Characterization theorem for best polynomial spline approximation with free knots, variable degree and fixed tails

Jean-Pierre Crouzeix   Nadezda Sukhorukova   Julien Ugon

February 14, 2017

Abstract

In this paper, we derive a necessary condition for a best approximation by piecewise polynomial functions of varying degree from one interval to another. Basing on these results, we obtain a characterization theorem for the polynomial splines with fixed tails, that is the value of the spline is fixed in one or more knots (external or internal). We apply nonsmooth nonconvex analysis to obtain this result, which is also a necessary and sufficient condition for inf-stationarity in the sense of Demyanov-Rubinov. This paper is an extension of a paper where similar condition were obtained for free tails splines. The main results of this paper are essential for the development of a Remez-type algorithm for free knot spline approximation.

1 Introduction

The problem of approximating a continuous function by a piecewise polynomial (polynomial spline) has been studied for over four decades [9]; yet, when the knots joining the polynomials are also variable, finding conditions for a best Chebyshev approximation remains an open problem [6, problem 1]. A necessary optimality condition for a best approximation were obtained in [5, 7] when only continuity is required. Then these conditions were improved in [13].

In this paper we are concentrating on necessary optimality conditions on the problems where the spline degree may change from one interval to another, namely, different polynomial pieces may have different degrees. These modifications enable one to reduce the dimension of the corresponding optimisation problems. Basing on these results, we obtain the characterization theorem for fixed tails polynomial splines, namely, the spline $s(t)$ approximates a continuous function at the interval $[\xi_0, \xi_N]$ and the spline value is fixed in at least one of the knot $\xi_i,\ i = 0, \ldots, N$.

These conditions are interesting from several points of view. First of all, they are important as a characterisation theorem, which provides necessary optimality conditions for such kind of splines. Second, these conditions can be used in the construction of sufficient optimality condition. Finally, these conditions are essential for the development of a Remez-like algorithm for a best polynomial spline approximation (see [8]).

There have been several attempts to develop a Remez-type approximation algorithm to the case of polynomial splines (fixed [5, 10, 12] and free [5] knots). In the case of fixed knots the problem is convex and nonsmooth, and a variety of optimisation methods can be applied to solve this problem ([1, 2, 3, 4]). In the case of free knot splines, the optimisation problems are nonconvex and nonsmooth and therefore no general optimisation method can be applied to solve it efficiently. Although methods exist to find a best free-knot spline approximation [5], their convergence, even towards a local best approximation, is not guaranteed [13].

The paper is organised as follows. In section 2 we introduce necessary definitions and relevant results from the area of polynomials spline approximation. Section 2.2.1 provides necessary results from the theory of quasidifferentials, developed by Demyanov and Rubinov [3, 4]. In section 4 we formulate and prove our main result. Finally, in section 5 we conclude and highlight future research directions.

2 Preliminaries

2.1 Definitions and formulations

Definition 2.1 (Polynomial Spline)**.** A polynomial spline is a piecewise polynomial. Each polynomial piece lies on an interval $[\xi_i, \xi_{i+1}],\ i = 0, \dots, N-1$. The points ξ_0 and ξ_N are the *external knots*, and the points ξ_i, $(i = 1, \dots, N-1)$ are the *internal knots* of the polynomial spline.

Generally, the spline is not infinitely differentiable at its knots. In the case examined in this paper, only continuity of the spline is required.

Definition 2.2. The difference between the spline and the function to approximate is called the *deviation*.

We denote the deviation function at point t by $\psi_t(s) \equiv \psi(s,t) \equiv s(t) - f(t)$.

Our aim is to minimize the maximal absolute deviation. This maximal deviation occurs at points in the interval $[\xi_0, \xi_N]$ which we call *extreme points*.

Local optimality of such spline approximation with the same polynomial degree m at each subinterval has been studied in [13]. These results are based on the notion of stable and unstable knots. The definition of stable and unstable extreme points is given in [13]. Essentially, a spline knot is unstable when it is a maximal absolute deviation point and one of the following two conditions satisfies.

1. The right derivative of the spline (as a function of t) at the knot is greater than the left (the spline can be locally approximated as a maximum of

two linear functions) and the function $f(t)$ is above the spline at this knot (that is, the deviation function is negative).

2. The right derivative of the spline (as a function of t) at the knot is smaller than the left (that is, the spline can be locally approximated as a minimum of two linear functions) and the function $f(t)$ is below the spline at this knot (that is, the deviation function is positive).

Remark 2.1. Knots may also be stable, neutral or knots that are not maximal absolute deviation points.

