function  $p (= \text{row}\{p_1, \dots, p_N\})$  such that

$$(3.17) \quad -\dot{p}(t) \in p(t) \cdot F - \partial_z \tilde{L}(t, D_0^{N-1}x_i(t), D^N x_i(t)), \quad \text{a.e. } t \in [s_i, t_i],$$

and such that  $D^N x_i(t)$  maximizes

$$u \rightarrow p_N(t)u - \tilde{L}(t, D_0^{N-1}x_i(t), u), \quad \text{a.e. } t \in [s_i, t_i].$$

Since  $\tilde{L}$  is convex in its third argument, this last property can be expressed as

$$(3.18) \quad p_N(t) \in \partial_w \tilde{L}(t, D_0^{N-1}x_i(t), D^N x_i(t)).$$

**Lemma 3.7.** *For some  $i$ , there exists a polynomial  $f$  (of degree at most  $N - 1$ ) such that*

$$(3.19) \quad f(t) \in \partial_w \tilde{L}(t, D_0^{N-1}x_i(t), D^N x_i(t)) + \varepsilon B, \quad \text{a.e. } t \in [s_i, t_i].$$

Here  $B$  is the open unit ball and  $\varepsilon$  is the positive number in (3.7).

*Proof.* We appeal to the necessary conditions (3.17) and (3.18). Note that, for each  $i$ ,  $p$  can be decomposed into the sum of two functions  $q (= \text{row}\{q_1, \dots, q_N\})$  and  $r$ ,

$$p(t) = q(t) + r(t),$$

where  $q$  satisfies the homogeneous equation

$$(3.20) \quad -\dot{q} = q \cdot F, \quad \text{a.e.,}$$

and  $r$  solves the inhomogeneous equation

$$(3.21) \quad -\dot{r} = r \cdot F - h, \quad \text{a.e.,}$$

together with the boundary condition

$$(3.22) \quad r(s_i) = 0.$$

Here  $h(t)$  is a measurable function which satisfies

$$h(t) \in \partial_z \tilde{L}(t, D_0^{N-1}x_i(t), D^N x_i(t)), \quad \text{a.e. } t \in [s_i, t_i].$$



By Proposition 3.5, parts (b) and (e), there exists a constant  $k_1$ , which does not depend on  $i$ , such that

$$\|h\|_{\infty, s_i, t_i} \leq k_1.$$

We deduce from the differential equation (3.21) and the associated boundary condition (3.22), along with Gronwall's inequality, that

$$\|r\|_{\infty, s_i, t_i} \leq \varepsilon,$$

provided  $i$  is chosen large enough. On the other hand the differential equation (3.20) implies that  $f := q_N$  is a polynomial of degree at most  $N - 1$ . It now follows from (3.18) that  $f$  has the desired properties.

Henceforth  $i$  is fixed at an index value for which the assertions of Lemma 3.7 are true.

**Lemma 3.8.**  $\|D^N x_i\|_{\infty, s_i, t_i} \leq R_1$ .

*Proof.* Assume to the contrary that

$$\|D^N x_i\|_{\infty, s_i, t_i} > R_1$$

on a set of positive measure. Let  $f$  be the polynomial of Lemma 3.7. We choose  $\bar{t}$  to be a point at which the inclusion (3.19) holds and

$$(3.23) \quad |D^N x_i(\bar{t})| > R_1.$$

By (3.19) and the subgradient inequality,

$$\tilde{L}(\bar{t}, D_0^{N-1} x_i(\bar{t}), 0) - \tilde{L}(\bar{t}, D_0^{N-1} x_i(\bar{t}), D^N x_i(\bar{t})) \geq -\langle f(\bar{t}), D^N x_i(\bar{t}) \rangle - \varepsilon |D^N x_i(\bar{t})|.$$

It follows that

$$\begin{aligned} |f(\bar{t})| \cdot |D^N x_i(\bar{t})| &\geq \tilde{L}(\bar{t}, D_0^{N-1} x_i(\bar{t}), D^N x_i(\bar{t})) - \tilde{L}(\bar{t}, D_0^{N-1} x_i(\bar{t}), 0) - \varepsilon |D^N x_i(\bar{t})| \\ &\geq \frac{1}{2} \theta (|D^N x_i(\bar{t})|) - L(\bar{t}, D_0^{N-1} x_i(\bar{t}), 0) - \varepsilon |D^N x_i(\bar{t})|, \end{aligned}$$

by Proposition 3.5, parts (c) and (d). By (3.5) and (3.23)

$$|f(\bar{t})| \geq \frac{\theta(r)}{2r} - \frac{c_0}{r} - \varepsilon$$

for some number  $r > R_1$ .

