JOURNAL OF APPROXIMATION THEORY 25, 142–152 (1979)

Perfect Splines and Nonlinear Optimal Control Theory

GUNNAR ARONSSON

Department of Mathematics, University of Uppsala, Uppsala, Sweden

Communicated by Carl de Boor

Received July 26, 1977

In this paper we use a method from nonlinear optimal control theory to establish the "perfect spline" properties of a solution to a certain extremum problem. The problem is to minimize the L^∞ norm of a nonlinear expression of the form $F(t, x(t), \dot{x}(t), \ddot{x}(t),..., x^{(n)}(t))$ over all sufficiently smooth functions $x(t)$ which satisfy given boundary conditions. Under suitable assumptions, we show that a solution $x_0(t)$ must be such that $F(t, x_0(t), \dot{x}_0(t),..., x_0^{(n)}(t))$ is constant, and $x_0^{(n)}(t)$ is piece-wise continuous with a finite number of jump discontinuities. This generalizes results by D. S. Carter, G. Glaeser, D. McClure, and others, who studied the same problem for linear differential expressions.

1. INTRODUCTION

It is well known that many types of extremum problems have solutions that are spline functions. We will discuss a few specific results for L^∞ problems, and state first some definitions.

A real-valued function $f(t)$ defined on an interval $a \leqslant t \leqslant b$ is said to be a *spline* of degree n, if there is a partition

$$a = t_0 < t_1 < \cdots < t_k < t_{k+1} = b$$

such that:

(i) f is a polynomial of degree $\leqslant n$ on (t_j, t_{j+1}) for $j = 0, 1,..., k$, and
(ii) $f \in C^{n-1}[a, b]$.

The points $\{t_j\}_1^k$ are the *knots* of f. Such a function f is called a *perfect spline* if, furthermore,

(iii) $|f^{(n)}(t)|$ is constant on $[a, b]$ for $t \neq t_1, t_2,..., t_k$.

The following two theorems by G. Glaeser [11, 1967] form a good starting-point for this paper:

THEOREM A. *If* $2n$ *real values* $x_0^{(\nu)}, x_1^{(\nu)}$ *for* $\nu = 1, 2,..., n$ *are specified,*

0021-9045/79/020142-11$02.00/0
Copyright © 1979 by Academic Press, Inc.
All rights of reproduction in any form reserved.

then the two-point Hermite interpolation problem $f^{(\nu)}(a) = x_0^{(\nu+1)}$, $f^{(\nu)}(b) = x_1^{(\nu+1)}$, *for* $\nu = 0, 1,..., n-1$ *has a unique solution* $f_0(t)$ *that is a perfect spline of degree n having at most* $n-1$ *knots in* (a, b).

THEOREM B. *The perfect spline* f_0 *of the above theorem is the unique function that minimizes* $\|f^{(n)}\|_{L^\infty}$ *among all functions f which satisfy* $f^{(\nu)}(a) = x_0^{(\nu+1)}$, $f^{(\nu)}(b) = x_1^{(\nu+1)}$ *for* $\nu = 0, 1,..., n-1$ *and are such that* $f^{(\nu)}(t)$ for $\nu = 0, 1,..., n-1$ *are absolutely continuous on* $[a, b]$.

These theorems by Glaeser have been extended by S. Karlin [14; 1973] and D. McClure [17; 1975]. We will consider more closely some results of McClure [17]. Before stating his results, we introduce some notations and definitions. We introduce the Sobolev space

$$\begin{aligned} W^{n,\infty} &= W^{n,\infty}[a, b] \\ &= \{f \in R[a, b] \mid f^{(\nu)} \text{ is absolutely continuous for } \nu = 0, 1,..., n-1 \\ &\qquad\qquad \text{and } \|f^{(n)}\|_{L^\infty} < \infty\}. \end{aligned}$$

If $x_0^{(\nu)}$, $x_1^{(\nu)}$ for $\nu = 1, 2,..., n$ are given real values we consider

$$U = \{f \in W^{n,\infty} \mid f^{(\nu)}(a) = x_0^{(\nu+1)}, f^{(\nu)}(b) = x_1^{(\nu+1)} \text{ for } \nu = 0, 1,..., n-1\}.$$