1. Neutral (that is, the slope is not changing). In this case the spline (as a function of t) is smooth at this knot and a maximal absolute deviation point.

2. Stable. It is a maximal absolute deviation point and one of the following two conditions satisfies.

   (a) The slope of the spline (as a function of t) is increasing (that is, the spline can be locally approximated as a maximum of two linear functions) and the function $f(t)$ is below the spline at this knot (that is, the deviation function is positive).

   (b) The slope of the spline (as a function of t) is decreasing (that is, the spline can be locally approximated as a minimum of two linear functions) and the function $f(t)$ is above the spline at this knot (that is, the deviation function is negative).

2.2 Existing results

Results obtained in [13] are local optimality conditions. More specifically, these conditions are necessary and sufficient conditions for a spline to be satationary in the sense of Demyanov-Rubinov. Demyanov-Rubinov stationarity is defined based on the notion of quasidifferential.

2.2.1 Quasidifferentiability

We start from the definition of quasidifferentiability. This notion can be considered as a possible way to generalise the concept of differentiability.

Definition 2.3. A function f defined on an open set Ω is *quasidifferentiable* [3, 4] at a point $x \in \Omega$ if it is locally Lipschitz continuous, directionally differentiable at this point and there exists compact, convex sets $\underline{\partial} f(x)$ and $\overline{\partial} f(x)$ such that the derivative of f at x in any direction g can be expressed as

$$f'(x,g) = \max_{\mu \in \underline{\partial} f(x)} \langle \mu, g \rangle + \min_{\nu \in \overline{\partial} f(x)} \langle \nu, g \rangle.$$

The sets $\underline{\partial} f(x)$ and $\overline{\partial} f(x)$ are called respectively the *sub-* and *superdifferential* of the function f at the point x. The pair $[\underline{\partial} f(x), \overline{\partial} f(x)]$ is called a *quasidifferential* of the function f at the point x.

At any local minimizer $x^* \in \Omega$ of a quasidifferentiable function f we have [3, 4]

$$-\overline{\partial} f(x^*) \subset \underline{\partial} f(x^*). \tag{2.1}$$

A point x^* satisfying condition (2.1) is an *inf-stationary* point.

2.2.2 Characterisation of free knots polynomial splines with constant degree

The main result of this paper is based on the following theorem (free tails, see [13] for details).

Theorem 2.1. *A spline satisfies condition* (2.1) *over the interval* $[\xi_0, \xi_n]$ *if and only if there exists a subinterval* $[\xi_p, \xi_q]$ *containing a sequence of* $m(q-p)+2+l$ *alternating extreme points of the deviation function, where* l *is the number of non-neutral internal knots inside* (ξ_p, ξ_q)*. The end-points* ξ_p *and* ξ_q *may be included in this sequence only if they are not unstable.*

Proof. The proof of this theorem can be found in [13]. The main structure of the proof is as follows.

- First, we show that the quasidifferential of function from [13] is expressed in terms of extreme points and their gradients.
- Second, we define an invertible linear transformation which simplifies the vectors from the sub- and superdifferentials.
- Third, we obtain the optimality conditions that are equivalent to the stationarity condition in the sense of Demyanov-Rubinov. □

The proof outlined in this paper will follow the same structure.

3 Formulation of the optimisation problem

In the current paper we extend the results from [13] to the following cases:

- the degree of the polynomial composing the spline may vary from interval to interval, that is, the polynomial degree in the $i-$th interval is $m_i,\ i = 1, \dots, n$;
- the value of the spline may be fixed at one or more knots $\xi_0, \dots, \xi_n$.

The second case in particular is useful in the construction of an optimal spline, to improve the polynomial spline on a specific subinterval which fixing it outside of this interval while guaranteeing continuity.

In [13], as in most derivations of necessary conditions [5], polynomial splines are formulated using the truncated power function [5, Appendix, p. 191]:

$$(t-\tau)_+^j = \begin{cases} 0, & \text{if } t < \tau \\ (t-\tau)^j, & \text{if } t \geq \tau \end{cases}.$$

Let $X = (a_{00}, x_0, \xi_1, x_1, \dots, \xi_{n-1}, x_{n-1}) \in \mathbb{R}^{(m+1)n}$, where

$$x_i = (a_{i1}, \dots, a_{im}) \in \mathbb{R}^m,\ i = 0, \dots, N-1$$

and

$$a = \xi_0 \leq \xi_1 \leq \dots \leq \xi_{n-1} \leq \xi_n = b, \tag{3.1}$$

then

$$s(t) = s[X](t) = a_{00} + \sum_{i=0}^{n-1} \sum_{j=1}^{m} a_{ij}(t - \xi_i)_+^{m+1-j}. \tag{3.2}$$

