Polynomial approximation on polytopes[1]

Vilmos Totik

February 9, 2013

[1]AMS Classification: 41A10, 41A17 Keywords: polynomials, several variables, approximation, moduli of smoothness, K-functionals, strong inequalities

Contents

Abstract

Polynomial approximation on convex polytopes in $\mathbf{R}^d$ is considered in uniform and L^p-norms. For an appropriate modulus of smoothness matching direct and converse estimates are proven. In the L^p-case so called strong direct and converse results are also verified. The equivalence of the moduli of smoothness with an appropriate K-functional follows as a consequence. The results solve a problem that was left open since the mid 1980's when some of the present findings were established for special, so called simple polytopes.

Part I

The continuous case

1 The result

We consider the problem of characterization of best polynomial approximation on polytopes in $\mathbf{R}^d$. To have a basis for discussion, first we briefly review the one-dimensional case.

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

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

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

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

is the r-th symmetric difference, and $\|\cdot\|_S$ denotes the supremum norm on a set S. In (1.1) it is agreed that $\Delta_h^r f(x) = 0$ if $[x - \frac{r}{2}h, x + \frac{r}{2}h] \not\subseteq [-1,1]$. Let

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

be the error of best approximation of f by polynomials p_n of degree at most n. Then (see [12, Theorem 7.2.1]) for $n \ge r$

$$E_n(f)_{[-1,1]} \le M\omega_\varphi^r\left(f, \frac{1}{n}\right) \tag{1.3}$$

and (see [12, Theorem 7.2.4])

$$\omega_\varphi^r\left(f, \frac{1}{n}\right) \le \frac{M}{n^r} \sum_{k=0}^{n} (k+1)^{r-1} E_k(f)_{[-1,1]}, \qquad n = 1,2,\ldots, \tag{1.4}$$

where M depends only on r.

(1.3)–(1.4) constitute what is usually called a characterization of the rate of best polynomial approximation in terms of moduli of smoothness, e.g. they give

$$E_n(f)_{[-1,1]} = O(n^{-\alpha}) \Longleftrightarrow \omega_\varphi^r(f,\delta) = O(\delta^\alpha)$$

for $\alpha < r$. This is precisely what we want to do for multidimensional polynomial approximation in $\mathbf{R}^d$. (1.3) is usually called the direct, or Jackson-type, while (1.4) is the converse, or Stechkin-type estimate. This latter (1.4) is a weak converse to (1.3), but that is natural, since $E_n(f)$ can tend to zero arbitrarily fast, but $\omega_\varphi^r(f, 1/n) \ge c/n^r$ unless f is a polynomial of degree at most $r-1$.

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 [12, Chapter 12], and to recall its definition we need to consider the function along lines in different directions. A direction e in $\mathbf{R}^d$ is just a unit vector $e \in \mathbf{R}^d$. Clearly, 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 $e \in S^{d-1}$ a direction. The line $l_{e,x}$ through x which is parallel with e intersects K in a segment $A_{e,x}B_{e,x}$. We call the minimum of the distances between x and $A_{e,x}, B_{e,x}$ the distance from x to the boundary of K in the direction of e:

$$d_K(e,x) = \min\{\mathrm{dist}(x, A_{e,x}), \mathrm{dist}(x, B_{e,x})\}, \tag{1.5}$$

while

$$\tilde{d}_K(e,x) = \sqrt{\mathrm{dist}(x, A_{e,x}) \cdot \mathrm{dist}(x, B_{e,x})} \tag{1.6}$$

could be called the normalized distance. Note that even if x lies on the boundary of K, it may happen that $d_K(e,x), \tilde{d}_K(e,x) > 0$; for example, if K is a cube of side length a, x is the midpoint of an edge and e is the direction of that edge, then $d_K(e,x) = \tilde{d}_K(e,x) = a/2$.

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

$$\Delta^r_{he} f(x) = \sum_{k=0}^{r} (-1)^k \binom{r}{k} f\left(x + (\frac{r}{2} - k)he\right) \tag{1.7}$$

with the agreement that this is 0 if $x + \frac{r}{2}he$ or $x - \frac{r}{2}he$ does not belong to K. Finally, define the r-th modulus of smoothness as (see [12, Section 12.2])

$$\omega^r_K(f,\delta) = \sup_{e \in S^{d-1},\ h \le \delta,\ x \in K} |\Delta^r_{h\tilde{d}_K(e,x)e} f(x)|, \tag{1.8}$$

which we shall often write in the form

$$\omega^r_K(f,\delta) = \sup_{e \in S^{d-1}} \sup_{h \le \delta} \|\Delta^r_{h\tilde{d}_K(e,x)e} f(x)\|_K, \tag{1.9}$$

i.e. $\omega^r_K(f,\delta)$ is the supremum of the directional moduli of smoothness

$$\omega^r_{K,e}(f,\delta) := \sup_{h \le \delta} \|\Delta^r_{h\tilde{d}_K(e,x)e} f(x)\|_K$$

for all directions. 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.

$$\omega^r_\varphi(f,\delta) = \omega^r_{[-1,1]}(f,\delta). \tag{1.10}$$

Another way to write the modulus of smoothness (1.8) is

$$\omega_K^r(f,\delta) = \sup_I \sup_{h\le\delta} \|\Delta^r_{h\tilde{d}_K(e,x)e} f(x)\|_I = \sup_I \omega_I^r(f,\delta), \tag{1.11}$$

where I runs through all chords of K, so $\omega_K^r(f,\delta)$ is just the supremum of all the moduli of smoothness $\omega_I^r(f,\delta)$ on chords of K, and here $\omega_I^r(f,\delta)$ is just the analogue (actually a transformed form) of the φ-modulus of smoothness ω_φ^r for the segment I.

It is also immediate that

$$\omega_\varphi^r(f,\delta) \equiv \omega_\varphi^r(f,1), \qquad \text{for all } \delta \ge 1, \tag{1.12}$$

and as a consequence,

$$\omega_K^r(f,\delta) \equiv \omega_K^r(f,1), \qquad \text{for all } \delta \ge 1. \tag{1.13}$$

We also set

$$E_n(f)_K = \inf_{P_n} \|f - P_n\|_K,$$

where the infimum is taken for all polynomials in d-variables of total degree at most n. This is the error in best polynomial approximation and this is what we would like to characterize.

The main result of this paper is

Theorem 1.1 *Let $K \subset \mathbf{R}^d$ be a d-dimensional convex polytope and $r = 1, 2, \ldots$. Then, for $n \ge rd$, we have*

$$E_n(f)_K \le M\omega_K^r\left(f, \frac{1}{n}\right), \tag{1.14}$$

where M depends only on K and r.

The matching weak converse

$$\omega_K^r\left(f, \frac{1}{n}\right) \le \frac{M}{n^r}\sum_{k=0}^{n}(k+1)^{r-1}E_k(f)_K, \qquad n = 1, 2, \ldots, \tag{1.15}$$

is an immediate consequence of (1.4) if we apply it on every chord (considered as $[-1,1]$) of K. See [12, Theorem 12.2.3,(12.2.4)], which proof goes over to our case without any change. Note also, that, exactly as in [12, Corollary 12.2.6], we get the following consequence of (1.14)–(1.15).

Corollary 1.2 *Let $\alpha > 0$ and let f be a continuous function on a d-dimensional convex polytope $K \subset \mathbf{R}^d$. If f can be approximated with error $n^{-\alpha}$ on any chord I of K by polynomials (of a single variable on I) of degree at most $n = 1, 2, \ldots$, then $E_n(f)_K \le Mn^{-\alpha}$, where M depends only on K and α.*

This Corollary tells us that $n^{-\alpha}$ rate of d-dimensional polynomial approximation is equivalent to $n^{-\alpha}$-rate of one-dimensional polynomial approximation along every segment of K (note that if we restrict any function/polynomial of d-variables to a chord I of K we get a function/polynomial of a single variable on I). This corollary is true only on polytopes, see [12, Proposition 12.2.7].

In Section 13 the same problem in L^p spaces will be considered, and in the second part of the paper we verify a complete analogue of Theorem 1.1 for L^p-approximation. In L^p spaces one can even do somewhat better (see Section 20), and we shall prove a stronger form of Theorem 1.1 and its converse (1.15).

There have been many works on polynomial approximation in several variables, for some of the recent ones see e.g. [4]–[1], [6]–[8], [11], [17] and [21]–[23] and the references there. In these works various moduli of smoothness are constructed for special sets like balls and spheres which solve the approximation problem there. Often the moduli are shown to be equivalent to a K-functional, and the approximation goes trough the use of that K-functional. These do not work on polytopes, and precisely the absence of the relevant K-functional what makes the problem of the present work difficult. We also mention the paper [13] where global approximation is characterized in terms of local ones.

For special polytopes Theorem 1.1 had a predecessor: call $K \subset \mathbf{R}^d$ a simple polytope if there are precisely d edges at every vertex of K. For example, simplices and cubes/parallelepipeds are simple polytopes. Now it was proven in Theorem [12, Theorem 12.2.3] that if K is a simple polytope, then

$$E_n(f)_K \le M\left(\omega_K^r\left(f, \frac{1}{n}\right) + n^{-r}\|f\|_K\right). \tag{1.16}$$

It has been an open problem in the last 25 or so years if this is true for non-simple polytopes (even for a single one!), and it is precisely what Theorem 1.1 claims in a slightly sharper form. The second term on the right of (1.16) is usually dominated by the first one, so the main improvement in Theorem 1.1 is not the dropping of this term (although we shall see that dropping that term is an important step in the proof), but the dropping of the "simple polytope" assumption. Why are simple polytopes easier to handle, i.e. why is (1.14) for simple polytopes substantially weaker then for general ones? The answer is that the crux of the matter is approximation around the vertices of the polytope. Now a vertex of a simple polytope looks like a vertex of a cube (modulo an affine transformation), and cubes are relatively easy to handle since they are products of segments (therefore, approximation on cubes can be reduced to approximation on $[-1, 1]$, as was done in [12]). This is no longer true if there are more than d edges at a vertex. Still, the simple polytope case will play an important role in the proof of Theorem 1.1.

Acknowledgement. This work has been done over several years (the L^∞ case preceding considerably the L^p case) during which the research was supported by the National Science Foundation DMS0968530 (Part I) and by the Europearn Research Council grant No. 267055 (Part II).

2 Outline of the proof

The first part of this paper is devoted to the proof of Theorem 1.1. The proof has several components, some of which are quite technical, therefore, in this section we give an outline. Recall that a polytope in $\mathbf{R}^d$ is simple if there are precisely d edges at every vertex.

Besides (1.8), we need another modulus of smoothness defined as

$$\overline{\omega}_K^r(f,\delta) = \sup_{e\in\mathcal{E}} \sup_{h\le\delta} \|\Delta^r_{h\tilde{d}_K(e,x)e} f(x)\|_K, \tag{2.1}$$

where $\mathcal{E} = \mathcal{E}_K$ is the direction of the edges of K. So the only difference in between $\omega_K^r(f,\delta)$ and $\overline{\omega}_K^r(f,\delta)$ is that in the former one we consider the symmetric differences in all directions, while in the latter only in the direction of the edges of K. We shall use the fact that $\omega_K^r(f,\delta)$ is invariant under affine transformations: if Φ is an affine mapping of $\mathbf{R}^n$ onto $\mathbf{R}^n$, if F is a continuous function on $\Phi(K)$ and $f = F(\Phi)$, then

$$\omega^r_{\Phi(K)}(F,\delta) = \omega^r_K(f,\delta), \qquad \delta > 0. \tag{2.2}$$

Indeed, it follows from the definitions that

$$\Delta^r_{h\tilde{d}_{\Phi(K)}(\Phi(e),\Phi(x))\Phi(e)}(F(\Phi(x)) = \Delta^r_{h\tilde{d}_K(e,x)e} f(x) \tag{2.3}$$

for all x.

In [12, Theorem 12.1.1] it was proved that if Q is a cube then

$$E_n(f)_Q \le M\overline{\omega}_Q^r\left(f,\frac{1}{n}\right), \tag{2.4}$$

and [12, Theorem 12.1.1] claimed

$$E_n(f)_K \le M\left(\overline{\omega}_K^r\left(f,\frac{1}{n}\right) + n^{-r}\|f\|_K\right) \tag{2.5}$$

for all simple polytopes. Our first step will be to get rid of the term $n^{-r}\|f\|_K$ on the right, i.e. to prove

$$E_n(f)_K \le M\overline{\omega}_K^r\left(f,\frac{1}{n}\right) \tag{2.6}$$

for simple polytopes, and this is achieved in Section 4 with the use of (2.4), which does not contain that term.

It has already been observed in [12, Chapter 12] that one can get from any polytope a simple polytope by cutting off small pyramids around every vertex, and putting back these cut off pyramids constitutes no problem from the point of view of approximation. So the main difficulty is proving the result (Theorem 1.1) for pyramids.

For a pyramid S with apex at P let aS be the dilation of S from P by the factor a, and set $K_a = aS \setminus (a/4)S$. For small a this is a tiny simple polytope close to P, and the main idea of the proof is to show that for this simple polytope we have

$$\overline{\omega}^r_{K_a}\left(f, \frac{1}{n\sqrt{a}}\right) \le M\omega^r_S\left(f, \frac{1}{n}\right) \tag{2.7}$$

(note that on the left we have the modulus (2.1), while on the right the modulus (1.9)). This will be done in Section 6. Now an application of (2.6) gives that for $a \ge \mathrm{const}/n^2$ there are polynomials $p_{n\sqrt{a}}$ of d variables of total degree at most $n\sqrt{a}$ such that

$$\|f - p_{n\sqrt{a}}\|_{K_a} \le M\overline{\omega}^r_{K_a}\left(f, \frac{1}{n\sqrt{a}}\right) \le M\omega^r_S\left(f, \frac{1}{n}\right),$$

and here M does not depend on n or a since all K_a are similar to one another. Note that there is a huge gain here: the degree of the polynomial $p_{n\sqrt{a}}$ on the left is much smaller than n, and still we get the required rate of approximation (on the small set K_a). Once we have these local approximants in Section 7, we will patch them together (in Section 8) by something like a polynomial partition of unity. This system of polynomials will be constructed from non-symmetric fast decreasing polynomials in Section 3. This patching works because even though $p_{n\sqrt{a}}$ approximates f only on a tiny set K_a and outside that set they can blow up, this blow-up can be controlled since the degree of $p_{n\sqrt{a}}$ is small compared to n.

So basically everything boils down to proving (2.7). So why is (2.7) true? K_a has edges that are parallel with the base edges of S, as well as with the edges of S that emanate from the apex P (called apex edges). Now in the direction of these last edges K_a is much shorter ($\sim$ a-times shorter) than S, and this gives $\tilde{d}_{K_a}(e,x) \le M\sqrt{a}\tilde{d}_S(e,x)$ in these directions. So, when taking the norms of the symmetric differences in (2.1) in these directions on K_a, we automatically get the improvement stated in (2.7). However, in the direction of the base edges of S the tiny set K_a is of the same "length" as S, and in these directions the improvement needed in (2.7) is obtained via the observation that any base direction is a linear combination of two apex edge directions, and, as we have just observed, apex edge directions behave nicely. Of course, to apply this idea somehow one needs to compare smoothness in the base edge directions with smoothness in the corresponding apex edge directions; this will be done in Section 5.