We next introduce a differential operator

$$A = D^n + \sum_{\nu=1}^{n} a_\nu(t)\, D^{n-\nu}$$

where $D = d/dt$. We assume that $a_\nu(t) \in C^{n-\nu}[a, b]$ for $\nu = 1, 2,..., n$, and the adjoint A^* is given by

$$A^*\phi = (-1)^n D^n\phi + \sum_{\nu=1}^{n} (-1)^{n-\nu} D^{n-\nu}(a_\nu\phi).$$

We can now consider the problem to minimize $\|Af\|_{L^\infty}$ over U. Put $\alpha_0 = \inf_{f\in U} \|Af\|_{L^\infty}$. Now that A is defined, we say that a function $f \in W^{n,\infty}$ is an *A-spline* if there is a subdivision $a = t_0 < t_1 < \cdots < t_k < t_{k+1} = b$ such that $(Af)(t)$ is constant on each (t_j, t_{j+1}) for $j = 0, 1,..., k$. If, furthermore, $|(Af)(t)| =$ constant on $[a, b]$ for $t \neq t_1, t_2,..., t_k$, then f is a *perfect A-spline*. The points t_j, $j = 1, 2,..., k$, are still called *knots*.

Now McClure introduces a disconjugacy condition, called Property T, for A^*. The operator A^* is said to possess the Property T, if any nontrivial solution ϕ of $A^*\phi = 0$ on $[a, b]$ has at most $n-1$ zeros in $[a, b]$. Observe that $A^* = (-1)^n D^n$ (Glaeser's case) has Property T. We can then state

THEOREM C. (D. McClure [17]). *There is a unique function* $f_0 \in U$ *such* *that* $\| Af_0 \|_{L^\infty} = \alpha_0$. *The function* f_0 *is a* perfect A-spline *on* $[a, b]$ *with knots* $\{t_j\}_1^k$. *Thus* $(Af_0)(t) = \pm\alpha_0$, *with a fixed sign on each sub-interval* (t_j, t_{j+1}), $j = 0, 1,..., k$. *Further, if* A^* *possesses Property T on* $[a, b]$, *then* $k \leqslant n - 1$.

Furthermore, McClure [17] proves

THEOREM D. *If* A^* *has Property T on* $[a, b]$, *then there exists a* unique *perfect A-spline* f_0 *in U with at most* $n - 1$ *knots in* (a, b).

This shows the analogy with Glaeser's theorems. McClure [17] proves three more theorems on perfect A-splines. His paper is based on linear control theory, whereas we will use here a quite different method from nonlinear control theory.

Although the results of Glaeser and McClure provide a very good background for our theorem, it seems necessary to mention the work of some others. Parts of the results of Glaeser and McClure are contained in earlier papers, and we will refer to a few of them.

First, we mention the paper by D. S. Carter [5; 1957]. In fact, Theorem 2 in [5], p. 140, implies Theorem C above, except that Carter gives no explicit bound for the number of knots. Carter's conditions on the differential operator A are weaker than the conditions imposed by McClure. It seems that Carter's paper has not received the attention that it deserves.

Next, the paper by W. T. Reid [19; 1962] treats minimum L^p norm $(1 < p \leqslant \infty)$ problems for linear differential operators. For $p = \infty$, the perfect spline properties of a solution can be concluded from [19] by some effort (see pp. 603–605).

Finally, we mention the papers [9; 1974] and [8; 1974] by S. D. Fisher and J. W. Jerome. In these papers, generalized splines are obtained as solutions to L^∞ extremum problems where interpolation conditions are prescribed also at given points, interior to the basic interval $[a, b]$. We will not discuss the relationship of [9] and [8] to McClure [17], but refer to the interesting survey book [10; 1975; in particular sections 6 and 7] by Fisher and Jerome.

So far, we have only given background material for the linear case. But our theorem treats the problem of minimizing (over U)

$$\| F(t, x(t), \dot{x}(t),..., x^{(n)}(t))\|_{L^\infty}$$

for a fairly general, nonlinear F. The literature on this problem is very sparse, but in Section 4 we will make a few comments on its relation to our present work.

Finally, we mention two more papers, which connect splines and optimal control, namely O. L. Mangasarian and L. L. Schumaker [16, 1969]; and I. J. Schoenberg [20; 1971]. These papers treat linear control systems and functionals representable by integrals.