However, these formulations do not allow for the degree of the polynomial pieces to vary, nor to fix the value of the spline at some knots. We give an alternative formulation, more flexible from several point of view. First of all, the degree of the polynomials does not have to be the same in different intervals. Second, the value of the spline at any of the knots (including internal knots) can be fixed. We construct spline $s(t)$ as follows:

$$s(t) \equiv s[X](t) = P_i(t) = S_i + \sum_{j=1}^{m_{i+1}} a_{ij}(t - \xi_1)^{m_i+1-j}, t \in [\xi_i, \xi_{i+1}],\ i = 1 \dots, n-1 \tag{3.3}$$

where

$$X = (S_0, a_{0m_1}, \dots a_{02}, S_1, \xi_1, a_{1m_2}, \dots, a_{12}, S_2, \dots, \xi_{n-1}, a_{n-1m_{n-1}}, \dots, a_{n-12}, S_n)^T \tag{3.4}$$

and $S_i,\ i = 0, \dots, n-1$ are value of the spline at $t = \xi_i,\ i = 0, \dots, n-1$. To ensure continuity, we need to have

$$S_i = S_{i-1} + \sum_{j=2}^{m_i} a_{i-1j}(\xi_i - \xi_{i-1})^{m_i+1-j} + a_{i-1j}(\xi_i - \xi_{i-1})^{m_i},\ i = 1, \dots, n$$

and therefore

$$a_{i-11} = \frac{S_i - S_{i-1} - \sum_{j=2}^{m_i} a_{i-1j}(\xi_i - \xi_{i-1})^{m_i+1-j}}{(\xi_i - \xi_{i-1})^{m_i}}.$$

In general, if the spline value is fixed at a knot $\xi_k,\ k = 0, \dots, n$ then S_k is a constant. On the l-th interval, between the knots ξ_{l-1} and ξ_l, the spline $s[X](t)$ is a polynomial of degree m_l. We assume that $\sum_{i=1}^{n} m_i = M$.

Summarising all the above, our optimisation problem can be reformulated as follows:

$$\text{minimize} \sup_{t \in [\xi_0, \xi_N]} \left| s(t) - f(t) \right|, \text{ subject to } X. \tag{3.5}$$

X is specified in (3.4).

To verify whether an internal knot is unstable or not, we need to compare the left and right derivatives of the spline at the knot. Consider the knot ξ_i $i = 1, \ldots, n-1$. The left derivative (as a function of t) is

$$\sum_{j=1}^{m_{i-1}} a_{i-10}(1-j)(\xi_i - \xi_{i-1})^{m_{i-1}-j}$$

and the right one is simply a_{im_i}.

Therefore, the following characteristics can be used to differentiate between different knot types:

1. ξ_i is neutral if it is a maximal absolute deviation point and

$$\sum_{j=1}^{m_{i-1}} a_{i-10}(1-j)(\xi_i - \xi_{i-1})^{m_{i-1}-j} = a_{im_i}.$$

2. ξ_i is stable if it is a maximal absolute deviation point and one of the following two conditions satisfies.

   (a) The deviation is positive and

$$\sum_{j=1}^{m_{i-1}} a_{i-10}(1-j)(\xi_i - \xi_{i-1})^{m_{i-1}-j} < a_{im_i}.$$

   (b) The deviation is negative and

$$\sum_{j=1}^{m_{i-1}} a_{i-10}(1-j)(\xi_i - \xi_{i-1})^{m_{i-1}-j} > a_{im_i}.$$

3. ξ_i is unstable if it is a maximal absolute deviation point and one of the following two conditions satisfies.

   (a) The deviation is negative and

$$\sum_{j=1}^{m_{i-1}} a_{i-10}(1-j)(\xi_i - \xi_{i-1})^{m_{i-1}-j} < a_{im_i}.$$

   (b) The deviation is positive and

$$\sum_{j=1}^{m_{i-1}} a_{i-10}(1-j)(\xi_i - \xi_{i-1})^{m_{i-1}-j} > a_{im_i}.$$

Note that the only points where the spline function can be nonsmooth are its knots.