It follows now from Lemma 3.2 that

$$|f(t)| \geq \frac{\theta(r)}{4r} - \frac{c_0}{2r} - \frac{\varepsilon}{2}$$

for all  $t \in S$ , where  $S$  is a subinterval of  $[s_i, t_i]$ , of length at least  $16^{-(N-1)} \cdot |t_i - s_i|$ .

Bearing in mind the definition of the function  $d$  (see (3.6)), we deduce from this bound and the inclusion (3.19) that

$$|D^N x_i(t)| \geq d(|f(s)| - \varepsilon) \geq d\left(\frac{\theta(r)}{4r} - \frac{c_0}{2r} - \frac{3\varepsilon}{2}\right)$$

for all  $t \in S$ .

Then for all  $t \in S$

$$\frac{1}{2}\theta(|D^N x_i(t)|) \geq \frac{1}{2}\theta \cdot d \left( \frac{\theta(r)}{4r} - \frac{c_0}{2r} - \frac{3\varepsilon}{2} \right),$$

since  $\theta$  is monotone,

$$(3.24) \quad > c_0(16)^{N-1},$$

since  $r \geq R_1$  and in view of (3.7).

We have

$$\int_{s_i}^{t_i} \tilde{L}(t, D_0^{N-1} x_i(t), D^N x_i(t)) dt \geq \frac{1}{2} \int_{s_i}^{t_i} \theta(|D^N x_i(t)|) dt$$

by Proposition 3.5, part (c),

$$\begin{aligned} &= \frac{1}{2} \int_{[s_i, t_i] \setminus S} \theta ds + \frac{1}{2} \int_S \theta ds \\ &> c_0(16)^{N-1} \cdot (16)^{-(N-1)} \cdot |t_i - s_i| = c_0 |t_i - s_i|, \end{aligned}$$

since  $S$  has length at least  $(16)^{-(N-1)} \cdot |t_i - s_i|$  and  $\theta$  is nonnegative valued, and by (3.24),

$$\geq \int_{s_i}^{t_i} L(t, D_0^{N-1} y_i(t), D^N y_i(t)) dt$$

since the interpolating polynomial satisfies

$$\|D_0^{N-1} y_i\|_{\infty, s_i, t_i} < M \quad \text{and} \quad \|D^N y_i\|_{\infty, s_i, t_i} \leq R_0,$$

and by definition of  $c_0$ ,

$$= \int_{s_i}^{t_i} \tilde{L}(t, D_0^{N-1} y_i(t), D^N y_i(t)) dt,$$

again by the properties of  $y_i$ , and Proposition 3.5, part (d). But this contradicts optimality of  $x_i$ . It follows that  $\|D^N x_i\|_{\infty, s_i, t_i} \leq R_1$ , as claimed.

We now show

$$(3.25) \quad \|D^N x\|_{\infty, s_i, t_i} \leq R_2,$$

which will mean that  $D^N x$  is locally essentially bounded on the interval  $[s_i, t_i]$ .

Suppose to the contrary that

$$|D^N x(t)| > R_2$$

on a subset  $T \subset [s_i, t_i]$  of positive measure. Since  $\|D_0^{N-1} x\|_{\infty, a, b} \leq M$  we know from Proposition 3.5, part (d) that

$$(3.26) \quad \tilde{L}(t, D_0^{N-1} x(t), D^N x(t)) < L(t, D_0^{N-1} x(t), D^N x(t)), \quad \text{for } t \in T.$$

Then

$$\int_{s_i}^{t_i} L(t, D_0^{N-1}x(t), D^N x(t)) dt \leq \int_{s_i}^{t_i} L(t, D_0^{N-1}x_i(t), D^N x_i(t)) dt,$$

by optimality of  $x$ ,

$$= \int_{s_i}^{t_i} \tilde{L}(t, D_0^{N-1}x_i(t), D^N x_i(t)) dt,$$

since  $\|D^N x_i\|_{\infty, s_i, t_i} \leq R_1$  and by Proposition 3.5, part (d),

$$\leq \int_{s_i}^{t_i} \tilde{L}(t, D_0^{N-1}x(t), D^N x(t)) dt,$$

by optimality of  $x_i$ ,

$$< \int_{s_i}^{t_i} L(t, D_0^{N-1}x(t), D^N x(t)) dt,$$

by (3.26), and since  $L$  majorizes  $\tilde{L}$ . This contradiction implies (3.25).