2. Perfect Spline Properties in the Nonlinear Case

We will consider the problem of minimizing ess $\sup_{a<t<b} | F(t, x(t),..., x^{(n)}(t))|$ over U. We omit the modulus sign, since we can just as well consider $G(\cdots) = F(\cdots)^2$. We thus study the functional

$$H(x) = \operatorname*{ess\,sup}_{a<t<b} F(t, x(t), \dot{x}(t),..., x^{(n)}(t)).$$

Naturally, we must impose conditions on the function $F = F(t, y_0, y_1,..., y_n)$, and besides smoothness we need conditions concerning its dependence on y_n.

We make the following assumptions:

(a) $F \in C^1([a, b] \times R^{n+1})$

(b) there is a function $\omega \in C([a, b] \times R^n)$ such that

$$\frac{\partial F}{\partial y_n} \text{ is } \begin{cases} > 0 & \text{if } \; y_n > \omega(t, y_0, y_1,..., y_{n-1}) \\ = 0 & \text{if } \; y_n = \omega(t, y_0,..., y_{n-1}) \\ < 0 & \text{if } \; y_n < \omega(t, y_0,..., y_{n-1}) \end{cases}$$

(c) $\lim_{|y_n|\to\infty} F(t, y_0, y_1,..., y_n) = +\infty$

for arbitrary fixed $(t, y_0,..., y_{n-1}) \in [a, b] \times R^n$.

These conditions *can* be relaxed, but they seem convenient here. Observe that our conditions are satisfied in the linear case (if all coefficients $a_\nu(t) \in C^1$), since we have then

$$F(t, y_0,..., y_n) \equiv \left(y_n + \sum_{\nu=0}^{n-1} a_\nu(t)\, y_\nu \right)^2,$$

and $\omega(t, y_0,..., y_{n-1}) = -\sum_{\nu=0}^{n-1} a_\nu(t)\, y_\nu$.

Introduce the "minimum function" $m(t, y_0, y_1,..., y_{n-1}) \equiv F(t, y_0 y_1,..., y_{n-1}, \omega(t, y_0, y_1,..., y_{n-1}))$. Clearly, $m \in C([a, b] \times R^n)$.

We can now state our main result.

Theorem. *Let $x_0(t)$ minimize $H(x)$ over U. Put $H(x_0) = M_0$. Assume that*

$$m(t, x_0(t), \dot{x}_0(t),..., x_0^{(n-1)}(t)) < M_0 \qquad (*)$$

for $a \leqslant t \leqslant b$.

Then there are a finite number of points $\{t_k\}_1^N$ such that

$$a < t_1 < t_2 < \cdots < t_N < b$$

and such that:

(i) $x_0(t) \in C^{n+1}$ *on* $[a, b]\backslash\{t_k\}_1^N$.

(ii) $F(t, x_0(t), \dot{x}_0(t), \ldots, x_0^{(n)}(t)) = M_0$ *on* $[a, b]\backslash\{t_k\}_1^N$.

(iii) $x_0^{(n)}(t)$ *has a jump discontinuity at each* t_k.

(By definition $x_0(t) \in C^{n-1}[a, b]$.*)*

We have thus good reasons for calling $x_0(t)$ a perfect spline, or a "perfect F-spline", with knots $\{t_k\}_1^N$.

The condition (*) is superfluous in the linear case, but essential here, as we will show by a trivial example.

Choose $F \equiv t + \sum_{\nu=0}^{n} y_\nu^2$, $a = 0$, $b = 1$, and all boundary data zero. Clearly, $\min_{x \in U} H(x) = 1$, and any $x(t) \in U$ such that

$$t + \sum_{\nu=0}^{n} (x^{(\nu)}(t))^2 \leqslant 1 \text{ a.e.}$$

is optimal. Thus, the theorem does not apply here, and the reason is that (*) does not hold at $t = 1$.

3. Proof of the Theorem

We divide the proof into four parts.

(I) *Transformation of the Problem into Convenient Control Form*

Let $\tilde{x}(t) \in W^{n,\infty}$. Introduce a vector $x = x(t) \in R^n$ by the identification $\tilde{x}^{(\nu)}(t) = x_{\nu+1}(t)$, $\nu = 0, 1, \ldots, n-1$. Then $\dot{x}_i = x_{i+1}$ for $i = 1, 2, \ldots, n-1$. We must represent $\dot{x}_n = \tilde{x}^{(n)}$ in a way that is convenient for our purpose.