Unfortunately, we cannot do exactly what was described, instead the above ideas give the necessary n-th degree approximation on a subset of S that misses a small strip (of width $\sim 1/n^2$) around the boundary of S. However, in Section 9 it will be shown that the appropriate rate of approximation on this strip around the boundary automatically follows from the approximation on the rest of S.

The proof is the same in all dimensions, but the language is simpler in $\mathbf{R}^3$, so first we prove Theorem 1.1 in $\mathbf{R}^3$, and in Section 11 we indicate the necessary changes in higher dimension.

There is a K-functional related to polynomial approximation on $[-1, 1]$, namely

$$K_{r,\varphi}(f,t) = \inf_g \left(\|f - g\|_{[-1,1]} + t\|\varphi^r g^{(r)}\|_{[-1,1]} \right). \tag{2.8}$$

[12, Theorem 2.1.1] gives that there is an absolute constant M such that

$$\frac{1}{M} K_{r,\varphi}(f, t^r) \le \omega_\varphi^r(f,t) \le M K_{r,\varphi}(f, t^r) \tag{2.9}$$

for all $0 < t \le 1$ ([12, Theorem 2.1.1] is a general result and there this is stated for $0 < t \le t_0$ with some t_0, but the proof works for all $t \le 1$, when $\varphi(x) = \sqrt{1-x^2}$; or note simply that for $t_0 \le t \le 1$ we have both $K_{r,\varphi}(f,t^r) \sim K_{r,\varphi}(f,t_0^r)$ and $\omega_\varphi^r(f,t) \sim \omega_\varphi^r(f,t_0)$). An analogue of (2.9) is not known for polytopes in $\mathbf{R}^d$, and precisely the lack of such a K-functional makes the proof of Theorem 1.1 complicated. In Section 12 we *shall prove* an analogue of (2.9) for polytopes, but that will actually be a consequence of Theorem 1.1.

Since

$$K_{r,\varphi}(f, (\lambda t)^r) \le \lambda^r K_{r,\varphi}(f, t^r), \qquad \lambda > 1, \tag{2.10}$$

we can deduce from (2.9) the inequality

$$\omega_\varphi^r(f, \lambda t) \le M \lambda^r \omega_\varphi^r(f,t), \qquad \lambda > 1. \tag{2.11}$$

Indeed, for $\lambda t \le 1$ this is a consequence of (2.10) and (2.9), and for $\lambda t \ge 1$ it follows from this case and from that fact that the modulus of continuity $\omega_\varphi^r(f,t)$ is constant on $[1,\infty)$ (see (1.12)). (2.11) can be easily transformed to any segment I (cf. (1.10)) by linear transformation, where it takes the form

$$\omega_I^r(f, \lambda t) \le M \lambda^r \omega_I^r(f,t), \qquad \lambda > 1. \tag{2.12}$$

Now this and the form

$$\omega_K^r(f,t) = \sup_{I \subset K} \omega_I^r(f,t), \tag{2.13}$$

where the supremum is taken for all chords I of K, gives

$$\omega_K^r(f, \lambda t) \le M \lambda^r \omega_K^r(f,t), \qquad \lambda > 1, \tag{2.14}$$

with an absolute constant M. In a similar manner, the modulus of smoothness (2.1) can be written as

$$\overline{\omega}_K^r(f,t) = \sup_{I \subset K} \omega_I^r(f,t), \tag{2.15}$$

where the supremum is now taken for all chords I of K that are parallel with the edges of K, hence we obtain as before

$$\overline{\omega}_K^r(f,\lambda t) \le M\lambda^r \overline{\omega}_K^r(f,t), \qquad \lambda > 1. \tag{2.16}$$

In what follows we shall encounter inequalities where numerous constants appear. Since in most cases we are not interested in the exact value of these constants, we introduce the notation

$$A \prec B \tag{2.17}$$

for $A \le CB$ with some constant C, the exact value of which is indifferent for us. Sometimes we will indicate on what parameters the constant in $\prec$ is depending on.

3 Fast decreasing polynomials

"Fast decreasing" or "pin" polynomials have been used in the past in various contexts. Their characteristic is that they decrease fast away from a given point, hence they are a sort of polynomial versions of Dirac deltas and they allow e.g. to patch local approximants together to get a global one. This is precisely how we will use them in this work; and they are one of the cornerstones of the method of this paper. Actually, we shall use their integral forms; namely fast decreasing polynomials go hand in hand with good polynomial approximants to the signum function (in the sense specified below), which can be obtained from fast decreasing polynomials by integration.

The "best" fast decreasing *symmetric* polynomials on $[-1,1]$ were found in [14]. However, symmetric polynomials are not suitable for us, therefore below we give a nonsymmetric construction that will suit our needs. We also mention that the idea of using polynomial approximants to the signum function on non-comparable intervals has already been proven useful in the theory of orthogonal polynomials with exponential weights. See [15, Theorem 7.5], where nonsymmetric fast decreasing polynomials of the sort we are going to discuss have been used. However, we shall need somewhat faster decrease than what is in [15, Theorem 7.5].

I. Symmetric fast decreasing polynomials on $[-1,1]$*.* We start from symmetric polynomials. Let Φ, $\Phi(0) \le 0$, be an arbitrary even function on $[-1,1]$ that is increasing on $[0,1]$; Φ may depend on parameters, as we shall see below. We are interested in polynomials P such that $P(0)=1$, P is even, P is decreasing and nonnegative on $[0,1]$, and $P(x) \le e^{-\Phi(x)}$, $x \in [-1,1]$ (so called "bell-shaped" polynomials). [14, Theorems 1, 2] give very precise estimates on the smallest possible degree n_Φ of such a polynomial P. In particular, if $\Phi(1/2) \ge 1$,

$\Phi^{-1}(1) \le L\Phi^{-1}(0)$ and $\Phi(1) \le L\Phi(1/2)$, then (see [14, Theorems 1, 2 and Corollary 2])

$$n_\Phi \le C_1 \int_{\Phi^{-1}(1)}^{1} \frac{\Phi(u)}{u^2} du \tag{3.1}$$

with some constant C_1 that depends only on L (here, for $v \ge 0$, we set

$$\Phi^{-1}(v) = \sup\{\tau \in [0,1] \,\big|\, \Phi(\tau) \le v\}.)$$

Let $A \ge 32$ be a constant that will be at our disposal when we use the construction below. For a natural number $n \ge 2$ and for $1/n^2 \le a \le 1/4$ we set

$$\Phi(x) = \Phi^A_{n,a}(x) = \begin{cases} 0 & \text{for } 0 \le |x| \le \sqrt{a}/16, \\ 4An\sqrt{a}\log(2Ax/\sqrt{a}) & \text{for } \sqrt{a}/16 \le |x| \le 1. \end{cases}$$

We have $\Phi^{-1}(0) = \Phi^{-1}(1) = \sqrt{a}/16$, $\Phi(1) \le 2\Phi(1/2)$, and (integrate by parts)

$$\int_{\sqrt{a}/16}^{1} \frac{4An\sqrt{a}\log(2Au/\sqrt{a})}{u^2} du \le \frac{4An\sqrt{a}\Big(1+\log(2A\frac{\sqrt{a}}{16}/\sqrt{a})\Big)}{\sqrt{a}/16} \le 64An\log A.$$

Therefore, by (3.1), there are (bell-shaped) polynomials $R_n^{(0)}$ of degree

$$\le 64C_1 An\log A \tag{3.2}$$

with some absolute constant C_1 such that $R_n^{(0)}$ is even, $R_n^{(0)}(0) = 1$, $0 \le R_n^{(0)} \le 1$ on $[-1,1]$, $R_n^{(0)}$ is increasing on $[-1,0]$ and decreasing on $[0,1]$, and

$$R_n^{(0)}(x) \le \exp\Big(-4An\sqrt{a}\log(2Ax/\sqrt{a})\Big) \qquad \text{for } \frac{\sqrt{a}}{16} \le x \le 1. \tag{3.3}$$

II. Symmetric fast decreasing polynomials on $[-2,2]$. The polynomial $R_n^{(1)}(x) = R_n^{(0)}(x/2)$ has similar properties on $[-2,2]$, except that instead of (3.3) we have

$$R_n^{(1)}(x) \le \exp\Big(-4An\sqrt{a}\log(Ax/\sqrt{a})\Big) \qquad \text{for } x \in [-2,2] \setminus \left[-\frac{\sqrt{a}}{8}, \frac{\sqrt{a}}{8}\right]. \tag{3.4}$$

III. Offsetting the peaking point. Set

$$R_n^{(2)}(x) = \frac{R_n^{(1)}\left(x - \sqrt{\frac{3a}{2}}\right) + R_n^{(1)}\left(x + \sqrt{\frac{3a}{2}}\right)}{1 + R_n^{(1)}\left(2\sqrt{\frac{3a}{2}}\right)}.$$

For this $R_n^{(2)}(\sqrt{3a/2}) = 1$, $R_n^{(2)}$ is even, and for $0 \le x \le 1$ we have

$$R_n^{(2)}(x) \le 2R_n^{(1)}\left(x - \sqrt{\frac{3a}{2}}\right) \tag{3.5}$$

Indeed, this is immediate, since, by the bell-shape form of $R_n^{(1)}$, we have for $0 \le x \le 1$

$$R_n^{(1)}\left(x+\sqrt{\frac{3a}{2}}\right) \le R_n^{(1)}\left(x-\sqrt{\frac{3a}{2}}\right).$$

IV. Nonsymmetric fast decreasing polynomials on $[0,1]$. Since $R_n^{(2)}$ is even, $R_n^{(3)}(x) = R_n^{(2)}(\sqrt{x})$ is a polynomial, $R_n^{(3)}\left(\frac{3a}{2}\right) = 1$ and (see (3.5)) $0 \le R_n^{(3)} \le 2$ for $x \in [0,1]$. If $\left|x - \frac{3a}{2}\right| \ge \frac{a}{2}$, then

a) either $x \ge 2a$, and then

$$\left|\sqrt{x} - \sqrt{\frac{3a}{2}}\right| \ge \frac{x - 3a/2}{\sqrt{x} + \sqrt{3a/2}} \ge \frac{x/4}{2\sqrt{x}} = \frac{\sqrt{x}}{8}, \tag{3.6}$$

b) or $0 \le x \le a$, and then

$$\left|\sqrt{x} - \sqrt{\frac{3a}{2}}\right| \ge \frac{|x - 3a/2|}{\sqrt{x} + \sqrt{3a/2}} \ge \frac{a/2}{2\sqrt{a}} = \frac{\sqrt{a}}{4}. \tag{3.7}$$

Hence, for $0 \le x \le a$ we get from (3.4), (3.5) and (3.7)

$$\begin{aligned} R_n^{(3)}(x) &\le 2R_n^{(1)}\left(\sqrt{x} - \sqrt{\frac{3a}{2}}\right) \le 2\exp\left(-4An\sqrt{a}\log\left(A\left|\sqrt{x} - \sqrt{\frac{3a}{2}}\right|\Big/\sqrt{a}\right)\right) \\ &\le 2\exp\Big(-4An\sqrt{a}\log(A\sqrt{a}/4\sqrt{a})\Big) \le 2\exp(-4An\sqrt{a}) \end{aligned} \tag{3.8}$$

since $A \ge 32$. On the other hand, for $2a \le x \le 1$ we get in a similar way from (3.4), (3.5) and (3.6)

$$R_n^{(3)}(x) \le 2\exp\Big(-4An\sqrt{a}\log(A\sqrt{x}/8\sqrt{a})\Big) \le 2\exp\left(-4An\sqrt{a}\log\frac{4\sqrt{x}}{\sqrt{a}}\right). \tag{3.9}$$

V. Approximation of a jump function. Let

$$\gamma_n = \int_0^1 R_n^{(3)}(u)du,$$

and first we estimate this quantity from below. In view of (3.8) and (3.9) and $An\sqrt{a} \ge 32$, we can see that

$$1 = R_n^{(3)}\left(\frac{3a}{2}\right) \le M := \max_{x\in[0,1]} R_n^{(3)}(x)$$

is attained at some $x_0 \in [a, 2a]$. Now apply Bernstein's inequality [5, Ch 4., Corollary 1.2] on $[0, 1]$ to conclude that for $t \in [a, 2a]$

$$\left| \left(R_n^{(3)} \right)' (t) \right| \le \frac{n}{\sqrt{t(1-t)}} M \le \frac{2n}{\sqrt{a}} M,$$

and so for $u \in [x_0 - \sqrt{a}/4n, x_0 + \sqrt{a}/4n] \cap [a, 2a]$ we have

$$R_n^{(3)}(u) \ge R_n^{(3)}(x_0) - \frac{2n}{\sqrt{a}} M |u - x_0| \ge M - \frac{M}{2} = \frac{M}{2}.$$

Now the interval $[x_0 - \sqrt{a}/4n, x_0 + \sqrt{a}/4n] \cap [a, 2a]$ has length $\ge \sqrt{a}/4n$, so we can conclude

$$\gamma_n \ge \frac{M}{2} \frac{\sqrt{a}}{4n} \ge \frac{\sqrt{a}}{8n}. \tag{3.10}$$

Set now

$$R_n^{(4)}(x) = R_{n,a}^{(4)}(x) = \frac{1}{\gamma_n} \int_x^1 R_n^{(3)}(t) dt, \qquad x \in [0, 1]. \tag{3.11}$$

Clearly, $0 \le R_n^{(4)}(x) \le 1$ and $R_n^{(4)}$ is decreasing on $[0, 1]$. We obtain from (3.9) and (3.10) for $2a \le x \le 1$ the inequality

$$\begin{aligned} R_n^{(4)}(x) &\le \frac{16n}{\sqrt{a}} \int_x^1 \exp\left(-4An\sqrt{a} \log \frac{4\sqrt{t}}{\sqrt{a}} \right) dt \\ &\le \frac{16n}{\sqrt{a}} a \int_{x/a}^{1/a} \exp\left(-4An\sqrt{a} \log(4\sqrt{u}) \right) du \\ &= 16n\sqrt{a} \int_{x/a}^{1/a} \left(\frac{1}{16u} \right)^{2An\sqrt{a}} du \le \frac{16n\sqrt{a}}{2An\sqrt{a} - 1} \frac{1}{(16x/a)^{2An\sqrt{a}-1}} \\ &\le \exp\left(-An\sqrt{a} \log(16x/a) \right). \end{aligned} \tag{3.12}$$

On the other hand, for $0 \le x \le a$ we get from (3.8) and (3.10)

$$\begin{aligned} 1 - R_n^{(4)}(x) &= \frac{1}{\gamma_n} \int_0^x R_n^{(3)}(t) dt \le \frac{1}{\gamma_n} \int_0^a 2e^{-4An\sqrt{a}} dt \le \frac{16n}{\sqrt{a}} a e^{-4An\sqrt{a}} \\ &= 16n\sqrt{a} \exp(-4An\sqrt{a}) \le \exp(-An\sqrt{a}). \end{aligned} \tag{3.13}$$

The polynomials $R_n^{(4)}$ of degree at most $(64C_1 An \log A)/2 + 1 \le C_{2,A} n$ are the ones that we need. Note that, by (3.12) and (3.13), they approximate on $[0, a] \cup [2a, 1]$ the function that is 1 on $[0, a]$ and 0 on $[2a, 1]$. In particular, our construction in proving the main theorem of this work will use the estimates (3.12) and (3.13).

