Polynomial approximation in several variables [*]

Vilmos Totik[†]

Dedicated to Zeev Ditzian

Abstract

This paper is about the characterization of the rate of best polynomial approximation on a domain in $\mathbf{R}^d$ by multivariate polynomials of degree n via a suitable modulus of continuity or smoothness. We shall be concerned with the situations: approximation on convex polytopes and approximation on (not necessarily convex) smooth algebraic domains. We shall introduce appropriate moduli of smoothness and prove the corresponding Jackson theorem. The usual weak converses are also true.

Contents

One of the primary questions of approximation theory is to characterize the rate of approximation by a given system in terms of some (computable) modulus of continuity or smoothness. In this spirit we set out the task to characterize the rate of best polynomial approximation on a domain in $\mathbf{R}^d$ by multivariate polynomials of degree n via a suitable modulus of continuity or smoothness.

We shall be concerned with the situations: approximation on convex polytopes and approximation on (not necessarily convex) smooth algebraic domains. We shall introduce appropriate moduli of smoothness and prove the corresponding Jackson theorem. The usual weak converses are also true, they actually automatically follow from classical converse results and from the way the moduli of smoothness are defined.

Part I
Approximation on polytopes

In this first part we shall be concerned with convex polytopes in any dimension.

1   The main result

Let f be a continuous function on $[-1, 1]$. With $\varphi(x) = \sqrt{1 - x^2}$ and $r = 1, 2, \ldots$ let

$$\omega_\varphi^r(f, \delta) = \sup_{0 < h \leq \delta, \ x \in [-1,1]} |\Delta_{h\varphi(x)}^r f(x)| \tag{1.1}$$

be its so called φ-modulus of smoothness of order r, where

$$\Delta_h^r f(x) = \sum_{k=0}^r (-1)^k \binom{r}{k} f\left(x + (\tfrac{r}{2} - k)h\right) \tag{1.2}$$

is the r-th symmetric difference. In (1.1) it is agreed that $\Delta_h^r f(x) = 0$ if $[x - \frac{r}{2}h, x + \frac{r}{2}h] \not\subseteq [-1, 1]$. See the book [6] for this kind of moduli of smoothness, for their properties and for their applications.

With this modulus of smoothness one can characterize the rate of best polynomial approximation on $[-1, 1]$. Indeed, if

$$E_n(f) = \inf_{P_n} \|f - P_n\|_{[-1,1]}$$

is the error (in the supremum norm) of the best polynomial approximation by polynomials of degree at most n, then the Jackson-type estimate

$$E_n(f) \leq C_r \omega_\varphi^r(f, 1/n),$$

as well as its weak converse

$$\omega_\varphi^r(f, 1/n) \leq C_r \frac{1}{n^r} \sum_{k=0}^n (k+1)^{r-1} E_k(f) \tag{1.3}$$

are true (see [6, Theorem 7.2.1] and [6, Theorem 7.2.4]). Our aim in this paper is to find the analogues of these in several variables.

In $\mathbf{R}^d$ we call a closed set $K \subset \mathbf{R}^d$ a convex polytope if it is the convex hull of finitely many points. K is d-dimensional if it has an inner point, which we shall always assume. The analogue of the φ-modulus of smoothness on K was defined in [6, Chapter 12], and to recall its definition we need to consider the given function along lines in different directions. A direction $\underline{e}$ in $\mathbf{R}^d$ is just a unit vector $\underline{e} \in \mathbf{R}^d$. Clearly, $\underline{e}$ can be identified with an element of the unit sphere S^{d-1}, so S^{d-1} is the set of all directions in $\mathbf{R}^d$. Let K be a convex polytope, $x \in K$ and $\underline{e} \in S^{d-1}$ a direction. The line $l_{\underline{e},x}$ through x which is parallel with $\underline{e}$ intersects K in a segment $A_{\underline{e},x} B_{\underline{e},x}$. The role of φ on that segment is taken over by

$$\tilde{d}_K(\underline{e}, x) = \sqrt{\mathrm{dist}(x, A_{\underline{e},x}) \cdot \mathrm{dist}(x, B_{\underline{e},x})}. \tag{1.4}$$

If f is a continuous function on K, then we define its r-th symmetric differences in the direction of $\underline{e}$ as

$$\Delta_{h\underline{e}}^r f(x) = \sum_{k=0}^r (-1)^k \binom{r}{k} f\left(x + (\frac{r}{2} - k)h\underline{e}\right) \tag{1.5}$$

with the agreement that this is 0 if $x + \frac{r}{2}h\underline{e}$ or $x - \frac{r}{2}h\underline{e}$ does not belong to K.

If $\mathcal{E}$ is a set of directions, then the r-th modulus of smoothness with these directions is defined as (see [6, Section 12.2])

$$\omega_K^r(f, \delta)_\mathcal{E} = \sup_{e \in \mathcal{E}, \ h \leq \delta, \ x \in K} |\Delta_{h\tilde{d}_K(\underline{e},x)\underline{e}}^r f(x)|. \tag{1.6}$$

If $\mathcal{E}$ is the direction of the edges of K, then we shall simply write $\omega_K^r(f, \delta)$ for $\omega_K^r(f, \delta)_\mathcal{E}$. These moduli of smoothness share many properties of classical moduli, in particular

$$\omega_K^r(f, \lambda\delta)_\mathcal{E} \leq C_r \lambda^r \omega_K^r(f, \delta)_\mathcal{E}$$

3

for any $\lambda \geq 1$. We shall also frequently use monotonicity in K: if $K_1 \subset K_2$ then $\omega_{K_1}^r(f,\delta)_{\mathcal{E}} \leq \omega_{K_2}^r(f,\delta)_{\mathcal{E}}$, which is immediate from the definitions.

Note that when $K = [-1,1]$, then there is only one direction (and its negative) and this modulus of smoothness takes the form (1.1), i.e.

$$\omega_\varphi^r(f,\delta) = \omega_{[-1,1]}^r(f,\delta). \qquad (1.7)$$

Let Π_n^d be the set of polynomials in d-variables of total degree at most n, and let the supremum norm on a set K be denoted by $\|\cdot\|_K$. For a continuous function $f \in C(K)$ we are interested in

$$E_n(f)_K = \inf_{P_n \in \Pi_n^d} \|f - P_n\|_K, \qquad (1.8)$$

which is the error of best approximation of f by polynomials of total degree at most n.

In the first part of this paper we prove

Theorem 1.1 *If K is a convex polytope in $\mathbf{R}^d$, then for $r = 1, 2, \ldots$*

$$E_n(f)_K \leq C\omega_K^r(f,1/n), \qquad n \geq rd. \qquad (1.9)$$

Here the constant C depends only on K and r.

Call a polytope K in $\mathbf{R}^d$ a simple polytope if at every vertex there are precisely d edges. Theorem 1.1 was verified in [6, Theorem 12.1.1] for cubes/parallelepipeds in all dimensions, and in [6, Theorem 12.2.3] (with an additional term) for all simple polytopes. The somewhat weaker result

$$E_n(f)_K \leq C\omega_K^r(f,1/n)_{S^{d-1}}, \qquad (1.10)$$

i.e. when all directions are used in the modulus of smoothness, was proven in [14, Chapters 1-10] for all polytopes in R^3. In [14, Chapter 11, Theorem 1.1], it was also stated that (1.10) holds also in $\mathbf{R}^d$, $d > 3$, but Andriy Prymak noticed that the proof (that reduced the problem to approximation on simple polytopes) does not give that, so the claim was wrong in the sense that the proof in [14, Chapter 11] yields (1.10) only for special polytopes K (which satisfy the property that the intersection of K with any hyperplane that cuts off a vertex is a simple $(d-1)$-dimensional polytope). Thus, in its generality, (1.9) (which also covers (1.10)) is new (and actually sharper than what has been previously proven in the special cases known in the literature).

The weak converse

$$\omega_K^r(f,1/n) \leq C_r \frac{1}{n^r} \sum_{k=0}^n (k+1)^{r-1} E_k(f)_K,$$

is an immediate consequence of (1.3) if we take into account how the modulus of smoothness on the left has been defined.

The Jackson-type estimate Theorem 1.1 and the weak converse (1.3) easily imply the following. If $K_1 \subset K$ is a segment and we consider f only on K_1, then

we obtain a continuous function on a segment/interval, and we can consider the approximation of this one-variable function by polynomials of the same (single) variable. Now it follows that if $\alpha > 0$ and the error of best approximation of f on any segment of K by (one variable) polynomials of degree at most n is $\leq n^{-\alpha}$, then on the whole K the function f can be approximated with error $\leq Cn^{-\alpha}$ by polynomials (of several variables) of degree at most n (where C depends only on K).

Remark 1.2 We need to say a few words on the degree of polynomials. In this work we use total degree (i.e. the degree of $x_1^{\alpha_1} \cdots x_d^{\alpha_d}$ is $\alpha_1 + \cdots + \alpha_d$), while the works [6] and [14] used maximal degree in the separate variables (in which case the degree of $x_1^{\alpha_1} \cdots x_d^{\alpha_d}$ is $\max \alpha_j$). This has not much effect on the corresponding Jackson theorem, only on the range of n for which the theorem holds. Indeed, let $E_n(f)$ be as in (1.8) where the infimum is taken for all polynomials of total degree at most n, and let

$$E_n^*(f)_K = \inf_{Q_n} \|f - Q_n\|_K,$$

where now the infimum is taken for all polynomials Q_n that are of degree at most n in each variable. It is clear that $E_{nd}(f)_K \leq E_n^*(f)_K \leq E_n(f)_K$, so a Jackson-type estimate for either of E_n or E_n^* yields a Jackson-type estimate for the other. Note however, that if f_0 is the polynomial $x_1^{r-1} \cdots x_d^{r-1}$ and K is the unit cube in $\mathbf{R}^d$, then $E_n^*(f_0) = 0$ for $n \geq r - 1$ while $E_n(f_0) = 0$ only for $n \geq d(r-1)$. In a similar manner, if we use edge directions in forming the modulus of smoothness, then $\omega^r(f_0, \delta)_K \equiv 0$, so when using $E_n(f)$ the Jackson-type estimate (1.9) cannot hold for $n < d(r-1)$. That is why we claimed (1.9) only for $n \geq rd$.

2   Outline of the proof of Theorem 1.1

In the proof of Theorem 1.1 a special role is played by pyramids. Let S_0 be a convex $(d-1)$-dimensional polytope in $\mathbf{R}^d$, and $P \in \mathbf{R}^d$ a point outside the hyperplane of S_0. The convex hull of S_0 and P determines a d-dimensional pyramid S with apex at P. We call S_0 the base of S, and the edges emanating from P the apex edges of S.

Let $K \subset \mathbf{R}^d$ be a convex d-dimensional polytope with vertices $\{Q_j\}_{j=1}^m$. If Q_j is a vertex of K and L is a hyperplane that intersects all the edges emanating from Q_j, then L cuts off a pyramid S_j from K with apex at Q_j (and with base $K \cap L$).

As has been mentioned, (1.10) was proven in [14] in $\mathbf{R}^3$. That proof works in any dimension for a polytope K which has the property that if we cut off a small pyramid S_j around every vertex Q_j, then what remains (say K_0) is a simple polytope (note that in $\mathbf{R}^2$ any polytope is simple, so the just mentioned property is true for any 3-dimensional polytopes). In that approach simple polytopes are used in two essential ways.

- 1) K_0 is simple, so for it the claim (say (1.10)) on the rate of approximation is true by [6, Theorem 12.2.3] (with the additional term in that theorem separately removed).

- 2) If S_j^λ is the λ-shrunk copy of S_j from Q_j, then the truncated pyramid $S_j \setminus S_j^\lambda$ is simple, so for it the claim on the rate of approximation is true again by [6, Theorem 12.2.3].

Note that if we are working in $\mathbf{R}^3$, then no assumption is needed, K_0 is always simple, so the just outlined approach yields the theorem. This is no longer true in $\mathbf{R}^d$, $d > 3$ (an observation by A. Prymak[1]), so the use of simple polytopes in 1)–2) must be avoided. The proof we give is by induction on the dimension, but for the induction we need the approximants to be given by linear operators, and the modulus of smoothness should involve only edge-directions.

A) In 2) the truncated pyramid is (for $\lambda < 1$ close to 1) the union of well-overlapped product sets (with the smaller base of the truncated pyramid $\times$ the edges), so induction and simple patching gives the result for the truncated pyramid. From there the proof in [14] gives the result for each of the pyramids S_j.

B) The problem in 1) is avoided by not looking at all on the remaining polytope K_0, but rather appropriately covering the original polytope by translated copies of the pyramids S_j and patching the approximants on them (which were established in A) together.

We start with the last tusk, i.e. by constructing an appropriate pyramidal covering.

3   Pyramidal covering and approximation on general polytopes

We shall write $a_n \prec b_n$ if there is a $C > 0$ such that for sufficiently large n we have $|a_n| \le Cb_n$. If we also specify a range $n \ge n_0$ for n, then $a_n \prec b_n$, $n \ge n_0$ means that there is a $C > 0$ for which $|a_n| \le Cb_n$ for all $n \ge n_0$.

We shall prove Theorem 1.1 using bounded linear operators mapping $C(K)$ into Π_n^d (since Π_n^d is finite dimensional, it does not matter what norm is used on Π_n^d). Therefore, we introduce the notation

$$E_n^L(f)_K \prec \omega(1/n) \tag{3.1}$$

to denote the fact that for all sufficiently large n (independent of f) there are bounded linear operators $L_n : C(K) \to \Pi_n^d$ such that

$$\|f - L_n f\|_K \le C\omega(1/n). \tag{3.2}$$

---

[1]Indeed, if V is a non-simple polytope in $\mathbf{R}^{d-1}$, $d > 3$, and K is the convex hull in $\mathbf{R}^d$ of V and the point $(0, \ldots, 0, 1) \in \mathbf{R}^d$, then for any hyperplane L in $\mathbf{R}^d$ that is parallel with the $x_d = 0$ hyperplane that intersects K in at least two points the intersection $K \cap L$ is similar to V, hence L cuts off a non-simple pyramid from K.

with some constant C independent of f and n. Sometimes we need to specify a subclass $\mathcal{C}$ of $C(K)$ for which (3.2) is required, in which case we write

$$E_n^L(f)_K \prec \omega(1/n), \qquad f \in \mathcal{C}. \tag{3.3}$$

(3.1) proves Theorem 1.1 for sufficiently large n. Small n's will be separately handled in Theorem 5.1 in Section 5.

In this section we show that if one can approximate in the correct order on pyramids, then the approximation on any polytope follows by a pyramidal covering of the polytope. Then, in the next section, we show how to handle pyramids.

We shall need some simple geometric facts.

3.1 Lemmas on geometry and polynomial partitions

Lemma 3.1 *For every convex body $S \subset \mathbf{R}^d$ there is a constant $\alpha > 0$ with the following property: if we represent S as the intersection of half spaces, then for any point P in $\mathbf{R}^d$ the distance from P to one of these half spaces is at least α times the distance from P to S.*

Proof.   Let us write a ball B_r of some radius r inside S about a point O of S, and let R be such that the ball B_R about O of radius R contains S in its interior. We show that for any $\beta < 1$ the number $\alpha = \beta r/R$ is suitable.

Indeed, let P be any point outside S. The OP segment intersects the boundary of S at a point Q (see Figure 1). Then $\overline{PQ} \geq \mathrm{dist}(P,S)$, so if we select the point q on the segment PQ close to Q, then $\overline{Pq} \geq \beta \mathrm{dist}(P,S)$ and at the same time $\overline{Oq} < R$ are satisfied (note that Q lies inside B_R, so we may select q so that it also lies inside that ball). There is a half space K among the given half spaces that contains S but does not contain the point q. Let H be the boundary of that half space. Then H intersects the OP segment in a point A (that lies necessarily on the segment Qq), and consider the normal line to H at the point A. This normal line and the line passing through O and P determines a plane L that intersects H in a line ℓ (see below for the case when these two lines coincide). Let B be the closest point of ℓ to O and C the closest point of ℓ to P. Since L is perpendicular to H and $P \in L$, the distance from P to C is precisely the distance from P to the half space K. Furthermore, $\overline{OB} \geq r$ (K contains the ball B_r), $\overline{OA} < R$ (A lies on the segment Qq which lies inside B_R), and from the similarity of the triangles ACP and ABO we obtain that $\overline{PC}/\overline{PA} = \overline{OB}/\overline{OA} > r/R$. Since $\overline{PA} \geq \overline{Pq} \geq \beta \, \mathrm{dist}(P,S)$ is also true, we finally obtain

$$\mathrm{dist}(P,K) = \overline{PC} > (r/R)\overline{PA} \geq (r/R)\beta \, \mathrm{dist}(P,S).$$

The just given proof used that the normal line to H at A and the line passing through O and P are different (so they form a plane). If they are the same, then $\mathrm{dist}(P,K) = \overline{PA} \geq \overline{Pq} \geq \beta \mathrm{dist}(P,S)$.

■
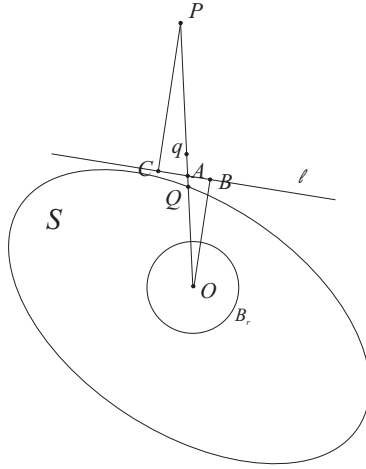
Figure 1:

Lemma 3.2 *Let S be a convex polytope lying in the unit ball $B_1(0)$ of $\mathbf{R}^d$, and let $\varepsilon > 0$. There is a constant $k = k(S, \varepsilon)$ such that for every $n \geq 1$ there are polynomials R_{kn} of degree at most kn for which*

a) $0 \leq R_{kn}(x) \leq 1$ *for $x \in B_1(0)$,*

b) $1 - R_{kn}(x) \leq 2^{-n}$ *for $x \in S$,*

c) $R_{kn}(x) \leq 2^{-n}$ *for $x \in B_1(0)$, $\mathrm{dist}(x, S) \geq \varepsilon$.*

Remark 3.3 1. The lemma is closely related to well localized polynomial partitions of unity. Indeed, if $S_1, \ldots, S_m$ are finitely many closed convex bodies in $B_1(0)$ and $U_1, \ldots, U_m$ are neighborhoods of them, then there is a constant $\theta > 0$ and for each large n there are polynomials $P_{1,n}, \ldots, P_{m,n}, P_{m+1,n}$ of degree at most n such that

- $0 \leq P_{j,n} \leq 1$ on $B_1(0)$ for all j,

- $\sum_{j=1}^{m+1} P_{j,n} \equiv 1$,

- $P_{j,n} \leq \theta^n$ on $B_1(0) \setminus U_j$ for all $1 \leq j \leq m$ and

- $P_{m+1,n} \leq \theta^n$ on $\cup_{j=1}^{m} S_j$.

In particular, if the S_j are pairwise disjoint, then (by choosing U_j to lie outside $\cup_{s \neq j} S_s$) we can also have (with a possibly large θ)

- $1 - \theta^n \leq P_j \leq 1$ on S_j for all $1 \leq j \leq m$.

8

Indeed, to see that, we may assume that each S_j is a convex polytope (just place such a polytope in between S_j and U_j). Then we can apply Lemma 3.2 to each S_j (with an ε for which the ε-neighborhood of S_j is part of U_j) to get polynomials $R_{j,n/m}$ of degree at most n/m as in Lemma 3.2 (just apply the lemma with n replaced by n/km). Now it is easy to see that

$$P_{j,n} = R_{j,n/m}(1 - R_{1,n/m}) \cdots (1 - R_{j-1,n/m}), \qquad j = 1, \ldots, m$$

and

$$P_{m+1,n} = (1 - R_{1,n/m}) \cdots (1 - R_{m,n/m})$$

satisfy the requirements. The θ depends on m, ε and on the numbers $k(S_j, \varepsilon)$ in Lemma 3.2.

2. To our knowledge in more than one variables polynomial partitions of unity were created for the fist time in [7] (that paper dealt with dimension 2). There the decrease was of power type away from the sets, and that decrease started right at the boundary of the sets S_j. In the just given construction $P_{j,n}$ tends to 0 geometrically fast outside U_j, though there is no information on its behavior in $U_j \setminus S_j$. In patching local polynomial approximants together to a global approximant one needs, in general, a geometric rate of convergence to 0 as just has been explained.

Proof of of Lemma 3.2. Let $\alpha > 0$ be the constant from Lemma 3.1 for the polytope S.

It is well known (see e.g. [12, Corollary VI.3.6]) that for every $\varepsilon > 0$ there are an s and for every n polynomials $\tilde{R}_{sn}(x_1)$ of a single variable x_1 of degree sn such that $0 \le \tilde{R}_{sn} \le 1$ on $[-3, 3]$, $1 - \tilde{R}_{sn} \le 3^{-n}$ on $[-3, 0]$ and $\tilde{R}_{sn} \le 3^{-n}$ on $[\alpha\varepsilon, 3]$. We regard $\tilde{R}_{sn}$ as a polynomial of degree sn of the variables $(x_1, \ldots, x_d)$.