Consider the curve in $[a, b] \times R^n$:

$$\gamma = \{(t, x_0(t), \dot{x}_0(t), \ldots, x_0^{(n-1)}(t)) \mid a \leqslant t \leqslant b\}.$$

According to (*) we have $m(t, x) < M_0$ on γ, and by continuity there is a neighbourhood V of γ such that $m(t, x) < M_0$ in V. Now take an arbitrary $(t, x) \in V \subset [a, b] \times R^n$. Then, because of our assumptions on the function F, the equation

$$F(t, x_1, x_2, \ldots, x_n, y) = M_0$$

has exactly two solutions $y = \psi_1(t, x)$ and $y = \psi_2(t, x)$. Let us agree that $\psi_1(t, x) < \psi_2(t, x)$. Then we have $\psi_1(t, x) < \omega(t, x) < \psi_2(t, x)$ and, clearly, ψ_1 and ψ_2 are in $C^1(V)$.

We now introduce a scalar control variable u and write our control system as follows:

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = x_3$$
$$------$$
$$\dot{x}_{n-1} = x_n$$
$$\dot{x}_n = \tfrac{1}{2}(\psi_2(t, x) + \psi_1(t, x)) + \tfrac{1}{2}(\psi_2(t, x) - \psi_1(t, x)) \cdot u \equiv f(t, x, u).$$

In vector form we write $\dot{x} = g(t, x, u)$. From now on, *we restrict our attention to* $(t, x) \in V$ and observe that $g(t, x, u) \in C^1(V \times R)$. Because of our construction, we have $F(t, x, f(t, x, u)) \leqslant M_0$ if and only if $|u| \leqslant 1$ and $F(\cdots) < M_0$ if and only if $|u| < 1$. Furthermore, possibly after reducing V, the following holds: for any ξ, $0 < \xi < 1$, we have

$$\sup\{F(t, x, f(t, x, u)) \mid (t, x) \in V, |u| \leqslant \xi\} < M_0 .$$

We draw the following important conclusion: if the control system $\dot{x} = g(t, x, u)$ can be steered between the prescribed endpoints, without leaving V, by a control function $u(\cdot)$, such that $\| u(\cdot)\|_{L^\infty} < 1$, then this will contradict the definition of M_0, since we would then obtain an $\tilde{x} \in U$ such that $H(\tilde{x}) < M_0$.

(II) *Proof that the Maximum Principle is Applicable Here*

We shall use the technique of "elementary perturbations" of the control. For this, we refer to Pontryagin et al. [18], Chap. 2, or (preferably) Lee and Markus [15], pp. 247–252. We shall use the terminology of [15]. Return to our optimal function $x_0(t)$ and let $\bar{x}(t)$ be the vector function $(x_0(t), \dot{x}_0(t),..., x_0^{(n-1)}(t))^T$. Then $\dot{\bar{x}}(t) = g(t, \bar{x}(t), \bar{u}(t))$, with a uniquely defined (a.e.) control function $\bar{u}(t)$. Further, we have $-1 \leqslant \bar{u}(t) \leqslant +1$ (a.e.), since $H(x_0) = M_0$. Put $\Omega = [-1, 1] =$ the control restraint set. We can now define "elementary perturbations" of $\bar{u}$ and $\bar{x}$, and consider the "tangent perturbation cone" K_t exactly as in [15], pp. 247–249. The fact that our system is not autonomous, has no importance here. The results of [15] carry over. We are interested in the final perturbation cone K_b. If K_b is *not* the whole tangent space at $\bar{x}(b)$, then it is contained in a halfspace bounded by a hyperplane through the origin, and then the maximum principle will follow; see [15], pp. 254–255.

Suppose then that K_b *is* the whole tangent space at $\bar{x}(b)$. But then there exists a control $u_{\xi_1}(t)$ such that $\| u_{\xi_1} \|_{L^\infty} \leqslant \xi_1 < 1$, and such that u_{ξ_1} steers our system $\dot{x} = g(t, x, u)$ between the prescribed endpoints without leaving V, and this gives a contradiction, according to (I). The existence of $u_{\xi_1}(t)$ follows by exactly the same arguments as in the author's own paper [4], pp. 61-63. We use a "shrinking trick", plus a topological covering argument, and ξ_1 can be thought of as the "shrinking parameter".