For later use let us also mention the following, which can be easily obtained from symmetric fast decreasing polynomials (e.g. from $R_n^{(0)}$ if A, a are properly chosen) by integration as in step V:

Lemma 3.1 *If $B \geq 2$, $\theta > 0$ are given, then there is an l depending only on B and θ such that for every n there are polynomials U_n of degree at most ln for which $0 \leq U_n \leq 1$ on $[-1, 1]$, $U_n \leq \theta^n$ on $[-1, -1/2B]$ and $1 - U_n \leq \theta^n$ on $[1/2B, 1]$.*

Indeed, the existence of such an U_n also follows immediately from [14, Theorem 3].

4 Approximation on simple polytopes

In this section we prove Theorem 1.1 for simple polytopes in the sharper form when $\omega_K^r(f, 1/n)$ is replaced by the smaller quantity $\overline{\omega}_K^r(f, 1/n)$ from (2.1). Recall that a polytope $K \subset R^d$ is simple if at each vertex there are precisely d edges. For example, cubes are simple polytopes.

Thus, we want to prove

$$E_n(f)_K \prec \overline{\omega}_K^r(f, n^{-1}) \tag{4.1}$$

for $n \geq rd$, where $\overline{\omega}$ is the modulus of smoothness (2.1) taken in edge directions of K. Recall the $\prec$ notation from (2.17), so (4.1) means that $E_n(f)$ is at most a constant times $\overline{\omega}_K^r(f, n^{-1})$.

The weaker inequality

$$E_n(f)_K \prec \overline{\omega}_K^r(f, n^{-1}) + n^{-r}\|f\|_K \tag{4.2}$$

was proved in [12, Theorem 12.2.3]. In general, this additional term is bounded by the first term, but getting rid of this term is much more important than aesthetical reasons would warrant. Indeed, we shall need to apply (4.1) to some small pieces K_a of a general polytope S and to some small $m \ll n$ instead of n for which

$$\overline{\omega}_{K_a}^r(f, m^{-1}) \prec \omega_S^r(f, n^{-1}),$$

and then we shall get $\omega_S^r(f, n^{-1})$ rate of approximation on K_a by much smaller degree polynomials than n. The weaker estimate (4.2) would completely destroy this method, for then the additional factor $m^{-r}\|f\|_{K_a}$ would be much larger than $\omega_S^r(f, n^{-1})$, so the improvement given in (4.1) is absolutely necessary for the proof in this paper.

Note however, that (4.1) was proved in [12, Theorem 12.1.1] for cubes, and hence, via an affine transformation, for all d-dimensional parallelepipeds. We shall deduce (4.1) from this special case by representing K as a union of d-dimensional parallelepipeds. We shall carry out the proof only for $d = 3$, the general case is completely similar.

We may assume that K lies in the unit ball $B_1(0)$.

Fix a small $\varepsilon > 0$. Consider a vertex V of K and mark on each edge emanating from V a point which is of distance ε from V. V and the marked points generate a 3-dimensional parallelepiped T^V, all edges of which are parallel with the edges adjacent to V and all edges of T^V are of length ε. Clearly, for small $\varepsilon > 0$ this T^V is part of K, and for sufficiently small $\varepsilon > 0$ the following is also true: for every $y \in K$ there is a 3-dimensional parallelepiped $T(y) \subset K$ containing y such that $T(y)$ is a translation of one of the T^V's. Indeed, with some large but fixed M and small $\varepsilon > 0$ we do the following:

A) when y is in the $M\varepsilon$-neighborhood of a vertex V, then we use as $T(y)$ a translation of T^V,

B) when y belongs to an edge V_1V_2, then we use as $T(y)$ a translation of either T^{V_1} or T^{V_2} depending on which endpoint is closer to y, and the same is the process if y is in the $M\varepsilon$-neighborhood of the edge V_1V_2 but not in the $M\varepsilon$-neighborhood of V_1 or V_2,

C) when y belongs to a face $V_1 \dots V_l$, then we use as $T(y)$ a translation of either one of $T^{V_1}, \dots, T^{V_l}$ depending on which endpoint is closer to y, and the same is the process if y lies in the $M\varepsilon$-neighborhood of the face $V_1 \dots V_l$ but not in the $M\varepsilon$-neighborhood of either of the edges of $V_1 \dots V_l$, and, finally,

D) when y is of distance $\geq M\varepsilon$ from all faces, then we use as $T(y)$ a translation of any of the T^V's (provided M is sufficiently large).

In addition, we also require

E)
- y is not a vertex of $T(y)$ unless y is a vertex of K,
- y does not lie on an edge of $T(y)$ unless y lies on an edge of K, and
- y does not lie on a face of $T(y)$ unless y lies on a face of K.

These can be easily achieved using the preceding procedure in A)–D).

Note that if, say, y is a vertex of K, then it is necessarily a vertex of $T(y)$.

Note also that this selection of $T(y)$ was made in such a way that for small $\varepsilon > 0$ we have

F) if $\lambda T(y)$ is the dilation of $T(y)$ by a factor λ made from its center, then $4T(y) \cap K$ is still a parallelepiped (of side lengths in between ε and 4ε, since $T(y) \subset 4T(y) \cap K$).

Call a parallelepiped T a K-parallelepiped if each edge of T is parallel with an edge of K. Clearly, by property F, $4T(y) \cap K$ and $2T(y) \cap K$ are K-parallelepipeds.

We have

$$\bigcup_{y \in K} T(y) = K,$$

so by compactness (see below) finitely many of these $T(y)$ cover K, say

$$K = \bigcup_{j=1}^{k} T(y_j). \tag{4.3}$$

Indeed, it is clear that the vertices of K are covered by the $T(y)$'s where y runs through the vertices of K. Next, if E is an edge with endpoints A, B, then there is a closed segment $E' \subset E$ not containing A and B for which $E \subset T(A) \cup T(B) \cup E'$. By property E) the interiors of the edges of the $T(y)$'s for $y \in E'$ cover E', so by compactness we can select finitely many which cover E'. If we do this for all edges we get a finite covering of the edges. Apply similar reasoning (using property E)) to cover the faces, and finally all of K.

We set

$$\widetilde{T}(y) = 2T(y) \cap K. \tag{4.4}$$

In view of F) this is a K-parallelepiped of side length in between ε and 2ε. We claim that we can form a sequence $T_1, \ldots, T_{2k^2}$ such that each T_i is one of the $\widetilde{T}(y_j)$, $j = 1, 2, \ldots, k$ (repetition allowed),

$$K = \bigcup_{i=1}^{2k^2} T_i \tag{4.5}$$

and $T_i \cap T_{i+1}$ has non-empty interior for each i. Indeed, consider the graph G the vertices of which are the $\widetilde{T}(y_j)$'s, $j = 1, 2, \ldots, k$, and $\widetilde{T}(y_j)$ and $\widetilde{T}(y_l)$ are connected if their interior is non-empty. If this graph of k vertices is connected, then there is a walk in it of length $\leq 2k^2$ going through all the points (just go from a designated point to all the points in the graph along a path and back; each such path is of length at most $2(k-1)$). Thus, it is sufficient to show that G is connected. Suppose this is not the case, and let H be the union of all the $\widetilde{T}(y_j)$'s that can be reached from $\widetilde{T}(y_1)$, i.e. H is the union of a connected component G_1 of G. Then H cannot cover the whole interior of K, since then every $\widetilde{T}(y_j)$ would intersect an element of this G_1 so that the intersection has non-empty interior, in which case necessarily $\widetilde{T}(y_j) \in G_1$, and we would get $G = G_1$, i.e. the connectedness of G. Hence, if G is not connected, then H must have a boundary point Y lying in the interior of K, say $Y \in \widetilde{T}(y_{j'})$, where $\widetilde{T}(y_{j'}) \in G_1$. But Y also belongs to one of the $T(y_j)$'s, say $Y \in T(y_{j''})$ (see (4.3)), and then it is clear that Y belongs to the interior of $\widetilde{T}(y_{j''})$ (use property

E) for Y). But then $\widetilde{T}(y_{j'}) \cap \widetilde{T}(y_{j''})$ has non-empty interior, so $\widetilde{T}(y_{j''}) \in G_1$, which shows that Y is actually an interior point of

$$H = \bigcup_{\widetilde{T}(y_j) \in G_1} \widetilde{T}(y_j).$$

This contradiction proves the claim regarding the connectedness of G, and with it concerning the existence of the sequence $T_1, \ldots, T_{2k^2}$.

So each one of $T_1, \ldots, T_{2k^2}$ is a 3-dimensional K-parallelepiped with side-lengths in between ε and 2ε, their union is K, and there is a $\delta > 0$ such that each $T_i \cap T_{i+1}$ contains a ball B_i of radius δ. Now we need

Lemma 4.1 *Let $U \subset K$ be a set, $T \subset K$ a K-parallelepiped with side-lengths in between ε and 2ε such that $U \cap T$ contains a ball B of radius δ. Then there is an l that depends only on ε, δ and K for which*

$$E_{ln}(f)_{U \cup T} \le 2E_n(f)_U + 2E_n(f)_{2T \cap K}. \tag{4.6}$$

In particular, if $U = T_1 \cup \cdots \cup T_j$ and $T = T_{j+1}$, $j = 1, \ldots, 2k^2 - 1$, then l depends only on K.

Recall that $2T$ is obtained from T by a dilation of factor 2 from its center.

If for $j = 1, \ldots, 2k^2 - 1$ we set $U = T_1 \cup \cdots T_j$ and $T = T_{j+1}$, then a repeated application of the lemma gives (c.f. (4.5))

$$E_{l^{2k^2}n}(f)_K \le 2^{2k^2} \sum_{j=1}^{2k^2} E_n(f)_{2T_j \cap K}. \tag{4.7}$$

We have already mentioned that (4.1) is true for each $2T_j \cap K$ by [12, Theorem 12.1.1]. More precisely, it was proved there that if H is a cube, then for every $n \ge r$ there is a polynomial Q_n of degree at most n *in each variable* such that

$$\|f - Q_n\|_H \le M\overline{\omega}_H^r(f, n^{-1}), \tag{4.8}$$

where M depends only on r and d. This gives

$$E_{dn}(f)_H \le M\overline{\omega}_H^r(f, n^{-1}) \tag{4.9}$$

because the total degree of Q_n is at most dn. Then the same is true (via an affine transformation) for all H which is a (3-dimensional) K-parallelepiped. Here we used the affine invariance of $\overline{\omega}_H^r$ mentioned in (2.2).

It is now important to observe that $2T_j \cap K$ is a K-parallelepiped, i.e. all its edges are parallel with some edges of K, hence (4.9) yields

$$E_{dn}(f)_{2T_j \cap K} \le M\overline{\omega}_{2T_j \cap K}^r(f, n^{-1}). \tag{4.10}$$

Indeed, we assumed that $4T(y) \cap K$ is a parallelepiped, so it is a K-parallelepiped since its edges are parallel with the edges of $T(y)$ and K. Now each T_j is a $2T(y) \cap K$ (see (4.4)), so (see below)

$$2T_j \cap K = 2(2T(y) \cap K) \cap K \subseteq 4T(y) \cap K, \tag{4.11}$$

hence it is again a K-parallelepiped. In (4.11) we used the fact that if $\tilde{T} \subseteq T$ are parallelepipeds with pairwise parallel edges, then

$$2\tilde{T} \subset 2T. \tag{4.12}$$

(Note that this is not absolutely trivial because the center of dilation for T may be different from the center of dilation for $\tilde{T}$, since these parallelepipeds can have different centers.) To prove (4.12), let Φ be an affine map that maps T into $[-1,1]^3$. Then $\Phi(2T) = 2\Phi(T) = [-2,2]^3$ and $\Phi(\tilde{T})$ is a rectangular cuboid with $\Phi(2\tilde{T}) = 2\Phi(\tilde{T})$, so it is enough to prove (4.12) for $T = [-1,1]^3$. Then

$$\tilde{T} = [a_1 - b_1, a_1 + b_1] \times [a_2 - b_2, a_2 + b_2] \times [a_3 - b_3, a_3 + b_3]$$

with some $a_i \in (-1,1)$ and $b_i \in (0,1)$. The center of $\tilde{T}$ is (a_1, a_2, a_3) and

$$2\tilde{T} = [a_1 - 2b_1, a_1 + 2b_1] \times [a_2 - 2b_2, a_2 + 2b_2] \times [a_3 - 2b_3, a_3 + 2b_3].$$

So all remains to be proven is that if $a_j \pm b_j \in [-1,1]$ then $a_j \pm 2b_j \in [-2,2]$, which is clear, since e.g. $a_j + 2b_j \le 2$ if $a_j \ge 0$, while if $a_j < 0$ then $b_j \le a_j + 1$, so $a_j + 2b_j \le a_j + 2(a_j+1) = 3a_j + 2 < 2$.

Since

$$\overline{\omega}^r_{2T_j \cap K}(f,t) \le \overline{\omega}^r_K(f,t)$$

by the definition of $\overline{\omega}_K$ in (2.1) (recall that, as we have just seen, $2T_j \cap K$ is a K-parallelepiped), we obtain from (4.7) (applied with dn instead of n) and (4.10)

$$E_{l^{2k^2}dn}(f)_K \prec \overline{\omega}^r_K(f, n^{-1}) \prec \overline{\omega}^r_K(f, (l^{2k^2}dn)^{-1}),$$

where, in the last inequality, we used (2.16). In view of the monotonicity of E_n and property (2.16) of $\overline{\omega}^r_K(f,t)$, this proves (4.1) for all large n, say for $n \ge n_0$. Thus, the proof of (4.1) for $n \ge n_0$ will be complete once we verify Lemma 4.1.

Proof of Lemma 4.1. We need two additional lemmas.

Lemma 4.2 *If B is a ball of radius ρ lying in the unit ball $B_1(0)$, then for any polynomial Q_n of degree at most n*

$$\|Q_n\|_{B_1(0)} \le \|Q_n\|_B (4/\delta)^n. \tag{4.13}$$

Proof. It is known (see [5, Proposition 4.2.3]) that if q_n is a polynomial of a single variable of degree at most n, then

$$\begin{aligned} |q_n(x)| &\leq \|q_n\|_{[-1,1]}\frac{1}{2}\left\{\left(|x|+\sqrt{x^2-1}\right)^n+\left(|x|-\sqrt{x^2-1}\right)^n\right\} \\ &\leq \|q_n\|_{[-1,1]}(2|x|)^n, \qquad x\in\mathbf{R}\setminus[-1,1]. \end{aligned} \tag{4.14}$$

As a consequence, for any interval $I=[\alpha-\delta,\alpha+\delta]$ we have for $x\in\mathbf{R}\setminus I$

$$|q_n(x)|\leq\|q_n\|_I(2\cdot\mathrm{dist}(x,\alpha)/\delta)^n. \tag{4.15}$$

Now let $B=B_\delta(A)$, i.e. A is the center and δ is the radius of B. If X is any point in the unit ball, then let l be the line through A and X. The polynomial Q_n in the lemma when restricted to l, is a polynomial q_n of a single variable of degree at most n, and on this line X lies from A closer than 2. Hence, on applying (4.15) we get

$$|Q_n(X)|\leq\|Q_n\|_{B\cap l}(2\cdot\mathrm{dist}(X,A)/\delta)^n\leq\|Q_n\|_B(4/\delta)^n, \tag{4.16}$$

which proves the lemma.