S is the intersection of some half-spaces $K_1, \ldots, K_m$. We may assume that the boundary of each K_j contains a point of S. If Φ_j is an isometry of $\mathbf{R}^d$ that maps K_j into the half-space $x_1 \le 0$, then $\tilde{R}_{sn}(\Phi_j(x))$ is a polynomial of degree at most sn, it is in between 0 and 1 on the unit ball (note that the unit ball is mapped into the strip $-3 \le x_1 \le 3$ by Φ_j), it is closer to 1 than 3^{-n} on S (S is mapped into $-3 \le x_1 \le 0$) and it is closer to 0 than 3^{-n} at any point of $B_1(0) \setminus K_j$ which lies of distance $\ge \alpha\varepsilon$ from K_j. However, by the choice of α, any point in the unit ball that is of distance $\ge \varepsilon$ from S lies of distance $\ge \alpha\varepsilon$ from one of these K_j, hence, the product of the polynomials $\tilde{R}_{sn}(\Phi_j(x))$, $j = 1, \ldots, m$, is appropriate in the lemma for sufficiently large n.

∎

Lemma 3.4 *Let $S \subset K$ be convex polytopes in $\mathbf{R}^d$, and let S^ε be the closed ε-neighborhood of S. Let furthermore $H \subset K$ be a closed set such that for some $\delta > 0$ the intersection $S \cap H$ contains a (d-dimensional) ball of radius δ. Then for every l_0 there is an l such that*

$$E_{l_0 n}^L(f)_{S^\varepsilon \cap K} \prec \omega(1/n), \qquad E_{l_0 n}^L(f)_H \prec \omega(1/n), \qquad f \in C(K) \qquad (3.4)$$

imply

$$E_{ln}^L(f)_{S \cup H} \prec \omega(1/n), \qquad f \in C(K). \tag{3.5}$$

Here the constant in the last $\prec$ does not depend on f, it depends only on the diameter of K, on δ and on the constants in the preceding two $\prec$. The quantity $\omega(1/n)$ is a generic term that tends to 0 as $n \to \infty$ — we shall use the lemma with $\omega(1/n)$ replaced by the $1/n$-value of a certain modulus of smoothness. Recall also our convention that the implication $(3.4) \Rightarrow (3.5)$ is not for individual functions, but should be read that if (3.4), i.e.

$$E_{l_0 n}^L(f)_{S^\varepsilon \cap K} \le C_0 \omega(1/n), \qquad E_{l_0 n}^L(f)_H \le C_0 \omega(1/n), \quad n \ge n_0$$

holds for all functions $f \in C(K)$ with some constants C_0, n_0, then (3.5), i.e.

$$E_{ln}^L(f)_{S \cup H} \le C_1 \omega(1/n), \quad n \ge n_1$$

is also true with some C_1, n_1 for all such f.

Proof. We may assume $K \subset B_1(0)$, i.e. that K lies in the unit ball of $\mathbf{R}^d$.
First we recall

Lemma 3.5 [14, Lemma 4.1] *If B is a ball of radius δ lying in the unit ball $B_1(0)$, then for any polynomial Q_n of degree at most n*

$$\|Q_n\|_{B_1(0)} \le \|Q_n\|_B (4/\delta)^n. \tag{3.6}$$

Let $P_{1,n}$ and $P_{2,n}$ be bounded linear operators on $C(K)$ mapping into $\Pi_{l_0 n}^d$ such that

$$\|f - P_{1,n}f\|_{S^\varepsilon \cap K} \le C\omega(1/n), \qquad \|f - P_{2,n}f\|_H \le C\omega(1/n)$$

for $f \in C(K)$. By assumption there is a ball $B \subseteq H \cap S$ of some radius $\delta > 0$, and on that ball we have

$$\|P_{1,n}f - P_{2,n}f\|_B \le \|f - P_{2,n}f\|_H + \|f - P_{1,n}f\|_{S^\varepsilon \cap K} \le 2C\omega(1/n),$$

hence, by Lemma 3.5

$$\|P_{1,n}f - P_{2,n}f\|_{B_1(0)} \le 2C(4/\delta)^{l_0 n}\omega(1/n). \tag{3.7}$$

Choose an integer b such that $(1/2)^b < (\delta/4)^{l_0}$ and with the polynomials R_{kn} from Lemma 3.2 for S and with $l = (l_0 + kb)$ set

$$L_{ln}f = R_{kbn}P_{1,n}f + (1 - R_{kbn})P_{2,n}f.$$

This is a bounded linear operator from $C(K)$ to Π_{ln}^d, and for it we have

- on $H \cap S^\varepsilon$

$$|f - L_{ln}f| \le R_{kbn}|f - P_{1,n}f| + (1 - R_{kbn})|f - P_{2,n}f| \le C\omega(1/n), \tag{3.8}$$

10

- on S (see (3.7))

$$
\begin{aligned}
|f - L_{ln}f| &= |f - P_{1,n}f + (1 - R_{kbn})(P_{1,n}f - P_{2,n}f)| \leq |f - P_{1,n}f| + \\
&\quad + (1 - R_{kbn})|P_{1,n}f - P_{2,n}f| \leq C\omega(1/n) + 2^{-bn}2C(4/\delta)^{l_0 n}\omega(1/n) \\
&\leq 3C\omega(1/n), \tag{3.9}
\end{aligned}
$$

- and on $H \setminus S^{\varepsilon}$

$$
\begin{aligned}
|f - L_{ln}f| &= |(f - P_{2,n}f - R_{kbn}(P_{1,n}f - P_{2,n}f)| \leq |f - P_{2,n}f| \\
&\quad + R_{kbn}|P_{1,n}f - P_{2,n}f| \leq C\omega(1/n) + 2^{-bn}2C(4/\delta)^{l_0 n}\omega(1/n) \\
&\leq 3C\omega(1/n). \tag{3.10}
\end{aligned}
$$

The inequalities (3.8)–(3.10) verify the lemma.

∎

As usual, we identify a point $Q \in \mathbf{R}^d$ with its position vector $\overrightarrow{OQ}$, where O is the origin.

Let $\Phi_{Q,\eta}$ be the homothetic transformation of $\mathbf{R}^d$ with center at Q and with dilation factor η, i.e. $\Phi_{Q,\eta}$ is the linear transformation of $\mathbf{R}^d$ for which $\Phi_{Q,\eta}(P) = Q + \eta(P - Q)$ for any $P \in \mathbf{R}^d$.

Let K be a convex polytope in $\mathbf{R}^d$ with vertices $Q_1, \ldots, Q_m$.

Lemma 3.6 K *is covered by the sets* $\Phi_{Q_j, d/(d+1)}K$, $j = 1, \ldots, m$.

Proof. By taking a simplicial decomposition[2] of K into d-dimensional simplices with vertices from the set $\{Q_1, \ldots, Q_m\}$, we may assume that K is a simplex (so $m = d + 1$). Then any point X in K is of the form $X = \alpha_1 Q_1 + \cdots \alpha_{d+1} Q_{d+1}$ with $\alpha_j \geq 0$, $\sum \alpha_j = 1$. Here one of the α_j's, say α_1, is at least $1/(d+1)$, hence from the representation

$$
X = (1 - \alpha_1)\Big(\frac{\alpha_2}{1-\alpha_1}Q_2 + \cdots + \frac{\lambda_{d+1}}{1-\alpha_1}Q_{d+1}\Big) + \alpha_1 Q_1 =: (1 - \alpha_1)Y + \alpha_1 Q_1
$$

we can see that X divides the segment $\overline{YQ_1}$ into segments with length-ratio $\alpha_1/(1 - \alpha_1) \geq 1/d$, which implies that X belongs to $\Phi_{Q_1, d/(d+1)}K$ because $Y \in K$.

∎

---

[2]Actually, all we need is Caratheodory's theorem that any point of the convex hull of a set H in $\mathbf{R}^d$ lies in the convex hull of some of its $(d+1)$ points, so if K is a polytope in $\mathbf{R}^d$, then it is the union of the simplices spanned by $(d+1)$ vertices of K.

Lemma 3.7 *Let K be as in the preceding lemma and let $0 < \theta < 1/(d+1)$ be fixed. If P is any point of $\Phi_{Q_j, 1-\theta}K$ and we translate $\Phi_{Q_j, \theta}K$ by the vector $\overrightarrow{Q_j P}$, then the translated polytope is part of K.*

Proof. Indeed, if R is any point of $\Phi_{Q_j, \theta}K$ and we denote the vectors $\overrightarrow{Q_j P}$ and $\overrightarrow{Q_j R}$ by $\underline{v}$ and $\underline{u}$, respectively, then $Q_j + (1-\theta)^{-1}\underline{v}$, $Q_j + \theta^{-1}\underline{u}$ belong to K, hence so does the point $(1-\theta)(Q_j + (1-\theta)^{-1}\underline{v}) + \theta(Q_j + \theta^{-1}\underline{u})$ on their connecting segment. But this point is $Q_j + \underline{u} + \underline{v} = R + \underline{u} = R + \overrightarrow{Q_j P}$, which means that $\Phi_{Q_j, \theta}K + \overrightarrow{Q_j P} \subset K$. This proof also gives that if P lies in the d-dimensional interior of $\Phi_{Q_j, 1-\theta}K$, then the translated polytope $\Phi_{Q_j, \theta}K + \overrightarrow{Q_j P}$ also lies in the interior of K.

$\blacksquare$

What we have just proven can be applied on every lower dimensional face of K. Indeed, let F_k be a k-dimensional face of K that contains Q_j. Since $F_k \cap \Phi_{Q_j, \theta}K = \Phi_{Q_j, \theta}F_k$, what we have just established can be applied to the face F_k and to $\Phi_{Q_j, \theta}F_k$. Thus, if $P \in \Phi_{Q_j, 1-\theta}K$ is on the face F_k, then the face $\Phi_{Q_j, \theta}F_k + \overrightarrow{Q_j P}$ of the translated polytope $\Phi_{Q_j, \theta}K + \overrightarrow{Q_j P}$ lies in F_k, and if P lies in the k-dimensional interior of $\Phi_{Q_j, 1-\theta}F_k$, then $F_k \cap (\Phi_{Q_j, \theta}K + \overrightarrow{Q_j P})$ lies in the k-dimensional interior of F_k. In the latter case if P' lies in the k dimensional interior of $F_k \cap (\Phi_{Q_j, \theta}K + \overrightarrow{Q_j P})$, then we can enlarge $\Phi_{Q_j, \theta}K + \overrightarrow{Q_j P}$ from P' by a factor $1 + \tau > 1$ with some small $\tau > 0$ so that the enlarged polytope is still part of K. Furthermore, this enlarged polytope then contains the intersection of a neighborhood of $\Phi_{Q_j, \theta}K + \overrightarrow{Q_j P}$ with F_k (note however, that it does not contain a d-dimensional neighborhood of $\Phi_{Q_j, \theta}K + \overrightarrow{Q_j P}$ since all points of $F_k \cap (\Phi_{Q_j, \theta}K + \overrightarrow{Q_j P})$ lie on the boundary — actually on the corresponding k-dimensional face — of the enlarged polytope).

3.2 Pyramidal coverings

In this section we describe a covering of a polytope K by pyramids K_s with the property that there is a listing of the elements K_s in the covering in which consecutive elements have relatively large intersections (the intersection of their interiors is non-empty). This covering will then allow us in the next subsection, by repeatedly applying the results from the preceding subsection, to obtain a global polynomial approximant on K (to a given function f) from approximants on the K_s. In this way we reduce the problem of polynomial approximation on K to approximation on (the) pyramids (K_s).

Since the construction is rather involved with many notations, for the sake of the reader we shall list the notations in a separate subsection at the end of this section.

Let K and Q_j as before, and fix $0 < \theta < 1/(d+1)$. Around each Q_j we cut off from K a small pyramid S_j with vertex at Q_j by a $(d-1)$-dimensional
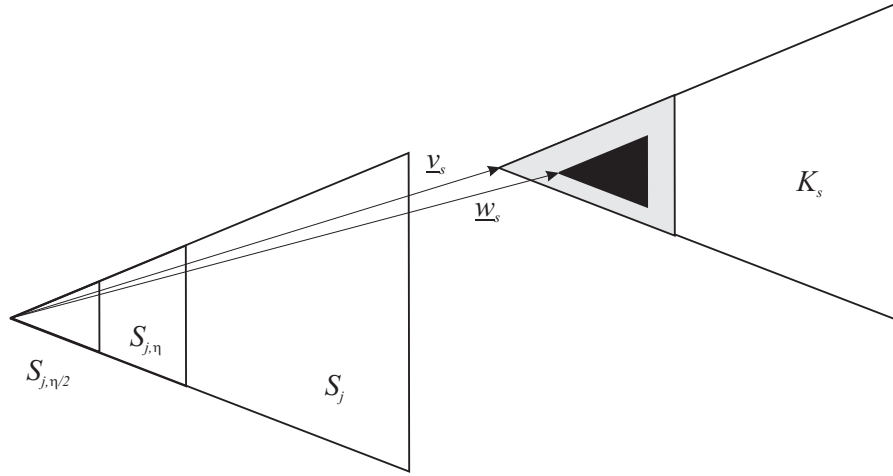
Figure 2: A schematic figure of the sets S_j, $S_{j,\eta}$, $S_{j,\eta/2}$, K_s, $K_{\eta,s}$ (lightly shaded figure) and $\mathcal{K}_{\eta/2,s}$ (dark shaded figure), as well as the vectors $\underline{v}_s, \underline{w}_s$.

hyperplane. We assume that for each j we have $S_j \subset \Phi_{Q_j,\theta}K$, and we fix these S_j in the following discussion.

For simplicity let us also write $S_{j,\eta}$ for the homothetic copy $\Phi_{Q_j,\eta}S_j$.

Let $\eta < 1/2$ be given. We shall deal with the following type of sets K_s, $K_{\eta,s}$ and $\mathcal{K}_{\eta/2,s}$, $1 \leq s \leq T$, where T will be given during the construction below.

- Each $K_s = S_j + \underline{v}_s$ is a translated copy of some S_j by a translation vector $\underline{v}_s$,

- $K_{\eta,s} \subset S_{j,1/2} + \underline{v}_s$ is the translation of the η-size copy $S_{j,\eta}$ of S_j by the same vector $\underline{v}_s$, and

- $\mathcal{K}_{\eta/2,s} \subseteq K_{\eta,s}$ is a translation of the smaller $S_{j,\eta/2}$ by a possibly different vector $\underline{w}_s$.

See Figure 2. For emphasis we use calligraphic letter in $\mathcal{K}_{\eta/2,s}$ to indicate that the vertex of $\mathcal{K}_{\eta/2,s}$ may be at a different point than the common vertices of K_s and $K_{\eta,s}$.

By an η-pyramid covering of K we mean a covering $K = \cup_{s=1}^{T} K_s$, where each K_s is a translated copy of one of the S_j's, and which has the following properties: there are an $\varepsilon > 0$ and related subpyramids $K_{\eta,s}$, $\mathcal{K}_{\eta/2,s}$ of K_s as just described such that

(i) each K_s is contained in K and each $\mathcal{K}_{\eta/2,s}$ is contained in $K_{\eta,s}$,

(ii) $K = \cup_{s=1}^{T} \mathcal{K}_{\eta/2,s}$,

(iii) the intersection of the interiors of $\mathcal{K}_{\eta/2,s}$ and $\mathcal{K}_{\eta/2,s+1}$ is non-empty for all $s = 1, \ldots, T-1$,

(iv) $\mathcal{K}_{\eta/2,s}^{\varepsilon} \cap K \subset K_{\eta,s}$, where $\mathcal{K}_{\eta/2,s}^{\varepsilon}$ is the closed ε-neighborhood of $\mathcal{K}_{\eta/2,s}$.

Proposition 3.8 *For every $\eta < 1/2$ there is an η-pyramid covering of K.*

Proof. We say that a finite set $X \subset K$ is τ-dense in K if every point of K lies closer to some point of X than τ.

We define for some small $0 < \tau$ a τ-dense set X on K as follows. First of all, let X_1 be the set of all vertices Q_j of K. Then for every edge F_1 of K extend $F_1 \cap X_1$ to a τ-dense set of F_1. X_1 and the added points for all edges form X_2. Then for every 2 dimensional face F_2 of K extend $F_2 \cap X_2$ to a τ-dense set of F_2, etc. keep extending the point sets X_k using higher and higher dimensional faces. This way we get a τ-dense set X_{d-1} on the boundary of K, and finally extend that to a τ-dense set on K.

First we define the sets $\mathcal{K}_{\eta/2,s}$. If $P \in X$ and P belongs to $\Phi_{Q_j,1-\theta}K$ (there must be such a j in view of Lemma 3.6) then let K_P be the translation of $S_{j,\eta/2} = \Phi_{Q_j,\eta/2}S_j$ by the vector $\overrightarrow{Q_jP}$. If for a P there are more than one j with $P \in \Phi_{Q_j,1-\theta}K$, then get a K_P for all such j. We show that the collection of all these K_P generate a suitable η-pyramid covering by repeating each K_P a few times (to meet the requirement in (iii)). According to Lemma 3.7 we have $S_j + \overrightarrow{Q_jP} \subset K$ (recall that $S_j \subset \Phi_{Q_j,\theta}K$), and since $\eta < 1/2$ we also have

$$K_P \subset S_{j,1/2} + \overrightarrow{Q_jP} \subset S_j + \overrightarrow{Q_jP} \subset K. \tag{3.11}$$

If τ is sufficiently small, then these K_P cover K. Indeed, from how X was constructed it is clear that the vertices, and hence small neighborhoods (relative to K) of them are covered (note that each Q_j belongs to X, so $S_{j,\eta/2}$ is one of the K_P's). Then each edge F_1, and hence a small neighborhood (relative to K) of F_1 is covered (this neighborhood depends only on η and not on τ), and so on, we continue with higher and higher dimensional faces and their small neighborhoods, and finally with all of K. This proves (ii).

We also get that for small $\tau > 0$ the interiors $\mathrm{Int}(K_P)$ of the pyramids K_P cover the interior $\mathrm{Int}(K)$ of K. At this point we need the following simple lemma.

Lemma 3.9 *Let $\mathcal{G}$ be a connected set and suppose that $\mathcal{G} = \cup_i \mathcal{G}_i$, where $\mathcal{G}_i$ are finitely many open sets. Let $\{\mathcal{G}_i\}_i$ be the vertices of a graph, and connect in that graph the vertices $\mathcal{G}_i$ and $\mathcal{G}_j$ if for the sets $\mathcal{G}_i, \mathcal{G}_j$ we have $\mathcal{G}_i \cap \mathcal{G}_j \neq \emptyset$. Then this graph is connected.*

This lemma says that in any covering $\mathcal{G} = \cup_i \mathcal{G}_i$ of a connected open set $\mathcal{G}$ by open sets $\mathcal{G}_i$, for any index pairs $k \neq l$ there are indices $i_0, i_j, \ldots, i_m$ with some m such that $i_0 = k$, $i_m = l$, and for any $0 \leq j < m$ we have $\mathcal{G}_{i_j} \cap \mathcal{G}_{i_{j+1}} \neq \emptyset$.

Proof. For clearer notation let us write $\mathcal{G}_j^V$ if we talk about $\mathcal{G}_j$ as a vertex of the graph. Let $V_1, \ldots, V_r$ be the vertex sets of the connected components of the graph (i.e. each V_s is a non-empty set of certain $\mathcal{G}_j^V$, $V_{s_1} \cap V_{s_2} = \emptyset$ for $s_1 \neq s_2$,

and each $\mathcal{G}_j^V$ belongs to one of the V_s). Each $\cup_{\mathcal{G}_j^V \in V_i} \mathcal{G}_j$ is open, and for different components V_i these sets are disjoint. Now the conclusion follows from the fact that a connected set cannot be the disjoint union of non-empty open sets.

∎

Since the just constructed graph for $\text{Int}(K)$ and $\{\text{Int}(K_P)\}_{P \in X}$ is connected, there is a path in it which goes through every vertex. In other words, there is a listing $\mathcal{K}_{\eta/2,s}$, $s = 1, \ldots, T$, of all the K_P's (maybe with repetition) in which $\mathcal{K}_{\eta/2,s} \cap \mathcal{K}_{\eta/2,s+1}$ contains a ball for all $s < T$. This proves (iii).