(III) *The Maximum Principle and the Adjoint System*

Thus the maximum principle holds for $\bar{x}(t), \bar{u}(t)$ and the control system $\dot{x} = g(t, x, u)$. Hence there is a nontrivial solution $\eta(t) = (\eta_1(t),..., \eta_n(t))$ of the system

$$\dot{\eta} = -\eta \frac{\partial g}{\partial x}(t, \bar{x}(t), \bar{u}(t)) \tag{1}$$

such that

$$\eta(t)\, g(t, \bar{x}(t), \bar{u}(t)) = \max_{|u|\leqslant 1} (\eta(t)\, g(t, \bar{x}(t), u)) \text{ a.e.}$$

(The system (1) is adjoint to the variational system for $\dot{x} = g(t, x, \bar{u}(t))$, along $\bar{x}(t)$.) From the form of the function $g(t, x, u)$ we conclude at once that

$$\bar{u}(t) = \operatorname{sign} \eta_n(t) \qquad \text{a.e. on the set, where } \eta_n(t) \neq 0.$$

We must therefore study the zeros of $\eta_n(t)$. Clearly, the adjoint system (1) has the form

$$\begin{aligned}
\frac{d\eta_1}{dt} &= -\frac{\partial f}{\partial x_1}(t, \bar{x}(t), \bar{u}(t))\, \eta_n \\
\frac{d\eta_2}{dt} &= -\eta_1 - \frac{\partial f}{\partial x_2}(t, \bar{x}(t), \bar{u}(t))\, \eta_n \\
\frac{d\eta_3}{dt} &= -\eta_2 - \frac{\partial f}{\partial x_3}(\cdots)\, \eta_n \\
&\text{-- -- -- -- -- -- -- -- -- --} \\
\frac{d\eta_n}{dt} &= -\eta_{n-1} - \frac{\partial f}{\partial x_n}(\cdots)\, \eta_n
\end{aligned}$$

Now let $\eta_n(t) > 0$ on some interval $I \subset [a, b]$. From $\bar{u} = \operatorname{sign} \eta_n$ we conclude that

$$x_0^{(n)}(t) = \psi_2(t, x_0(t), \dot{x}_0(t),..., x_0^{(n-1)}) \text{ on } I.$$

If $\eta_n(t) < 0$, we find $x_0^{(n)} = \psi_1(\cdots)$ on I. Since $\psi_1, \psi_2 \in C^1(V)$ it follows in both cases that $x_0(\cdot) \in C^{n+1}(I)$, and $F(t, x_0(t),..., x_0^{(n)}(t)) = M_0$ on I. Thus the theorem will follow if we can prove that $\eta_n(t)$ has only a finite number of zeros.

(IV) *Proof that the Number of Switches is Finite*

We consider now $\eta(t)$ as a column vector and write the adjoint system as

$$\dot{\eta} = -A_0\eta - \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t))\, \eta_n$$

where

$$A_0 = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ - & - & - & - & - \\ 0 & 0 & 0 & \cdots 1 & 0 \end{bmatrix}$$

Assume that $\eta_n(t)$ has an infinity of zeros on $[a, b]$. After reversing the t-axis we may write

$$\dot{\eta} = A_0\eta + \varphi(t)\, \eta_n$$

where $\varphi(t) \in L^\infty$ is an $(n \times 1)$ vector function. We can assume that the zeros of $\eta_n(t)$ cluster at $t = 0$. We have

$$\eta(t) = e^{A_0 t}\eta(0) + \int_0^t e^{A_0(t-s)}\varphi(s)\, \eta_n(s)\, ds.$$

Now the matrix A_0 is nilpotent and $e^{A_0 t}$ is easily computed. In fact, we have (see [12], p. 99)

$$e^{A_0 t} = \begin{bmatrix} 1 & 0 & 0 & & & \\ t & 1 & 0 & & & \\ \frac{t^2}{2} & t & 1 & & & \\ - & - & - & - & - & - \\ \frac{t^{n-1}}{(n-1)!} & \frac{t^{n-2}}{(n-2)!} & \frac{t^{n-3}}{(n-3)!} & \cdots \frac{t^2}{2} & t & 1 \end{bmatrix}$$

By taking the n:th component in the above equation for $\eta(t)$ we find

$$\eta_n(t) = P(t) + \int_0^t (e^{A_0(t-s)}\varphi(s))_n\, \eta_n(s)\, ds \tag{2}$$

where the polynomial $P(t) \not\equiv 0$, since $\eta(0) \neq 0$. Let $c_k t^k$ be the lowest order term in $P(t)$. Thus $|\eta_n(t)| \leqslant c_k' \,|t|^k + d_k \,|\int_0^t |\eta_n(s)|\, ds|$ for some constants $c_k' > 0$ and d_k. From the generalized Gronwall inequality ([12], p. 36) we infer that $\eta_n(t) = O(t^k)$. But then the integral in (2) is $O(t^{k+1})$. Hence it follows from (2) and our choice of k that the zeros of η_n can *not* cluster at $t = 0$.