■

Lemma 4.3 *Let $T\subset B_1(0)$ be a 3-dimensional K-parallelepiped such that its side-lengths lie in between some ε and 2ε and let $2T$ be its dilation from its center by a factor 2. If $\eta>0$, then there is an L depending only on η, ε and K such that for every n there is a polynomial R_n of degree at most Ln for which $0\leq R_n(x)\leq 1$ if $x\in B_1(0)$, $R_n(x)\leq\eta^n$ if $x\in B_1(0)\setminus 2T$ and $1-R_n(x)\leq\eta$ if $x\in T$.*

Proof. Let Φ be an affine transformation that maps T into $[-1,1]^3$. Then $\Phi(2T)=[-2,2]^3$, and there is an $A\geq 2$ depending only on T and ε such that $\Phi(B_1(0))\subset[-A,A]^3$ (recall that T is a K-parallelepiped, so its shape is dictated by the geometry of K).

By Lemma 3.1 (with $B=2A$, $\theta=\eta/6$, and set $Q_n(x)=U_n(x/2A)$) there is a polynomial $Q_n(x)$ of degree ln (with some l depending only on η and A) such that $0\leq Q_n(t)\leq 1$ for $t\in[-2A,2A]$, $Q_n(t)\leq\frac{1}{6}\eta^n$ if $t\in[-2A,-1/2]$, and $1-Q_n(t)\leq\frac{1}{6}\eta^n$ if $t\in[1/2,2A]$. Then, for

$$\widetilde{Q}_n(t)=Q_n(t+3/2)(1-Q_n(t-3/2)),$$

we have $0\leq\widetilde{Q}_n(t)\leq 1$ on $[-A,A]$, for $t\in[-1,1]$

$$1-\widetilde{Q}_n(t)=1-Q_n(t+3/2)+Q_n(t+3/2)Q_n(t-3/2)\leq\frac{\eta^n}{6}+\frac{\eta^n}{6}=\frac{\eta^n}{3},$$

and for $t \in [-A, A] \setminus [-2, 2]$

$$\widetilde{Q}_n(t) = Q_n(t + 3/2)(1 - Q_n(t - 3/2)) \leq \frac{\eta^n}{6}.$$

Hence, for the polynomial

$$\widetilde{R}_n(X_1, X_2, X_3) = \widetilde{Q}_n(X_1)\widetilde{Q}_n(X_2)\widetilde{Q}_n(X_3)$$

we have $0 \leq \widetilde{R}_n(X) \leq 1$ for $X \in [-A, A]^3$,

$$\widetilde{R}_n(X) \leq \eta^n, \qquad X \in [-A, A]^3 \setminus [-2, 2]^3,$$

and

$$1 - \widetilde{R}_n(X) \leq \eta^n, \qquad X \in [-1, 1]^3.$$

Since $\Phi(B_1(0)) \subset [-A, A]^3$, it is clear that then the polynomials $R_n(x) = \widetilde{R}_n(\Phi(x))$ are suitable in the lemma.

■

After these we return to the proof of Lemma 4.1. Let P_1 and P_2 be polynomials of degree n such that

$$\|f - P_1\|_{2T\cap K} \leq E_n(f)_{2T\cap K}, \qquad \|f - P_2\|_U \leq E_n(f)_U.$$

On the ball $B \subseteq U \cap T$ we have

$$\|P_1 - P_2\|_B \leq \|f - P_2\|_U + \|f - P_1\|_{2T\cap K} \leq E_n(f)_U + E_n(f)_{2T\cap K},$$

hence, by Lemma 4.2

$$\|P_1 - P_2\|_{B_1(0)} \leq (4/\delta)^n (E_n(f)_U + E_n(f)_{2T\cap K}). \tag{4.17}$$

With $\eta = \delta/4$ choose the polynomials R_n as in Lemma 4.3, and set $P = R_n P_1 + (1 - R_n)P_2$. This is a polynomial of degree at most $Ln + n$, and for it we have on $U \cap 2T = U \cap (2T \cap K)$

$$|f - P| \leq R_n|f - P_1| + (1 - R_n)|f - P_2| \leq E_n(f)_U + E_n(f)_{2T\cap K}, \tag{4.18}$$

on T

$$\begin{aligned} |f - P| &= |f - P_1 + (1 - R_n)(P_1 - P_2)| \leq |f - P_1| + (1 - R_n)|P_1 - P_2| \\ &\leq E_n(f)_{2T\cap K} + \eta^n (4/\delta)^n \big(E_n(f)_U + E_n(f)_{2T\cap K}\big) \\ &\leq 2E_n(f)_{2T\cap K} + E_n(f)_U \end{aligned} \tag{4.19}$$

(see (4.17)), and on $U \setminus 2T$

$$\begin{aligned} |f-P| &= |(f-P_2-R_n(P_1-P_2)| \leq |f-P_2| + R_n|P_1-P_2| \\ &\leq E_n(f)_U + \eta^n (4/\delta)^n \big(E_n(f)_U + E_n(f)_{2T\cap K}\big) \\ &\leq 2E_n(f)_U + E_n(f)_{2T\cap K}. \end{aligned} \tag{4.20}$$

Since $U\cap 2T$, $U\setminus 2T$ and T cover $U\cup T$, (4.18)–(4.20) verify the lemma.

■

So far we have verified (4.1) for all sufficiently large n, say $n \geq n_0$. To get (4.1) for all degree $n \geq rd$ all we need is to apply the above procedure and Lemma 4.4 below (for $n = rd, rd+1, \ldots, n_0-1$) instead of Lemma 4.1.

Lemma 4.4 *Let $U \subset K$ be a set, $T \subset K$ a K-parallelepiped with side-lengths in between ε and 2ε such that $U \cap T$ contains a ball B of radius δ. Then there is a C that depends only on ε, δ, n and K for which*

$$E_n(f)_{U\cup T} \leq C\Big(E_n(f)_U + E_n(f)_{2T\cap K}\Big). \tag{4.21}$$

In particular, if $U = T_1 \cup \cdots \cup T_j$ (see (4.5)) and $T = T_{j+1}$, $j = 1, \ldots, 2k^2-1$, then C depends only on n and K.

Proof. Let, as before, P_1 and P_2 be polynomials of degree n such that

$$\|f-P_1\|_{2T\cap K} \leq E_n(f)_{2T\cap K}, \qquad \|f-P_2\|_U \leq E_n(f)_U.$$

On the ball $B \subseteq U\cap T$ we have again

$$\|P_1-P_2\|_B \leq \|f-P_1\|_{2T\cap K} + \|f-P_2\|_U \leq E_n(f)_U + E_n(f)_{2T\cap K}.$$

Then, by Lemma 4.2,

$$\|P_1-P_2\|_{B_1(0)} \leq (4/\delta)^n (E_n(f)_U + E_n(f)_{2T\cap K}), \tag{4.22}$$

and so

$$\|f-P_1\|_{U\cup T} \leq C(E_n(f)_U + E_n(f)_{2T\cap K})$$

since

$$|f-P_1| \leq \min\left(|f-P_1|, |f-P_2|\right) + |P_1-P_2|.$$

■

5 Polynomial approximants on rhombi

We start with the following result in one variable. Let $g \in C[-1,1]$ be a continuous function on $[-1,1]$ and P_n its best approximant by polynomials of degree n. Then, according to [12, Theorem 7.3.1], for any integer $r > 0$ we have

$$\|\varphi^r P_n^{(r)}\|_{[-1,1]} \le M_r n^r \omega_\varphi^r(g, n^{-1}), \tag{5.1}$$

where $\varphi(t) = \sqrt{1-t^2}$, and

$$\omega_\varphi^r(g,\delta) = \sup_{0<h\le\delta} \|\Delta_{h\varphi(t)}^r g(t)\|_{[-1,1]} \tag{5.2}$$

is the standard φ-modulus of smoothness (1.1) of g, and the constant M_r depends only on r. If P_n is any polynomial of degree at most n, then clearly for $g = P_n$ the polynomial P_n is the best approximant among polynomials of degree at most n, hence we obtain

$$\|\varphi^r P_n^{(r)}\|_{[-1,1]} \le M_r n^r \omega_\varphi^r(P_n, n^{-1}). \tag{5.3}$$

Now if $g \in C[-1,1]$, ω is an increasing function for which $\omega_\varphi^r(g, n^{-1}) \prec \omega(n^{-1})$ and P_n is a polynomial (not necessarily best approximating g) for which $\|g - P_n\|_{[-1,1]} \prec \omega(n^{-1})$, then

$$\|\varphi^r P_n^{(r)}\|_{[-1,1]} \prec n^r \omega(n^{-1}). \tag{5.4}$$

Indeed, this is immediate from (5.3), since

$$\omega_\varphi^r(P_n, n^{-1}) \prec \|g - P_n\|_{[-1,1]} + \omega_\varphi^r(g, n^{-1}) \prec \omega(n^{-1}).$$

Recall now (1.10), that is the fact that the modulus of smoothness $\omega_\varphi^r(g,\delta)$ is the same as $\omega_{[-1,1]}^r(g,\delta)$ as defined in (1.8). We can transform the interval $[-1,1]$ by a linear map to any interval $[\overline{A}, \overline{B}]$, and via this transformation we obtain from (5.4): if $g \in C[\overline{A}, \overline{B}]$, $\omega_{[\overline{A},\overline{B}]}^r(g, n^{-1}) \prec \omega(n^{-1})$ and P_n is a polynomial for which $\|g - P_n\|_{[\overline{A},\overline{B}]} \prec \omega(n^{-1})$, then for $t \in [\overline{A}, \overline{B}]$

$$|P_n^{(r)}(t)| \prec \frac{1}{\{|t-\overline{A}| \cdot |t-\overline{B}|\}^{r/2}} n^r \omega(n^{-1}), \tag{5.5}$$

and here $\prec$ depends only on r and on the $\prec$ in the two assumptions.

Let now $T = ABCD \subset \mathbf{R}^2$ be a rhombus with diagonals AC and BD and with side-directions e_1 and e_2 (see Figure 5.1). We assume $\mathrm{dist}(A,C) = 1$. Then $\mathrm{dist}(A,B)$ and $\mathrm{dist}(B,D)$ depend only on e_1 and e_2. Let F be a continuous function on T, and suppose that there is a polynomial Q_n of two variables of degree at most n such that

$$\omega_T^r(F, n^{-1}) \prec \omega(n^{-1}), \qquad \|F - Q_n\|_T \prec \omega(n^{-1}) \tag{5.6}$$

with some increasing function ω. We are going to show

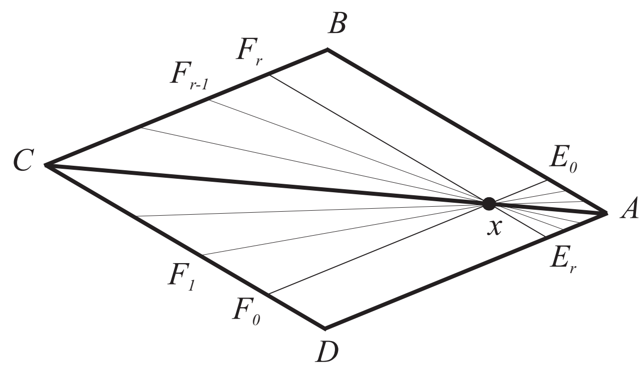

Figure 5.1:

Proposition 5.1 *Under the condition* (5.6)*, if x is a point on the diagonal AC with* $\mathrm{dist}(x, A) \leq \mathrm{dist}(A, C)/4$*, then for any direction e (in* $\mathbf{R}^2$*) we have*

$$\left| \frac{\partial^r Q_n(x)}{\partial e^r} \right| \prec \frac{1}{\mathrm{dist}(x, A)^{r/2}} n^r \omega(n^{-1}). \tag{5.7}$$

The same is true if the length of the diagonal AC is a number in the interval $[1/2, 1]$*, provided the side-directions are the fixed vectors e_1 and e_2.*

Remark. This is a fairly nontrivial estimate, since the standard Bernstein inequality [5, Ch 4., Corollary 1.2] would only give

$$\left| \frac{\partial^r Q_n(x)}{\partial e^r} \right| \prec \frac{1}{\mathrm{dist}(x, A)^{r}} n^r \omega(n^{-1})$$

for example if e is the direction of the other diagonal BD. The improvement of $1/\mathrm{dist}(x, A)^r$ to $1/\mathrm{dist}(x, A)^{r/2}$ is exactly what is needed below; this is one of the key steps in the proof of Theorem 1.1.

Proof. Let E_0F_0 and E_rF_r be the two chords of T that go through x and which are parallel with the sides AD and AB, respectively. Divide the angle E_0xE_r into r equal angles by the chords E_jF_j, $j = 1, \ldots, r-1$ of T (see Figure 5.1), and let $\overline{e}_j$ be the direction of E_jF_j. Clearly $\overline{e}_0 = e_1$ and $\overline{e}_r = e_2$ (or vice versa). It is clear that if $d = \mathrm{dist}(x, A)$, then $\mathrm{dist}(x, E_j) \geq d/2$ and

$$\mathrm{dist}(x, F_j) \geq \mathrm{dist}(x, F_0) \geq \mathrm{dist}(A, D)/2 \geq (\mathrm{dist}(A, C)/2)/2 \geq 1/4$$

(recall that $\mathrm{dist}(A, C) = 1$).

Now suppose (5.6), say

$$\omega_T^r(F, n^{-1}) \leq \omega(n^{-1}), \qquad \|F - Q_n\|_T \leq \omega(n^{-1}). \tag{5.8}$$

On applying (5.5) on the segment E_jF_j (the restriction of Q_n to that segment is a polynomial of a single variable of degree at most n) we obtain

$$\left|\frac{\partial^r Q_n(x)}{\partial \overline{e}_j^r}\right| \prec \frac{1}{d^{r/2}} n^r \omega(n^{-1}). \tag{5.9}$$

If $\overline{e}_j = \alpha_j e_1 + \beta_j e_2$, then, with the agreement $\beta_0^0 = \alpha_r^0 = 1$ (i.e. with $0^0 = 1$, since $\beta_0 = 0$ and $\alpha_r = 0$), this takes the form

$$\sum_{i=0}^{r} \binom{r}{i} \alpha_j^i \beta_j^{r-i} \frac{\partial^r Q_n(x)}{\partial e_1^i \partial e_2^{r-i}} =: \theta_j d^{-r/2} n^r \omega(n^{-1}) \tag{5.10}$$

with some $|\theta_j| \le L$, where L depends only on r and the angle in between e_1 and e_2. It is clear that $\alpha_j/\beta_j \ne \alpha_k/\beta_k$ if $j \ne k$ (otherwise $\overline{e}_j$ and $\overline{e}_k$ would point in the same direction). Note also that $\beta_0 = \alpha_r = 0$, so we can develop the determinant of the system (5.10) according to its first and last columns and we get that the determinant is

$$\left|\binom{r}{i}\alpha_j^i \beta_j^{r-i}\right|_{i,j=0}^{r} = (-1)^r(-1)^{r-1}\left(\prod_{i=1}^{r-1}\binom{r}{i}\right)\left(\prod_{j=1}^{r-1}\beta_j^r\right)\left|(\alpha_j/\beta_j)^i\right|_{i=1,j=1}^{r-1,r-1}.$$

Now this is not zero, since the last factor is $\prod_1^{r-1}(\alpha_j/\beta_j)$ times a Vandermonde-determinant with different α_j/β_j's. Note also that this determinant depends only on r, e_1 and e_2. Hence, we can solve the system of equations (5.10) for $\partial^r Q_n(x) \Big/ \partial e_1^i \partial e_2^{r-i}$, and we get

$$\left|\frac{\partial^r Q_n(x)}{\partial e_1^i \partial e_2^{r-i}}\right| \prec d^{-r/2} n^r \omega(n^{-1}),$$

where $\prec$ depends only on r, e_1 and e_2. But if e is any direction then $e = \alpha e_1 + \beta e_2$ with some $|\alpha|, |\beta| \prec 1$, and so

$$\left|\frac{\partial^r Q_n(x)}{\partial e^r}\right| = \left|\sum_{i=0}^{r}\binom{r}{i}\alpha^i\beta^{r-i}\frac{\partial^r Q_n(x)}{\partial e_1^i \partial e_2^{r-i}}\right| \prec d^{-r/2} n^r \omega(n^{-1}),$$

which is (5.7).