We still need to define the sets $K_{\eta,s}$ and K_s, see Figure 2. Consider a $\mathcal{K}_{\eta/2,s} = K_P$, say it is $S_{j,\eta/2} + \overrightarrow{Q_j P}$ (i.e. the vector $\underline{w}_s$ in the definition of pyramidal covering is $\overrightarrow{Q_j P}$ in this case). If P is the vertex Q_j, then enlarge K_P from P by a factor 2 to get $K_{\eta,s}$. However, if P is not a vertex of K, then there is a smallest dimensional face F_k of K such that P lies in the interior $\text{Int}_k(F_k)$ of F_k (if F_k is k-dimensional, then $\text{Int}_k(F_k)$ denotes its k-dimensional interior), where we consider K itself as the d-dimensional face of K. Hence, according to what was said after Lemma 3.7, $F_k \cap (S_{j,\eta} + \overrightarrow{Q_j P})$ also lies in $\text{Int}_k(F_k)$. Let $Q \in \text{Int}_k(F_k \cap S_{j,\eta})$ be a point lying in the k-dimensional interior of $F_k \cap S_{j,\eta}$. Then $Q + \overrightarrow{Q_j P}$ lies in the k-dimensional interior $\text{Int}_k(F_k \cap (S_{j,\eta} + \overrightarrow{Q_j P}))$ of the face $F_k \cap (S_{j,\eta} + \overrightarrow{Q_j P})$ of $S_{j,\eta} + \overrightarrow{Q_j P}$, and the translation of $\text{Int}_k(F_k \cap (S_{j,\eta} + \overrightarrow{Q_j P}))$ by the vector $\overrightarrow{QQ_j}$ will contain $F_k \cap K_P$ if Q lies sufficiently close to Q_j (and at the same time this translation is contained in F_k). Thus, if we set

$$K_{\eta,s} = S_{j,\eta} + \overrightarrow{QP}$$

(i.e. the vector $\underline{v}_s$ in the pyramidal covering is $\overrightarrow{QP}$), then this $K_{\eta,s}$ is such that it contains the intersection of a neighborhood of $\mathcal{K}_{\eta/2,s} = K_P$ with K, proving (iv).

Finally, with the previous notations we set $K_s = S_j + \overrightarrow{QP}$, which is part of K (see Lemma 3.7), so (i) holds.

∎

List of notations used in the pyramidal covering construction

For the sake of the reader we repeat here in a concised form the notations used in the pyramidal covering.

$\Phi_{Q,\eta}$ – homothetic transformation with center at Q and dilation factor η

Q_j – vertices of K

S_j $(\subset K)$ – small pyramids cut off from K with vertex at S_j

$S_{j,\eta}$ – the homothetic copy $\Phi_{Q_j,\eta} S_j$ of S_j

K_s – an $S_j + \underline{v}_s$ with a translation vector $\underline{v}_s$

$K_{\eta,s}$ – $S_{j,\eta} + \underline{v}_s$ with the same translation vector $\underline{v}_s$

$\mathcal{K}_{\eta/2,s}$ $(\subset K_{\eta,s})$ – $S_{j,\eta/2} + \underline{w}_s$ with another translation vector $\underline{w}_s$

3.3 Patching approximants on a pyramidal covering together to a global approximant

Let $\tilde{\omega}_{S_j}^r(f, 1/n)$ be the modulus of smoothness on S_j when only the directions in the apex edges of S_j are used. We shall use again the notation $S_{j,\eta}$ for $\Phi_{Q_j,\eta} S_j$.

Proposition 3.10 *If for each $j = 1, \ldots, m$ we have*

$$E_n^L(f)_{S_{j,2\eta}} \prec \tilde{\omega}_{S_j}^r(f, 1/n), \qquad f \in C(S_j), \tag{3.12}$$

then there is an l such that

$$E_{ln}^L(f)_K \prec \omega_K^r(f, 1/n), \qquad f \in C(K).$$

Recall that here, on the right hand side, the modulus of smoothness uses the edge directions of K.

Proof. Let $K = \cup_{s=1}^T K_s$ be an η-pyramid covering of K as in Proposition 3.8, and set $H_t = \cup_{s=1}^t \mathcal{K}_{\eta/2,s}$. With induction we are going to prove that for every $1 \le t \le T$ there is an l for which

$$E_{ln}^L(f)_{H_t} \prec \omega_K^r(f, 1/n). \tag{3.13}$$

The claim then follows by setting $t = T$.

For $t = 1$ the set $H_1 = \mathcal{K}_{\eta/2,1}$ is an $S_{j,\eta/2} + \underline{w}_1$ for some j and some vector $\underline{w}_1$, $K_1 = S_j + \underline{v}_1$ with the same j and with another vector $\underline{v}_1$, and the corresponding $K_{\eta,1}$ is $S_{j,\eta} + \underline{v}_1$. Note that $K_{\eta,1}$ is contained in $S_{j,1/2} + \underline{v}_1$. By the assumption (make a homothecy with factor 2 and a translation by $\underline{v}_1$) there are linear operators $L_n : C(S_{j,1/2} + \underline{v}_1) \to \Pi_n^d$ such that

$$\|f - L_n f\|_{S_{j,\eta} + \underline{v}_1} \prec \tilde{\omega}_{S_{j,1/2} + \underline{v}_1}^r(f, 1/n). \tag{3.14}$$

Since $\mathcal{K}_{\eta/2,1} = S_{j,\eta/2} + \underline{w}_1$ is part of $S_{j,\eta} + \underline{v}_1$ (see property (i) of the η-pyramid covering), and the right-hand side in (3.14) is at most $\omega_K^r(f, 1/n)$ (note that all apex-edge directions of $S_j + \underline{v}_1$ are also edge directions of K), we obtain

$$\|f - L_n f\|_{\mathcal{K}_{\eta/2,1}} = \|f - L_n f\|_{S_{j,\eta/2} + \underline{w}_1} \prec \omega_K^r(f, 1/n), \tag{3.15}$$

and hence the claim follows in this case.

The induction step is a consequence of Lemma 3.4. Indeed, if (3.13) is known for $t - 1$ ($t > 1$), say

$$E_{l_0 n}^L(f)_{H_{t-1}} \prec \omega_K^r(f, 1/n), \tag{3.16}$$

then let $H = H_{t-1}$, $S = \mathcal{K}_{\eta/2,t}$ in Lemma 3.4, and set $\omega(1/n) = \omega_K^r(f, 1/n)$ there (note that by property (iii) of the η-pyramidal covering $H \cap S = H_{t-1} \cap \mathcal{K}_{\eta/2,t} \supseteq \mathcal{K}_{\eta/2,t-1} \cap \mathcal{K}_{\eta/2,t}$ contains a ball). As we have just seen (see the argument leading to (3.14)–(3.15)), (3.12) implies

$$E_n^L(f)_{K_{\eta,t}} \prec \omega_K^r(f, 1/n).$$

16

Furthermore, by property (iv) of an η-pyramid covering, for some $\varepsilon > 0$ the relative (with respect to K) neighborhood $\mathcal{K}_{\eta/2,t}^{\varepsilon} \cap K$ of $\mathcal{K}_{\eta/2,t}$ is part of $K_{\eta,t}$, hence we obtain

$$E_n^L(f)_{\mathcal{K}_{\eta/2,t}^{\varepsilon} \cap K} \prec \omega_K^r(f, 1/n),$$

and together with this also

$$E_{l_0 n}^L(f)_{\mathcal{K}_{\eta/2,t}^{\varepsilon} \cap K} \prec \omega_K^r(f, 1/n).$$

Now the claim follows for t (i.e. for $H_t = H_{t-1} \cup \mathcal{K}_{\eta/2,t}$) in view of this estimate, (3.16) and Lemma 3.4.

■

4 Approximation on pyramids and the proof of Theorem 1.1 for large degrees

Let S_0 be a $(d-1)$-dimensional convex polytope lying in the hyperplane H of $\mathbf{R}^d$ and S the convex hull of $S_0 \cup \{P\}$, where P is a point outside H (so S is a d dimensional pyramid with apex at P). Recall that the edges of S emanating from P are called apex edges.

Assume that $P = 0$. Let λH be the hyperplane that we obtain from H by a dilation with factor $0 < \lambda < 1$ and center $P = 0$. Then λH cuts S into two parts, the pyramid λS and the polytope $U_\lambda := S \setminus \lambda S$ such that S_0 and λS_0 are $(d-1)$-dimensional similar faces of U_λ. Let $e_1, \ldots, e_m$ be the segments connecting the corresponding points of S_0 and λS_0. These segments lie on the apex edges of S, on each apex edge lying one of them. If E is the endpoint of e_j lying in λS_0, then the product set $(\lambda S_0) \times e_j$ is defined as the union of all the translations $e_j + \overrightarrow{EM}$, where M runs through the points of λS_0 (see Figure 3).

Proposition 4.1 *If $\lambda > (d-1)/d$, then the product sets $(\lambda S_0) \times e_j$, $1 \le j \le m$ have a common interior and their union is U_λ.*

Proof. Let the vertices of S_0 be $Q_1, \ldots, Q_m$. Then the vertices of λS are $\lambda Q_1, \ldots, \lambda Q_m$, and we may assume that $e_j = \overline{Q_j(\lambda Q_j)}$. As before, let $\Phi_{Q_j,\mu} S_0$ be the homothetic image of S_0 with factor μ and dilation center at Q_j.

Let Z be any point in U_λ, and ℓ_Z be the half-line from the origin through Z. This ℓ_Z intersects S_0 in a point X, and by Lemma 3.6 $X \in \Phi_{Q_j,(d-1)/d} S_0$ for some $1 \le j \le m$. If $Z = \sigma X$, then $\lambda \le \sigma \le 1$, and $Z \in \Phi_{\sigma Q_j,(d-1)/d}(\sigma S_0)$. However, this latter set is a $\sigma(d-1)/d$ size copy $(\Phi_{Q_j,\sigma(d-1)/d}(S_j))$ of S_0 translated by the vector $\overrightarrow{Q_j(\sigma Q_j)}$, so it is contained in the $\lambda > \sigma(d-1)/d$ size copy $\Phi_{Q_j,\lambda} S_0$ of S_0 translated by the same vector $\overrightarrow{Q_j(\sigma Q_j)}$.

The translation of λS_0 by the vector $\overrightarrow{(\lambda Q_j)Q_j}$ is $\Phi_{Q_j,\lambda} S_0$, and the product set $(\lambda S_0) \times e_j$ is precisely the set of points that lie in one of the translations of λS_0 by a vector $\overrightarrow{(\lambda Q_j)(\tau Q_j)}$, $\lambda \le \tau \le 1$. But then, in view of $\overrightarrow{Q_j(\sigma Q_j)} =$
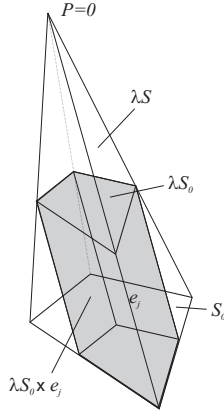
Figure 3: The pyramid S, its base S_0 and their λ-dilations from the apex $P = 0$. The shaded slab is the product set $\lambda S_0 \times e_j$.

$\overrightarrow{Q_j(\lambda Q_j)} + \overrightarrow{(\lambda Q_j)(\sigma Q_j)}$, what we have just proven means that Z is in this product set $(\lambda S_0) \times e_j$.

The claim that the product sets $(\lambda S_0) \times e_j$ have a common interior point is trivial, since if Y lies in the $(d-1)$-dimensional interior of λS_0, then all points of $U_\lambda = S \setminus \lambda S$ that lie sufficiently close to Y lie in all the product sets $(\lambda S_0) \times e_j$.
∎

Remark 4.2 *Easy modification of the proof gives that if $1 > \lambda > (d-1)/d$ and $\mu < 1$, $\mu\lambda > (d-1)/d$, then for μ sufficiently close to 1 even the product sets $\Phi_{\lambda Q_j,\mu}(\lambda S_0) \times e_j$, $1 \le j \le m$ have a common interior and their union is U_λ.[3] For these sets it is clearly true that for small $\varepsilon > 0$*

$$\left(\Phi_{\lambda Q_j,\mu}(\lambda S_0) \times e_j\right)^\varepsilon \cap U_\lambda \subset (\lambda S_0) \times e_j, \tag{4.1}$$

where $\cdot^\varepsilon$ denotes ε-neighborhood.

Proposition 4.3 *Suppose that S_0 is a polytope in $\mathbf{R}^{d-1}$ for which there are bounded linear operators $V_n : C(S_0) \to \Pi_n^{d-1}$ such that*

$$\|f - V_n f\|_{S_0} \prec \omega_{S_0}^r(f, 1/n)_{\mathcal{E}},$$

---

[3]Actually, the claim is true for all $\mu < 1$ with $\mu\lambda > (d-1)/d$. Indeed, in view of Proposition 4.1 we need to prove only that the convex sets $\Phi_{\lambda Q_j,(d-1)/d}\lambda S_0$ have non-empty intersection. For this the following simple argument was communicated to the author by János Kincses. In view of Helly's theorem all these sets have non-empty intersection if any d of them have non-empty intersection (note that we are working now in a $d-1$ dimensional hyperplane of $\mathbf{R}^d$). However, if $j_1, \ldots, j_d$ are different indices, then the center of mass of the $((d-1)$-dimensional) simplex Δ determined by $\lambda Q_{j_1}, \ldots, \lambda Q_{j_d}$ lies in all $\Phi_{\lambda Q_{j_s},(d-1)/d}\lambda S_0$, $s = 1, \ldots, d$, because the center of mass divides the segment from any vertex Q_{j_s} through the center of mass to the opposite face of Δ in the ration $1 : (d-1)$.

where $\omega^r_{S_0}(f,\delta)_{\mathcal{E}}$ is a modulus of smoothness on S_0 using some directions $\mathcal{E}$. Then there is a sequence of bounded linear operators $L_n : C(S_0 \times [0,1]) \to \Pi_n^d$ such that

$$\|f - L_n f\|_{S_0 \times [0,1]} \prec \omega^r_{S_0 \times [0,1]}(f, 1/n)_{\mathcal{E} \cup \{e_d\}},$$

where e_d is the d-th unit vector in $\mathbf{R}^d$ (perpendicular to $\mathbf{R}^{d-1}$).

Proof. In view of Remark 1.2 it is sufficient to construct the linear operators L_n with the required properties except that $L_n f$ can be a polynomial of degree at most n in each variable (as opposed to the claim where $L_n f$ has total degree at most n).

This is basically [6, Section 12.1]. Indeed, it was proven there that for every r there are $U_n : C[0,1] \to \Pi_n^1$ such that U_n are uniformly bounded linear operators from $C[0,1]$ to $C[0,1]$ with the property that

$$\|f - U_n f\|_{[0,1]} \prec \omega^r_{[0,1]}(f, 1/n)$$

(recall also (1.7)).

Let $\underline{y} = (y_1, \ldots, y_{d-1})$. For a function $f(\underline{y}, x)$ define $f_x(\underline{y}) = f(\underline{y}, x)$ and $f^{\underline{y}}(x) = f(\underline{y}, x)$. Then

$$\omega_{S_0 \times [0,1]}(f, \delta)_{\mathcal{E} \cup \{e_d\}} = \max \left(\sup_{x \in [0,1]} \omega_{S_0}(f_x, \delta)_{\mathcal{E}}, \sup_{\underline{y} \in S_0} \omega_{[0,1]}(f^{\underline{y}}, \delta) \right).$$

When we apply V_n in the variable $\underline{y} \in S_0$, then we denote it by $V_n^{\underline{y}}$. Let now $L_n f = U_n^x \circ V_n^{\underline{y}} f_x$. The function $V_n^{\underline{y}} f_x$ is of the form $\sum_{k_i \le n} a_{\underline{k},n}(x) y_1^{k_1} \cdots y_{d-1}^{k_{d-1}}$ where $\underline{k} = (k_1, \ldots, k_{d-1})$, and here each $a_{\underline{k},n}(x)$ is continuous because V_n are bounded linear operators. Hence, the linearity of U_n yields that $L_n f$ is of the form $\sum_{k_i \le n} U_n^x(a_{\underline{k},n}(x)) y_1^{k_1} \cdots y_{d-1}^{d_{d-1}}$, which is a polynomial of degree at most n in each of the variables $y_1, \ldots, y_{d-1}, x$. Now

$$|L_n f(\underline{y}, x) - f(\underline{y}, x)| \le |U_n^x \circ V_n^{\underline{y}} f_x(\underline{y}) - U_n^x f_x(\underline{y})| + |U_n^x f_x(\underline{y}) - f_x(\underline{y})|.$$

With some fixed C, C_1, C_2 for all large n the first term on the right is

$$
\begin{aligned}
|U_n^x \circ V_n^{\underline{y}} f_x(\underline{y}) - U_n^x f_x(\underline{y})| &= |U_n^x(V_n^{\underline{y}} f_x(\underline{y}) - f_x(\underline{y}))| \\
&\le C \sup_x |V_n^{\underline{y}} f_x(\underline{y}) - f_x(\underline{y}))| \\
&\le C_1 \sup_x \omega_{S_0}(f_x, 1/n)_{\mathcal{E}} \le C_1 \omega_{S_0 \times [0,1]}(f, 1/n)_{\mathcal{E} \cup \{e_d\}},
\end{aligned}
$$

while for each fixed $\underline{y} \in S_0$ the second term is

$$|U_n^x f^{\underline{y}}(x) - f^{\underline{y}}(x)| \le C_2 \omega_{[0,1]}(f^{\underline{y}}, 1/n) \le C \omega_{S_0 \times [0,1]}(f, 1/n)_{\mathcal{E} \cup \{e_d\}}.$$

∎

The preceding proposition says that if we can prove a Jackson-type inequality on a set K_0 via bounded linear operators, then a similar Jackson-type theorem holds on the product space $K_0 \times [0, 1]$. By applying an affine transformation we obtain that the same is true for product sets of the form $K_0 \times e$, where e is not necessarily perpendicular to the hyperspace of K_0 (but lies outside it). In particular, we can apply this to all the product sets $(\lambda S_0) \times e_j$ from Proposition 4.1 above, and we obtain that if

$$E_n^L(f)_{S_0} \prec \omega_{S_0}^r(f, 1/n), \tag{4.2}$$

(which implies the same for the set λS_0 replacing S_0 since λS_0 is a homothetic image of S_0) where $\omega_{S_0}^r(f, \delta)$ is the modulus of smoothness on S_0 using the edge directions S_0, then for each $j = 1, \ldots, m$ we have

$$E_n^L(f)_{(\lambda S_0) \times e_j} \prec \omega_{(\lambda S_0) \times e_j}^r(f, 1/n).$$

Let us return now to the pyramid S and to its decomposition $S = \lambda S \cup U_\lambda$ from the beginning of this section. If we combine the just discussed relation with the fact (4.1) and apply Lemma 3.4 repeatedly with $\omega(1/n) = \omega_{U_\lambda}^r(f, 1/n)$, $H = \cup_{j=1}^{t-1} \Phi_{\lambda Q_j, \mu}(\lambda S_0) \times e_j$, $S = \Phi_{\lambda Q_t, \mu}(\lambda S_0) \times e_t$, $t = 1, \ldots, m$, and $K = U_\lambda$, then we can conclude that (4.2) implies that there is an l such that

$$E_{ln}^L(f)_{U_\lambda} \prec \omega_{U_\lambda}^r(f, 1/n).$$

If this is true for all n with $ln \geq n_r$ (where n_r is some number), then we have proved

Proposition 4.4 *Under the assumption (4.2) there is an n_r such that*

$$E_n^L(F)_{U_\lambda} \prec \omega_{U_\lambda}^r(F, 1/n), \qquad n \geq n_r.$$

Recall our convention that this means that for all functions F continuous on U_λ we have

$$E_n^L(F)_{U_\lambda} \leq C_0 \omega_{U_\lambda}^r(F, 1/n), \qquad n \geq n_r$$

with some constant C_0 independent of F.

Since both sides are homothetic invariant, we also have

$$E_{bn}^L(F)_{qU_\lambda} \prec \omega_{qU_\lambda}^r(F, 1/bn), \qquad bn \geq n_r.$$

for any $q > 0$ and $b > 0$.

In particular, if $\mathcal{E}$ is the set of the apex edge directions of S and for some function F and numbers $q, b > 0$

$$\omega_{qU_\lambda}^r(F, 1/bn) \prec \omega_S^r(f, 1/n)_{\mathcal{E}} \tag{4.3}$$

also holds, then we can conclude

$$E_{bn}^L(F)_{qU_\lambda} \prec \omega_S^r(f, 1/n)_{\mathcal{E}}, \qquad bn \geq n_r. \tag{4.4}$$

Now we are ready to prove Theorem 1.1 for large degrees n.

We shall prove it in the form

20

Theorem 4.5 *If K is a convex polytope in $\mathbf{R}^d$, then*

$$E_n^L(f)_K \prec \omega_K^r(f, 1/n).$$

Recall our convention from Section 3 that $a_n \prec b_n$ means that there is a $C > 0$ such that *for sufficiently large n* we have $|a_n| \leq Cb_n$.

Proof of Theorem 4.5. We use induction with respect to d, the $d = 1$ case being covered by [6, Section 12.1]. Suppose the claim is true in $\mathbf{R}^{d-1}$, and let K be a polytope in $\mathbf{R}^d$.

Let Q be a vertex of K, and cut off a small pyramid S from K by a hyperplane going close to Q. We may assume that S is as small as in the discussion after Lemma 3.6, i.e. $S \subset \Phi_{Q,\theta}K$ (where $\theta < 1/(d+1)$ is a fixed number). The apex of S is at Q and its base S_0 is in the hyperplane of the cut. It was proved in [14, Chapter 11] part A) that the cut can be made so that

a) no base edge e of S is parallel with any $(d-2)$-dimensional base face not containing e,

b) the height of S (i.e. the perpendicular segment to S_0 from Q) lies in the interior of S (except for its two endpoints).