In [13] a similar problem was studied, where all the polynomials have the same degree m and all the spline values are free at any knot. In the proof of the main result of [13] (Theorem 3.1, formula (7.3) and (7.4)) it was observed that the inf-stationarity in the sense of Demyanov-Rubinov (see also section 2.2.1) is equivalent to the following:

$$\mathbf{0}_{M+1} \in \bigcap_{P \in E} P, \tag{3.6}$$

where E is a family of polytopes whose vertices are constructed from the gradients of the deviation function at the extreme points. Essentially, this means that there exists a subinterval

$$[\xi_p, \xi_q] \subset [\xi_0, \xi_n],$$

such that the convex hull of the gradients that correspond to the maximal deviation points in this interval $[\xi_p, \xi_q]$ contains zero. If $p \neq 0$ $(q \neq n)$, then the gradient at ξ_p (ξ_q) can only be included if this point is not only one of the maximal absolute deviation points, but also if it is not unstable [13].

There are several approaches that enables to work with nonsmooth functions (local optimisation). In the next section we describe one of them.

4 Characterization through quasidifferentiability

4.1 Quasidifferential of the objective function

A very extensive analysis of the quasidifferentiability of the objective function for the case of free tails is given in [13]. In the case of fixed tails and flexible degrees all the reasonings are the same and the construction of the corresponding quasidifferentials and confined quasidifferentials are the same. The only differences are

- the polynomial degrees may vary from one interval to another;
- the corresponding $S_i,\ i = 0, \dots, n$ should be treated as constants if the spline value at $\xi_i,\ i = 0, \dots, n$ is fixed.

We recall following notations and basic results from [13]. Let

$$\beta_{j_t}(X, t) = \mathrm{sign}(\Psi(s[X], t)) \cdot \nabla_X P_{j_t}(X, t). \tag{4.1}$$

To the smooth extreme points (non internal knots), neutral knots and stable knots we associate the set

$$\mathcal{S} = \{\beta_{j_t}(X, t) : t \in E_{\mathrm{smooth}} \cup E_{\mathrm{neutral}} \cup (K^+ \cap E_{\max}) \cup (K^- \cap E_{\min})\}.$$

To the extreme points coinciding with unstable knots we associate the following sets:

$$\Delta_{j_t} = \mathrm{co}\{\beta_{j_t}(X,t), \beta_{j_t+1}(X,t)\};$$
$$C_\Delta = -\sum_{\Delta_{j_t}\neq\Delta} \Delta_{j_t}.$$

Define $\mathcal{U} = \{\Delta_{j_t} : t \in (K^+ \cap E_{\min}) \cup (K^- \cap E_{\max})\}$. The quasidifferential of the function Ψ is $\overline{\partial}\Psi = \sum_{\Delta_{j_t}\in\mathcal{U}} \Delta_{j_t}$ and $\underline{\partial}\Psi = \mathrm{co}\{\mathcal{S} - \overline{\partial}\Psi, \cup_{\Delta\in\mathcal{U}}(\mathbf{0} + C_\Delta)\}$.

Definition 4.1. An interval $[\xi_p, \xi_q]$ is *stationary* if

$$-\overline{\partial_p^q}\Psi(s) \subset \underline{\partial_p^q}\Psi(s) \tag{4.2}$$

Proposition 4.1. *The interval $[\xi_p, \xi_q]$ is stationary if and only if*

$$\mathbf{0} \in \mathrm{co}\big\{\mathcal{S}_p^q, \mathcal{C}\big\}, \forall \mathcal{C} \in \coprod_{\Delta\in\mathcal{U}_p^q} \Delta, \tag{4.3}$$

Corollary 4.1. *A spline $s[X]$ is an inf-stationary solution to the problem* (3.5) *if and only if there exists a stationary subinterval.*

These results are independent from the formulation of the polynomial spline, and therefore are applicable in the case of formulation (3.3).

4.2 Characterization using alternating extreme points

Let us introduce the following definitions (see [13]).

Definition 4.2. Subintervals $[\xi_{i-1}, \xi_i], i = 1, \ldots, N$ are called *unit subintervals.* Intervals delimited by non-neutral external knots and whose internal knots are neutral are called *block subintervals.*

Definition 4.3. If the value of the spline is fixed at one or more knots

$$\xi_i,\ i = 1, \ldots, N$$

then these knots are called fixed points and we say that this is a polynomial spline with fixed tails.