The last statement is an immediate consequence of the first one if we apply a dilation.

■

6 Pyramids and local moduli on them

In this section we give an estimate on the moduli of smoothness in question on small parts of a given pyramid. The estimates in this section are somewhat technical, but they form a central part of the proof.

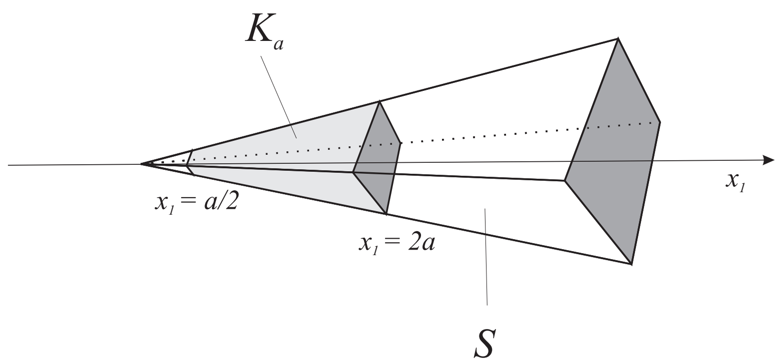

Figure 6.1: The pyramid S and the sets K_a

As we have already mentioned, first we work in $\mathbf{R}^3$. For a point $x \in \mathbf{R}^3$, x_1 will always denote the first coordinate of x, i.e. $x = (x_1, \cdot, \cdot)$.

Let Z be a convex polygon in a plane $\mathcal{L}$ lying in $\mathbf{R}^3$ and P a point outside that plane. The convex hull S of $Z \cup \{P\}$ is called a pyramid with apex at P and with base Z. P is connected by an edge to every vertex of Z, these are called the apex edges of S; while the edges of S that are also edges of Z are called base edges. The height of S is the segment that connects P with the orthogonal projection of P onto $\mathcal{L}$ (the plane of the base), i.e. it is a segment from P to a point on the plane of the base that is orthogonal to that plane.

We shall consider pyramids S that have the following two properties:

a) no two base edges of S are parallel,

b) the height of S lies in the interior of S (except for its two endpoints).

Without loss of generality we may assume that S is placed in R^3 so that its apex is at the origin 0, $S \setminus \{0\}$ lies in the half-space $x_1 > 0$ and the base of S lies in the plane $\{x \,|\, x_1 = 2\}$. Then the height of S is the segment $\{x \,|\, 0 \le x_1 \le 2\}$, i.e. the height lies on the x_1-axis. For $0 < a \le 1$ let (see Figure 6.1)

$$S_a = aS = S \cap \{x \,|\, 0 \le x_1 \le a\}, \quad K_1 = S \setminus S_{1/4}, \quad K_a = aK_1 = S_a \setminus S_{a/4}. \tag{6.1}$$

Then K_1, and hence each K_a, is a simple polytope (i.e. there are precisely 3 edges at every vertex), and K_1 has the same edge directions as S has.

Let $\mathcal{E}$ be the direction of the base edges of S, and let $e_1, \ldots, e_m$ be the direction of the apex edges, where the orientation of each e_j is such that it points from the apex 0 towards the base. Note that every base edge has endpoints on two apex edges, so every $e \in \mathcal{E}$ is a linear combination of two of the $e_1, \ldots, e_m$. The direction of the edges of K_1 (and hence of all K_a) are $\mathcal{E} \cup \{e_1, \ldots, e_m\}$, but, as we have just seen, all these directions are linear combinations of two-two members of the apex edge directions $\{e_1, \ldots, e_m\}$. In what follows, let $e \in \mathcal{E}$ be a base edge direction, and, by appropriate labeling, we may assume that $e = \alpha e_1 + \beta e_2$. Of course, then the base edge E with direction e lies in the plane $\langle E_1, E_2 \rangle$ spanned by the apex edges E_1, E_2 in the direction of e_1 and e_2 (see property a), according to which there is only one base edge E with direction e).

Recall now the definition of the distances $d_K(e, x)$ from (1.5). A crucial observation is

Proposition 6.1 *For $0 \le x_1 \le 1$, $x \in S$, we have*

$$\min\{d_S(e_1, x), d_S(e_2, x)\} \ge c_0 d_S(e, x) \tag{6.2}$$

with a c_0 depending only on S.

We remark that this proposition is false when there are two base edges parallel with e. Neither is the conclusion true for x lying close to the base, i.e. for points for which the first coordinate x_1 is close to 2.

Proof. The claim is clear when x lies on the plane $\langle E_1, E_2 \rangle$ spanned by E_1, E_2, so in what follows we assume that this is not the case.

Let S^∞ be the infinite cone with vertex at 0 determined by S (formally $S^\infty = \cup_{n=1}^\infty nS$). Since the base of S is on the plane $\{x \mid x_1 = 2\}$, we have for $x_1 \le 1$ the equalities

$$d_S(e, x) = d_{S^\infty}(e, x), \qquad d_S(e_j, x) = d_{S^\infty}(e_j, x), \quad j = 1, 2,$$

because the line through x, which is parallel with e, intersects both ∂S and ∂S^∞ in the same segment, while the line trough x that is parallel with e_j has a common intersection point Q with ∂S and ∂S^∞ in the domain $\{u \in \mathbf{R}^3 \mid 0 \le u_1 \le 1\}$ and the other intersection point of this line with ∂S is of distance $\ge \mathrm{dist}(x, Q)$ from x (this is due to the fact that $x_1 \le 1$, while the base of S lies on $\{u \mid u_1 = 2\}$). Therefore, by the homothecy-invariance of S^∞, it is enough to verify the claim for $x_1 = 1$.

Let H_x be the plane through x which is parallel with the plane $\langle E_1, E_2 \rangle$, and set $V_x^\infty = S^\infty \cap H_x$. Then

$$d_S(e, x) = d_{V_x^\infty}(e, x), \qquad d_S(e_j, x) = d_{V_x^\infty}(e_j, x), \quad j = 1, 2,$$

since the lines trough x in the direction of e, e_1, e_2 all lie in H_x. V_x^∞ is an infinite polygon with two infinite edges parallel with E_1 and E_2, respectively, and since the line of the edge E intersects E_1 and E_2, it follows that the line through x in the direction of e intersects V_x^∞ in a segment AB (see Figure 6.2). We claim that no edge of V_x^∞ is parallel with AB. Indeed, suppose to the contrary that an edge ab of V_x^∞ was parallel with AB, and hence with e. Enlarge the edge ab from the origin by a homothecy Φ so that $\Phi(b)$ becomes a point on the base plane $\{u \mid u_1 = 2\}$. Then $\Phi(ab)$ is parallel with the edge E which lies in the

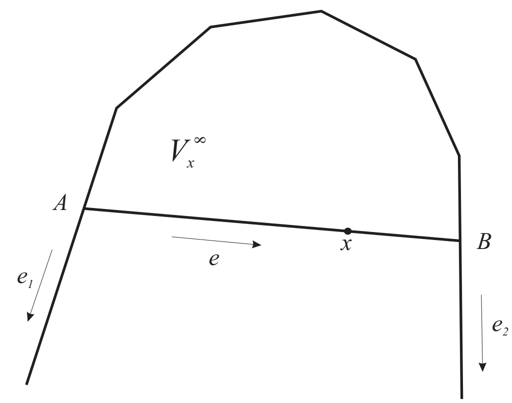

Figure 6.2:

plane $\{u \mid u_1 = 2\}$, so the whole segment $\Phi(ab)$ must lie in $\{u \mid u_1 = 2\}$. Since a and b were lying on two apex edges of S, the same is true of $\Phi(a)$ and $\Phi(b)$, and we can conclude that $\Phi(ab)$ is a base edge of S. But then S would have two parallel base edges E and $\Phi(ab)$, which is not the case by property a) above, and this contradiction proves the claim.

If $y \in S^\infty \setminus \langle E_1, E_2 \rangle$ is another point in S^∞ not lying on the face $\langle E_1, E_2 \rangle \cap S^\infty$ of S^∞, then H_y can be obtained from H_x by a dilation from the origin, and hence the same is true of V_y^∞ and V_x^∞. As a consequence, the possible (smaller) angles $\varphi_1, \ldots, \varphi_m$ in between the edges of V_x^∞ and the segment AB is independent of $x \in S^\infty$, and they are all different from 0. Hence, if ϕ_0 is the smallest of all these possible edges (see Figure 6.3), then $0 < \phi_0 < \pi_0$, and if draw the triangle ABC depicted in Figure 6.3, then this triangle lies in V_x^∞. Let d_j, $j = 1, 2$ be the length of the segment that the line through x and parallel with e_j cuts out of the triangle ABC. Then

$$d_{ABC}(e, x) \leq C \min\{d_1, d_2\} \tag{6.3}$$

is clear since any line through x cuts the triangle ABC in a segment of length

$$\geq (\sin \phi_0) \min\{\operatorname{dist}(x, A), \operatorname{dist}(x, B)\} = (\sin \phi_0) d_{ABC}(e, x).$$

Finally, (6.2) is a consequence of (6.3), since

$$d_S(e, x) = d_{ABC}(e, x), \qquad d_S(e_j, x) \geq d_j, \quad j = 1, 2.$$

■

We have already remarked that the set of edge directions of each K_η agrees with $\mathcal{E} \cup \{e_1, \ldots, e_m\}$. Let $e \in \mathcal{E}$ be a base edge direction and assume, as before,

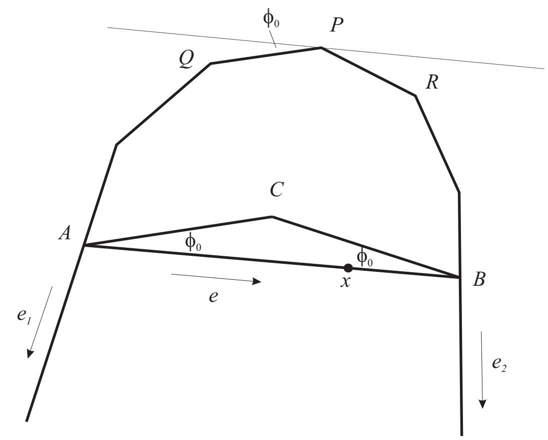

Figure 6.3:

that $e = \alpha e_1 + \beta e_2$. Let $x \in S$, $x_1 = b \le 1$. Consider again the plane H_x through x which is parallel with $\langle E_1, E_2 \rangle$ (the plane spanned by E_1 and E_2, see the previous proof), and consider (see Figure 6.4) the points $A_j = x - c_0 d_S(e, x) e_j$, $j = 1, 2$, where c_0 is the constant from the preceding proposition. According to that proposition these points belong to $V_x := S \cap H_x$, and hence, by convexity, so does their middle point $A = x - c_0(d_S(e, x)e_1 + d_S(e, x)e_2)/2$. Consider the rhombus $T = ABCD$ defined as follows (see Figure 6.4):

- A is a vertex of T,
- T contains x,
- the direction of the sides of T are e_1, e_2,
- the diagonal AC of T has length 1.

Note that x lies on this diagonal and $\mathrm{dist}(x, A) \sim d_S(e, x)$. Furthermore, $T \subset V_x^\infty$ is clear, and since A lies in the half-space $\{u \mid u_1 \le 1\}$, which is of distance 1 from the base of S, it follows that $T \subset S$. If F is a continuous function on S, then, by (4.1), there are polynomials Q_n of two variables of degree at most $n \ge 2r$ such that

$$\|F - Q_n\|_T \prec \overline{\omega}_T^r(F, n^{-1}), \tag{6.4}$$

and here $\prec$ is independent of T (by the affine invariance of $\overline{\omega}_T^r$).

Since we have $\overline{\omega}_T^r(F, \delta) \le \omega_S^r(F, \delta)$, with $\omega(\delta) = \omega_S^r(F, \delta)$ we get

$$\omega_T^r(F, n^{-1}) \prec \omega(n^{-1}), \qquad \|F - Q_n\|_T \prec \omega(n^{-1}), \tag{6.5}$$

i.e. (5.6) is satisfied. So we can apply Proposition 5.1 to conclude that if $\mathrm{dist}(x, A) \le 1/8$, then

$$\left| \frac{\partial^r Q_n(x)}{\partial e^r} \right| \prec \frac{1}{\mathrm{dist}(x, A)^{r/2}} n^r \omega(n^{-1}) \sim \frac{1}{d_S(e, x)^{r/2}} n^r \omega(n^{-1}). \tag{6.6}$$

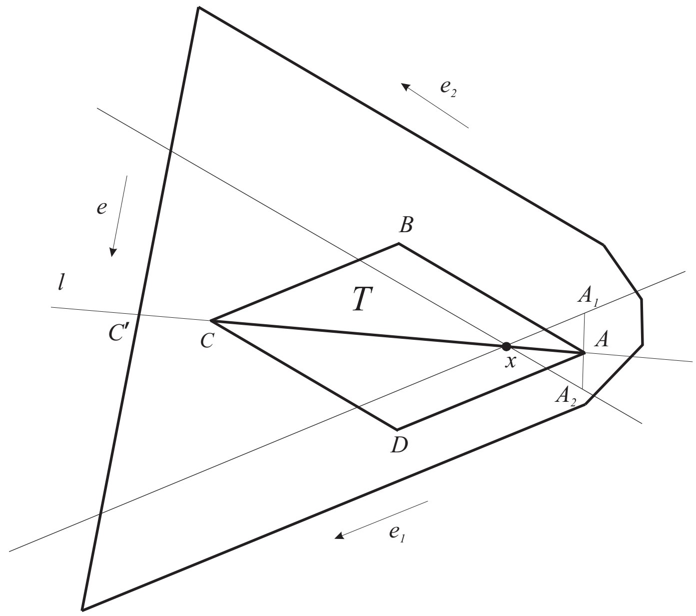

Figure 6.4:

The condition $\mathrm{dist}(x, A) \leq 1/8$ certainly holds for sufficiently small b, say for $b \leq b_0$ (recall that $x_1 = b$, and A lies in the region $\{u \,|\, 0 \leq u_1 \leq x_1\}$).