We may assume that $Q = 0$, and that S_0 lies in the plane $x_1 = 2$ (implying that the height is the segment $0 \leq x_1 \leq 2$, $x_2 = \cdots x_d = 0$), and let $K_a = aS \setminus (a/4)S$. Properties a)-b) are enough to prove (see [14, Chapter 6]) that for $a \leq 1/4$ we have

$$\omega_{K_a}^r(F, 1/n\sqrt{a}) \prec \omega_S^r(f, 1/n)_{\mathcal{E}} \tag{4.5}$$

where $\mathcal{E}$ is the set of the apex edge directions of S and where

$$F(x_1, \ldots, x_d) = f(x_1 + L/n^2, x_2, \ldots, x_d) \tag{4.6}$$

with some large but fixed L (in [14, (6.18)] this was stated with a different modulus of smoothness on the right which was needed there, but the proof also gives the inequality with $\omega_S^r(f, \delta)_{\mathcal{E}}$). As a consequence, we get the same conclusion if K_a is replaced by $aS \setminus (\lambda a)S = aU_\lambda$ with any $1/4 \leq \lambda < 1$, where U_λ is the set from Proposition 4.4, i.e.

$$\omega_{aU_\lambda}^r(F, 1/n\sqrt{a}) \prec \omega_S^r(f, 1/n)_{\mathcal{E}} \tag{4.7}$$

also follows.

Now this is then a situation when (4.4) can be applied with $q = a$, $b = \sqrt{a}$ since (4.2) is guaranteed by our induction hypothesis and (4.7) is the same as (4.3), so we can conclude

$$E_{n\sqrt{a}}^L(F)_{aU_\lambda} \prec \omega_S^r(f, 1/n)_{\mathcal{E}} \qquad n\sqrt{a} \geq n_r. \tag{4.8}$$

Recall that here F and f are connected via (4.6). These local approximants can be patched together to a global approximant exactly as in [14, Chapter 8]. Namely, if

$$S_n^* = S \cap \left\{ M/n^2 \leq x_1 \leq 1/16 \right\}$$

with some appropriate large but fixed M, then (4.8) implies that for some l we have

$$E_{ln}^L(F)_{S_n^*} \prec \omega_S^r(f, 1/n)_{\mathcal{E}}.$$

The only difference here compared to [14, Chapter 8] is that there $\lambda = 1/4$ was used (in which case $aU_\lambda = K_a$ with the K_a from (4.5)), but for any $\frac{1}{4} \leq \lambda < 1$ the argument is the same.

Since $\omega_S^r(f, 1/n)_{\mathcal{E}} \prec \omega_S^r(f, 1/ln)_{\mathcal{E}}$, it follows that

$$E_n^L(f)_{(S/64)+((M+L)/n^2)\underline{v}} \prec \omega_S^r(f, 1/n)_{\mathcal{E}}, \tag{4.9}$$

where $\underline{v} = (1, 0, \dots, 0)$ is the unit vector pointing in the x_1-direction. This conclusion is immediate, since $E_n^L(F)_{S_n^*} = E_n^L(f)_{S_n^*+(L/n^2)\underline{v}}$, and the set $S_n^* + (L/n^2)\underline{v}$ contains $(S/64) + ((M+L)/n^2)\underline{v}$ for large n.

Now we invoke

Proposition 4.6 *(4.9) implies*

$$E_n^L(f)_{S/128} \prec \omega_S^r(f, 1/n)_{\mathcal{E}}. \tag{4.10}$$

The proof of this will be given after we complete the proof of Theorem 4.5. Thus, we can see that

$$E_n^L(f)_{S_{2\eta}} \prec \omega_S^r(f, 1/n)_{\mathcal{E}} \tag{4.11}$$

with any $2\eta < 1/128$ (recall that $S_{2\eta}$ is the dilation of S from its apex by the factor 2η), and this is precisely (3.12) for the vertex Q and the pyramid S attached to Q.

Since this can be done for all vertices Q of K, the claim in the theorem follows from Proposition 3.10.

∎

We still need to prove Proposition 4.6.

Proof. Recall that $\mathcal{E}$ is the direction of the apex edges of S and $\underline{v}$ is the vector $(1, 0, \dots, 0)$. It is enough to show that if p_n is any polynomial such that

$$\|f - p_n\|_{(S/64)+(M/n^2)\underline{v}} \leq C_0 \omega_S^r(f, 1/n)_{\mathcal{E}} \tag{4.12}$$

with some fixed M, then automatically

$$\|f - p_n\|_{S/128} \leq C_1 \omega_S^r(f, 1/n)_{\mathcal{E}} \tag{4.13}$$

with some C_1 that is independent of n (in the Proposition apply this with $M+L$ instead of M and with $p_n = L_n f$). As before, let S_0 be the base of S, $Q_1, \dots, Q_m$ the vertices of S_0 and $\underline{e}_j = \overrightarrow{0Q_j}$ the vectors from the apex 0 of S (which is at the origin) to the base vertices. The vector pointing from the origin to the apex of $S/64 + (M/n^2)\underline{v}$ is $(M/n^2)\underline{v} = (M/n^2, 0, \dots, 0)$, and first we express this in
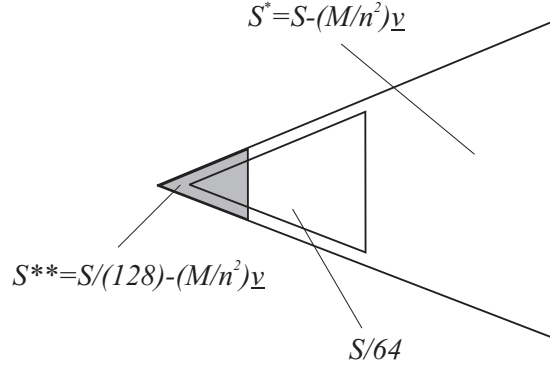
22

Figure 4: A schematic figure of the sets $S/64$, $S^* := S - (M/n^2)v$ and $S^{**} := S/128 - (M/n^2)v$ (shaded area)

terms of the vectors $\underline{e}_j$. Recall that the base of S is on the $x_1 = 2$ hyperplane, and the point $(2, 0, \ldots, 0)$ is a point of the base (because the height of S was lying inside S). Thus, with some numbers $\alpha_j \in [0, 1]$ with total sum 1 we have

$$\alpha_1 \underline{e}_1 + \cdots + \alpha_m \underline{e}_m = (2, 0, \ldots, 0),$$

hence

$$(M\alpha_1/2n^2)\underline{e}_1 + \cdots + (M\alpha_m/2n^2)\underline{e}_m = (M/n^2, 0, \ldots, 0). \qquad (4.14)$$

Thus, if we translate $S/64$ successively by the vectors $(M\alpha_j/2n^2)\underline{e}_j$, $j = 1, \ldots, m$, then we obtain the pyramid $S/64 + (M/n^2)\underline{v}$, and during these translations we never leave S.

For simpler notation we shift everything by $-(M/n^2)\underline{v}$. Under this shift $(S/64) + (M/n^2)\underline{v}$ becomes $S/64$, S becomes $S^* := S - (M/n^2)\underline{v}$ and $S/128$ becomes $S^{**} := (S/128) - (M/n^2)\underline{v}$ (see Figure 4).

What we have just established can also be formulated in the following way: if we translate $S/64$ successively by the vectors $-(M\alpha_j/2n^2)\underline{e}_j$, $j = m, \ldots, 1$, then we obtain a pyramid (which is twice the size of S^{**}) with apex at the common apex of S^* and S^{**}, and during these translations we never leave S^*.

If we also define $F(x) = f(x + (M/n^2)\underline{v})$ on S^* and $P_n(x) = p_n(x + (M/n^2)\underline{v})$, then (4.12) is the same as

$$\|F - P_n\|_{S/64} \le C_0 \omega_{S^*}^r(F, 1/n)_{\mathcal{E}}, \qquad (4.15)$$

and from here we want to conclude (4.13), which is the same as

$$\|F - P_n\|_{S^{**}} \le C_1 \omega_{S^*}^r(F, 1/n)_{\mathcal{E}}. \qquad (4.16)$$

It will be proven in Lemma 7.3 (see also [14, Proposition 9.1]) that if $L > 0$, then for any function $g \in C([0, 1])$ and for any polynomial q_n of a single variable

and of degree at most n with $n^2 \geq 2L$, we have

$$\|g - q_n\|_{[0,1]} \leq C \left(\|g - q_n\|_{[L/n^2, 1]} + \omega_{[0,1]}^r(g, 1/n) \right), \qquad (4.17)$$

where C depends only on L and r. This statement implies that a given rate of Jackson-type polynomial approximation, say

$$\|g - q_n\|_{[0,1]} \leq C_0 \omega_{[0, 1+D/n^2]}^r(g, 1/n) \qquad (4.18)$$

on the interval $[0, 1]$ automatically holds on a $1 + D/n^2$-dilation of the interval, i.e.

$$\|g - q_n\|_{[0, 1+D/n^2]} \leq D_1 C_0 \omega_{[0, 1+D/n^2]}^r(g, 1/n),$$

with some D_1 depending on D, provided in the Jackson estimate (4.18) the modulus of smoothness is measured on the larger interval as is indicated in (4.18). On applying a linear transformation from $[0, 1]$ to a segment, we get from (4.15) that if $I = [A, B]$ is a segment in $S/64$ parallel to one of the apex edges of S and $I' := \Phi_{A, 1+D/n^2} I$ is its dilation from A with factor $1 + D/n^2$, then we have

$$\|F - P_n\|_{I' \cap S^*} \leq D_1 C_0 \omega_{S^*}^r(f, 1/n)_{\mathcal{E}}, \qquad (4.19)$$

and here D_1 depends only on D. This allows us to extend the rate of approximation on $S/64$ given in (4.15) to points lying outside $S/64$. The set $S^{**} \setminus (S/64)$ in (4.16) consists of points lying close to $S/64$ (they are of distance $O(1/n^2)$ from $S/64$), and we are going to use the just described extension process to cover these points. However, in this extension we must use the directions of the apex edges, and this creates problem, since not every point of S^{**} is directly reachable from $S/64$ using segments in the apex edge directions (for example, the apex of S^{**} cannot be reached). Therefore, we shall need to use repeated extensions in different directions.

We shall use the product sets from Proposition 4.1. Let us denote $S/64$ by S' and its base by S_0'. If $Q_1, \dots, Q_m$ are the vertices of S_0 (the base of S), then $Q_j' = Q_j/64$ are the vertices of $S_0' = S_0/64$. Let $1 > \lambda > d/(d+1)$ be fixed so that $\lambda^m > 1/2$. With an edge e_j' that connects the corresponding vertices of S_0' and $\lambda S_0'$ consider the product set $(\lambda S_0') \times e_j'$ (so e_j' is in the direction of the apex edge $\underline{e}_j$ and its length is $(1-\lambda)|\underline{e}_j|/64$). This is part of S' (see Proposition 4.1), which means that from every point A of $\Phi_{Q_j', \lambda} S_0'$ we can draw a segment I_A parallel with $\underline{e}_j'$ to a boundary point B of S', and the length of this segment is $\geq |e_j'| \geq (1-\lambda)|\underline{e}_j|/64 \geq (1-\lambda)/32$ (because $\underline{e}_j$ points from the origin to a point of the hyperplane $x_1 = 2$, so its length is at least 2). Since I_A lies within S', we can use (4.19) and conclude that if I_A' is the dilation of I_A from A by a factor $1 + D/n^2$, then on $I_A' \cap S^*$ we have the bound

$$\|F - P_n\|_{I_A' \cap S^*} \leq D_2 C_0 \omega_{S^*}^r(f, 1/n)_{\mathcal{E}}, \qquad (4.20)$$

where D_2 depends only on D.

Since $|e_j'| \geq (1-\lambda)/32$, it is immediate that all these I_A', $A \in \Phi_{Q_j', \lambda} S_0'$ cover the translation of $\lambda S'$ by the vector $-(D(1-\lambda)/32n^2)\underline{e}_j/|\underline{e}_j|$ and for

24

sufficiently large D (depending on λ and M), this will cover the translation of $\lambda S'$ by the vector $-(M\alpha_j/2n^2)\underline{e}_j$ from (4.14). In later steps we shall use smaller copies (namely $\lambda^k S'$, $k = 1, \ldots, m$) of S' instead of S' itself, and for them the corresponding translation vectors will be $-(D(1-\lambda)\lambda^k/32n^2)\underline{e}_j/|\underline{e}_j|$, but for sufficiently large D (depending on λ and M), this is still at least as long as $-(M\alpha_j/2n^2)\underline{e}_j$ from (4.14) because we selected λ so that $\lambda^m \geq 1/2$.

We apply this first with $j = m$, and recall that we said in connection with (4.14) that the translation $S' - (M\alpha_m/2n^2)\underline{e}_m$ lies inside S^*, so (4.19) yields

$$\|F - P_n\|_{\lambda S' - (M\alpha_m/2n^2)\underline{e}_m} \leq D_2 C_0 \omega^r_{S^*}(f, 1/n)_{\mathcal{E}}.$$

Let us denote the transformation from S' to $\lambda S' - (M\alpha_j/2n^2)\underline{e}_j$ by $\Psi_j(S')$. What we got can be written as

$$\|F - P_n\|_{\Psi_m(S')} \leq D_2 C_0 \omega^r_{S^*}(f, 1/n)_{\mathcal{E}}.$$

Note that here $\Psi_m(S')$ is a λ-size copy of S' with apex at $-(M\alpha_m/2n^2)\underline{e}_m$. Now apply the same procedure to $\Psi_m(S')$, but with the edge $\underline{e}_m$ replaced by $\underline{e}_{m-1}$, to conclude

$$\|F - P_n\|_{\Psi_{m-1}(\Psi_m(S'))} \leq D_2^2 C_0 \omega^r_{S^*}(f, 1/n)_{\mathcal{E}},$$

and here $\Psi_{m-1}(\Psi_m(S'))$ is a λ^2-size copy of S' with apex at $-(M\alpha_m/2n^2)\underline{e}_m - (M\alpha_{m-1}/2n^2)\underline{e}_{m-1}$. Iterate this further with the edge directions $\underline{e}_{m-2}, \ldots, \underline{e}_1$. After m iterations we obtain for $H = \Psi_1(\cdots(\Psi_m(S')\cdots)$

$$\|F - P_n\|_H \leq D_2^m C_0 \omega^r_{S^*}(f, 1/n)_{\mathcal{E}}, \tag{4.21}$$

and here H is a λ^m-size copy of S' with apex at

$$-(M\alpha_m/2n^2)\underline{e}_m - (M\alpha_{m-1}/2n^2)\underline{e}_{m-1} - \cdots - (M\alpha_1/2n^2)\underline{e}_1.$$

In view of (4.14) this last point is precisely the apex of S^*, so H is a homothetic copy of S^* of size $\lambda^m/64$ (recall that S' was a $1/64$ size copy of S) placed so that its apex coincides with the apex $(-M/n^2, 0, \ldots, 0)$ of S^*. Since $\lambda^m \geq 1/2$, this H then covers $S^{**} = S/128 + (-M/n^2, 0, \ldots, 0)$, hence (4.16) is a consequence of (4.21).

∎

In the next section we use the following modification of Proposition 4.6.

Corollary 4.7 *Let $\mathcal{E}^*$ be any set of directions that include the apex edge directions of S and let $S \subset K$. Then for any fixed M*

$$E_n^L(f)_{S+(M/n^2)\underline{v}} \prec \omega^r_K(f, 1/n)_{\mathcal{E}^*} \tag{4.22}$$

implies

$$E_n^L(f)_{S/2} \prec \omega^r_K(f, 1/n)_{\mathcal{E}^*}. \tag{4.23}$$

Furthermore, in this conclusion $\underline{v}$ can be any fixed vector pointing from the apex of S to an inner point of S.

The proof of this is the same that was given for Proposition 4.6 (just replace in that proof $S/64$ by S), since in that proof we only used apex edge directions of S and the fact that the vector $\underline{v}$ was a convex combination of the corresponding apex edge vectors ($\underline{e}_j$ in the proof of Proposition 4.6).

5 Small degrees

In this section we show that the Jackson type estimate that we are discussing on convex polytopes holds for every fixed $n \geq rd$ (by a bounded linear operator), so in the previous part of the paper we could concentrate only on large degree polynomial approximation.

Theorem 5.1 *If K is a convex polytope in $\mathbf{R}^d$, then for every $n \geq rd$ there is a bounded linear operator $L_n : C(K) \to \Pi_n^d$ and a constant C independent of f such that*

$$\|f - L_n f\|_K \leq C\omega_K^r(f, 1/n). \tag{5.1}$$

Here C may depend on n (and of course, on K).

Proof. Take a vertex of K and $d+1$ independent edge directions $\underline{e}_1, \ldots, \underline{e}_{d+1}$ at this vertex. With these directions form a parallelepiped $H \subset K$ as its edge directions inside K. We may assume that the origin is in the interior of H, and let θK be a dilation of K from the origin with some small dilation factor θ such that θK lies in the interior of H.

The claim is true on parallelepipeds, see [6, Theorem 12.1.1] (there the degree of a polynomial was the maximal degree in each variables, and the result was for all polynomials with that degree $\geq r$; but now we use total degree, so we can claim the result for parallelepipeds for polynomials of total degree $\geq rd$). Thus, we have for some L_n

$$\|f - L_n f\|_H \leq C_0\omega_H^r(f, 1/n) \leq C_0\omega_K^r(f, 1/n),$$

hence

$$\|f - L_n f\|_{\theta K} \leq C_0\omega_K^r(f, 1/n), \tag{5.2}$$

is even more true. Thus, all we need to show is that (5.1) is a consequence of (5.2).

For simpler notations we write K instead of θK and $(1/\theta)K$ instead of K, and what we are going to prove is that if P_n is any polynomial of the fixed degree n and

$$\|f - P_n\|_K \leq C_0\omega_{K/\theta}^r(f, 1/n), \tag{5.3}$$

then

$$\|f - P_n\|_{K/\theta} \leq C_1\omega_{K/\theta}^r(f, 1/n). \tag{5.4}$$

To this end it is sufficient to show that there is a $1 < \lambda_0 < 1/\theta$ depending on K such that for every $1 \leq \lambda \leq \lambda_0$

$$\|f - P_n\|_K \leq C_0\omega_{\lambda K}^r(f, 1/n) \tag{5.5}$$

implies

$$\|f - P_n\|_{\lambda K} \le C_2 \omega^r_{\lambda K}(f, 1/n). \tag{5.6}$$

Indeed, then choose such a λ with $\lambda^k = 1/\theta$ with some k, and iterate (5.5) $\Rightarrow$ (5.6) k-times (with K replaced by $\lambda^l K$, $l = 0, 1, \ldots, k-1$, successively) to obtain (5.4). (The iteration is legitimate because any $\lambda^l K$ is a homothetic copy of K and $\omega_{\lambda^l K}(f, 1/n) \le \omega_{K/\theta}(f, 1/n)$ for $l \le k$).

For the implication (5.5) $\Rightarrow$ (5.6) we have done most of the work in the proof of Proposition 4.6. Let S be a small pyramid with apex at a point Q such that the apex edge directions of S are among the edge directions of K, and let S_η be the η-dilation of S from its apex. Let also $\tilde{S}_{6/7}$ be a translation of $S_{6/7}$ so that it lies in the interior of S. In the proof of Proposition 4.6 and Corollary 4.7 we verified that if $\tilde{K}$ is a polytope such that $S \subset \tilde{K}$ and the apex edge directions of S are among the edge directions of $\tilde{K}$, then

$$\|f - P_n\|_{\tilde{S}_{6/7}} \le C_0 \omega^r_{\tilde{K}}(f, 1/n) \tag{5.7}$$

implies

$$\|f - P_n\|_{S_{3/7}} \le C_1 \omega^r_{\tilde{K}}(f, 1/n) \tag{5.8}$$

with some C_1 that depends on C_0 and n. In fact, in that proof we worked with translations of size $1 + D/n^2$, but if n is fixed, then this can be of any size if D is sufficiently large. So Proposition 4.6 and its proof gives, indeed, the implication (5.7) $\Rightarrow$ (5.8).

Take now an η-pyramidal covering of K (with some small but fixed η) as in Section 3. Recall that this was a covering by some pyramids $\mathcal{K}_{\eta/2,s}$, where $\mathcal{K}_{\eta/2,s}$ is a 1/2-scaled version of a larger pyramid $K_{\eta,s}$ (still lying in K) scaled from its apex. Fix for each $K_{\eta,s}$ a vector $\underline{u}_s$ from its apex pointing to an inner point of $K_{\eta,s}$. Let $\lambda > 1$. Then $\{\lambda \mathcal{K}_{\eta/2,s}\}$ is an η-pyramidal covering of λK, and from the construction of η-pyramidal covering in Section 3 it follows that if $\lambda > 1$ is sufficiently close to 1, then each $\lambda \mathcal{K}_{\eta/2,s} + \underline{u}_s$ lies in K. Imagine now that the above $\tilde{S}_{6/7}$ is one of these $\lambda K_{\eta,s} + \underline{u}_s$ and $S_{3/7}$ is precisely $\lambda \mathcal{K}_{\eta/2,s}$. (5.7) for $\tilde{K} = \lambda K$ is then a consequence of (5.5), and in the conclusion (5.8) the right hand side is at most as large as $C_1 \omega^r_{\tilde{K}}(f, 1/n) = C_1 \omega^r_{\lambda K}(f, 1/n)$, so

$$\|f - P_n\|_{\lambda \mathcal{K}_{\eta/2,s}} \le C_1 \omega^r_{\lambda K}(f, 1/n) \tag{5.9}$$

follows. Since the sets $\lambda \mathcal{K}_{\eta/2,s}$ for all s cover λK, we obtain (5.6).