To show that  $x$  satisfies the Euler inclusion, we note that the state function  $\xi = D_0^{N-1}x$  and associated control function  $u = D^N x$  (restricted to  $[s_i, t_i]$ ) solve a modified form of the optimal control problem (CP<sub>*i*</sub>) introduced earlier, in which  $L$  replaces  $\tilde{L}$  as cost integrand. Now apply the maximum principle [3, Theorem 5.2.3] at  $(\xi, u)$ . (The reader may check that the relevant hypotheses are satisfied; the essential boundedness of  $u$  is used here.) This tells us there exists a Lipschitz continuous function  $p$  ( $= \text{row}(p_1, \dots, p_N)$ ) which satisfies

$$-p \in p \cdot F - \partial_z L, \quad \text{a.e. on } [s_i, t_i]$$

and

$$(3.27) \quad p_N \in \partial_w L, \quad \text{a.e. on } [s_i, t_i],$$

$F$  is as in (3.16). These relationships imply (2.1).

Under the extra continuity and strict convexity hypotheses imposed in part (ii) of the theorem, the function

$$g(t) := \arg \min\{p_N(t) \cdot v - L(t, \xi(t), v) : |v| \leq R_2\}$$

is continuous on  $[s_i, t_i]$ . However strict convexity permits us to replace (3.27) by the statement that the functions  $u$  and  $g$  coincide a.e. on  $[s_i, t_i]$ . Thus  $u$  ( $= D^N x$ ) is continuous as claimed, following adjustment on a nullset.

Consider now the assertions in part (iii) of the theorem. These are proved by an adaptation of Weierstrass' classical argument, to accomodate the higher derivatives. Under the extra hypotheses the Euler inclusion is expressible as the following integral equation

$$(3.28) \quad \begin{aligned} L_w(t, D_0^{N-1}x(t), D^N x(t)) &= q(t) + \int_{s_i}^t L_{z_{N-1}}(\sigma, D_0^{N-1}x, D^N x) d\sigma \\ &- \int_{s_i}^t \int_{s_i}^{\sigma_1} L_{z_{N-2}}(\sigma_2, D_0^{N-1}x, D^N x) d\sigma_2 d\sigma_1 + \dots \end{aligned}$$

in which  $q$  is a polynomial function. The hypotheses ensure that the right side of (3.28) and also  $(t, w) \rightarrow L(t, D_0^{N-1}x(t), w)$  are  $C^1$  functions. Since  $L_{ww} > 0$ , the implicit function theorem tells us that  $D^N x$  is  $C^1$ . We may differentiate across equation (3.28) then and solve for  $D^{N+1}x$ ;

$$(3.29) \quad D^{N+1}x(t) = L_{ww}^{-1} \left[ -L_{t,w} - L_{z_0, \dots, z_{N-1}, w} \cdot D_1^N x(t) + Dq(t) + L_{z_{N-1}} - \int_{s_i}^t L_{z_{N-2}} d\sigma + \dots \right].$$

Suppose  $x$  is known to be  $C^{\alpha-1}$  for  $N + 1 \leq \alpha \leq r$ , then the right side of (3.29) is  $C^{\alpha-1-N}$  and hence so is  $D^{N+1}x$ , i.e.,  $x$  is  $C^\alpha$ . However  $x$  is  $C^N$  by part (ii) and it follows by induction that  $x$  is  $C^r$ .

#### 4. HYPOTHESES RESTRICTING POINTS OF BAD BEHAVIOUR AND A CONJECTURE REGARDING AUTONOMOUS PROBLEMS

Theorem 2.1 gives information about points at which a minimizer  $x$  is badly behaved, i.e. points in a neighbourhood of which  $D^N x$  is not locally essentially bounded; we are told that bad points are confined to a closed set of zero measure. It is interesting to know when we can further restrict the points of bad behaviour. Of special interest are hypotheses under which the set  $\Omega$  of Corollary 2.2. is the set  $[a, b]$ , i.e. we can eliminate bad points altogether. In such circumstances  $x$  has the property that  $D^N x$  is essentially bounded on the whole interval  $[a, b]$ , and  $x$  satisfies a strong form of the Euler inclusion, in which a single set of Lipschitz continuous functions  $p_i$  serves for the entire interval  $[a, b]$ .