The contradiction completes the proof.

4. Various Remarks

(a) *The Linear Case*

Consider the problem to minimize

$$\operatorname*{ess\,sup}_{a<t<b} \left(x^{(n)}(t) + \sum_{\nu=0}^{n-1} a_\nu(t)\, x^{(\nu)}(t) \right)^2$$

over U. Let all $a_\nu(t) \in C^1[a, b]$. Here, we can apply standard arguments to the control formulation and deduce that a minimizing function $x_0(\cdot)$ exists (this need not hold in the general nonlinear case). See McClure [17], pp. 229, 235.

If $M_0 = 0$, then $x_0^{(n)}(t) + \sum_{\nu=0}^{n-1} a_\nu(t)\, x_0^{(\nu)}(t) = 0$, and x_0 is obviously unique.

Now let $M_0 > 0$. Then, clearly, the condition (*) holds, and our theorem is applicable. The functions ψ_1 and ψ_2 (in part I of the proof) are:

$$\psi_1(t, x) = - \sum_{\nu=0}^{n-1} a_\nu(t)\, x_{\nu+1} - (M_0)^{1/2}$$

and

$$\psi_2(t, x) = - \sum_{\nu=0}^{n-1} a_\nu(t)\, x_{\nu+1} + (M_0)^{1/2}.$$

On an interval where $\bar{u}(t) = +1$, we have $f(t, \bar{x}(t), \bar{u}(t)) = \psi_2(t, \bar{x}(t))$, and if $\bar{u}(t) = -1$, then $f(\cdots) = \psi_1(\cdots)$.

Consequently, the adjoint system for η in both cases has the form

$$\begin{aligned}
\dot{\eta}_1 &= a_0(t)\, \eta_n \\
\dot{\eta}_2 &= -\eta_1 + a_1(t)\, \eta_n \\
\dot{\eta}_3 &= -\eta_2 + a_2(t)\, \eta_n \\
&\cdots\cdots\cdots\cdots \\
\dot{\eta}_{n-1} &= -\eta_{n-2} + a_{n-2}(t)\, \eta_n \\
\dot{\eta}_n &= -\eta_{n-1} + a_{n-1}(t)\, \eta_n
\end{aligned}$$

Assume that $a_\nu(t) \in C^\nu[a, b]$ for $\nu = 2, 3, \ldots, n-1$. Clearly $\eta(t) \in C^1[a, b]$. The last equation implies $\eta_n \in C^2$, and

$$\eta_n^{(2)} = \eta_{n-2} - (a_{n-2}\eta_n) + D(a_{n-1}\eta_n)$$

by the equation for $\dot{\eta}_{n-1}$. Hence $\eta_n \in C^3$, and

$$\eta_n^{(3)} = -\eta_{n-3} + (a_{n-3}\eta_n) - D(a_{n-2}\eta_n) + D^2(a_{n-1}\eta_n)$$

by the equation for $\dot{\eta}_{n-2}$. Since all $a_\nu \in C^\nu$, we can continue in this way and find, eventually,

$$\eta_n^{(n)} + \sum_{\nu=0}^{n-1} (-1)^{n-\nu} D^\nu(a_\nu \eta_n) = 0.$$

Now, if $A \equiv D^n + \sum_{\nu=0}^{n-1} a_\nu(t) D^\nu$, then this equation is (not surprising!) the adjoint equation for η_n.

We thus obtain Theorem C (by McClure) of Section 1 again, except that the uniqueness of $x_0(\cdot)$ does not follow from our theorem.

If the disconjugacy condition "Property T" is omitted from Theorem C, then no explicit bound can be given for the number of knots of $x_0(\cdot)$. To warrant that "Property T" holds, one can try various disconjugacy criteria. See e.g. Coppel [6], or Hartman [13].

(b) *The Nonlinear Case*

Here, in contrast to the linear case the adjoint system for $\eta(t)$ will depend on the sought function $x_0(\cdot)$. Thus, the situation is more complicated and it is more difficult to estimate the number of switches.