This completes the proof of the implication (5.5) $\Rightarrow$ (5.6) for all $\lambda > 1$ lying sufficiently close to 1, and that finishes the proof of Theorem 5.1.

$\blacksquare$

Part II

Approximation on algebraic domains

In the second part of this paper we shall deal with domains with smooth (actually algebraic) boundary — see Definition 6.3 for the precise concept. So far in the literature the rate of approximation has been characterized in details via appropriate moduli of smoothness only on balls, see [3] and [4] by F. Dai and Y. Xu. Balls exhibit a large degree of symmetry, and in this paper we shall break this symmetry, and prove the characterization for a large family of domains.

6   Domains and moduli of smoothness

In this section we shall be primarily concerned how to define appropriate moduli of continuity/smoothness on smooth domains and how the geometry influences the construction.

6.1   Domains in $\mathbf{R}^2$

We start with two-dimensional domains. Let $G \subset \mathbf{R}^2$ be the closure of a bounded, finitely connected domain (connected open set) with C^2 boundary Γ. For simpler discussion we shall sometimes identify $\mathbf{R}^2$ with the complex plane. Since Γ is C^2 smooth, there is an r_0 such that a circle of radius $8r_0$ can be drawn (from the inside) to every point of Γ which stays inside G together with its interior.

The modulus of smoothness on a circle $C_\rho(z_0) \subset G$ about a center z_0 and of radius ρ is defined as in classical trigonometric approximation theory, namely

$$\omega^1_{C_\rho(z_0)}(f,\delta) := \sup_{\theta \leq \delta, \ \varphi} |f(z_0 + \rho e^{i(\varphi+\theta/2)}) - f(z_0 + \rho e^{i\varphi - \theta/2})| \qquad (6.1)$$

(here and in what follows in such expressions, $\theta \leq \delta$ means that $0 \leq \theta \leq \delta$), and the higher order moduli of smoothness as

$$\omega^r_{C_\rho(z_0)}(f,\delta) \quad := \quad \sup_{\theta \leq \delta, \ \varphi} |\Delta^r_\theta f(z_0 + \rho e^{i(\varphi + \cdot)})| \qquad (6.2)$$

$$= \quad \sup_{\theta \leq \delta, \ \varphi} \left| \sum_{j=0}^r (-1)^j \binom{r}{j} f\left(z_0 + \rho e^{i(\varphi + (r/2-j)\theta)}\right) \right|.$$

If $I = \overline{AB} \subset G$ is a segment, then the modulus of smoothness is defined on I as in (1.6):

$$\omega^r_I(f,\delta) \quad = \quad \sup_{h \leq \delta, \ z \in I} \|\Delta^r_{h\tilde{d}_I(\underline{e},z)\underline{e}} f(z)|$$

28

$$= \sup_{h \le \delta, \; z \in I} \left| \sum_{j=0}^{r} (-1)^j \binom{r}{j} f \left(z + (r/2 - j) h \tilde{d}_I(\underline{e}, z) \right) \right|, \quad (6.3)$$

where $\underline{e}$ is the direction of $I = \overline{AB}$ and $\tilde{d}_I(\underline{e}, z) = \sqrt{|z - A| |z - B|}$. Recall the usual convention that here $\Delta^r_{h \tilde{d}_I(\underline{e}, z) \underline{e}} f(z)$ is defined to be 0 if either of the points $z \pm (r/2) h \tilde{d}_I(\underline{e}, z)$ does not belong to I. Finally, we set

$$\omega^r(f, \delta)_G = \max \left(\sup_{C_\rho(z_0)} \omega^r_{C_\rho(z_0)}(f, \delta), \sup_I \omega^r_I(f, \delta) \right), \quad (6.4)$$

where the suprema are taken for all circles $C_\rho(z_0)$ that lie in G, and for all horizontal or vertical segments $I \subset G$.

We shall not deal with the various properties of these moduli, but we shall frequently use that

$$\omega^r(f, \lambda \delta)_G \le C \lambda^r \omega^r(f, \delta)_G, \qquad \lambda \ge 1.$$

Indeed, this inequality is well known (see [5, Chapter 2, (7.8)] and [6, Theorem 4.1.2]) for both ω^r_I and $\omega^r_{C_\rho(z_0)}$, from which the claim follows.

We shall also often use that if $G_1 \subset G$, then $\omega^r(f, \delta)_{G_1} \le \omega^r(f, \delta)_G$, which is clear from the definition.

Remark 6.1 1. One could request the second supremum in (6.4) only for (horizontal or vertical) maximal segments I that are contained in G, the value of $\omega^r(f, \delta)_G$ would not change.

2. In (6.4) the supremum is take for all circles $C_\rho(z_0)$ that lie in G. For first order, i.e. for modulus of continuity, it would be sufficient to take the supremum in (6.4) only for circles C_ρ of the fixed radius $\rho = r_0$, but it is not clear if that is sufficient when proving Jackson type theorems via higher order moduli. A careful examination of the proof of Theorem 6.5 below shows, however, that in the case of r-th order modulus of smoothness there is a $\rho_0 \le r_0$ such that it is sufficient to take the supremum in (6.4) only for circles C_ρ lying in G with $\rho_0/2 \le \rho \le \rho_0$.

3. The way they were defined, the moduli of continuity/smoothness in (6.4) depend on the position of G in the coordinate system (since horizontal/vertical segments are used in them). If the second supremum in (6.4) is taken for all segments in G, then the so obtained (larger) moduli would already be isometry invariant, and everything in what follows would work for these modified moduli exactly as for the $\omega^r(f, \delta)_G$.

4. On domains in $\mathbf{R}^2$ for a different modulus of smoothness based on averages see the paper [8] by K. G. Ivanov. With that modulus the correct order Jackson theorem was announced in [8] for domains with piecewise C^2 boundary, as well a corresponding weak converse in some special cases. One should also mention the works [7] and [9] that contain the first results on approximation on domains by polynomials in several variables.

6.2 Domains in $\mathbf{R}^d$

Now let $G \subset \mathbf{R}^d$ be the closure of a bounded, finitely connected domain (in short "closed domain") with C^2 boundary Γ. Since the boundary Γ is C^2 smooth, there is an r_0 such that a $(d-1)$-dimensional sphere of radius $4dr_0$ can be drawn (from the inside) to every point of Γ which stays inside G together with its interior.

We say that a circle C_ρ is parallel with a coordinate plane if its plane is parallel with one of the planes spanned by two of the coordinate axes. The modulus of smoothness on such a circle $C_\rho \subset G$ is defined as before in (6.2), and the modulus of continuity on a segment $I \subset G$ as in (6.3). Finally let

$$\omega^r(f,\delta)_G = \max\left(\sup_{C_\rho} \omega^r_{C_\rho}(f,\delta), \sup_I \omega^r_I(f,\delta)\right),$$

where the suprema are taken for all circles $C_\rho \subset G$ (of arbitrary radius $\rho > 0$) which are parallel with a coordinate plane and for all segments $I \subset G$ that are parallel with one of the coordinate axes.

Remark 6.2 Saying that C_ρ is parallel with a coordinate plane and using the modulus of smoothness on C_ρ is pretty much the same as using the so called Euler angles as in [3], [4], so our $\omega^r(f,\delta)_G$ is close to the moduli of smoothness used on the unit ball in $\mathbf{R}^d$ in those works.

We shall work with domains in $\mathbf{R}^d$ bounded by algebraic surfaces. So let $\Gamma_1, \ldots, \Gamma_k$ be the connected components of $\Gamma = \partial G$, and we assume that there are polynomials $\Phi_j(x_1, \ldots, x_d)$ of d variables such that Γ_j is part of the surface $\Phi_j(x_1, \ldots, x_d) = 0$. We shall also assume that on Γ_j the gradient $\nabla \Phi_j$ does not vanish. Then this gradient is a normal to Γ_j.

Definition 6.3 *We say that G is an algebraic domain if all these properties are satisfied.*

Examples include balls, domains enclosed by thori, etc., but there are many more exotic examples like the one[4] in Figure 5. Note also that if G is as explained and we remove from the interior of G disjoint simply connected domains $G_1, \ldots, G_k$ satisfying similar conditions, then $G \setminus (\mathrm{Int}(G_1) \cup \cdots \cup \mathrm{Int}(G_k))$ also satisfies the conditions set forth.

Remark 6.4 We emphasized that we do not assume that Γ_j is the whole surface $\Phi_j(x_1, \ldots, x_d) = 0$, it can be just as well one of its components.

6.3 The Jackson theorem and its converse

Let G be a closed domain as described before and let

$$E_n(f)_G = \inf_{P_n \in \Pi^d_n} \|f - P_n\|_G$$

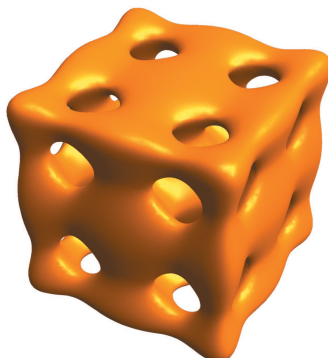[4]Courtesy of János Karsai, who used Mathematica 11.3 to create the figure

Figure 5: The domain enclosed by the so called Banchoff surface $x^2(3-4x^2)^2 + y^2(3-4y^2)^2 + z^2(3-4z^2)^2 = 3/2$

be the error of best polynomial approximation to $f \in C(G)$ by polynomials of d-variables of total degree at most n. We have the following Jackson theorem[5]

Theorem 6.5 *If G is a (closed) algebraic domain, then for $r = 1, 2, \ldots$*

$$E_n(f)_G \leq C\omega^r(f, 1/n)_G, \qquad n \geq rd, \qquad (6.5)$$

with a constant C independent of f and n.

The standard weak converse is true:

Theorem 6.6 *With the assumptions of Theorem 6.5*

$$\omega^r(f, 1/n)_G \leq \frac{C}{n^r} \sum_{k=0}^{n} (k+1)^{r-1} E_k(f)_G, \qquad (6.6)$$

with a constant C independent of $n \geq rd$.

Since the restriction of a polynomial $P_n(x_1, \ldots, x_d)$ to any line segment is an algebraic polynomial of a single variable, and its restriction to any circle C_ρ, say to $C_\rho(z_0) = \{(x, y, \zeta_2, \ldots, \zeta_d)\}$, $z_0 = (x_0, y_0, \zeta_2, \ldots, \zeta_d)$, $x = x_0 + \rho\cos\theta$, $y = y_0 + \rho\sin\theta$, $x_3 = \zeta_3, \ldots, x_d = \zeta_d$, is a trigonometric polynomial of degree at most n in θ, the inequality (6.6) is an immediate consequence of the corresponding inequalities on $[-1, 1]$ and on the unit circle (see [6, 7.2.4] and [5, Theorem 7.3.1]) if we take into account the definition of the modulus of smoothness ω_G^r. Therefore, in what follows we shall only concentrate on (6.5).

We also state

---

[5]See also the note at the end of the paper

Corollary 6.7 *Let $0 < \alpha$. If $f \in C(G)$ can be approximated with error $\leq 1/n^\alpha$ by algebraic polynomials of degree at most n on every segment of G parallel to one of the coordinate axes, and on every circle lying in G parallel with one of the coordinate planes, then f, on the whole of G, can be approximated with error $\leq C/n^\alpha$ by polynomials of d-variables of degree at most n.*

This corollary says that in certain cases d-variable polynomial approximation can be reduced to ordinary one-variable trigonometric and algebraic approximation.

The major steps in the proof of Theorem 6.5 are as follows. First of all, in the next section it is shown that all difficulties arise around the boundary of the domain: we shall decompose a relative neighborhood of this boundary into small pieces, and if approximation can be done in the correct order on the individual pieces, then the global result follows. The different pieces can be handled the same way, so the problem will be to prove the correct rate of approximation on a small part around the boundary. This small part is the image of a cube under a polynomial mapping, and the problem will be transformed into a corresponding approximation problem on the cube. The case of a cube is well-known and classical, however, the transformation used is not linear, and a major effort will be to establish how the moduli of smoothness change under the given transformation. This is a highly non-trivial problem (to be done in Sections 8 and 10) because the distance from the boundary plays an important role in the definition of the moduli of smoothness. The $r = 1$ case shows all the difficulties coming from the geometry and of the just mentioned transformation, and this case will be dealt with separately in Section 8. The general $r > 1$ case will involve estimates for derivatives of composed functions (detailed in Section 10), as well as the special case when the domain is a disk (detailed in Section 9). This approach gives the required estimate for the moduli of smoothness of the transformed function on the whole domain except for points that lie too close to the boundary (closer than C/n^2 for n-th degree approximation). Finally, Section 11 establishes that that is sufficient to have the same estimate on the whole domain, and that section contains also the completion of the proof.

7 Reduction to approximation close to the boundary

In this section we start the proof of Theorem 6.5. We show that if we can prove a Jackson-type bound on suitable sets attached to the boundary (with the modulus of smoothness on the whole domain), then Theorem 6.5 follows.

The Jackson-type estimate (6.5) can be proven for each fixed n by a constant that may depend on n (we do not give details, they are similar — actually simpler — to what was used in Section 5), so in what follows we may concentrate only on approximation for large n.

Let $\underline{e}_1, \ldots, \underline{e}_d$ be the unit vectors in $\mathbf{R}^d$ pointing to the positive directions of the corresponding coordinate axis. If P is a point on Γ and $\underline{n}_P$ is the unit inner normal to Γ at P, then there is a j for which the dot product $|\underline{n}_P \cdot \underline{e}_j|$ is maximal. Then we must have $|\underline{n}_P \cdot \underline{e}_j| \geq 1/\sqrt{d}$ (the square of the left hand sides for all j sums up to 1), and for all other k the inequality $|\underline{n}_P \cdot \underline{e}_k| \leq 1/\sqrt{2}$ is a consequence of the definition of $\underline{e}_j$. Then there is a small neighborhood of P such that for all $Q \in \Gamma$ in that neighborhood we have $|\underline{n}_Q \cdot \underline{e}_j| \geq 1/2\sqrt{d}$, and for all other k the inequality $|\underline{n}_Q \cdot \underline{e}_k| \leq 3/4$ is true, and call the part of Γ (more precisely the part of the connected component of Γ that contains P) that lies in that neighborhood an x_j-surface. Since x_j-surfaces for different j can be similarly handled, in what follows we shall assume that $j = d$, and that close to P the domain G lies above Γ (in the direction of the x_d-axis). Around P the boundary Γ is part of a level surface $\Phi_s = 0$, hence the normal vector to Γ is just the normalized gradient (or its negative) $\nabla \Phi_s$. Therefore, $|\partial \Phi_s / \partial x_d| \geq c/2\sqrt{d}$ in the neighborhood in question with some $c > 0$. Thus, by the implicit function theorem, there is a cube $I_P \subset \mathbf{R}^{d-1}$ such that in a neighborhood of P the boundary Γ is given by the graph of a function $F_P(x_1, \ldots, x_{d-1})$, $(x_1, \ldots, x_{d-1}) \in I_P$. Call this graph Γ_P — it is part of Γ. By shrinking I_P if necessary, we may assume that P is the image of the center of I_P under the mapping F_P. Since we are dealing with algebraic domains with nonzero gradient of the functions that define the boundary, it is easy to see that this I_P can be chosen so that its diameter is bigger than a positive number independent of P: $\mathrm{diam}(I_P) \geq \delta_0 > 0$.

Consider the set

$$
\begin{aligned}
G_P \;=\; & \{(x_1, \ldots, x_d) \,|\, (x_1, \ldots, x_{d-1}) \in I_P, \\
& x_d \in [F_P(x_1, \ldots, x_{d-1}), F_P(x_1, \ldots, x_{d-1}) + r_0]\}
\end{aligned}
\tag{7.1}
$$

(recall the definition of r_0 from Section 6), i.e. we place a segment of length r_0 parallel with the x_d-axis "above" every point $Q \in \Gamma_P$, see Figure 6. For a $Q \in \Gamma_P$ let S_Q be the sphere of radius $4dr_0$ which lies on the side of Γ_P as G (in a small neighborhood of Γ_P) and touches Γ at the point Q (i.e. S_Q goes through Q and at Q it has the same tangent hyperplane as Γ). Using the fact that S_Q lies inside G together with its interior (this follows from the definition of r_0), we can conclude that G_P is a subset of G. Indeed, the condition $|n_Q \cdot \underline{e}_d| \geq 1/2\sqrt{d}$ is the same as saying that the projection of the segment in between Q and the center of S_Q onto the line through Q in the $\underline{e}_d$ direction is at least $4dr_0/2\sqrt{d} > 2r_0$, so the $\underline{e}_d$ directional segment of length r_0 placed to Q lies inside the sphere S_Q.

We claim that f can be approximated on G_P in the order $\omega^r(f, 1/n)_G$ by polynomials of d variables of degree at most n, i.e.

$$
E_n(f)_{G_P} \leq C \omega^r(f, 1/n)_G.
\tag{7.2}
$$

This is the hard part of the proof, and in this section we only show how the proof of Theorem 6.5 can be completed from (7.2).

Let I_P^* be the dilation of I_P from its center by a factor $1/2$, and let Γ_P^* be graph of the function $\Gamma(x_1, \ldots, x_{d-1})$ over I_P^*. Then Γ_P^* is part of Γ_P.. Over Γ_P^*
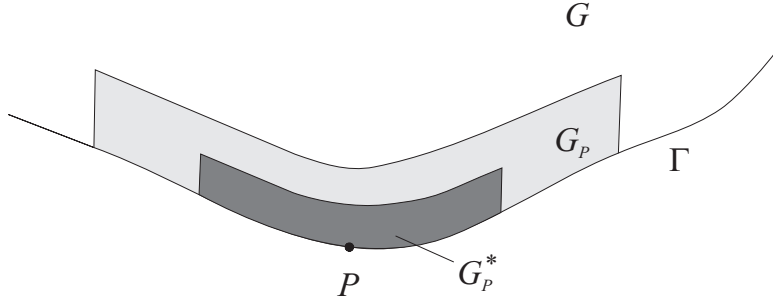
Figure 6: The domains G, G_P and G_P^*

form the corresponding set G_P^*, with the difference that now use $\underline{e}_d$-directional segments over the points of Γ_P^* of length $r_0/2$ (see Figure 6). Then G_P contains a relative neighborhood of G_P^* with respect to G. Since $\operatorname{diam}(I_P^*) \geq \delta_0/2$ and the functions F_P are uniformly C^2 (actually, C^∞) smooth, there is a $\delta_1 > 0$ independent of $P \in \Gamma$ such that G_P^* covers the relative δ_1-neighborhood of P (relative with respect to G), i.e. $B_{\delta_1}(P) \cap G \subseteq G_P^*$, where $B_\delta(P)$ is the ball of radius δ about P. Let $\{P_i\}_{i=1}^k \subset \Gamma$ be a finite set such that every point of Γ is of distance $\leq \delta_1/4$ from some P_i. It is clear that $\cup \Gamma_{P_i}^*$ covers Γ, furthermore $\cup_i B_{\delta_1}(P_i)$, and hence also $\cup_i G_{P_i}^*$, covers the (open) $\delta_1/2$-neighborhood O of Γ. Cover the rest of G by the interiors of finitely many d-dimensional closed cubes U_i^*, $1 \leq i \leq l$, so that if U_i is the dilation of U_i^* from its center by some factor $(1 + \theta)$ with some small $\theta > 0$, then even U_i lies in G. We can choose these U_i so that their edges are parallel with the coordinate axes. On each U_i we have the analogue of (7.2):

$$E_n(f)_{U_i} \leq C\omega^r(f, 1/n)_G. \tag{7.3}$$

Indeed, this follows from [6, Theorem 12.1.1] (or from Theorem 1.1 in the first part of this paper) and from how $\omega^r(f, 1/n)_G$ has been defined. Since the interiors of $G_{P_i}^*$, $i = 1, 2, \ldots, k$ and of U_i^*, $i = 1, 2, \ldots, l$ cover the interior of G, we can invoke Lemma 3.9 and conclude that from (possibly repeated) copies of the U_i^* and from (possibly repeated) copies of the sets $G_{P_i}^*$ we can form a sequence of sets K_s^*, $1 \leq s \leq T$, and the corresponding sequence of sets K_s (which is U_j if $K_s^* = U_j^*$ and it is G_{P_j} if $K_s^* = G_{P_j}^*$) that satisfy

(i) each K_s is contained in G,

(ii) $G = \cup_{s=1}^T K_s^*$,

(iii) the intersection of the interiors of K_s^* and K_{s+1}^* is non-empty for all $s = 1, \ldots, T - 1$,

(iv) for some $\varepsilon > 0$ we have $(K_s^*)^\varepsilon \cap G \subset K_s$, where $(K_s^*)^\varepsilon$ is the closed ε-neighborhood of K_s^*.

Note that (iv) is the consequence of (i) and the fact that each K_s contains a (relative to G) neighborhood of K_s^*.

This is now the analogue of the η-pyramid covering from Proposition 3.8, and note that on each K_s we have

$$E_n(f)_{K_s} \leq C\omega^r(f, 1/n)_G \tag{7.4}$$

by (7.2)–(7.3). So we can combine local approximants as in Section 3, and to that we shall only need the following analogue of Lemma 3.4.