In the case  $N = 1$ , many hypotheses are known which restrict the points of bad behaviour. A variety of hypotheses of this type are given in [5 and 6]. For example, if  $N = 1$  and the hypotheses (H1)–(H3) are supplemented by the requirement that  $L$  is polynomial, then points of bad behaviour are confined to a countable set having, at most, a finite number of cluster points. (See [8]). Concerning hypotheses which are known to eliminate the possibility of points of bad behaviour in the  $N = 1$  case we have the following.

**Theorem 4.1.** *Let  $x$  be a minimizer for (P). Suppose (in addition to hypotheses (H1)–(H3)) that  $N = 1$ , and either condition (i) or (ii) below is satisfied. Then  $\Omega = [a, b]$ .*

- (i) (*Morrey-type conditions*). *There exist some  $\alpha \in L^1[a, b]$  such that  $|\partial_z L(t, x(t), Dx(t))| \leq \alpha(t)$ , a.e.  $t \in [a, b]$ .*
- (ii) (*The autonomous case*).  *$L(t, z, w)$  is independent of the  $t$  variable.*

These results are proved in [5].

It is important to appreciate the role theorems such as Theorem 2.1 can play in establishing refined regularity properties of minimizers, under additional hypotheses. [5] provides numerous illustrations in the  $N = 1$  case. The regularity

information about  $x$  supplied by Theorem 2.1 can be put to work alongside the Euler inclusion (2.1) and extra hypotheses, to generate further regularity properties. Proof of the following theorem, which imposes Morrey type conditions to eliminate points of bad behaviour, illustrates how Theorem 2.1 can be used to obtain refined regularity results for problems where  $N > 1$ .

**Theorem 4.2.** *Let  $x$  be a minimizer for (P). Suppose (in addition to (H1)–(H3)) that there exists some  $\alpha \in L^1[a, b]$  such that*

$$|\partial_z L(t, D_0^{N-1}x(t), D^N x(t))| \leq \alpha(t), \quad \text{a.e. } t \in [a, b].$$

Then  $\Omega = [a, b]$ .

*Proof.* Take  $\bar{t} \in \Omega \cap (a, b)$ . We define

$$t_{\max} := \sup\{t > \bar{t} : D^N x \text{ is essentially bounded on } [\bar{t}, t]\}.$$

We show presently that  $D^N x$  is essentially bounded on  $[\bar{t}, t_{\max}]$ . This implies that  $t_{\max}$  is a regular point of  $x$ ; it follows from Theorem 2.1 that  $t_{\max} = b$  and  $x$  is essentially bounded on  $[\bar{t}, b]$ . Similar arguments applied to intervals to the left of the point  $\bar{t}$  establish that  $D^N x$  is essentially bounded on  $[a, \bar{t}]$  also. Putting these conclusions together, we deduce that  $D^N x$  is essentially bounded on  $[a, b]$ , i.e.,  $\Omega = [a, b]$ .

We turn now to the task of showing that  $D^N x$  is essentially bounded on  $[\bar{t}, t_{\max}]$ . Notice that  $t_{\max} > \bar{t}$  by Theorem 2.1. Let  $\{t_i\}$  be an increasing sequence in  $(\bar{t}, t_{\max}]$  such that  $t_i \rightarrow t_{\max}$ . Then there exists  $\varepsilon > 0$  such that  $[\bar{t}, \bar{t} + \varepsilon] \subset \Omega$  and  $\bar{t} + \varepsilon \in [\bar{t}, t_i]$  for all  $i$ .