Theorem 4.1. *If a polynomial spline with one or more fixed points is optimal then at least one of the following conditions should satisfies.*

1. *One or more of the fixed points is a maximal absolute deviation point.*

2. *There exists a subinterval $[\xi_p, \xi_q]$, containing a sequence of*

$$2 + l - r + \sum_{i=p}^{q-1} m_i$$

alternating extreme points of the deviation function, where l is the number of non-neutral internal knots inside (ξ_p, ξ_q) and r is the number of fixed points of the optimal spline in the interval $[\xi_p, \xi_q]$. The end-points ξ_p and ξ_q may be included in this sequence only if they are not unstable. This condition is also equivalent to inf-stationarity in the sense of Demyanov-Rubinov.

Proof. The first condition is obvious. Indeed, if at least one fixed point is included in the set of maximal deviation points, then the maximal deviation can not be improved. This condition is also a sufficient optimality condition.

The second condition is equivalent to the free tails stationarity condition [13] if none of the fixed points is located in the interval $[\xi_p, \xi_q]$, but the number of alternating points should be adjusted, since the corresponding degree is varying from interval to interval.

Similar to [13], assume that $[\xi_p, \xi_q]$ is the shortest stationary subinterval. Since l is the number of non-neutral internal knots inside (ξ_p, ξ_q), the number of block subintervals inside this interval is $l + 1$.

The equation (3.6) is equivalent to the fact that all the systems in a sequence of homogeneous linear systems

$$A_i X = \mathbf{0}_{M+1},\ i = 1, 2^l \tag{4.4}$$

has a nonnegative solution with at least one strictly positive component. The columns of $A_i,\ i = 1, \dots, 2^l$ the deviation function graduates, calculated at the maximal absolute deviation points. The structure of $A_i,\ i = 1, \dots, 2^l$ is as follows:

$$\begin{pmatrix}
1 - K_1^{m_1}(t) & 0 & 0 & \dots & 0 \\
M_1(t) & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\
K_1^{m_1}(t) & 1 - K_2^{m_2}(t) & 0 & \dots & 0 \\
\Sigma_1(t) & \Theta_2(t) & 0 & \dots & 0 \\
\mathbf{0} & M_2(t) & \mathbf{0} & \dots & \mathbf{0} \\
0 & K_2^{m_2}(t) & 1 - K_3^{m_3}(t) & \dots & 0 \\
0 & \Sigma_2(t) & \Theta_3(t) & \dots & 0 \\
\mathbf{0} & \mathbf{0} & M_3(t) & \dots & \mathbf{0} \\
0 & 0 & K_3^{m_3}(t) & \dots & 0 \\
0 & 0 & \Sigma_3(t) & \dots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \dots & \Sigma_n(t) \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & M_n(t) \\
0 & 0 & 0 & \dots & K_n^{m_n}(t)
\end{pmatrix},$$

where

$$K_i(t) = \frac{t - \xi_{i-1}}{\xi_i - \xi_{i-1}},\ t \in [\xi_{i-1}, \xi_i],\ i = 1, \dots, n;$$

$$M_i(t) = \begin{pmatrix} (t-\xi_{i-1}) - (\xi_i - \xi_{i-1})K_i^{m_i}(t) \\ (t-\xi_{i-1})^2 - (\xi_i - \xi_{i-1})^2 K_i^{m_i}(t) \\ \vdots \\ (t-\xi_{i-1})^{m_i-1} - (\xi_i - \xi_{i-1})^{m_i-1} K_i^{m_i}(t) \end{pmatrix},\ t \in [\xi_{i-1}, \xi_i],\ i = 1, \dots, n;$$

$$\Sigma_i(t) = \sum_{j=2}^{m_i} a_{i-1j}(j-1)(\xi_i - \xi_{i-1})^{m_i-j} K_i^{m_i}(t),\ i = 1, \dots n-1;$$

$$\begin{aligned} \Theta_i(t) = &- \sum_{j=2}^{m_i} a_{i-1j}((t-\xi_{i-1})^{m_i-j} - (\xi_i - \xi_{i-1})^{m_i-j} K_i^{m_i})(t) \\ &+ \sum_{j=2}^{m_i} a_{i-1j}(\xi_i - \xi_{i-1})^{m_i-j+1} m_i K_i^{m_i-1(t)} \frac{(t-\xi_{i-1}) - (\xi_i - \xi_{i-1})}{(\xi_i - \xi_{i-1})^2} \\ = &- \sum_{j=2}^{m_i} a_{i-1j}((t-\xi_{i-1})^{m_i-j} - (\xi_i - \xi_{i-1})^{m_i-j}) K_i^{m_i}(t) \\ &+ \sum_{j=2}^{m_i} a_{i-1j}(\xi_i - \xi_{i-1})^{m_i-j} m_i K_i^{m_i}(t) \\ &- \sum_{j=2}^{m_i} a_{i-1j}(\xi_i - \xi_{i-1})^{m_i-j} m_i K_i^{m_i-1}(t), \\ &t \in [\xi_i - \xi_{i-1}],\ i = 2, \dots, n. \end{aligned}$$

Note that we do not include $\Theta_1(t)$, since ξ_0 is fixed and $\Sigma_n(t)$, since ξ_n is fixed.