Lemma 7.1 *Let S^* be one of the sets K_s^* and S the corresponding set K_s, and let $H \subset G$ be a closed set such that for some $\delta > 0$ the intersection $S^* \cap H$ contains a ball of radius δ. Then for every l_0 there is an l such that*

$$E_{l_0 n}(f)_S \prec \omega^r(f, 1/n)_G, \qquad E_{l_0 n}(f)_H \prec \omega^r(f, 1/n)_G, \qquad f \in C(G) \tag{7.5}$$

imply

$$E_{ln}(f)_{S^* \cup H} \prec \omega^r(f, 1/n)_G, \qquad f \in C(G). \tag{7.6}$$

If we repeatedly apply this with $H = \cup_{j=1}^{s-1} K_j^*$ and $S^* = K_s^*$, $S = K_s$, then for $s = T$ we can conclude from (7.6) the Jackson estimate (6.5) for large n.

Proof of Lemma 7.1. We may assume that G lies in the unit ball $B_1(0)$ of $\mathbf{R}^d$, and follow the proof of Lemma 3.4. It is clear that all we need to prove is the following variant of Lemma 3.2.

Lemma 7.2 *Let S^* be one of the K_s^* and S the corresponding set K_s. There is a constant $k = k(S)$ such that for every $n \geq 1$ there are polynomials R_{kn} of degree at most kn for which*

a) $0 \leq R_{kn}(x) \leq 1$ *for* $x \in B_1(0)$,

b) $1 - R_{kn}(x) \leq 2^{-n}$ *for* $x \in S^*$,

c) $R_{kn}(x) \leq 2^{-n}$ *for* $x \in G \setminus S$.

Proof. If K_s is a cube U_i, then this follows from Lemma 3.2, so let K_s be one of the sets G_{P_j}, say G_P.

Cover $S^* = G_P^*$ by cuboids $V_1^*, \ldots, V_m^*$ as depicted in Figure 7 so that for some $\varepsilon > 0$ it is true that if V_j is the $(1 + \varepsilon)$-enlarged copy of V_j^* enlarged from its center, then V_j also lies in $B_1(0)$ and $V_j \cap G$ lies in G_P. By Lemma 3.2 for some l and all $n \geq 1$ there are polynomials $r_{ln,j}$ of degree at most $\leq ln$ that satisfy

1) $0 \leq r_{ln,j}(x) \leq 1$ for $x \in B_1(0)$,

2) $1 - r_{ln,j}(x) \leq 3^{-n}$ for $x \in V_j^*$,

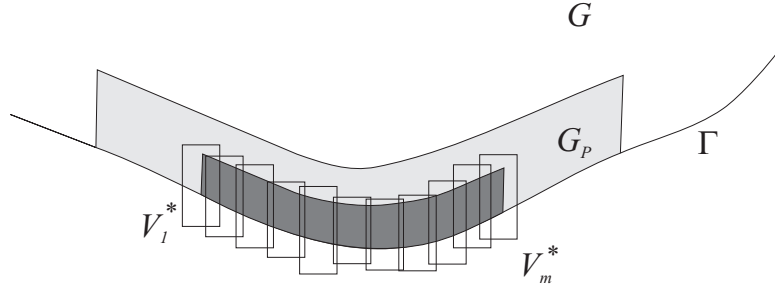**3)** $r_{ln,j}(x) \leq 3^{-n}$ for $x \in B_1(0) \setminus V_j$.

Figure 7: The sets V_j^*

With $P_0 \equiv 0$, for $j = 1, \ldots, m$ set $P_j(x) = P_{j-1}(x)(1 - r_{ln,j}(x)) + r_{ln,j}(x)$. These are polynomials of degree at most kn with $k = lm$, and induction on j gives that

A) $0 \leq P_j(x) \leq 1$ for $x \in B_1(0)$,

B) $1 - P_j(x) \leq 3^{-n}$ for $x \in \cup_{s=1}^{j} V_s^*$,

C) $P_j(x) \leq j3^{-n}$ for $x \in B_1(0) \setminus \cup_{s=1}^{j} V_s$.

On applying this for $j = m$ we obtain a)–c) for all sufficiently large n for the polynomials $R_{kn}(x) := P_m(x)$.

∎

In some cases we cannot directly do the approximation up to the boundary, just if we stay off it at least L/n^2 distance (for n-th degree approximation). However, in such situations the missing range (closer than L/n^2 to the boundary) is automatically covered by the same approximants. This will follow from the next lemma.

Lemma 7.3 *Let $g \in C[0,1]$ and let $L > 1$ be fixed. Then for any polynomial q_n of a single variable and of degree at most n with $n^2 \geq 2L$, we have*

$$\|g - q_n\|_{[0,1]} \leq C \left(\|g - q_n\|_{[L/n^2, 1]} + \omega_{[0,1]}^r(g, 1/n) \right), \qquad (7.7)$$

where C depends only on L and r.

Remark 7.4 The lemma implies for any fixed $0 < a \leq 1/2$ and for all q_n of degree at most n the inequality

$$\|g - q_n\|_{[0,1]} \leq C \left(\|g - q_n\|_{[a,1]} + \omega_{[0,1]}^r(g, 1/n) \right), \qquad (7.8)$$

with a C that depends on a and n. By iteration we obtain from here (7.8) for all $0 < a < 1$.

Lemma 7.3 is Proposition 9.1 in [14], but since it plays a vital role in some of our reasonings, we give a new proof for it.

Proof. For all $Lr/2n^2 \le x \le 1 - Lr/2n^2$ the r-th difference $\Delta^r_{L/n^2} g(x)$ is defined and

$$|\Delta^r_{L/n^2} g(x)| \le \omega^r_{[0,1]}(g, \sqrt{4L}/n) \le C\omega^r_{[0,1]}(g, 1/n)$$

(note that on the right we have the φ-modulus of smoothness) since $L/n^2 \le \sqrt{4L}\sqrt{x(1-x)}/n$. Thus, for $2Lr/n^2 \le x \le 1 - 2Lr/n^2$ we have

$$
\begin{aligned}
|\Delta^r_{L/n^2} q_n(x)| &\le |\Delta^r_{L/n^2}(q_n - g)(x)| + |\Delta^r_{L/n^2} g(x)| \\
&\le 2^r \|g - q_n\|_{[L/n^2, 1]} + C\omega^r_{[0,1]}(g, 1/n) =: M, \qquad (7.9)
\end{aligned}
$$

so the norm of the left hand side over $[2Lr/n^2, 1 - 2Lr/n^2]$ is bounded by the right hand side, i.e. by M. But $\Delta^r_{L/n^2} q_n(x)$ is a polynomial of degree at most n, so it follows from Remez' inequality [11] ([1]) that then

$$\|\Delta^r_{L/n^2} q_n\|_{[0,1]} \le C_1 M$$

is also true with some $C_1 \ge 1$ that depends only on L, and so for $Lr/2n^2 \le x \le 1/2$ we obtain

$$|\Delta^r_{L/n^2}(g - q_n)(x)| \le 2C_1 M.$$

But

$$
\begin{aligned}
\Delta^r_{L/n^2}(g - q_n)(x) &= (-1)^r (g(x - Lr/2n^2) - q_n(x - Lr/2n^2)) \\
&+ \sum_{j=0}^{r-1} (-1)^j \binom{r}{j} (g(x - L(r/2 - j)/n^2) - q_n(x - L(r/2 - j)/n^2)),
\end{aligned}
$$

and since all arguments $x - L(r/2 - j)/n^2$, $j = 0, 1, \ldots, r-1$, in the sum belong to $[L/n^2, 1]$, we can conclude (see also the definition of M in (7.9))

$$|g(x - Lr/2n^2) - q_n(x - Lr/2n^2)| \le 2C_1 M + M,$$

which completes the proof, since $x - Lr/2n^2$ can be any point in $[0, L/n^2]$ by appropriately selecting $x \in [Lr/2n^2, 1/2]$.

∎

Thus, Theorem 6.5 follows if we can prove (7.2). Our approach will be to map G_P into a cube, approximate on that cube (this was done already in 1987 in [6]) by some polynomials $Q_n(x_1, \ldots, x_{d-1}, u)$, and then go back to G_P via the mapping

$$(x_1, \ldots, x_{d-1}, u) \to (x_1, \ldots, x_{d-1}, \Phi(x_1, \ldots, x_d)).$$

37

Since Φ is a polynomial, this latter mapping will result in a polynomial approximant on the original G_P.

Before turning to details in connection with (7.2), we would like to make one more remark which simplifies things later. First of all, without loss of generality we may assume that the cube I_P in the construction of Γ_P is $[0,1]^{d-1}$ (this can always be achieved by rescaling). We have defined G_P by attaching an $\underline{e}_d$ directional segment of fixed length r_0 to any point of Γ_P, but it is clear that the only thing we need of G_P is that it covers a neighborhood (relative to G) of Γ_P^*. So instead of fixed length segments we can define G_P by attaching an $\underline{e}_d$ directional segment to any $Q \in \Gamma_P$ of some positive length $l_Q \leq r_0$, where l_Q is a continuous function of Q. This modified G_P can serve the same purpose as the originally defined G_P, in particular, it will be enough to prove (7.2) for any such domain G_P.

It will be convenient to select this G_P as follows. Let Φ be the polynomial that describes the boundary Γ around $P \in \Gamma$ via $\Phi(x_1, \ldots, x_d) = 0$. We have assumed that in a neighborhood of P, the domain G lies "above" Γ in the sense that if $(x_1, \ldots, x_{d-1}, x_d) \in \Gamma_P$, then for small $\varepsilon > 0$ the point $(x_1, \ldots, x_{d-1}, x_d + \varepsilon)$ belongs to G (and then $(x_1, \ldots, x_{d-1}, x_d - \varepsilon)$ does not belong to G). We have also seen that $\partial\Phi/\partial x_d \neq 0$ on Γ_P, and we may assume that $\partial\Phi/\partial x_d > 0$ (otherwise replace Φ by its negative). Then in a neighborhood of P the domain G is given by the inequality $\Phi \geq 0$, so for small $\delta > 0$ all the points $(x_1, \ldots, x_d)$ with $0 \leq x_1, \ldots, x_{d-1} \leq 1$ and $0 \leq \Phi(x_1, \ldots, x_d) \leq \delta$ belong to G, and we choose G_P as the set of all these points:

$$ G_P = \{(x_1, \ldots, x_d) \,|\, (x_1, \ldots, x_{d-1}) \in [0,1]^{d-1}, \ 0 \leq \Phi(x_1, \ldots, x_d) \leq \delta\}. \quad (7.10) $$

This choice works just as well as the original one in (7.1), and actually in it $\delta > 0$ can be as small as we wish.

8   The $r = 1$ case

In this section we prove (7.2) in the special case $r = 1$. Although later we shall give the proof for higher order moduli of smoothness, we feel it is instructive to separately give the proof for this simplest case because in that proof we may concentrate on the geometry involved in the problem, and the $r = 1$ case holds all the necessary geometric reasoning without the technical details needed for higher order moduli. Later we will not repeat this geometric part in the more technical proof involving higher order smoothness.

So let G_0 be the part of G which contains the points $(x_1, \ldots, x_d)$ with $(x_1, \ldots, x_{d-1}) \in [0,1]^{d-1}$ and $0 \leq \Phi(x_1, \ldots, x_d) \leq \delta$ (with the polynomial Φ that defines the boundary Γ) for some δ. By replacing Φ by Φ/δ we may assume $\delta = 1$. As we have seen, in this domain $\partial\Phi/\partial x_d > 0$, and we have the natural mapping $X = (x_1, \ldots, x_d) \to Y = (x_1, \ldots, x_{d-1}, u)$ from G_0 onto $[0,1]^d$ where $u = \Phi(x_1, \ldots, x_d)$. We can go from Y to X by setting $x_d = \chi_{x_1, \ldots, x_{d-1}}(u)$ for which

$$ \Phi(x_1, \ldots, x_{d-1}, \chi_{x_1, \ldots, x_{d-1}}(u)) = u. \quad (8.1) $$

Clearly (use the inverse function theorem), this χ is real analytic in $x_1, \ldots, x_{d-1}, u$ with strictly positive derivative for $u \in [0, 1]$. In fact, the assumed properties of Φ and the inverse function theorem imply positive upper and lower bounds for the derivative $\partial\chi_{x_1,\ldots,x_{d-1}}(u)/\partial u$ of $\chi_{x_1,\ldots,x_{d-1}}$ that are independent of $x_1, \ldots, x_{d-1}$. Another way to see that is to differentiate the equation (8.1) with respect to u.

If for fixed $\zeta_1, \ldots, \zeta_{d-1}$ we denote $\ell_{\zeta_1,\ldots,\zeta_{d-1}}$ the line $\{x_1 = \zeta_1, \ldots, x_{d-1} = \zeta_{d-1}\}$ in the x_d-direction, then the mapping $Y \to X$ maps the segment

$$J_{\zeta_1,\ldots,\zeta_{d-1}} = [0, 1]^d \cap \ell_{\zeta_1,\ldots,\zeta_{d-1}}$$

onto the segment

$$I_{\zeta_1,\ldots,\zeta_{d-1}} = G_0 \cap \ell_{\zeta_1,\ldots,\zeta_{d-1}}.$$

We also set $F(Y) = f(X)$, so F is defined on $[0, 1]^d$. On $J_{\zeta_1,\ldots,\zeta_{d-1}}$ we have

$$F(\zeta_1, \ldots, \zeta_{d-1}, u) = f(\zeta_1, \ldots, \zeta_{d-1}, \chi_{\zeta_1,\ldots,\zeta_{d-1}}(u)),$$

i.e.

$$F(x_1, \ldots, x_{d-1}, \Phi(x_1, \ldots, x_d)) = f(x_1, \ldots, x_d),$$

so we obtain from the from the aforementioned properties of $\chi_{\zeta_1,\ldots,\zeta_{d-1}}$ that

$$\omega^1_{J_{\zeta_1,\ldots,\zeta_{d-1}}}(F, \delta) \sim \omega^1_{I_{\zeta_1,\ldots,\zeta_{d-1}}}(f, \delta)$$

uniformly in $(\zeta_1, \ldots, \zeta_{d-1}) \in [0, 1]^{d-1}$, i.e., the u-directional modulus of continuity of F is bounded by the x_d-directional modulus of continuity of f. We emphasize that this is no longer true for higher order moduli of smoothness, for a nonlinear substitution does not preserve the size of those moduli (for example, the second order modulus of smoothness of $g(t) = t$ is identically 0, but if we make any nonlinear substitution into g then the second order modulus will not be 0).

Consider now another coordinate direction, say the x_1-directional modulus of continuity of F. Note that the mapping $Y \to X$ does not change the 2nd, 3rd,... $(d-1)$st coordinates of a point. For fixed $\zeta_2, \ldots, \zeta_{d-1}, \zeta \in [0, 1]$ the image of the segment $\{(x, \zeta_2, \ldots, \zeta_{d-1}, \zeta) \mid x \in [0, 1]\}$ under the mapping $Y \to X$ is the intersection of G_0 with the polynomial curve

$$\gamma_{\zeta_2,\ldots,\zeta_{d-1},\zeta} := \{(x_1, x_d) \mid \Phi(x_1, \zeta_2, \ldots, \zeta_{d-1}, x_d) = \zeta\}$$

in the plane $\{x_2 = \zeta_2, \ldots, x_{d-1} = \zeta_{d-1}\}$, and the mapping in between them is given by $x \to (x, \zeta_1, \ldots, \zeta_{d-1}, \chi_{x,\zeta_1,\ldots,\zeta_{d-1}}(\zeta))$. There is a $c_0 > 0$ such that the distance from $\gamma_{\zeta_2,\ldots,\zeta_{d-1},\zeta} \cap G_0$ to the boundary Γ of G is $\geq c_0\zeta$ (because $\partial\Phi/\partial x_d > 0$). Since $Y \to X$ is a Lipshitz map, it follows that if Y_1, Y_2 are two points of distance $\leq 1/n$ apart on the segment $\{(x, \zeta_2, \ldots, \zeta_{d-1}, \zeta) \mid x \in [0, 1]\}$ and X_1, X_2 correspond to them under the mapping $Y \to X$, then X_1, X_2 are of distance $\leq D/n$, and if $\zeta \geq L/n^2$, then their distance from the boundary of G is $\geq c_0L/n^2$. Let r_0 be the radius from the beginning of Section 6 and C_{r_0} the circle
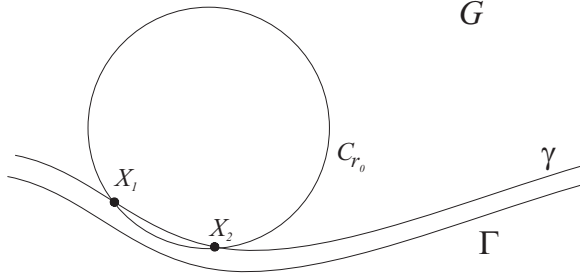
Figure 8: The points X_1, X_2 on the curve $\gamma = \gamma_{\zeta_2,\ldots,\zeta_{d-1},\zeta}$ and the circle C_{r_0} through them

in the plane $\{x_2 = \zeta_2, \ldots, x_{d-1} = \zeta_{d-1}\}$ of radius r_0 which passes through X_1 and X_2 and which has its center above the curve $\gamma_{\zeta_2,\ldots,\zeta_{d-1},\zeta}$ in that plane. We claim that for sufficiently large L the circle C_{r_0} lies inside G. Indeed, if σ_{X_1,X_2} is its shorter arc in between X_1 and X_2, then σ_{X_1,X_2} is of distance $\leq D_1/n^2$ — with some constant D_1 independent of n and $\zeta_2,\ldots,\zeta_{d-1},\zeta$ — from the arc of $\gamma_{\zeta_2,\ldots,\zeta_{d-1},\zeta}$ lying in between X_1 and X_2. On the other hand, this latter arc lies of distance $\geq c_0 L/n^2$ from the boundary, so for large L the arc σ_{X_1,X_2} lies inside G, see Figure 8. Furthermore, for large n the other (longer) arc of C_{r_0} in between X_1 and X_2 lies inside G by the definition of r_0 from Section 6, and these prove that C_{r_0} lies, indeed, inside G.

Thus, by the definition of $\omega^1(f, \delta)_G$, we have

$$|F(Y_1) - F(Y_2)| = |f(X_1) - f(X_2)| \leq \omega^1_{C_{r_0}}(f, D_2/n) \prec \omega^1(f, 1/n)_G.$$

This can be done with the other, $x_2, \ldots, x_{d-1}$ directions, as well, and it follows that $\omega^1_{[0,1]^{d-1} \times [L/n^2, 1]}(F, 1/n)_{\{\underline{e}_j\}_{j=1}^d}$, the directional modulus of continuity of F on $[0,1]^{d-1} \times [L/n^2, 1]$ in the coordinate directions, is bounded by an absolute constant times $\omega^1(f, 1/n)_G$. Now we can invoke [6, Theorem 12.1.1] for the rectangular parallelepiped $[0,1]^{d-1} \times [L/n^2, 1]$, according to which there are polynomials $Q_n(x_1, \ldots, x_{d-1}, u)$ of degree at most n such that

$$\|F - Q_n\|_{[0,1]^{d-1} \times [L/n^2, 1]} \leq C\omega^1_{[0,1]^{d-1} \times [L/n^2, 1]}(F, 1/n)_{\{\underline{e}_j\}_{j=1}^d} \leq C\omega^1(f, 1/n)_G.$$

On substituting here $u = \Phi(x_1, \ldots, x_d)$ we obtain

$$|f(x_1, \ldots, x_d) - Q_n(x_1, \ldots, x_{d-1}, \Phi(x_1, \ldots, x_d))| \leq C\omega^1(f, 1/n)_G \qquad (8.2)$$

on the part of G_0 which satisfies $\Phi(x_1, \ldots, x_d) \geq L/n^2$. Here the function $Q_n(x_1, \ldots, x_{d-1}, \Phi(x_1, \ldots, x_d))$ is a polynomial of degree $\leq n \cdot \deg(\Phi)$. On every x_d-directional segment $I = I_{\zeta_1,\ldots,\zeta_{d-1}} \subset G_0$ we have $u = \Phi(x_1, \ldots, x_d) \geq L/n^2$ for all points but those lying on a segment of the "lower" part of I (in the x_d direction) of length $\leq C_0 L/n^2$ (with some C_0 independent of L and n), so (8.2) is true for all such points, and from that we can conclude from Lemma 7.3 that

(8.2) pertains to hold on the whole of G_0. This proves (7.2) in the special case when $r = 1$.

∎

9 Approximation on disks

In this section we consider the special case when G is the disk

$$G := \{(x_1, x_2) \,|\, R_0^2 - (x_1 - 1/2)^2 - x_2^2 \geq 0\},$$

i.e., $d = 2$ and $\Phi(x_1, x_2) = R_0^2 - (x_1 - 1/2)^2 - x_2^2$, and prove Theorem 6.5, more precisely (7.2), in this case.

This section is devoted solely to disks. This special domain will be used in an essential way when proving the general case. Furthermore, several estimates on the moduli of smoothness of transformed functions that we do in this section work also for general domains and they will not be repeated (just mentioned with reference to this section) in final part of the proof.