On  $[\bar{t}, \bar{t} + \varepsilon]$ ,  $D^N x$  is essentially bounded and consequently the Euler inclusion is available to us. The inclusion may be expressed as follows: for each  $i$  there exists a Lipschitz continuous  $nN$  row vector function  $p^i = \text{row}(p_1^i, \dots, p_N^i)$  and an essentially bounded  $nN$  row vector function  $q_i$  such that

$$(4.1) \quad q_i(t) \in \partial_z L(t, D_0^{N-1}x(t), D^N x(t)), \quad -\dot{p}^i(t) = p^i(t) \cdot F - q_i(t),$$

$$(4.2) \quad p_N^i(t) \cdot w - L(t, D_0^{N-1}x(t), w) \leq p_N^i(t) \cdot D^N x(t) - L(t, D_0^{N-1}x(t), D^N x(t))$$

for all  $w \in \mathbf{R}^n$ , a.e.  $t \in [\bar{t}, t_i]$ . ( $F$  is as in (3.16).)

Since  $L$  is convex in its final argument, (4.2) implies

$$(4.3) \quad p^i(t) \cdot G (= p_N^i(t)) = r_i(t), \quad \text{a.e. } t \in [\bar{t}, t_i]$$

for some essentially bounded  $n$  row vector function  $r_i$  satisfying

$$r_i(t) \in \partial_w L(t, D_0^{N-1}x(t), D^N x(t)), \quad \text{a.e. } t \in [\bar{t}, t_i]$$

( $G$  is as in (3.16)).

From (4.1) and the variation of constants formula, we deduce that

$$p^i(t) = p^i(\bar{t}) \cdot e^{-F(t-\bar{t})} + s_i(t), \quad \text{all } t \in [\bar{t}, t_i],$$

where  $s_i(t) := \int_{\bar{t}}^t q_i(s)e^{-F(t-s)} ds$ . Postmultiplying the penultimate relation by  $GH(t)$ , where

$$H(t) := G'e^{-F'(t-\bar{t})},$$

integrating over  $[\bar{t}, \bar{t} + \varepsilon]$ , and taking note of (4.3), we obtain

$$(4.4) \quad p^i(\bar{t}) \cdot W(\bar{t}, \bar{t} + \varepsilon) = \int_{\bar{t}}^{\bar{t}+\varepsilon} (r_i(t) - s_i(t)G)H(t) dt.$$

Here  $W(\bar{t}, \bar{t} + \varepsilon)$  is the controllability grammian of  $(-F, G)$ , namely

$$W(\bar{t}, \bar{t} + \varepsilon) := \int_{\bar{t}}^{\bar{t}+\varepsilon} e^{-Fs} GG'e^{-F's} ds.$$

However the right side of (4.4) is norm bounded by a number  $c$  which does not depend on  $i$ , since  $r_i$  and  $q_i$ , hence  $s_i$ , are expressible in terms of selectors of  $\partial_z L$  and  $\partial_w L$ , and these generalized gradients are bounded on  $[\bar{t}, \bar{t} + \varepsilon]$ . Concerning the left side, we note that  $(-F, G)$  is controllable, whence  $W(\bar{t}, \bar{t} + \varepsilon)$  is nonsingular. We conclude then

$$|p^i(\bar{t})| \leq \| [W(\bar{t}, \bar{t} + \varepsilon)]^{-1} \| c$$

( $\|W^{-1}\|$  denotes the operator norm). Bearing in mind now the integrable bound on  $\partial_z L$ , and since  $q_i$  is a selector of  $\partial_z L$ , we deduce from (4.1) and Gronwall's lemma that

$$\|p^i\|_{\infty, \bar{t}, t_i} < c_1.$$

Here  $c_1$  is another number which does not depend on  $i$ .

Now (4.2) implies

$$L(t, D_0^{N-1}x(t), D^N x(t)) - p_N^i(t) \cdot D^N x(t) \leq |p_N^i| + c_2$$

where

$$c_2 = \sup\{L(t, D_0^{N-1}x(t), w) : t \in [a, b], |w| \leq 1\}.$$

By hypothesis (H3) and (4.5) we have

$$-\alpha M + \theta(|D^N x(t)|) - c_1 |D^N x(t)| \leq c_1 + c_2,$$

a.e.  $t \in [\bar{t}, t_i]$  where  $M$  is an upper bound for  $|D_0^{N-1}x(t)|$ . It follows that

$$\|D^N x\|_{\infty, \bar{t}, t_i} \leq R,$$

where  $R$  is a number such that

$$\theta(r)/r > c_1 + (c_1 + c_2 + \alpha M)/r$$

whenever  $r > R$ . But since  $t_i \rightarrow t_{\max}$  and  $R$  does not depend on  $i$ ,

$$\|D^N x\|_{\infty, \bar{t}, t_{\max}} \leq R.$$

We have exhibited the desired bound.