Let now $y = x + \lambda e$ with $|\lambda| \leq d_T(e, x)/2$. Then the line through y which is parallel with AC intersects T in a segment $A'C'$ of length at least $1/2$, so the rhombus T' with $A'C'$ as its diagonal and with side directions e_1 and e_2 lies within T, see Figure 6.5. It is also clear that $\mathrm{dist}(y, A') \geq \mathrm{dist}(x, A)/2$, and simple geometry shows that $\mathrm{dist}(y, A') \leq \mathrm{dist}(A', C')/4$. On applying Proposition 5.1 again, but this time to y and T', we can see that

$$\left| \frac{\partial^r Q_n(y)}{\partial e^r} \right| \prec \frac{1}{\mathrm{dist}(y, A')^{r/2}} n^r \omega(n^{-1}) \sim \frac{1}{d_S(e, x)^{r/2}} n^r \omega(n^{-1}). \tag{6.7}$$

In what follows, we shall work with r-th central differences (see (1.7)) and with the moduli of smoothness (1.8) and (2.1). We shall frequently use that, if for an $x \in S$, for some $d > 0$ and for a direction e we have $d \leq \lambda \tilde{d}_S(e, x)$ (here $\tilde{d}_S$ is the normalized distance from (1.6)), then for $0 \leq h \leq \delta/\lambda$

$$|\Delta^r_{hde} F(x)| \leq \omega^r_S(F, \delta),$$

and if e is an edge-direction of S then

$$|\Delta^r_{hde} F(x)| \leq \overline{\omega}^r_S(F, \delta).$$

Consider now the normalized distance $\tilde{d}_S(e, x)$ from (1.6). For $x_1 = b$ and for the base edge direction e we have

$$\tilde{d}_S(e, x) \prec \sqrt{b d_S(e, x)} \tag{6.8}$$

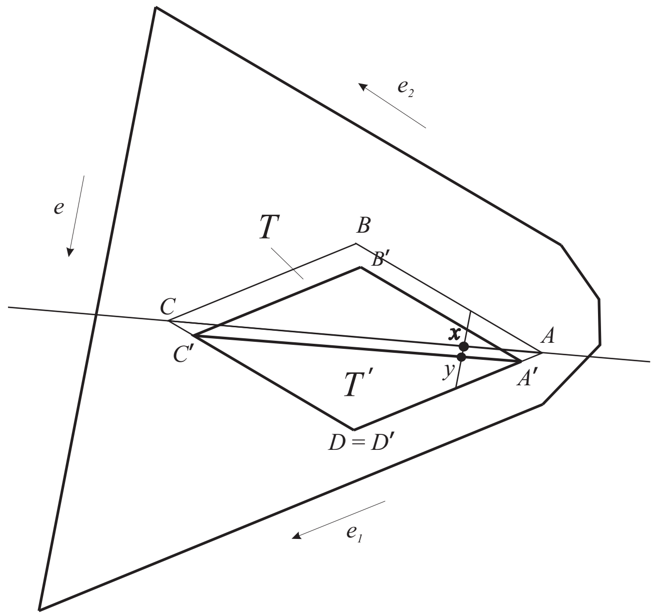

Figure 6.5:

because the $x_1 = b$ cross section of S is a homothetic copy of the base with dilation factor $b/2$, and x lies in that cross section. Hence, (6.7) implies

$$\tilde{d}_S(e,x)^r \left| \frac{\partial^r Q_n(y)}{\partial e^r} \right| \prec b^{r/2} n^r \omega(n^{-1}) \tag{6.9}$$

for all $y = x + \lambda e$ with $|\lambda| \le d_T(e,x)/2$, and note that here $d_T(e,x) \sim d_S(e,x)$, say $d_T(e,x) \ge c_1 d_S(e,x)$, by the construction of the rhombus T. Since

$$\begin{aligned} \left| \Delta^r_{h\tilde{d}_S(e,x)e} Q_n(x) \right| \quad &\prec \quad (h\tilde{d}_S(e,x))^r \times \\ &\times \max_{y \in [x - \frac{1}{2} rh\tilde{d}_S(e,x)e,\, x + \frac{1}{2} rh\tilde{d}_S(e,x)e]} \left| \frac{\partial^r Q_n(y)}{\partial e^r} \right|, \end{aligned}$$

(see [12, (2.4.5)]), it follows from (6.9) that for

$$rh\tilde{d}_S(e,x) \le c_1 d_S(e,x) \le d_T(e,x)$$

we have

$$\left| \Delta^r_{h\tilde{d}_S(e,x)e} Q_n(x) \right| \prec h^r b^{r/2} n^r \omega(n^{-1}).$$

Together with this we also get

$$\left| \Delta^r_{h\tilde{d}_S(e,x)e} F(x) \right| \prec h^r b^{r/2} n^r \omega(n^{-1}) + \omega(n^{-1})$$

because of the second relation in (6.5). Thus,

$$\sup_{h \le 1/n\sqrt{b}} \left| \Delta^r_{h\tilde{d}_S(e,x)e} F(x) \right| \prec \omega(n^{-1}) = \omega^r_S(F, n^{-1}) \tag{6.10}$$

provided

$$\frac{r\tilde{d}_S(e,x)}{n\sqrt{b}} \le c_1 d_S(e,x). \tag{6.11}$$

Because of (6.8), this latter condition is satisfied whenever

$$d_S(e,x) \ge \frac{M}{n^2} \tag{6.12}$$

with some sufficiently large, but fixed M.

All these under the assumption $x_1 = b \le b_0$, where b_0 was selected after (6.6). On the other hand, for $b \ge b_0$ (6.10) is automatic, since

$$\sup_{h\le 1/n\sqrt{b}} \left|\Delta^r_{h\tilde{d}_S(e,x)e}F(x)\right| \le \omega^r_S\left(F, \frac{1}{n\sqrt{b_0}}\right) \prec \omega^r_S\left(F, \frac{1}{n}\right)$$

because of (2.14). Thus, (6.10) is true for all $x \in S$ with $0 \le x_1 \le 1$ for which (6.12) holds.

Let now f be a continuous function on the pyramid S. Let $v_n = (-L/n^2, 0, 0)$ with some large fixed L, and apply what we have obtained to the pyramid

$$\mathbf{S}^{(n)} = S + v_n \quad \text{and to the function} \quad F(x) = F_n(x) = f(x - v_n) \text{ on } \mathbf{S}^{(n)}. \tag{6.13}$$

Since we are translating S in the direction of its height and towards the apex, it follows that $\mathbf{S}^{(n)}$ contains $S/2$ (at least for $n^2 > 2L$), and for any $x \in S/2$ its distance from the boundary of $\mathbf{S}^{(n)}$ is at least M/n^2 if L is sufficiently large, where M is the constant in (6.12). In particular, for any base edge direction e

$$d_{\mathbf{S}^{(n)}}(e,x) \ge \frac{M}{n^2}, \qquad x \in S/2, \tag{6.14}$$

i.e. the condition (6.12) is satisfied for $\mathbf{S}^{(n)}$ (i.e. $d_{\mathbf{S}^{(n)}}(e,x) \ge M/n^2$) provided $x \in S$, $x_1 = b \le 1/2$. Hence, by (6.10),

$$\sup_{h\le 1/n\sqrt{b}} \left|\Delta^r_{h\tilde{d}_{\mathbf{S}^{(n)}}(e,x)e}F(x)\right| \prec \omega^r_{\mathbf{S}^{(n)}}(F, n^{-1})$$

is true. But here

$$\omega_{\mathbf{S}^{(n)}}(F, n^{-1}) = \omega_S(f, n^{-1}) \quad \text{and} \quad \tilde{d}_S(e,x) \le \tilde{d}_{\mathbf{S}^{(n)}}(e,x),$$

and so

$$\sup_{h\le 1/n\sqrt{b}} \left|\Delta^r_{h\tilde{d}_S(e,x)e}F(x)\right| \prec \omega^r_S(f, n^{-1}) \tag{6.15}$$

also follows. This is true for all base direction $e \in \mathcal{E}$ and for all $x \in S$ with $x_1 \le 1/2$. Note also that for such edges $\tilde{d}_S(e,x) = \tilde{d}_{K_a}(e,x)$ provided $x \in K_a$ (see (6.1) for the definition of K_a). Therefore, (6.15) gives for $x \in K_a$, $a \le 1/4$

$$\sup_{h\le 1/n\sqrt{2a}} \left|\Delta^r_{h\tilde{d}_{K_a}(e,x)e}F(x)\right| \prec \omega^r_S(f, n^{-1}) \tag{6.16}$$

since for $x \in K_a$ we have $b = x_1 \le 2a$ in (6.15).

On the other hand, if e_j is an apex edge direction, then $\tilde{d}_S(e_j, x) \sim \sqrt{d_S(e_j, x)}$, while

$$\tilde{d}_{K_a}(e_j, x) \prec \sqrt{a d_{K_a}(e_j, x)} \prec \sqrt{a d_S(e_j, x)}.$$

As a consequence, $\tilde{d}_{K_a}(e_j, x) \le C_1 \sqrt{a} \tilde{d}_S(e_j, x)$ with some C_1. But then

$$\sup_{h \le 1/2nC_1\sqrt{a}} \left| \Delta^r_{h\tilde{d}_{K_a}(e_j,x)e_j} F(x) \right| \le \sup_{h \le 1/2n} \left| \Delta^r_{h\tilde{d}_S(e_j,x)e_j} F(x) \right| \le \omega^r_S(f, n^{-1}). \tag{6.17}$$

The last inequality is due to the fact that

$$\Delta^r_{h\tilde{d}_S(e_j,x)e_j} F(x) = \Delta^r_{h\tilde{d}_S(e_j,x)e_j} f(x - v_n),$$

and $\tilde{d}_S(e_j, x) \le 2\tilde{d}_S(e_j, x - v_n)$ since $d_S(e_j, x) \le d_S(e_j, x - v_n)$ (see also (6.13)).

If we take the supremum on the left of (6.16)–(6.17) for all $x \in K_a$ and the maximum for all $e \in \mathcal{E}$ (base edge directions) and for all e_j, $1 \le j \le m$ (apex edge directions), (6.16) and (6.17) yield (we may assume $C_1 \ge 1$)

$$\overline{\omega}^r_{K_a}\left(F, \frac{1}{2nC_1\sqrt{a}}\right) \prec \omega^r_S\left(f, \frac{1}{n}\right),$$

where the modulus of smoothness on the left-hand side is the one from (2.1), i.e. it is created via r-th differences in the edge directions of K_a and the modulus of smoothness on the right is the one from (1.8), i.e. it is created via r-th differences in all directions. On applying (2.16) we finally obtain for $a \le 1/4$

$$\overline{\omega}^r_{K_a}\left(F, \frac{1}{n\sqrt{a}}\right) \prec \omega^r_S\left(f, \frac{1}{n}\right), \tag{6.18}$$

which is the main result of this section.

7 Local approximation on the sets K_a

With the notations of the preceding section, let f be a continuous function on the pyramid S, $v_n = (-L/n^2, 0, 0)$ with the large but fixed L for which (6.14) is true, and

$$F(x) = F_n(x) = f(x - v_n) \quad \text{on} \quad \mathbf{S}^{(n)} = S + v_n. \tag{7.1}$$

We have seen that for $a \le 1/4$ the inequality (6.18) holds. Recall also the notations K_1, S_η, K_η from (6.1) which are constructed from the original pyramid S.

Set now $F^*(x^*) = F(ax^*)$. Then F^* is a continuous function on K_1 such that

$$\omega^r_{K_a}(F, \delta) \equiv \omega^r_{K_1}(F^*, \delta), \qquad \overline{\omega}^r_{K_a}(F, \delta) \equiv \overline{\omega}^r_{K_1}(F^*, \delta),$$

because both ω_K^r and $\overline{\omega}_K^r$ are invariant under homothetic transformations. Since K_1 is a simple polytope (at each vertex there are 3 edges), we can apply (4.1) from Section 4 to conclude that for $n\sqrt{a} \geq 3r$, i.e. for $a \geq 9r^2/n^2$, there are polynomials $P^*_{n\sqrt{a}} = P^*_{a,n\sqrt{a}}$ of 3 variables of degree at most $n\sqrt{a}$ such that

$$\|F^* - P^*_{n\sqrt{a}}\|_{K_1} \prec \overline{\omega}_{K_1}^r\left(F^*, \frac{1}{n\sqrt{a}}\right) = \overline{\omega}_{K_a}^r\left(F, \frac{1}{n\sqrt{a}}\right) \prec \omega_S^r\left(f, \frac{1}{n}\right),$$

where, in the last step, we used (6.18), and here $\prec$ is independent of n and $a \geq 9r^2/n^2$. With

$$p_{n\sqrt{a}}(x) = p_{a,n\sqrt{a}}(x) = P^*_{n\sqrt{a}}(x/a)$$

this is the same as

$$\|F - p_{n\sqrt{a}}\|_{K_a} \prec \omega_S^r\left(f, \frac{1}{n}\right). \tag{7.2}$$

For the polynomials

$$q = p_{n\sqrt{a}} - p_{n\sqrt{2a}} = p_{a,n\sqrt{a}} - p_{2a,n\sqrt{2a}}$$

of degree at most $n\sqrt{2a}$ this yields

$$\|q\|_{K_a\cap K_{2a}} \leq \|F - p_{n\sqrt{a}}\|_{K_a} + \|F - p_{n\sqrt{2a}}\|_{K_{2a}} \prec \omega_S^r\left(f, \frac{1}{n}\right),$$

and here

$$K_a \cap K_{2a} = S \cap \{x \,|\, a \leq x_1 \leq 2a\}.$$

Let now $x \in S$ not lying in $K_a \cap K_{2a}$, i.e. either $2a \leq x_1 \leq 2$ or $0 \leq x_1 \leq a$. Let ℓ be the line through 0 and x, and for $t \in \ell$ let t_1 be its first coordinate. Then $\tilde{q}(t_1) = q(t)$, $t \in \ell$, is a polynomial of degree at most $n\sqrt{2a}$ in the variable t_1, for which, as we have just seen,

$$\|\tilde{q}\|_{[a,2a]} \leq \|q\|_{K_a\cap K_{2a}} \prec \omega_S^r\left(f, \frac{1}{n}\right).$$

Then (4.15) gives

$$|q(x)| = |\tilde{q}(x_1)| \leq \left(\frac{|x_1 - 3a/2|}{a/2}\right)^{n\sqrt{2a}} \|\tilde{q}\|_{[a,2a]} \prec \left(\frac{|x_1 - 3a/2|}{a/2}\right)^{n\sqrt{2a}} \omega_S^r\left(f, \frac{1}{n}\right). \tag{7.3}$$

For $0 \leq x_1 \leq a$ this yields

$$|p_{n\sqrt{a}}(x) - p_{n\sqrt{2a}}(x)| \prec 3^{n\sqrt{2a}}\omega_S^r\left(f, \frac{1}{n}\right) \prec e^{3n\sqrt{a}}\omega_S^r\left(f, \frac{1}{n}\right), \tag{7.4}$$

while for $2a \leq x_1 \leq 2$ we obtain

$$|p_{n\sqrt{a}}(x) - p_{n\sqrt{2a}}(x)| \prec \left(\frac{8x_1}{a}\right)^{n\sqrt{2a}} \omega_S^r\left(f, \frac{1}{n}\right) \prec e^{2n\sqrt{a}\log(8x_1/a)}\omega_S^r\left(f, \frac{1}{n}\right). \tag{7.5}$$

All these for $a \leq 1/8$ and $n\sqrt{a} \geq 3r$, i.e. $a \geq 9r^2/n^2$.