Now we need to point out the differences between $A_i,\ i = 1, \dots, 2^l$. For each non-neutral (and therefore nonsmooth) knot (which has to be unstable due to Corollary 7.1 in [13]) we can assign two gradients, each corresponds to its adjusted interval. Since there are l non-neutral knots, there are also 2^l distinct possibilities to choose one and only one gradient for each knot in this sequence of knots. Each sequence gives rise to its own $A_i,\ i = 1, \dots, 2^2$.

Recall that the interval $[\xi_p, \xi_q]$ is a subinterval of the original interval $[\xi_0, \xi_n]$, such that (3.6) holds and there is no shorter subinterval inside $[\xi_p, \xi_q]$, such that (3.6) holds for this interval as well.

Consider the block that consists of non-zero coordinates of the gradients that correspond to maximal absolute deviation points of the second last interval

$[\xi_{q-2}, \xi_{q-1}]$ and the last interval $[\xi_{q-1}, \xi_q]$:

$$\begin{pmatrix} 1 - K_{q-2}^{m_{q-2}}(t) & 0 \\ \Theta_{q-2}(t) & 0 \\ M_{q-2}(t) & \mathbf{0} \\ K_{q-2}^{m_{q-2}}(t) & 1 - K_{q-1}^{m_{q-1}}(t) \\ \Sigma_{q-2}(t) = \sum_{j=2}^{m_{q-2}} a_{q-3j}(j-1)(\xi_{q-3} - \xi_{q-2})^{m_{q-2}-j} K_{q-2}^{m_{q-2}}(t) & \Theta_{q-1}(t) \\ \mathbf{0} & M_{q-1}(t) \\ 0 & K_{q-1}^{m_{q-1}}(t) \end{pmatrix}. \tag{4.5}$$

Using the row of this block (matrix) that corresponds to the first non-zero row of the last interval, one can obtain an equivalent matrix (multiplying this row by a constant and subtracting it from the next one):

$$\begin{pmatrix} 1 - K_{q-2}^{m_{q-2}}(t) & 0 \\ \Theta_{q-2}(t) & 0 \\ M_{q-2}(t) & \mathbf{0} \\ K_{q-2}^{m_{q-2}}(t) & 1 - K_{q-1}^{m_{q-1}}(t) \\ 0 & \Theta_{q-1}(t) - \sum_{j=2}^{m_{q-2}} a_{q-3j}(j-1)(\xi_{q-3} - \xi_{q-2})^{m_{q-2}-j}(1 - K_{q-1}^{m_{q-1}}(t)) \\ \mathbf{0} & M_{q-1}(t) \\ 0 & K_{q-1}^{m_{q-1}} \end{pmatrix}.$$

Using the last row, one can simplify the first two non-zero coordinates of the last interval $[\xi_{q-1}, \xi_q]$ gradients

$$\begin{pmatrix} 1 - K_{q-2}^{m_{q-2}}(t) & 0 \\ \Theta_{q-2}(t) & 0 \\ M_{q-2}(t) & \mathbf{0} \\ K_{q-2}^{m_{q-2}}(t) & 1 \\ 0 & \Theta_{q-1}(t) - \sum_{j=2}^{m_{q-2}} a_{q-3j}(j-1)(\xi_{q-3} - \xi_{q-2})^{m_{q-2}-j} \\ \mathbf{0} & M_{q-1}(t) \\ 0 & K_{q-1}^{m_{q-1}}(t) \end{pmatrix}.$$

Consider now

$$\begin{aligned}
\tilde{\Theta}_{q-1}(t) =& \Theta_{q-1}(t) - \sum_{j=2}^{m_{q-2}} a_{q-3j}(j-1)(\xi_{q-3}-\xi_{q-2})^{m_{q-2}-j} \\
=& -\sum_{j=2}^{m_{q-1}} a_{q-2j}((t-\xi_{q-2})^{m_{q-2}-j} - (\xi_{q-1}-\xi_{q-2})^{m_{q-2}-j} K_{q-1}^{m_{q-1}}(t)) \\
&+ \sum_{j=2}^{m_{q-1}} a_{q-2j}(\xi_i - \xi_{i-1})^{m_{q-1}-j} m_{q-1} K_{q-1}^{m_{q-1}}(t) \\
&- \sum_{j=2}^{m_{q-1}} a_{q-2j}(\xi_i - \xi_{i-1})^{m_{q-1}-j} m_{q-1} K_{q-1}^{m_{q-1}-1}(t) \\
&- \sum_{j=2}^{m_{q-2}} a_{q-3j}(j-1)(\xi_{q-3}-\xi_{q-2})^{m_{q-2}-j}(1-K_{q-1}^{q-1})(t).
\end{aligned}$$