We note the following corollary of Theorem 4.2, which is of interest because the hypotheses involved are directly verifiable.

**Corollary 4.3.** *Let  $x$  be a minimizer for (P). Suppose (in addition to (H1)–(H3)) that there exist an integrable function  $\gamma$ , a locally bounded function  $r$  and a constant  $c_1$  such that*

$$(4.5) \quad |\partial_z L(t, z, w)| \leq c_1(L(t, z, w) + |w|) + \gamma(t)r(z)$$

a.e.  $t \in [a, b]$ , for  $(z, w) \in \mathbf{R}^{n \times N} \times \mathbf{R}^n$ . Then  $\Omega = [a, b]$ .

*Proof.* We have merely to note that the hypotheses are strengthened versions of those in Theorem 4.2. In particular, the right side of (4.5) evaluated along a minimizer  $x$  gives

$$c_1(L(t, D_0^{N-1}x(t), D^N x(t)) + |D^N x(t)|) + \gamma(t)r(D_0^{N-1}x(t)).$$

All the terms here are integrable, by the properties of minimizers. This deals with the last hypothesis in Theorem 4.2.

We have in Theorems 2.1 and 4.2 two instances where previously known regularity results treating the  $N = 1$  case have counterparts for problems with  $N > 1$ , and they encourage us to look for others. However we now advance strong evidence that, at least in one important respect, our quest will be unsuccessful, and that problems where  $N > 2$  can exhibit pathologies not present when  $N = 1$ .

Consider the class of autonomous problems which satisfy the hypotheses (H1)–(H3). We conjecture that minimizers can have points of bad behaviour when  $N > 1$ . Recall that, by contrast, minimizers for autonomous problems have no points of bad behaviour when  $N = 1$ . (See Theorem 4.1.)

The following proposition falls somewhat short of proving this conjecture. It does however give evidence of it by exhibiting an extremal with a bad point at  $t = 0$  for an autonomous problem where hypotheses (H1)–(H3) are in force and where  $N = 2$ . Extremals (i.e. arcs satisfying necessary conditions of optimality, in this case the Euler equation (2.2) and the boundary conditions) are not guaranteed to be minimizers, but are strong candidates for being so.

We offer as an open research problem verification that the extremal involved is a minimizer (for  $\varepsilon > 0$  appropriately chosen). Techniques of [7] should be helpful. Application of field theory to this problem appears to be problematic however, because of the complications introduced by the second order derivative.

We define the function  $f_\varepsilon: \mathbf{R} \times \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}$  to be

$$f_\varepsilon(z_0, z_1, w) := |z_0^2 - z_1^2|^2 |w|^{22} + \varepsilon |w|^2$$

for  $\varepsilon > 0$ . The problem below, formulated in terms of  $f_\varepsilon$ , is a second-order extension of the Ball-Mizel problem which provided a significant example in the case  $N = 1$  (see [7]).

**Proposition 4.4.** *There exists a number  $\varepsilon^* > 0$  and a  $C^\infty$  function  $k: (-\varepsilon^*, \varepsilon^*) \rightarrow \mathbf{R}$  with the following properties. For  $\varepsilon > 0$ , take the problem  $P(\varepsilon)$  to be*

$$\left\{ \begin{array}{l} \text{Minimize } \int_0^1 f_\varepsilon(y(t), \dot{y}(t), \ddot{y}(t)) dt \\ \text{over arcs } y \in W^{2,1}([0, 1]; \mathbf{R}) \\ \text{which satisfy} \\ x(0) = 0, \quad \dot{x}(0) = 0, \quad x(1) = k(\varepsilon), \quad \dot{x}(1) = (5/3)k(\varepsilon). \end{array} \right.$$

Then for each  $\varepsilon \in (0, \varepsilon^*)$ ,  $P(\varepsilon)$  satisfies (H1)–(H3) and the function

$$(4.6) \quad x(t) = k(\varepsilon)t^{5/3}$$

satisfies both the boundary conditions and the equation

$$(4.7) \quad \left( L_{z_0} - \frac{d}{dt}L_{z_1} + \frac{d^2}{dt^2}L_w \right) (x(t), \dot{x}(t), \ddot{x}(t)) = 0, \quad \text{for } 0 < t \leq 1,$$

where  $L = f_\varepsilon$ .