A typical boundary domain from Section 7 for this G is

$$\{(x_1, x_2) \,|\, a \leq x_1 \leq b, \ R_0^2 - \delta \leq (x_1 - 1/2)^2 + x_2^2 \leq R_0^2, \ x_2 \leq 0\}$$

with some $1/2 - R_0 < a < b < 1/2 + R_0$ and small $\delta > 0$, and it is sufficient to show the possibility of Jackson type approximation on such a domain. For simplicity we set $[a, b] = [0, 1]$ and $\delta = 1$ (the reasoning in the general case would be the same), and again for convenience we assume $R_0 \geq 1$. So let

$$G_0 := \{(x_1, x_2) \,|\, 0 \leq x_1 \leq 1, \ R_0^2 - 1 \leq (x_1 - 1/2)^2 + x_2^2 \leq R_0^2, \ x_2 \leq 0\},$$

see Figure 9.

Under the mapping $X \to Y$, $(x_1, x_2) \to (x_1, u)$, $u = R_0^2 - (x_1 - 1/2)^2 - x_2^2$, discussed in Section 8 this G_0 is mapped into $[0, 1]^2$, and if we set $F(Y) = f(X)$, then we are going to show that

$$\omega_{[0,1]^2}^r(F, 1/n)_{\{\underline{e}_j\}_{j=1}^2} \prec \omega^r(f, 1/n)_G. \tag{9.1}$$

Recall that on the left the modulus is the directional modulus of smoothness (defined in (1.6) on $[0, 1]^2$, see also (1.7) and (1.1)) using the x_1 and $x_2 (\equiv u)$ directions on $[0, 1]^2$, while on the right $\omega^r(f, 1/n)_G$ is the modulus of smoothness defined in Section 6. First we discuss the directional modulus $\omega^r(F, 1/n)_{\underline{e}_1}$ of F in the x_1 direction.

For simplicity, in what follows we write $\omega(1/n)$ for $\omega^r(f, 1/n)_G$.

Let C_R be the circle about $(1/2, 0)$ and of radius $R \leq R_0$ with $R^2 \geq R_0^2 - 1$. By the definition of the modulus of smoothness, the 2π-periodic function

$$g(\theta) = f(1/2 + R \sin\theta, -R \cos\theta)$$

has r-th modulus of smoothness $\leq \omega^r(f, \delta)_G$, hence it can be approximated by trigonometric polynomials T_n of degree at most n in the order $\omega(1/n)$: $\|g -$
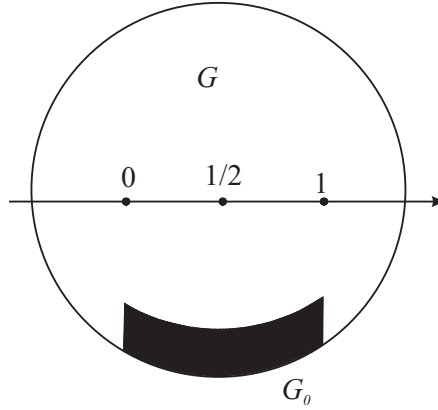
Figure 9:

$T_n\| \prec \omega(1/n)$ (see [5, Theorem 7.2.3]). Then we also have $\omega^r(T_n, 1/n) \prec \omega(1/n)$ (here ω^r is the classical modulus of smoothness of the trigonometric polynomial T_n), and since, by Stechkin's theorem [13], $\|T_n^{(r)}\| \prec n^r \omega^r(T_n, 1/n)$, it follows that $\|T_n^{(r)}\| \prec n^r \omega^r(1/n)$.

If $r - 1 \geq 1$, then $T_n^{(r-1)}$, being a trigonometric polynomial without a constant term, vanishes somewhere, say $T_n^{(r-1)}(\theta_0) = 0$, $\theta_0 \in [0, 2\pi]$. But then for all other $\theta \in [0, 2\pi]$ we have

$$|T_n^{(r-1)}(\theta)| = \left| \int_{\theta_0}^{\theta} T_n^{(r)}(u) du \right| \leq 2\pi \|T_n^{(r)}\| \prec n^r \omega(1/n).$$

Repeating this argument using $T_n^{(r-1)}$ in place of $T_n^{(r)}$ we get $\|T_n^{(r-2)}\| \prec n^r \omega(1/n)$, and so on, we can infer for all $1 \leq j \leq r$

$$\|T_n^{(j)}\| \prec n^r \omega(1/n). \tag{9.2}$$

Now consider $F(x_1, u)$ on the horizontal segment $u = R_0^2 - R^2$ of $[0, 1] \times [0, 1]$. It has the form

$$F(x_1, R_0^2 - R^2) = f(x_1, -\sqrt{R^2 - (x_1 - 1/2)^2}) = g(\arcsin((x_1 - 1/2)/R)).$$

Thus,

$$|\Delta_h^r F(x_1, R_0^2 - R^2)| \leq |\Delta_h^r T_n(\arcsin((x_1 - 1/2)/R))| + 2^r \|g - T_n\|,$$

which can be estimated as (see e.g. [5, Chapter 2, (7.10)])

$$|h|^r \left\| \left(T_n(\arcsin((x_1 - 1/2)/R)) \right)^{(r)} \right\|_{[0,1]} + O(\omega(1/n)).$$

42

Since, by Faà di Bruno's formula (cf. [10, Theorem 1.3.2]) the norm on the right is at most $C \sum_{j=1}^{r} \|T_n^{(j)}\|$, for $h \leq 1/n$ we obtain from (9.2) that

$$|\Delta_h^r F(x_1, R_0^2 - R^2)| \prec \omega(1/n),$$

which is stronger than

$$\omega_{[0,1]^2}^r (F, 1/n)_{\underline{e}_1} \prec \omega(1/n) \tag{9.3}$$

(recall that the modulus of smoothness on the left of (9.3) was defined in (1.6) and it involves a quantity smaller than 1 in the step-size $\cdot$ of $\Delta^r F$).

Before going on, let us record here the following consequence:

Corollary 9.1 *Let G be any domain in $\mathbf{R}^d$ as discussed in Section 6, and let C_R be any circle that lies in G together with its interior which is parallel with a coordinate plane. If P_1 and P_2 are any two points of C_R, then for any $n = 1, 2, \ldots,$*

$$|f(P_1) - f(P_2)| \leq C_R n^r \omega^r(f, 1/n)_G,$$

and here the constant C_R depends only on G and the radius of the circle in question.

Proof. We have essentially proved this above: choose a T_n as before with $\|f - T_n\| \prec \omega^r(f, 1/n)_G$ on C_R, and use that

$$|T_n(P_1) - T_n(P_2)| \leq C_R n^r \omega^r(f, 1/n)_G$$

is a consequence of $\|T_n'\| \prec n^r \omega^r(f, 1/n)_G$ (see (9.2)).

∎

Next, we consider the modulus of smoothness of F in the $\underline{e}_2$ (i.e., $u = x_2$) direction. For a ζ_1 let ℓ_{ζ_1} be the $x_1 = \zeta_1$ line, and, as in Section 8, let $J = J_{\zeta_1} = \ell_{\zeta_1} \cap [0,1]^2$ and $I = I_{\zeta_1} = \ell_{\zeta_1} \cap G_0$, and in what follows we keep the notations established there. We have seen that the mapping $Y \to X$, when restricted to J, is given by a real analytic function $\chi(u) = \chi_{\zeta_1}(u)$ with positive derivative in u (which is uniformly bounded from below and above in ζ_1).

By definition, $\omega_I^r(f, 1/n) \leq \omega^r(f, 1/n)_G$, and by the equivalence of the φ-modulus of smoothness with the K-functional (see Theorem [6, Theorem 2.1.1]) there is a g defined on I such that

$$\omega_I^r(f, 1/n) \prec \|f - g\|_I + \frac{1}{n^r} \|\varphi_I^r g^{(r)}\|_I \prec \omega_I^r(f, 1/n) \prec \omega(1/n), \tag{9.4}$$

where φ_I is the φ function for I (i.e., if the endpoints of I are (ζ_1, a) and (ζ_1, b), then $\varphi_I(\zeta_1, u) = \sqrt{|u - a||u - b|}$) and the differentiation in $g^{(r)}$ is understood in the x_2-direction, so $g^{(r)} \equiv \partial^r g / \partial x_2^r$. Furthermore, the $\prec$ in (9.4) is uniform in I. This g will be used only on I, so we shall think of it as a function of a single variable.

Let $\tau_1, \ldots, \tau_r$ be equidistant points in the middle third of I of consecutive distance $|I|/3r$. Any two consecutive points τ_j, τ_{j+1}, lie on a circle of radius $|I|/6r$ that lies in G, so, by Corollary 9.1, we have $|f(\tau_{j+1}) - f(\tau_j)| \prec n^r \omega(1/n)$. Then, in view of (9.4), the same is true for g: $|g(\tau_{j+1}) - g(\tau_j)| \prec n^r \omega(1/n)$. But then the mean value theorem gives that $g' = \partial g / \partial x_2$ is $\prec n^r \omega(1/n)$ somewhere in between τ_1 and τ_2. Using higher order mean value theorems (i.e. that $\Delta_h^j w = h^r w^{(j)}(\xi)$ for some ξ) we can conclude the same way that for all $1 \le j \le r$ there is a ξ_j in the middle third of I such that

$$g^{(j)}(\xi_j) \prec n^r \omega(1/n).$$

Now if $\xi \in I$ lies below $v \in I$ (in the x_2 direction) and at the same time v lies below ξ_{r-1}, then $\varphi_I(\xi) \le C\varphi_I(v)$ with an absolute constant (actually, $C = 3/2\sqrt{2}$ works). Here we used that ξ_{r-1} belongs to the middle third of I. Therefore, we obtain for ξ lying below ξ_{r-1}

$$
\begin{aligned}
\varphi_I(\xi)^r |g^{(r-1)}(\xi)| &\le \varphi_I(\xi)^r \left| \int_\xi^{\xi_{r-1}} g^{(r)}(v) dv \right| + \varphi_I(\xi)^r |g^{(r-1)}(\xi_{r-1})| \\
&\prec \left| \int_\xi^{\xi_{r-1}} \varphi_I(v)^r |g^{(r)}(v)| dv \right| + n^r \omega(1/n) \prec n^r \omega(1/n),
\end{aligned}
$$

where we used (9.4), and the choice of ξ_{r-1}. A similar estimate is true if ξ lies above ξ_{r-1}. Since the same reasoning can be given for $g^{(r-2)}$ (using $g^{(r-2)}(\xi_{r-2}) \prec n^r \omega(1/n)$ and the just proven relation), we obtain $\varphi_I^r |g^{(r-2)}| \prec n^r \omega(1/n)$, and in general, for all $j = 1, 2, \ldots, r$ we have

$$\varphi_I^r |g^{(j)}| \prec n^r \omega(1/n). \tag{9.5}$$

On the segment J the function F is given by $f(\zeta_1, \chi(u)))$, so it is natural to consider $G(u) = g(\chi(u))$. For it we have

$$\|F - G\|_J = \|f - g\|_I. \tag{9.6}$$

Since $\varphi_J(u) \sim \varphi_I(\chi(u))$, it follows from the Faà di Bruno formula ([10, Theorem 1.3.2]) that

$$\left\| \varphi_J^r \frac{\partial^r G}{\partial u^r} \right\|_J \prec \sum_{j=1}^r \left\| \varphi_I(\chi)^r g^{(j)}(\chi) \right\|_J \prec n^r \omega(1/n), \tag{9.7}$$

where the last relation follows from (9.5). Thus, using the equivalence relation between the K-functional and the φ-moduli of smoothness (see [6, Theorem 2.1.1]) we obtain from (9.4), (9.6) and (9.7)

$$\omega_J^r(F, 1/n)_{\underline{e}_2} \prec \|F - G\|_J + \frac{1}{n^r} \left\| \varphi_J^r \frac{\partial^r G}{\partial u^r} \right\|_J \prec \omega(1/n).$$

Here $J = J_{\zeta_1}$ with $\zeta_1 \in [0, 1]$ arbitrary, so we can conclude

$$\omega_{[0,1]^2}^r(F, 1/n)_{\underline{e}_2} \prec \omega(1/n). \tag{9.8}$$

44

Remark 9.2 Before going on with the discussion of approximation on the disk let us remark that the preceding proof for estimating $\omega^r_{[0,1]^2}(F, 1/n)_{\underline{e}_2}$ did not use that the boundary was a disk, it works for any polynomial curve as in Section 8 (just note that the small circles through the points τ_j are available also in that case, as well). What is more, the dimension need not be 2, either. Indeed, if the dimension is d, then the segments $I_{\zeta_1,\ldots,\zeta_{d-1}}$ and $J_{\zeta_1,\ldots,\zeta_{d-1}}$ (see Section 8) lie in the plane $\{x_2 = \zeta_2, \ldots, x_{d-1} = \zeta_{d-1}\}$, and the preceding proof can be carried out in that plane using circles and directional moduli belonging to that plane. So we have (9.8) in general:

$$\omega^r_{[0,1]^d}(F, 1/n)_{\underline{e}_d} \prec \omega^r(f, 1/n)_G. \tag{9.9}$$

Now we return to the proof of (7.2) for the disk that we are considering in this section. From (9.3) and (9.8) we can conclude, using Theorem [6, Theorem 12.1.1] (or Theorem 1.1 from the first part of this paper) for the square $[0,1]^2$, that there are polynomials $Q_n(x_1, u)$ of degree at most n such that for all $(x, u) \in [0,1]^2$

$$|F(x, u) - Q_n(x, u)| \leq C \left(\omega^r_{[0,1]^2}(F, 1/n)_{\underline{e}_1} + \omega^r_{[0,1]^2}(F, 1/n)_{\underline{e}_2} \right) \prec \omega(1/n).$$

On substituting here $u = \Phi(x_1, x_2) = R_0^2 - (x_1 - 1/2)^2 - x_2^2$, we obtain

$$|f(x_1, x_2) - Q_n(x_1, R_0^2 - (x_1 - 1/2)^2 - x_2^2)| \prec \omega^r(f, 1/n)_G$$

on G_0. According to Section 7, this is enough to conclude $E_n(f)_G \prec \omega^r(f, 1/n)_G$.

∎

10   A detour on the derivative of composite functions

Let $P(x, y) = P_n(x, y)$ be a polynomial of two variables and $\sigma(t)$ a one-variable infinitely many times differentiable function. For $j = 0, 1, \ldots$ we set

$$D_j(P, t) = D_j(P, \sigma, t) = \sum_{s=0}^{j} \binom{j}{s} \frac{\partial^j P(t, \sigma(t))}{\partial x^{j-s} \partial y^s} (\sigma'(t))^s,$$

where $D_0(P, t) = P(t)$. For these straightforward differentiation yields

$$\frac{dD_j(P, t)}{dt} = D_{j+1}(P, t) + jD_{j-1}(\frac{\partial P}{\partial y}, t)\sigma''(t), \tag{10.1}$$

and hence the derivative with respect to t of an expression of the form

$$D_j(\frac{\partial^l P}{\partial y^l}, t)(\sigma''(t))^{k_2}(\sigma'''(t))^{k_3} \cdots (\sigma^{(r)}(t))^{k_r} \tag{10.2}$$

is the sum of the terms

$$D_{j+1}(\frac{\partial^l P}{\partial y^l}, t)(\sigma''(t))^{k_2}(\sigma'''(t))^{k_3}\cdots(\sigma^{(r)}(t))^{k_r},$$

$$jD_{j-1}(\frac{\partial^{l+1} P}{\partial y^{l+1}}, t)(\sigma''(t))^{k_2+1}(\sigma'''(t))^{k_3}\cdots(\sigma^{(r)}(t))^{k_r}$$

and

$$D_j(\frac{\partial^l P}{\partial y^l}, t)\Big(k_2(\sigma''(t))^{k_2-1}(\sigma'''(t))^{k_3+1} \quad \cdots \quad (\sigma^{(r)}(t))^{k_r}$$
$$+\cdots+(\sigma''(t))^{k_2}(\sigma'''(t))^{k_3} \quad \cdots \quad k_r(\sigma^{(r)}(t))^{k_r-1}\sigma^{(r+1)}(t)\Big).$$

Since

$$\frac{dP(t,\sigma(t))}{dt} = D_1(P,t)$$

is also true, it follows by induction on $m = 1, 2, \ldots$ that every derivative $d^m P(t,\sigma(t))/dt^m$ is a linear combinations of terms of the form (10.2) with $j + 2l \leq m$. Furthermore, the only terms of the form (10.2) with $j + 2l = m$ that appear in $d^m P(t,\sigma(t))/dt^m$ are the terms (see (10.1))

$$D_m(P,t), \quad D_{m-2}(\frac{\partial P}{\partial y}, t)\sigma''(t), \quad D_{m-4}(\frac{\partial^2 P}{\partial y^2}, t)(\sigma''(t))^2, \cdots,$$

and since there is no cancellation above in the differentiation (c.f. (10.1)), these terms do appear with some positive coefficients.

Suppose now that for some $r \geq 1$ we have, besides σ, infinitely many times differentiable functions $\sigma_1, \ldots, \sigma_r$ with the property that at some point t_0 we have $\sigma_\tau'(t_0) = \sigma'(t_0)$ for all $\tau = 1, \ldots, r$, but the second derivatives of the σ_τ's at $t = t_0$ are all different: if $\sigma_\tau''(t_0) = \rho_\tau$, then $\rho_\tau \neq \rho_\nu$ if $\tau \neq \nu$. Suppose also that for some polynomials $P_n(x,y)$, $n = 1, 2, \ldots$ we have the bounds

$$\frac{d^m P_n(t,\sigma_\tau(t))}{dt^m}\Big|_{t=t_0} \prec \Omega_n \tag{10.3}$$

for all $\tau = 1, \ldots, r$ and $1 \leq m \leq r$ with some sequence $\{\Omega_n\}$. We claim that under these assumptions we also have

$$\frac{d^r P_n(t,\sigma(t))}{dt^r}\Big|_{t=t_0} \prec \Omega_n, \tag{10.4}$$

and here the $\prec$ depends only the $\prec$ in (10.3) and on the numbers $\rho_1, \ldots, \rho_r$.

Indeed, since we have $\sigma_\tau'(t_0) = \sigma'(t_0)$, the expressions $D_j(\frac{\partial^l P_n}{\partial y^l}, \sigma_\tau, t_0)$, $\tau = 0, \ldots, r$, are all the same if we set $\sigma_0 = \sigma$. Therefore, in view of what has been said before, to prove (10.4) it is sufficient to show that for all j, l with $j + 2l \leq r$ we have

$$D_j(\frac{\partial^l P_n}{\partial y^l}, t_0) \prec \Omega_n, \tag{10.5}$$

which we are going to prove by induction on $j + 2l$. First of all, when $j + 2l = 1$ then necessarily $j = 1$ and $l = 0$, and in this case (recall that $\sigma = \sigma_0$)

$$
\begin{aligned}
D_j\left(\frac{\partial^l P_n}{\partial y^l}, t_0\right) = D_1(P_n, t_0) \quad &= \quad D_1(P_n, \sigma_0, t_0) = D_1(P_n, \sigma_1, t_0) \\
&= \quad \left.\frac{dP_n(t, \sigma_1(t))}{dt}\right|_{t = t_0} \prec \Omega_n
\end{aligned}
$$

by the assumption (10.3).

Suppose now that (10.5) is true for all $j + 2l < m \ (\leq r)$. As we have seen,

$$
\begin{aligned}
\left.\frac{d^m P_n(t, \sigma_\tau(t))}{dt^m}\right|_{t = t_0} \quad &= \quad D_m(P_n, t) + c_{m,1} D_{m-2}\left(\frac{\partial P_n}{\partial y}, 0\right)\sigma_\tau''(t_0) \\
&+ \quad c_{m,2} D_{m-4}\left(\frac{\partial^2 P_n}{\partial y^2}, 0\right)(\sigma_\tau''(t_0))^2 + \cdots + \text{terms with } j + 2l < m,
\end{aligned}
$$

where the coefficients $c_{m,s}$ are non-zero. By the induction hypothesis here all the terms with $j + 2l < m$ are $\prec \Omega_n$, and by assumption the left hand side is also $\prec \Omega_n$ for $\tau = 1, \ldots, r$. Therefore, in view of the notation $\sigma_\tau''(t_0) = \rho_\tau$, we have

$$
D_m(P_n, t_0) + c_{m,1} D_{m-2}\left(\frac{\partial P_n}{\partial y}, t_0\right)\rho_\tau + c_{m,2} D_{m-4}\left(\frac{\partial^2 P_n}{\partial y^2}, t_0\right)\rho_\tau^2 + \cdots \prec \Omega_n, \qquad \tau = 1, \ldots, r.
$$

Using the first[6] $[m/2] + 1$ of these relations we obtain an $([m/2] + 1) \times ([m/2] + 1)$ size linear system of equations for $c_{m,s} D_{m-2s}\left(\frac{\partial^s P_n}{\partial y^s}, 0\right)$, $s = 0, \ldots, [m/2]$, with non-zero determinant (a Vandermonde determinant in $\rho_1, \rho_2, \ldots, \rho_{[m/2]+1}$, which numbers are different by assumption) and with right hand side consisting of numbers $\prec \Omega_n$. Solving this system we obtain that all the terms $c_{m,s} D_{m-2s}\left(\frac{\partial^s P_n}{\partial y^s}, t_0\right)$ are $\prec \Omega_n$. Since $c_{m,s} \neq 0$, this completes the induction step, and the proof of (10.4) is complete.