Using the last row, $\tilde{\Theta}_{q-1}(t)$ can be updated to the following

$$\begin{aligned}
&\tilde{\tilde{\Theta}}_{q-1}(t) = \\
&\qquad = -\sum_{j=2}^{m_{q-1}} a_{q-2j}(t-\xi_{q-2})^{m_{q-1}-j} \\
&\qquad\quad + 0 \\
&\qquad\quad - \sum_{j=2}^{m_{q-1}} a_{q-2j}(\xi_i - \xi_{i-1})^{m_{q-1}-j} m_{q-1} K_{q-1}^{m_{q-1}-1}(t) \\
&\qquad\quad - \sum_{j=2}^{m_{q-2}} a_{q-3j}(j-1)(\xi_{q-3}-\xi_{q-2})^{m_{q-2}-j}.
\end{aligned}$$

The rows of $M_{q-1}(t)$ block (matrix) can be updated to the following (multiplying the last row by the corresponding factor and adding to the rows of $M_{q-1}(t)$):

$$\tilde{M}_{q-1}(t) = \begin{pmatrix} t-\xi_{q-2}) \\ (t-\xi_{q-2})^2 \\ \vdots \\ (t-\xi_{i-1})^{m_{q-2}-1} \end{pmatrix}.$$

Finally, using the rows of $\tilde{M}_{q-1}(t)$ one can obtain

$$\begin{aligned}
\tilde{\tilde{\tilde{\Theta}}}_{q-1}(t) = \\
&= -\, a_{q-2m_{q-1}}(t) \\
&\quad + 0 \\
&\quad + 0 \\
&\quad - \sum_{j=2}^{m_{q-2}} a_{q-3j}(j-1)(\xi_{q-3} - \xi_{q-2})^{m_{q-2}-j} \\
&= -\, a_{q-2m_{q-1}} - \sum_{j=2}^{m_{q-2}} a_{q-3j}(j-1)(\xi_{q-3} - \xi_{q-2})^{m_{q-2}-j} \\
&= \sum_{j=2}^{m_{q-2}} a_{q-3j}(1-j)(\xi_{q-3} - \xi_{q-2})^{m_{q-2}-j} - a_{q-2m_{q-1}}
\end{aligned}$$

Therefore, $\tilde{\tilde{\tilde{\Theta}}}_{q-1}(t) = 0$ if and only if ξ_{q-2} is a smooth knot. If ξ_{q-2} is not smooth, then the block from (4.5) can be transformed to

$$\begin{pmatrix} 1 - K_{q-2}^{m_{q-2}}(t) & 0 \\ \Theta_{q-2}(t) & 0 \\ M_{q-2}(t) & \mathbf{0} \\ K_{q-2}^{m_{q-2}}(t) & 0 \\ 0 & 1 \\ \mathbf{0} & \tilde{M}_{q-1}(t) \\ 0 & K_{q-1}^{m_{q-1}}(t) \end{pmatrix}. \tag{4.6}$$

If ξ_{q-2} is smooth, then the block from (4.5) can be transformed to

$$\begin{pmatrix} 1 - K_{q-2}^{m_{q-2}}(t) & 0 \\ \Theta_{q-2}(t) & 0 \\ M_{q-2}(t) & \mathbf{0} \\ K_{q-2}^{m_{q-2}}(t) & 1 \\ 0 & 0 \\ \mathbf{0} & \tilde{M}_{q-1}(t) \\ 0 & K_{q-1}^{m_{q-1}}(t) \end{pmatrix}. \tag{4.7}$$

Note the following.

1. Each equivalent transformation with rows can be also obtained through multiplication by the corresponding nonsingular matrix. Therefore, all the equivalent row transformations, used above, do not affect the existence of a nonnegative solution with at least one strictly positive coordinate to the original problem (4.4).

2. In the case of smooth knots (inside a block subinterval) the structure of the gradient is the same as it is in the case of fixed knots, since the coordinate that corresponds to the knot is zero. Recall that block subintervals are bounded by nonsmooth knots and all the internal knots of such subintervals are smooth.