The case $n = 1$ of the nonlinear problem has been treated in detail by the present author in [1, 2, 3]. The author has also treated the case $n = 1$ for a vector-valued function in [4]. The case $n = 1$ of our present theorem is contained in the theorem in [3], p. 509. Observe that, in this case, there is obviously no switch. (Compare [3].)

In contrast to the linear case, the problem of the existence of a minimizing function in U is no longer trivial. This is illustrated by a counterexample in [2], p. 410. See further [10], pp. 12–22.

S. D. Fisher [7] approaches the nonlinear problem via functional analysis. The approach is interesting, but the results obtained are quite implicit, and he does not establish the spline properties of a minimizing function. See further [10], Sect. 3.

Examples for the case $n = 1$ are given in the author's papers [1, 2, 3] and in Fisher [7].

Finally, we remark that our theorem is *not* formulated in its greatest possible generality. For instance, most of the conditions on F need only hold locally, and it is not necessary that $x_0(\cdot)$ should give a global minimum over U.

References

1. G. Aronsson, Minimization problems for the functional $\sup_x F(x, f(x), f'(x))$, I, *Ark. Mat.* **6** (1965), 33–53.
2. G. Aronsson, Minimization problems for the functional $\sup_x F(x, f(x), f'(x))$, II, *Ark. Mat.* **6** (1966), 409–431.

3. G. ARONSSON, Minimization problems for the functional $\sup_x F(x, f(x), f'(x))$, III, *Ark. Mat.* **7** (1968), 509–512.
4. G. ARONSSON, Pontryagin's maximum principle and a minimax problem, *Math. Scand.* **29** (1971), 55–71.
5. D. S. CARTER, A minimum–maximum problem for differential expressions, *Canad. J. Math.* **9** (1957), 132–140.
6. W. A. COPPEL, "Disconjugacy," Lecture Notes in Mathematics, No. 220, Springer-Verlag, Berlin/New York, 1971.
7. S. D. FISHER, Solutions of some non-linear variational problems in L^∞ and the problem of minimum curvature, *Arch. Rational Mech. Anal.* **61** (1976), 291–305.
8. S. D. FISHER AND J. W. JEROME, Perfect spline solutions to L^∞ extremal problems, *J. Approximation Theory* **12** (1974), 78–90.
9. S. D. FISHER AND J. W. JEROME, The existence, characterization and essential uniqueness of solutions to L^∞ extremal problems, *Trans. Amer. Math. Soc.* **187** (1974), 391–404.
10. S. D. FISHER AND J. W. JEROME, "Minimum Norm Extremals in Function Spaces," Lecture Notes in Mathematics No. 479, Springer-Verlag, Berlin/New York, 1975.
11. G. GLAESER, "Prolongement Extremal de Fonctions Differentiables," Publications de la section de Mathématique de la Faculté des Sciences de Rennes 1967, Université de Rennes, France.
12. J. K. HALE, "Ordinary Differential Equations," Wiley, New York, 1969.
13. P. HARTMAN, On disconjugacy criteria, *Proc. Amer. Math. Soc.* **24** (1970), 374–381.
14. S. KARLIN, Some variational problems on certain Sobolev spaces and perfect splines, *Bull. Amer. Math. Soc.* **79** (1973), 124–128.
15. E. B. LEE AND L. MARKUS, "Foundations of Optimal Control Theory," Wiley, New York, 1967.
16. O. L. MANGASARIAN AND L. L. SCHUMAKER, Splines via optimal control, *in* "Approximations with Special Emphasis on Spline Functions" (I. J. Schoenberg, Ed.), Academic Press, New York, 1969.
17. D. E. MCCLURE, Perfect spline solutions of L^∞ extremal problems by control methods, *J. Approximation Theory* **15** (1975), 226–242.
18. L. S. PONTRYAGIN, V. G. BOLTYANSKII, R. V. GAMKRELIDZE, AND E. F. MISHCHENKO, "The Mathematical Theory of Optimal Processes," Wiley, New York, 1962.
19. W. T. REID, Ordinary linear differential operators of minimum norm, *Duke Math. J.* **29** (1962), 591–606.
20. I. J. SCHOENBERG, The perfect B-splines and a time optimal control problem, *Israel J. Math.* **10** (1971), 275–291.