What we have just proven, we shall use in Section 11 in the following situation. Suppose that D_{r_0} is a disk on the plane. Let $\sigma := (t, \sigma(t))$ be a $\mathbf{C}^\infty$ curve that lies in D_{r_0}, and suppose that $D_\rho \subset D_{r_0}$ is a disk that touches (i.e. its boundary touches) σ at some point $(t_0, \sigma(t_0))$. For a given $r \geq 1$ fix r different radii $\rho/2 \leq R_\tau \leq \rho$, $1 \leq \tau \leq r$, and let C_{R_τ} be the circle of radius R_τ that lies in D_ρ and touches σ at the point $(t_0, \sigma(t_0))$. If C_{R_τ} is the circle $(x - x_\tau)^2 + (y - y_\tau)^2 = R_\tau^2$, then in a neighborhood of the point $(t_0, \sigma(t_0))$ it has parametrization $(t, \sigma_\tau(t))$ with

$$
\sigma_\tau(t) = y_\tau - \sqrt{R_\tau^r - (t - x_\tau)^2}
$$

(note that by the form of σ, the common tangent line to σ and D_ρ at $(t_0, \sigma(t_0))$ is not vertical) with possibly $+$ replacing $-$ in front of the square root sign, but

---

[6][·] denotes integral part

in what follows we shall keep the $-$ sign. Now

$$\sigma_\tau'(t_0) = \frac{t - x_\tau}{\sqrt{R_\tau^2 - (t - x_\tau)^2}},$$

$$\sigma_\tau''(t_0) = \frac{R_\tau^2}{(R_\tau^2 - (t - x_\tau)^2)^{3/2}} = \frac{1}{R_\tau}(1 + \sigma_\tau'(t_0)^2)^{3/2}. \tag{10.6}$$

These σ_τ touch σ at $(t_0, \sigma(t_0))$, so they have the same derivative at t_0 as σ: $\sigma'(t_0) = \sigma_\tau'(t_0)$ (this follows from the fact that the tangent direction to σ_τ at $\sigma_\tau(t_0)$ is $(1, \sigma_\tau'(t_0))$). But then (10.6) shows that the second derivatives are different at t_0: $\sigma_\tau''(t_0) \neq \sigma_\tau''(t_0)$ because the R_τ, $\tau = 1, \ldots, r$ are different.

Let now f be a function on a domain G that contains D_{r_0}, and let $P_n(x, y)$ be polynomials (of degree $n = 1, 2, \ldots$) of two variables such that $\|f - P_n\|_{D_{r_0}} \prec \omega(1/n)$, where $\omega(1/n)$ is the r-th modulus of smoothness of f on G. In Section 9 it was shown that for every $r \geq 1$ there are such polynomials P_n. The polynomial P_n, when restricted to either of the circles C_{R_τ}, gives a trigonometric polynomial

$$T_n(\theta) = P_n(x_\tau + R_\tau \sin \theta, y_\tau - R_\tau \cos \theta).$$

For this T_n it was shown in Section 9 (see (9.2)) that

$$\|T_n^{(j)}\| \prec n^r \omega(1/n) \qquad \text{for all } j = 1, 2, \ldots, r, \tag{10.7}$$

and here $\prec$ depends only on R_τ. The expression $P_n(t, \sigma_\tau(t))$ is $T_n(\arcsin((t - x_\tau)/R_\tau)$, so from (10.7) and from Faà di Bruno's formula (cf. [10, Theorem 1.3.2]) we obtain

$$\frac{d^m P_n(t, \sigma_\tau(t))}{dt^m}\bigg|_{t = t_0} \prec n^r \omega(1/n) \tag{10.8}$$

for all $1 \leq m \leq r$. Since this is true uniformly for all $\tau = 1, \ldots, r$, we are in the situation of (10.3) with $\Omega_n = n^r \omega(1/n)$, so from (10.4) we can conclude

$$\frac{d^r P_n(t, \sigma(t))}{dt^r}\bigg|_{t = t_0} \prec n^r \omega(1/n). \tag{10.9}$$

11 Proof of Theorem 6.5

According to the general setup in Section 7, it is sufficient to prove (7.2), i.e. the possibility of Jackson-rate approximation on the sets G_P. So let G_0 be such a set, and without loss of generality we may assume that G_0 is positioned as in the discussion leading to (7.2), i.e. it is the union of some x_d-directional segments over a small part of the boundary G. As in Section 8, the closed domain G_0 corresponds to $[0, 1]^d$ under the mapping $X \to Y$, and f corresponds to $F(Y) = f(X)$.

As in Sections 8–9 it is sufficient to show that the directional modulus of smoothness

$$\omega_{[0,1]^d}^r(F, \delta)_{\{\underline{e}_i\}_{i=1}^d} = \max_{1 \leq i \leq d} \omega_{[0,1]^d}^r(F, \delta)_{\underline{e}_i}$$

of F satisfies

$$\omega^r_{[0,1]^d}(F, 1/n)_{\{\underline{e}_i\}_{i=1}^d} \leq C\omega^r(f, 1/n)_G,$$

for then we can use Theorem [6, Theorem 12.1.1] (or Theorem 1.1 from the first part of this work) to find polynomials $Q_n(x_1, \ldots, x_{d-1}, u)$ with

$$\|F - Q_n\|_{[0,1]^d} \leq C\omega^r(f, 1/n)_G,$$

and then the substitution $u = \Phi(x_1, \ldots, x_d)$ will lead to

$$\|f - \tilde{Q}_n\|_{G_0} \leq C\omega^r(f, 1/n)_G$$

for the polynomial $\tilde{Q}_n(x_1, \ldots, x_d) = Q(x_1, \ldots, x_{d-1}, \Phi(x_1, \ldots, x_d))$.

Thus, everything boils down to proving

$$\omega^r(F, 1/n)_{\underline{e}_i} \leq C\omega^r(f, 1/n)_G \tag{11.1}$$

for each $i = 1, 2, \ldots, d$. For $i = d$ this was done in Section 9, see Remark 9.2, so here we shall only consider $i = 1, 2, \ldots, d-1$ (note that because of the assumed positioning of G_0, the directions $\underline{e}_1, \ldots, \underline{e}_{d-1}$ have a different role with respect to G_0 than the direction $\underline{e}_d$), and since these cases are completely similar to each other, we shall assume $i = 1$. Thus, let $\zeta_2, \ldots, \zeta_{d-1}, \zeta \in [0, 1]$ be fixed, and consider the segment $V = \{(x_1, \zeta_2, \ldots, \zeta_{d-1}, \zeta) \mid x_1 \in [0, 1]\}$. We need to estimate the r-th difference of F on that segment. As we have seen in Section 8, under the mapping $Y \to X$ the segment V is mapped into the intersection of G_0 with the polynomial curve

$$\gamma_{\zeta_2, \ldots, \zeta_{d-1}, \zeta} := \{(x_1, \zeta_2, \ldots, \zeta_{d-1}, x_d) \mid \Phi(x_1, \zeta_2, \ldots, \zeta_{d-1}, x_d) = \zeta\} \tag{11.2}$$

lying in the plane $\{x_2 = \zeta_2, \ldots, x_{d-1} = \zeta_{d-1}\}$, and the mapping is given by $(x, \zeta_2, \ldots, \zeta_{d-1}, \zeta) \to (x, \zeta_1, \ldots, \zeta_{d-1}, \chi_{x, \zeta_1, \ldots, \zeta_{d-1}}(\zeta))$ with a real analytic function $\chi_{x, \zeta_1, \ldots, \zeta_{d-1}}(\zeta)$. If we write $\sigma(t)$ for $\chi_{t, \zeta_1, \ldots, \zeta_{d-1}}(\zeta)$, then we need to estimate the r-th difference Δ_h^r of the composed function

$$f(t, \zeta_1, \ldots, \zeta_{d-1}, \sigma(t)).$$

We shall show that this can be done using the modulus of smoothness of f in the intersection of G_0 with the plane $\{x_2 = \zeta_2, \ldots, x_{d-1} = \zeta_{d-1}\}$, which is clearly smaller than the modulus of smoothness of f on the whole G. Thus, everything is happening in the plane $x_2 = \zeta_2, \ldots, x_{d-1} = \zeta_{d-1}$, so we may suppress the fixed coordinates $\zeta_2, \ldots, \zeta_{d-1}$ and we may assume that $d = 2$, and G is a domain on the plane. Let us also write $\gamma = \gamma_\zeta$ for the curve (11.2) in this case. This is the same as the curve $\{(t, \sigma(t)) \mid t \in [0, 1]\}$.

The information on the r-th modulus of smoothness on f is given on circles and straight segments, and we need to bound

$$\sup_{h \leq 1/n} \sup_{x \in [0,1]} \Delta_h^r f(x, \sigma(x)) \tag{11.3}$$

49

To be more precise, we should bound

$$\sup_{h \le 1/n} \sup_{x \in [0,1]} \Delta^r_{h\sqrt{x(1-x)}} f(x, \sigma(x)),$$

but that is at most as large as the expression in (11.3), so it is sufficient to deal with (11.3). Thus, we need to prove that

$$\Delta^r_h f(x, \sigma(x)) = \sum_{j=0}^r (-1)^j \binom{r}{j} f(x + (r/2 - j)h, \sigma(x + (r/2 - j)h))$$

are $\prec \omega^r(f, 1/n)_G$ for $0 \le h \le 1/n$ and $x \in [0,1]$. Here the points $U_j = (x + (r/2 - j)h, \sigma(x + (r/2 - j)h))$, $j = 0, \ldots, r$, lie on the curve γ and only their x-coordinates are equidistant. Our strategy will be the following. First of all, we shall use some polynomial approximant P_n to f on some large disk D that contains the points U_j which satisfies $\|f - P_n\|_D \prec \omega^r(f, 1/n)_G$. Then it is enough to consider $\Delta^r_h P_n(x, \sigma(x))$, which can be bounded by h^r times an estimate on the r-th derivative of P_n on the smallest arc of γ that contains all the points U_j. Thus, we have to find a bound for the r-th derivative of the composed function $P_n(x, \sigma(x))$, which was done in Section 10.

Thus, we have to estimate $\Delta^r_h F(x, \zeta) = \Delta^r_h(f(x, \gamma_\zeta(x))$ for $0 \le h \le 1/n$ with a fixed $\zeta \in [0,1]$. For simplicity we do the estimate for $h = 1/n$ (the $h < 1/n$ case can be similarly handled). Recall that $\gamma = \gamma_\zeta$ with $\zeta \in [0,1]$, and first assume that $\zeta \ge L/n^2$ for some fixed, but large L. Let $x \in [0,1]$ be such that the points $x + (r/2 - j)/n$ belong to $[0,1]$ for $j = 0, \ldots, r$. These points correspond to the points $U_j = (x + (r/2 - j)/n, \sigma(x + (r/2 - s)/n))$, $j = 0, \ldots, r$, on γ, which are of distance $\ge cL/n^2$ from the "lower" boundary of G, which is just the curve $(t, \gamma_0(t))$. Recall now the r_0 from the definition of the domain G from the beginning of Section 6.2, let C_{r_0} be the circle of radius r_0 that lies inside G together with its interior and touches the boundary curve γ_0 at the point $(x, \gamma_0(x))$, and let D_{r_0} be the closed disk enclosed by C_{r_0}. It is easy to see (see Figure 10) that if L is sufficiently large, then (for large n) the points U_j belong to D_{r_0} (recall also that G_0 was a small part of G). The points U_j lie on the part $\{(u, \gamma(u)) \mid x - r/2n \le u \le x + r/2n\}$ of γ, and there is a $\rho > 0$ depending only on r_0 and Φ such that to every point $(u, \gamma(u))$ of that portion of γ we can draw a disk of radius ρ that lies in D_{r_0} and touches γ at the given point.

According to what has been shown in Section 9, for $n = 1, 2, \ldots$ there are polynomials $P_n(x_1, x_2)$ of degree at most n such that they approximate f on D_{r_0} with error $\le C\omega^r(f, 1/n)_G$. Then, for fixed $\zeta \in [0,1]$,

$$|\Delta^r_{1/n}(F(x, \zeta))| \le |\Delta^r_{1/n}(F_n(x, \zeta))| + 2^r C\omega^r(f, 1/n)_G,$$

where $F_n(Y) = P_n(X)$ (i.e. $F_n(x, \zeta) = P_n(x, \gamma_\zeta(x))$). The first term on the right is bounded by (see [5, Chap. 2, (7.12)])

$$\frac{1}{n^r} \max_{x - r/2n \le u \le x + r/2n} \left| \frac{\partial^r F_n}{\partial u^r}(u, \zeta) \right|,$$
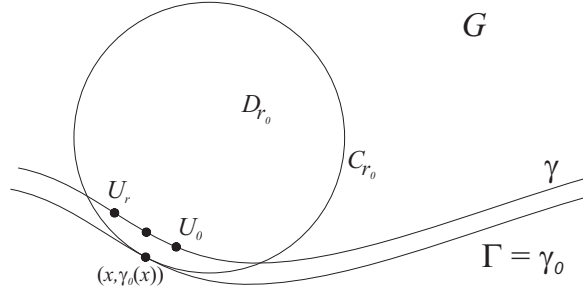
50

Figure 10: The points $U_0, \ldots, U_r$ on the curve $\gamma = \gamma_{\zeta_2, \ldots, \zeta_{d-1}, \zeta}$

so it is sufficient to show that for each u in the specified range

$$\left| \frac{\partial^r F_n}{\partial u^r}(u, \zeta) \right| \leq C n^r \omega^r(f, 1/n)_G. \tag{11.4}$$

But on the left we have with $\sigma(t) = \gamma_\zeta(t)$

$$\left| \frac{d^r P(t, \sigma(t))}{dt^r} \right|_{t=u}, \tag{11.5}$$

and at the end of Section 10 in (10.9) it was shown that this is $\prec n^r \omega(f, 1/n)$ (set in (10.9) $t_0 = u$). Furthermore, here the $\prec$ depends only on the radius ρ and on the radii $\rho/2 \leq R_1 < R_2 < \cdots < R_r \leq \rho$ used in the proof of (10.8), and it is easy to see that these can be selected independently of $x - r/2n \leq u \leq x + r/2n$, $0 \leq x \leq 1$, if L is sufficiently large (recall that now we are dealing with the situation $\zeta \geq L/n^2$). Therefore, we can conclude that this $\prec$ is independent of $u \in [x - r/2n, x + r/2n]$, x, $x \pm r/2n \in [0, 1]$. Thus, we have

$$\left| \frac{d^r P_n(u, \gamma_\zeta(u))}{du^r} \right| \prec n^r \omega(1/n),$$

which is the same as (11.4). This completes the proof of the estimate (11.1) for $i = 1$.

In this reasoning we have assumed $\zeta \geq L/n^2$. Now we show that the required estimate for $\Delta_{1/n}^r F(x, \zeta)$ in the missing range follows from what we have proven before. Indeed, for a $\zeta \in [0, L/n^2]$, with some large but fixed Λ, let us choose w so that $w - r\Lambda\sqrt{w(1-w)}/2n = \zeta$. Then ζ is the smallest point in the set $\{w_s\}_{s=0}^r$, $w_s := w + (r/2 - s)\Lambda\sqrt{w(1-w)}/n$, (we have $\zeta = w_r$), and if Λ is sufficiently large, then all these w_s, except ζ itself, are bigger than L/n^2 (note that $w \sim 1/n^2$). Now for each $x + (r/2 - jh)/n$, $0 \leq j \leq r$, consider the points $(x + (r/2 - j)/n, w_s)$, $s = 0, \ldots, r$, on the vertical segment $I_{x+(r/2-j)/n}$. By definition,

$$\left| \sum_{s=0}^r (-1)^s \binom{r}{s} F(x + (r/2 - j)/n, w_s) \right| \prec \omega^r(F, \Lambda/n)_{\underline{e}_2},$$

51

and above we have already seen that the right hand side is $\prec \omega^r(f, 1/n)_G$ (see (9.8)). Furthermore, for each $s < r$

$$\left| \sum_{j=0}^{r} (-1)^j \binom{r}{j} F(x + (r/2 - j)/n, w_s) \right|$$

is the $\Delta_{1/n}^r F(x, w_s)$ that we have just estimated (for $s < r$ we have $w_s > L/n^2$), so it is $\prec \omega^r(f, 1/n)$. But $\Delta_{1/n}^r F(x, \zeta)$ can be expressed by these:

$$
\begin{aligned}
(-1)^r \Delta_{1/n}^r F(x, \zeta) &= (-1)^r \sum_{j=0}^{r} (-1)^j \binom{r}{j} F(x + (r/2 - j)/n, \zeta) \\
&= \sum_{j=0}^{r} (-1)^j \binom{r}{j} \sum_{s=0}^{r} (-1)^s \binom{r}{s} F(x + (r/2 - j)/n, w_s) \\
&\quad - \sum_{s=0}^{r-1} (-1)^s \binom{r}{s} \sum_{j=0}^{r} (-1)^j \binom{r}{j} F(x + (r/2 - j)/n, w_s),
\end{aligned}
$$

so we have a bound $\prec n^r \omega(1/n)$ for it, as well. With this the proof of (11.1), and together with it also the proof of Theorem 6.5, is complete.

∎

$$********$$

Note added before final production. Feng Dai and Andriy Prymak have recently prepared the manuscript [2] on polynomial approximation on C^2-domains. In it they defined a new type of modulus of smoothness with which they proved Jackson and converse theorems (exactly as Theorems 6.5 and 6.6) for polynomial approximation on C^2 domains in any dimension. Their definitions and results are valid for L^p spaces, as well. Their modulus is close in spirit to the average moduli of Ivanov [8] mentioned in Remark 6.1,4, and in particular Dai and Prymak were able to deduce both the direct and the converse theorems on any C^2 domains that were announced in [8]. The main method of [2] is to use Whitney-type local approximation by polynomials of the fixed degree $(r-1)d$ (when one works with r-th order of smoothness in $\mathbf{R}^d$) in conjunction with polynomial partitions of unity similar to the one in the paper [7]. The results and methods in the second part of the present paper allows one to get similar quasi-Whitney local approximants (of fixed degree $\leq 2(r-1)d$) involving the moduli of smoothness of the present paper, and from there the procedure used in [2] gives Theorem 6.5 for all C^2 domains, not just for algebraic ones.

References

[1] B. Bojanov, Elementary proof of the Remez inequality, *The American Mathematical Monthly*, **100**(1993), 483–485.

[2] F. Dai and A. Prymak, Polynomial approximation on C^2-domains, arXiv:1910.11719

[3] F. Dai and Y. Xu, Moduli of smoothness and approximation on the unit sphere and the unit ball, *Adv. Math.*, **224**(2010), 1233–1310.

[4] F. Dai and Y. Xu, *Approximation theory and harmonic analysis on spheres and balls*. Springer Monographs in Mathematics, Springer, New York, 2013.

[5] R. A. DeVore and G.G. Lorentz, *Constructive approximation*, Grundlehren der mathematischen Wissenschaften, **303**, Springer-Verlag, Berlin, 1993.

[6] Z. Ditzian and V. Totik, *Moduli of smoothness*, Springer Series for Computational Mathematics, **9**, Springer Verlag, New York 1987.

[7] V. K. Dzjadyk and V. N. Konovalov, A method of partition of unity in domains with piecewise smooth boundary into a sum of algebraic polynomials of two variables that have certain kernel properties, *Ukrain. Mat. Z.*, **25**(1973), 179–192, 285 (Russian).

[8] K. G. Ivanov, Approximation of functions of two variables by algebraic polynomials. I, Anniversary volume on approximation theory and functional analysis (Oberwolfach, 1983), *Internat. Schriftenreihe Numer. Math.*, vol. **65**, Birkhäuser, Basel, 1984, pp. 249–255.

[9] V. N. Konovalov, On some constructive characteristics of some classes of functions of several variables. Dissertation, Kiev, 1972. (Russian)

[10] S. G. Krantz and H. R. Parks, *A primer of real analytic functions*, Second edition. Birkhäuser Advanced Texts: Basler Lehrbücher. Birkhäuser Boston, Inc., Boston, MA, 2002.

[11] E. J. Remez, Sur une propriété des polynômes de Tchebyscheff, *Comm. Inst. Sci. Kharkow.* **13**(1936), 93-95.

[12] E.B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Grundlehren der mathematischen Wissenschaften, **316**, Springer Verlag, Berlin, Heidelberg, 1997.

[13] S. B. Stechkin, Generalization of some inequalities of S. N. Bernstein, *Dokl. Akad. Nauk.*, **60**(1948), 1511-1514.

[14] V. Totik, *Polynomial Approximation on Polytopes*, Memoirs of the American Mathematical Society, **212**(2014), American Mathematical Society, Providence, Rhode Island, 2014.

MTA-SZTE Analysis and Stochastics Research Group
Bolyai Institute, University of Szeged
Szeged, Aradi v. tere 1, 6720, Hungary

and

Department of Mathematics and Statistics, University of South Florida
4202 E. Fowler Ave, CMC342, Tampa, FL 33620-5700, USA

totik@mail.usf.edu