8 Global approximation of $F = F_n$ on $S_{1/32}$ excluding a neighborhood of the apex

We use the preceding estimates (7.4) and (7.5) with $a = a_k = 2^k/n^2$, $k = 9r, \dots, m$, where m is chosen so that $1/16 \le 2^m/n^2 < 1/8$. Then $a_{9r-1} \ge 9r^2/n^2$. We combine the polynomials $p_{n\sqrt{a}} = p_{a,n\sqrt{a}}$ with the fast decreasing polynomials

$$R_{n,a}(x) := R^{(4)}_{n,a}(x_1) \tag{8.1}$$

where $R^{(4)}_{n,a}(x_1)$ is the polynomial of the single variable x_1 (the first coordinate of x) from (3.11) with $A = 4$ (recall that in Section 3 the parameter A was a free parameter), and set

$$P_n = \sum_{k=9r}^{m} \left(R_{n,a_k} - R_{n,a_{k-1}}\right) p_{n\sqrt{a}_k} + R_{n,a_{9r-1}} p_{n\sqrt{a_{9r}}} + (1 - R_{n,a_m}) p_{n\sqrt{a_m}}. \tag{8.2}$$

This is a polynomial of degree at most Cn with some universal constant C. We claim that this approximates $F(x) = F_n(x) = f(x - v_n)$ (see (7.1)) in the order $\omega^r_S(f, 1/n)$ on the set

$$S^*_n := S \cap \left\{ x \,\middle|\, \frac{2^{9r+1}}{n^2} \le x_1 \le \frac{1}{16} \right\}. \tag{8.3}$$

First of all, we have

$$\sum_{k=9r}^{m} \left(R_{n,a_k} - R_{n,a_{k-1}}\right) + R_{n,a_{9r-1}} + (1 - R_{n,a_m}) = 1,$$

hence

$$\begin{aligned} P_n - F &= \sum_{k=9r}^{m} \left(R_{n,a_k} - R_{n,a_{k-1}}\right)(p_{n\sqrt{a_k}} - F) \\ &+ R_{n,a_{9r-1}}(p_{n\sqrt{a_{9r}}} - F) + (1 - R_{n,a_m})(p_{n\sqrt{a_m}} - F). \end{aligned} \tag{8.4}$$

For $x \in \{x \,|\, a_{k_0} \le x_1 \le a_{k_0+1}\}$ with $9r + 1 \le k_0 \le m - 1$ the first sum on the right-hand side can be written in the form

$$\begin{aligned} &\sum_{k=9r}^{k_0-1} R_{n,a_k}(p_{n\sqrt{a_k}} - p_{n\sqrt{a_{k+1}}}) + R_{n,a_{k_0}}(p_{n\sqrt{a_{k_0}}} - F) - R_{n,a_{9r-1}}(p_{n\sqrt{a_{9r}}} - F) \\ &+ \sum_{k=k_0+1}^{m-1} (R_{n,a_k} - 1)(p_{n\sqrt{a_k}} - p_{n\sqrt{a_{k+1}}}) + (R_{n,a_m} - 1)(p_{n\sqrt{a_m}} - F) \\ &-(R_{n,a_{k_0}} - 1)(p_{n\sqrt{a_{k_0+1}}} - F) =: A_1 + A_2 - A_3 + A_4 + A_5 - A_6. \end{aligned} \tag{8.5}$$

Here $-A_3$ cancels the second term, while A_5 cancels the third term on the right of (8.4). Since

$$\{x \,|\, a_{k_0} \le x_1 \le a_{k_0+1}\} \cap S \subseteq K_{a_{k_0}+1},\ K_{a_{k_0}},$$

we obtain from (7.2)

$$|A_2| + |A_6| \prec \omega_S^r\left(f, \frac{1}{n}\right). \tag{8.6}$$

Now for $9r+1 \le k \le k_0 - 1$ the estimate (7.5) gives

$$|p_{n\sqrt{a_k}}(x) - p_{n\sqrt{2a_k}}(x)| \prec e^{2n\sqrt{a_k}\log(8x_1/a_k)}\omega_S^r\left(f, \frac{1}{n}\right),$$

while, by (3.12),

$$R_{n,a_k}(x) = R_n^{(4)}(x_1) \le e^{-4n\sqrt{a_k}\log(16x_1/a_k)}.$$

So

$$|A_1| \le \sum_{k=9r+1}^{k_0-1} e^{-2n\sqrt{a_k}\log(16x_1/a_k)}\omega_S^r\left(f, \frac{1}{n}\right) \prec \left(\sum_{k=9r+1}^{k_0-1} e^{-\sqrt{2^k}}\right)\omega_S^r\left(f, \frac{1}{n}\right). \tag{8.7}$$

In a similar manner, for $k_0 + 1 \le k \le m-1$ we get from (7.4) and (3.13)

$$|p_{n\sqrt{a_k}}(x) - p_{n\sqrt{2a_k}}(x)| \prec e^{3n\sqrt{a_k}}\omega_S^r\left(f, \frac{1}{n}\right)$$

and

$$0 \le 1 - R_{n,a_k}(x) = 1 - R_n^{(4)}(x_1) \le e^{-4n\sqrt{a_k}},$$

so

$$|A_4| \prec \sum_{k=k_0+1}^{m-1} e^{-n\sqrt{a_k}}\omega_S^r\left(f, \frac{1}{n}\right) = \left(\sum_{k=k_0+1}^{m-1} e^{-\sqrt{2^k}}\right)\omega_S^r\left(f, \frac{1}{n}\right). \tag{8.8}$$

Collecting the estimates from (8.4) to (8.8) we can see that for $x \in \{x \,|\, a_{k_0} \le x_1 \le a_{k_0+1}\} \cap S$

$$|P_n - F| \prec \left(\sum_{k=0}^{\infty} e^{-\sqrt{2^k}}\right)\omega_S^r\left(f, \frac{1}{n}\right) \prec \omega_S^r\left(f, \frac{1}{n}\right).$$

Since every point of S_n^* (see (8.3)) belongs to one of the sets $x \in \{x \,|\, a_{k_0} \le x_1 \le a_{k_0+1}\} \cap S$, $9r+1 \le k_0 \le m-1$, we can finally conclude

$$\|P_n - F\|_{S_n^*} \prec \omega_S^r\left(f, \frac{1}{n}\right). \tag{8.9}$$

This argument works for all large n, actually for all $n^2 > 2^{9r+6} + 2L$, where $L \ge 1$ is the number for which (6.14) is true (cf. the discussion before (6.14)).

9 Global approximation of f on $S_{1/64}$

We saw in (8.9) that

$$\|P_n - F_n\|_{S_n^*} \prec \omega_S^r\left(f, \frac{1}{n}\right),$$

where $F_n(x) = f(x - v_n)$, $v_n = (-L/n^2, 0, 0)$ and S_n^* is the set in (8.3). For the polynomial $p_n(x) = P_n(x + v_n)$ this gives

$$\|p_n - f\|_{S_n^* - v_n} \prec \omega_S^r\left(f, \frac{1}{n}\right). \tag{9.1}$$

This estimate is valid for sufficiently large n; actually, as we have mentioned at the end of the preceding section, for all $n^2 > 2^{9r+6} + 2L$.

Note that $n^2 > 2L$ also holds, so $S_n^* - v_n$ is a subset of S, and it covers a substantial part of $S_{1/32} = S/32$, the only points in $S/32$ that do not lie in $S_n^* - v_n$ are the points lying close to the boundary $\partial S \cap (S/32)$. Based on this, we are going to show that (9.1) automatically implies

$$\|p_n - f\|_{S/64} \prec \omega_S^r\left(f, \frac{1}{n}\right) \tag{9.2}$$

for sufficiently large n, say for $n \geq n_S$, depending on S (actually, $n^2 > 64(2^{9r+1} + L)$ will do).

Indeed, let $y \in (S/64) \setminus (S_n^* - v_n)$, see Figure 9.1. Then $y \in S/64$, and either

$$0 \leq y_1 \leq \frac{2^{9r+1}}{n^2} + \frac{L}{n^2}$$

(which is the case when $y + v_n \in S \setminus S_n^*$, see (8.3)) or the segment $[y + v_n, y]$ intersects the boundary of S. Let AB be the chord of $S/32$ which is parallel with the x_1-axis and which goes through the point y; i.e. $y \in AB$, $A \in \partial S$, the first coordinate of B is $1/16$ and AB is parallel with the x_1-axis. Let C be the point of intersection of the chord AB with $\partial(S_n^* - v_n)$, i.e. $CB = AB \cap (S_n^* - v_n)$. Then $\mathrm{dist}(A, B) \geq 1/32$ (this is due to the fact that the x_1-coordinate of A is at most $1/32$ since $y \in S_{1/64} = S/64$, while the x_1-coordinate of B is $1/16$) and

$$\mathrm{dist}(A, C) \leq \frac{2^{9r+1}}{n^2} + \frac{L}{n^2} < \frac{1}{64}$$

for $n^2 > 64(2^{9r+1} + L)$. Hence, for such n we have $\mathrm{dist}(C, B) \geq 1/64$.

By (9.1)

$$\|f - p_n\|_{[C,B]} \prec \omega_S^r\left(f, \frac{1}{n}\right),$$

and clearly we also have

$$\omega_{[A,B]}^r\left(f, \frac{1}{n}\right) \leq \omega_S^r\left(f, \frac{1}{n}\right).$$

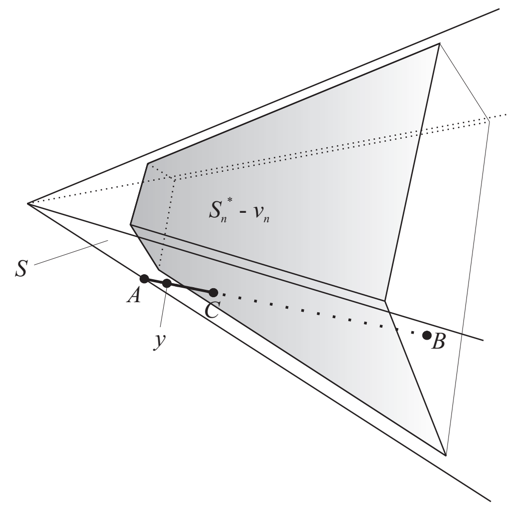

Figure 9.1: The set $S_n^* - v_n$ and the position of y, A, B, C

Therefore,

$$|p_n(y) - f(y)| \prec \omega_S^r(f, 1/n),$$

follows from the following proposition if we make a linear transformation from $[A, B]$ onto $[0, 1]$ (under this transformation $\omega_{[A,B]}^r(f, \delta)$ becomes $\omega_{[0,1]}^r(g, \delta)$ — with the same δ! — if $f\Big|_{[A,B]}$ becomes $g \in C[0, 1]$). Since here $y \in (S/64) \setminus (S_n^* - v_n)$ is arbitrary, this will prove (9.2).

Proposition 9.1 *Let $g \in C[0, 1]$ and let $\Lambda > 0$ be fixed. Then for any polynomial q_n of a single variable and of degree at most n with $n^2 \geq 2\Lambda$, we have*

$$\|g - q_n\|_{[0,1]} \leq C\left(\|g - q_n\|_{[\Lambda/n^2,1]} + \omega_{[0,1]}^r(g, 1/n)\right), \tag{9.3}$$

where C depends only on Λ and r.

Proof. With $\psi(x) = \sqrt{x(1-x)}$ let

$$\omega_\psi^r(g, \delta) = \sup_{h \leq \delta} \|\Delta_{h\psi(x)}^r g(x)\|$$

be the standard φ-modulus of smoothness (see [12]). This is the same as (1.1) but for the interval $[0, 1]$ rather than for $[-1, 1]$. Exactly as in (1.10) the two moduli $\omega_{[0,1]}^r(g, \delta)$ and $\omega_\psi^r(g, \delta)$ are the same:

$$\omega_{[0,1]}^r(g, \delta) = \omega_\psi^r(g, \delta),$$

hence (9.3) is equivalent to

$$\|g - q_n\|_{[0,1]} \leq C\left(\|g - q_n\|_{[\Lambda/n^2,1]} + \omega_\psi^r(g, 1/n)\right). \tag{9.4}$$

In proving this we may assume $\Lambda \geq 8$ (and, as assumed, $n^2 > 2\Lambda$). Let

$$x^* = \frac{x - \Lambda/n^2}{1 - \Lambda/n^2}, \qquad q_n^*(x^*) := q(x), \quad g^*(x^*) := g(x).$$

As x runs through $[\Lambda/n^2, 1]$, the point x^* runs through the interval $[0, 1]$. For $x \in [\Lambda/n^2, 1]$

$$\psi(x^*) = \sqrt{\frac{x - \Lambda/n^2}{1 - \Lambda/n^2} \frac{1 - x}{1 - \Lambda/n^2}} \leq \frac{\sqrt{x(1-x)}}{1 - \Lambda/n^2} \leq 2\sqrt{x(1-x)},$$

and hence for $0 \leq h \leq 1/n$

$$\begin{aligned} \left|\Delta^r_{h\psi(x^*)} q_n^*(x^*)\right| &\leq 2^r \|g^* - q_n^*\|_{[0,1]} + \left|\Delta^r_{h\psi(x^*)} g^*(x^*)\right| \\ &\leq 2^r \|g - q_n\|_{[\Lambda/n^2,1]} + \omega^r_\psi(g, 2/n) \prec \theta_n, \end{aligned}$$

where

$$\theta_n := \|g - q_n\|_{[\Lambda/n^2,1]} + \omega^r_\psi(g, 1/n). \tag{9.5}$$

This shows that

$$\omega^r_\psi(q_n^*, 1/n) \prec \theta_n,$$

and then (5.3) (more precisely its variant for the interval $[0, 1]$) gives

$$\|\psi^r (q_n^*)^{(r)}\|_{[0,1]} \leq M_r n^r \omega^r_\psi(q_n^*, n^{-1}) \prec n^r \theta_n, \tag{9.6}$$

and so

$$\|(q_n^*)^{(r)}\|_{[1/n^2,1-1/n^2]} \prec n^{2r}\theta_n \tag{9.7}$$

(since $\psi(x^*) \geq 1/2n$ for $x^* \in [1/n^2, 1 - 1/n^2]$). It also follows from (9.6) that

$$\|\psi^r (q_n)^{(r)}\|_{[(\Lambda+1)/n^2,1]} \prec n^r \theta_n, \tag{9.8}$$

since for $x \in [(\Lambda + 1)/n^2, 1]$ we have $\psi(x) \sim \psi(x^*)$. Use now Remez' inequality [16] (cf. (14.24) in Section 14) or the first inequality in (4.14) to conclude from (9.7)

$$\|(q_n^*)^{(r)}\|_{[-\Lambda/n^2,1]} \prec n^{2r}\theta_n,$$

i.e.

$$\|q_n^{(r)}\|_{[0,1]} \prec n^{2r}\theta_n.$$

This gives

$$\|\psi^r q_n^{(r)}\|_{[0,(\Lambda+1)/n^2]} \prec n^r \theta_n,$$

then (see (9.8))

$$\|\psi^r q_n^{(r)}\|_{[0,1]} \prec n^r \theta_n,$$

and finally (see [12, (2.4.4)] or apply (2.9))

$$\omega_\psi^r(q_n, 1/n) \prec \theta_n.$$

Together with this

$$\omega_\psi^r(g-q_n, 1/n) \prec \theta_n$$

also follows (see the definition of θ_n in (9.5)), and then

$$\omega_\psi^r(g-q_n, \Lambda/n) \prec \theta_n \tag{9.9}$$

is an immediate consequence (see (2.14)).