*Proof.* The data of problem  $P(\varepsilon)$  clearly satisfies hypotheses (H1)–(H3) for  $\varepsilon > 0$ . We see that  $x(t)$  given by (4.6) satisfies the boundary conditions, regardless of the manner in which we choose  $k(\varepsilon)$ .

Substitution of  $x(t) = kt^{5/3}$  into the left side of (4.7) yields

$$c(F(k) + \varepsilon)t^{-7/3}.$$

Here  $c$  is a nonzero constant and  $F$  is a polynomial. Set  $\bar{k} = (\frac{3}{5})^{5/3}$ . We find that

$$F(\bar{k}) = 0 \quad \text{and} \quad \frac{d}{dk}F(\bar{k}) \neq 0.$$

By the implicit function theorem there exists  $\varepsilon^* > 0$  and a  $C^\infty$  function  $k: (-\varepsilon^*, +\varepsilon^*) \rightarrow \mathbf{R}$  for which

$$F(k(\varepsilon)) + \varepsilon = 0, \quad \text{for all } \varepsilon \in (-\varepsilon^*, +\varepsilon^*).$$

Making this choice of function  $k(\varepsilon)$ , we see that  $x(t)$  given by (4.6) satisfies equation (4.7) for each  $\varepsilon \in (-\varepsilon^*, +\varepsilon^*)$ , as required.

*A final remark.* The conjecture that  $x(t) = k(\varepsilon)t^{5/3}$  is a minimizer for  $P(\varepsilon)$  is not inconsistent with Theorem 4.2. Notice that, for problem  $P(\varepsilon)$ ,

$$L_{z_0} \sim t^{-7/3}$$

along  $x$ , i.e.,  $L_{z_0}$  is not integrably bounded. Thus the hypotheses of Theorem 4.2 are violated, and so Theorem 4.2 does not rule out the possible existence of points of bad behaviour.



## REFERENCES

1. L. D. Berkovitz, *Lower semicontinuity of integral functionals*, Trans. Amer. Math. Soc. **192** (1974), 51–57.
2. L. Cesari, *Optimization—Theory and applications*, Springer-Verlag, New York, 1983.
3. F. H. Clarke, *Optimization and nonsmooth analysis*, Wiley Interscience, New York, 1983. (Reprinted 1989 by CRM, University of Montreal, POB 6128-A, Montreal (Qc) Canada H3C 3J7.)
4. F. H. Clarke and P. D. Loewen, *An intermediate existence theory in the calculus of variations*, Centre de Recherches Mathématiques, Report CRM-1418, April 1987, Ann. Scuola Norm. Sup. (to appear).
5. F. H. Clarke and R. B. Vinter, *Regularity properties of solutions to the basic problem in the calculus of variations*, Trans. Amer. Math. Soc. **289** (1985), 73–98.
6. —, *Existence and regularity in the small in the calculus of variations*, J. Differential Equations **59** (1985), 336–354.
7. —, *On conditions under which the Euler equation or the maximum principle hold*, Appl. Math. Optim. **12** (1984), 73–79.
8. —, *Regularity of solutions to variational problems with polynomial Lagrangians*, Bull. Polish Acad. Sci. **34** (1986), 73–81.
9. R. T. Rockafellar, *Existence theorems for general control problems of Bolza and Lagrange*, Adv. in Math. **15** (1975), 312–333.
10. D. R. Smith, *Variational methods in optimization*, Prentice-Hall, Englewood Cliffs, N. J., 1974.
11. L. Tonelli, *Sur une méthode directe du calcul des variations*, Rend. Circ. Math. Palermo **39** (1915), 233–264. Also appears in *Opere scelte*, vol. 2, Cremonese, Rome, 1961, pp. 289–333.
12. —, *Fondamenti di calcolo delle variazioni*, vols. 1, 2, Zanichelli, Bologna, 1921, 1923.

CENTRE DE RECHERCHES MATHÉMATIQUES (CRM), UNIVERSITÉ DE MONTRÉAL, C. P. 6128-A, MONTRÉAL, QUÉBEC, CANADA H3C 3J7

DEPARTMENT OF ELECTRICAL ENGINEERING, IMPERIAL COLLEGE, EXHIBITION ROAD, LONDON, SW7 2BT, ENGLAND