Continue the process. Finally, ignoring zero coordinates that correspond to the subintervals outside of $[\xi_p, \xi_q]$, one can obtain a block-diagonal system, where each block corresponds to its block interval. Then for (4.4) to satisfy one needs to prove that the convex hull of the gradients that correspond to each block (block subintervals) contains zero vector and therefore (characterisation theorem for fixed knots, see [11] for details) each block interval $[\xi_w, \xi_v]$ contains a subinterval $[\xi_{w_0}, \xi_{v_0}]$, such that the number of alternating points is $2+\sum_{i=w}^{v-1} m_i$ (also due to our assumption that there is no subinterval inside $[\xi_p, \xi_q]$, such that (3.6) holds for this interval as well). Each fixed point reduces the number of alternating points by one (due to the fact that the corresponding component of the gradient is zero, see [11, 10, 12]). Then the total number of alternating points is

$$l - r + \sum_{i=p}^{q-1} m_i.$$

□

This result resembles the fixed knots optimality results [11, 10, 12]. Namely, the number of alternance points is reducing by one when one of the tails is fixed (by two when both tails are fixed).

5 Discussion and conclusive remarks

In this paper we obtained a characterization theorem for the approximation of continuous functions by continuous polynomial splines with free knots. The value of the optimal spline at one or more knots is fixed and the corresponding polynomial degree is not necessary the same for different intervals. This result is equivalent to the Demyanov-Rubinov stationarity if none of the fixed points is a maximal deviation point.

These results are essential from several point of view. First of all, they are interesting on its own as a characterization theorem for such kind of splines. Second, these results are essential for sufficient optimality conditions construction. Third, they can be used in the development of a Remez-type algorithm for free knot spline approximation. The last two research directions are in our future research plan.

The main difficulty in solving such kind of problems is that the corresponding optimisation problems are nonconvex and nonsmooth. This essential characteristic restricts the choice of optimisation methods that can handle this problem efficiently and therefore, in order to solve this problem, one needs to develop a specific method (for example Remez-type) to solve this problem.

References

[1] Adil Bagirov, Napsu Karmitsa, and Marko M. Mäkelä, *Introduction to nonsmooth optimization: Theory, practice and software*, Springer Publishing Company, Incorporated, 2014.

[2] Gleb Beliakov and Julien Ugon, *Implementation of novel methods of global and nonsmooth optimization: Ganso programming library*, Optimization **56** (2007), no. 5–6, 543–546.

[3] V.F. Demyanov and A.M. Rubinov, *Constructive nonsmooth analysis*, Peter Lang, Frankfurt am Main, 1995.

[4] V.F Demyanov and A.M. Rubinov (eds.), *Quasidifferentiability and related topics*, Nonconvex Optimization and Its Applications, vol. 43, Kluwer Academic, Dordrecht/Boston/London, 2000.

[5] G. Nürnberger, *Approximation by spline functions*, Springer-Verlag, 1989.

[6] G. Nürnberger, *Bivariate segment approximation and free knot splines: Research Problems 96-4*, Constructive Approximation **12** (1996), no. 4, 555–558. MR 1412199 (97e:41020)

[7] G. Nürnberger, L. Schumaker, M. Sommer, and H. Strauss, *Uniform approximation by generalized splines with free knots*, Journal of Approximation Theory **59** (1989), no. 2, 150–169.

[8] E.Ya Remez, *General computational methods of Chebyshev approximation*, Atomic Energy Translation **4491** (1957).

[9] L. Schumaker, *Uniform approximation by chebyshev spline functions. II: free knots*, SIAM Journal of Numerical Analysis **5** (1968), 647–656.

[10] N. Sukhorukova, *A generalisation of Remez algorithm to the case of polynomial splines*, Ph.D. thesis, St-Petersburg State University, 2006, 134 pp. (in Russian).

[11] Nadezda Sukhorukova, *Uniform approximation by the highest defect continuous polynomial splines: necessary and sufficient optimality conditions and their generalisations*, Journal of Optimization Theory and Applications **147** (2010), no. 2, 378–394.

[12] ———, *Vallée Poussin theorem and Remez algorithm in the case of generalised degree polynomial spline approximation*, Pacific Journal of Optimization **6** (2010), no. 1, 103–114.

[13] Nadezda Sukhorukova and Julien Ugon, *Characterization theorem for best polynomial spline approximation with free knots*, Transactions of the American Mathematical Society, In Press, accepted in October 2015.