For an $x \in [0, \Lambda/n^2]$ let now the point $y \in [0,1]$ be defined by

$$y - \frac{r}{2}\frac{\Lambda}{n}\psi(y) = x.$$

Since

$$\frac{\Lambda}{n^2} - \frac{r}{2}\frac{\Lambda}{n}\sqrt{\frac{\Lambda}{n^2}\left(1-\frac{\Lambda}{n^2}\right)} < 0$$

for $\Lambda \geq 8$, we have $y \geq \Lambda/n^2$, and then $\psi(y) \geq 1/n$, $(\Lambda/n)\psi(y) > \Lambda/n^2$. Hence, for $j \geq 1$ we have

$$y - \left(\frac{r}{2} - j\right)\frac{\Lambda}{n}\psi(y) \geq x + \frac{\Lambda}{n}\psi(y) \geq x + \frac{\Lambda}{n^2} \geq \frac{\Lambda}{n^2}.$$

As a consequence, for $j = 1, 2, \ldots, r$

$$\left|\Big(g-q_n\Big)\big(y-(r/2-j)(\Lambda/n)\psi(y)\big)\right| \leq \|g-q_n\|_{[\Lambda/n^2,1]}.$$

Therefore,

$$\begin{aligned}
&|g(x)-q_n(x)| = \\
&\quad\left|\Delta^r_{(\Lambda/n)\psi(y)}(g-q_n)(y) - \sum_{j=1}^{r}(-1)^{r+j}\binom{r}{j}\Big(g-q_n\Big)\big(y-(r/2-j)(\Lambda/n)\psi(y)\big)\right| \\
&\qquad \leq \omega_\psi^r(g-q_n, \Lambda/n) + 2^r\|g-q_n\|_{[\Lambda/n^2,1]} \prec \theta_n,
\end{aligned}$$

where, in the last step we used (9.9) and the definition of θ_n in (9.5). This proves (9.4).

∎

The preceding proof for (9.2) covers all n with $n \geq n_S$. We still need to prove (9.2) for $3r \leq n \leq n_S$. First of all, we mention that the following variant of Proposition 9.1 is the $\Lambda = n^2/2$ special case of the proposition itself.

Proposition 9.2 *Let $g \in C[0,1]$. Then for any polynomial q_n of a single variable and of degree at most $n = 1, 2, \dots$ we have*

$$\|g - q_n\|_{[0,1]} \le C\left(\|g - q_n\|_{[1/2,1]} + \omega^r_{[0,1]}(g, 1/n)\right), \tag{9.10}$$

where C depends only on n and r.

Now for $3r \le n \le n_S$ we can do the following: fix a cube Q inside S. For this we know

$$\|p_n - f\|_Q \prec \omega^r_S\left(f, \frac{1}{n}\right) \tag{9.11}$$

with some polynomials p_n of degree at most n (this is weaker than (4.1), and, as we have already mentioned, it was proved in [12, Theorem 12.1.1]; recall also that $\omega^r_Q(f,\delta) \le \omega^r_S(f,\delta)$). Application of Proposition 9.2 (more precisely its scaled version as was discussed before Proposition 9.1) along lines going through the center of Q gives an estimate similar to (9.11) but on $2Q$ (obtained by enlarging Q from its center):

$$\|p_n - f\|_{(2Q)\cap S} \prec \omega^r_S\left(f, \frac{1}{n}\right).$$

Repeating this process we obtain

$$\|p_n - f\|_{(2^kQ)\cap S} \prec \omega^r_S\left(f, \frac{1}{n}\right),$$

where $\prec$ depends on S, Q and k. If k is such that 2^kQ covers S, then (9.2) follows.

Thus, (9.2) has been verified for all $n \ge 3r = 3d$.

10 Completion of the proof of Theorem 1.1

Let $K \subset \mathbf{R}^3$ be an arbitrary convex polytope with vertices $V_1, \dots, V_m$. We cut off from K a small pyramid with apex V_j by a plane σ_j in the following way. Let Σ_j be a supporting plane to K at V_j and let ℓ_j be the line perpendicular to Σ_j and going through V_j. Let σ_j be a plane parallel with Σ_j intersecting K. If σ_j lies sufficiently close to V_j, then σ_j intersects all the edges of K emanating from V_j. Consider the pyramid S_j spanned by V_j (as apex) and $\sigma_j \cap S$ (as base). It is clear that

$$\omega^r_{S_j}(f,\delta) \le \omega^r_K(f,\delta), \qquad \delta > 0.$$

We want to apply the approximation result (9.2) proven in Section 9 to S_j, but the whole consideration for pyramids used the two assumptions that

a) no two base edges of S_j are parallel,

b) the height of S_j lies in the interior of S_j (except for its two endpoints).

We claim that Σ_j, σ_j can be chosen so that these two properties are satisfied. Indeed, any base edge E of S_j lies in σ_j and it is obtained by intersecting σ_j by a face F of K (necessarily containing V_j). Now if the base edges E and E' of S_j are parallel, then they must be parallel with the line $\ell_{F,F'}$ obtained by intersecting the planes of F and F', and then $\ell_{F,F'} \subset \Sigma_j$. Thus, if Σ_j does not contain any of the lines $\ell_{F,F'}$ where F, F' run through the different faces of K containing V_j, then property a) holds. Property b) is also easy to fulfill, since all we need is that the line ℓ_j (which goes through V_j and is perpendicular to Σ_j) contains an interior point of K.

Let $S_j/64$ be the pyramid obtained by shrinking S_j from V_j by a factor 64. According to (9.2), for every $n \geq 3r$ there are polynomials $p_{n,j}$ of degree at most n such that

$$\|p_{n,j} - f\|_{S_j/64} \prec \omega^r_{S_j}\left(f, \frac{1}{n}\right) \leq \omega^r_K\left(f, \frac{1}{n}\right). \tag{10.1}$$

Also, since

$$K^* = K \setminus \bigcup_{j=1}^{m} (S_j/128)$$

is a simple polytope (obtained by cutting off all the pyramids $S_j/128$), by (4.1) there is a polynomial Q_n of degree at most n such that

$$\|Q_n - f\|_{K^*} \prec \overline{\omega}^r_{K^*}\left(f, \frac{1}{n}\right) \leq \omega^r_K\left(f, \frac{1}{n}\right). \tag{10.2}$$

Now we put back one-by-one the cut off pyramids $S_j/128$ and apply the following lemma to conclude that there is a polynomial P_n of degree $\prec n$ with

$$\|P_n - f\|_K \prec \omega^r_K\left(f, \frac{1}{n}\right). \tag{10.3}$$

This is enough to conclude the theorem, since to make P_n to have degree at most n all we have to do is to apply what we have just discussed to n/C with some appropriate C instead of n, and make use (2.14).

Lemma 10.1 *Let H be a polytope lying in $\{x \mid 0 \leq x_1 \leq b\}$ with some $b > 0$, and assume that for some $0 \leq a \leq b/2$ for both*

$$H_1 = H \cap \{x \mid 0 \leq x_1 \leq 2a\} \qquad \text{and} \qquad H_2 = H \cap \{x \mid a \leq x_1 \leq b\}$$

there are polynomials $P_{n,1}$ and $P_{n,2}$ of degree at most Ln such that

$$\|f - P_{n,j}\|_{H_j} \leq \theta_n, \qquad j = 1, 2$$

with some numbers θ_n. Then there is an M depending only on L, H, a, b such that there is a polynomial P_n of degree at most Mn for which

$$\|f - P_{n,j}\|_H \le 2\theta_n. \tag{10.4}$$

Recall that x_1 is the first coordinate of x.

Proof. There is a C (see Lemma 4.2) such that for any polynomial q_n of degree at most Ln we have

$$\|q_n\|_H \le C^n \|q_n\|_{H_1 \cap H_2}. \tag{10.5}$$

Consider

$$R_{n,a}(x) := R_{n,a}^{(4)}(x_1) \tag{10.6}$$

where $R_{n,a}^{(4)}(x_1)$ is the polynomial of the single variable x_1 (the first coordinate of x) from (3.11) with some big A (recall that A was a free parameter in Section 3). If A is sufficiently big, then (cf. (3.12)–(3.13))

$$|R_{n,a}(x)| \le \frac{1}{2C^n}, \qquad x_1 \in [2a, b] \tag{10.7}$$

and

$$|1 - R_{n,a}(x)| \le \frac{1}{2C^n}, \qquad x_1 \in [0, a] \tag{10.8}$$

Set now

$$Q_n = R_{n,a}Q_{n,1} + (1 - R_{n,a})Q_{n,2}.$$

On $H_1 \cap H_2$ we have

$$|f - Q_n| \le |f - Q_{n,1}| + |f - Q_{n,2}| \le 2\theta_n.$$

On $H_2 \setminus H_1$ (which lies in $\{x \,|\, 2a \le x_1 \le b\}$

$$\begin{aligned} |f - Q_n| = |(f - Q_{n,2}) + R_{n,a}(Q_{n,2} - Q_{n,1})| &\le \theta_n + \|Q_{n,2} - Q_{n,1}\|_H |R_{n,a}| \\ &\le \theta_n + C^n(2\theta_n)\frac{1}{2C^n} = 2\theta_n, \end{aligned}$$

where we used (10.7) and (10.5) for $q_n = Q_{n,2} - Q_{n,1}$ together with the fact that

$$\|q_n\|_{H_1 \cap H_2} \le \|f - Q_{n,1}\|_{H_1} + \|f - Q_{n,2}\|_{H_2} \le 2\theta_n.$$

In a similar way (10.8) gives on $H_1 \setminus H_2$

$$\begin{aligned} |f - Q_n| &= |(f - Q_{n,1}) + (1 - R_{n,a})(Q_{n,1} - Q_{n,2})| \\ &\le \theta_n + \|Q_{n,2} - Q_{n,1}\|_H |1 - R_{n,a}| \le \theta_n + C^n(2\theta_n)\frac{1}{2C^n} = 2\theta_n. \end{aligned}$$

The last three estimates prove the lemma.

∎

11 Approximation in $\mathbf{R}^d$

In this section we prove Theorem 1.1 in $\mathbf{R}^d$. Above we gave the proof for $\mathbf{R}^3$, and in the $\mathbf{R}^d$ case we follow that proof.

In $\mathbf{R}^d$ a k-dimensional affine subspace is a translation of a k-dimensional subspace. Any set $H \subset \mathbf{R}^d$ generates a minimal affine subspace $\langle H \rangle$ which is a subset of all affine subspaces containing H. Indeed, if $P \in H$ is any point then clearly $\langle H \rangle = \mathrm{Span}(H - P) + P$ (where Span denotes linear span). The dimension of $\mathrm{Span}(H - P)$ is called the dimension of $\langle H \rangle$.

In $\mathbf{R}^d$ a convex polytope S is called a (d-dimensional) pyramid if S is the convex hull of a $(d-1)$-dimensional convex polytope B lying in some hyperplane $\mathcal{L}$ and of a point $V \notin \mathcal{L}$. B is called the base of S and V is its apex. V is connected to every vertex of B by an edge — these are called the apex edges of S. Besides these the edges of B are also edges of S — these are called the base edges.

A hyperplane L is called a supporting hyperplane to S if $L \cap S \neq \emptyset$ and S lies in one of the two (closed) half-spaces determined by L. F is called a face of S if there is a supporting hyperplane L with $F = L \cap S$. If $\langle F \rangle$ is of k-dimension, then we say that F is a k-dimensional face of S.

Every base edge E has its endpoints on two apex edges E_1, E_2, and EE_1E_2 forms a 2-dimensional face of S. For us it will be crucial (just as it was in $\mathbf{R}^3$) that then for the directions e, e_1, e_2 of E, E_1, E_2 the vector e is a linear combination of the vectors e_1 and e_2.

A base edge E and a base face F are called parallel if F contains a segment which is parallel with E. We shall work with pyramids in $\mathbf{R}^d$ satisfying the following two properties:

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

b) the height of S lies in the interior of S (except for its two endpoints).

Of course, just as in $\mathbf{R}^3$, the height of S is the segment from the apex to the base which is perpendicular to the base. In $\mathbf{R}^3$ these are precisely the two conditions a), b) set forth for S in $\mathbf{R}^d$ (see the beginning of Section 6), namely in $\mathbf{R}^3$ a $(d-2)$-dimensional base face is a base edge.

Let now K be a convex polytope in $\mathbf{R}^d$ for which we want to prove Theorem 1.1. The proof in $\mathbf{R}^3$ was based on the following two facts:

A) We can cut off around every vertex V_j of K a d-dimensional pyramid S_j with properties a)–b).

B) Suppose $S \subset \mathbf{R}^d$ is a d-dimensional pyramid with apex at the origin and with base lying in the hyperplane $\{x \mid x_1 = 2\}$ such that a)–b) hold for S.

Then for any $f \in C(S)$ and for any $n \geq rd$ there is a polynomial p_n (of $x = (x_1, \ldots, x_d)$) of degree at most n such that

$$\|F_n - p_n\|_{S_n^*} \prec \omega_S^r \left(f, \frac{1}{n} \right),$$

where $F_n(x) = f(x - v_n)$,

$$v_n = (-L/n^2, 0, \ldots, 0), \quad \text{and} \quad S_n^* = S \cap \left\{ x \;\middle|\; \frac{2^{9r+1}}{n^2} \leq x_1 \leq \frac{1}{16} \right\}.$$

If we can show that A) and B) holds (perhaps with a larger constant than 2^{9r+1}, say with 2^{d^2r+1} in the definition of S_n^*), then the 3-dimensional proof goes over to any $\mathbf{R}^d$ with minor changes.

11.1 Proof of A)

Let K be a convex polytope (with interior point) in $\mathbf{R}^d$ and V a vertex of S. We may assume that $V = 0$ and that K lies in the half-space $\{x \mid x_1 > 0\}$ except for its vertex 0. Let Σ be a supporting hyperplane to K at 0, and cut off a small pyramid S from K by the hyperplane $\sigma = \Sigma + (b, 0, \ldots, 0)$ with some small $b > 0$ (the apex of S is 0). If b is sufficiently small, then σ intersects all edges of S emanating from 0, so S is a d-dimensional pyramid with base $K \cap \sigma$. We claim that Σ (and a small b) can be selected so that properties a) and b) above hold for S.

Consider first property a). What does it mean that a) is not true, i.e. there is a base edge E of S which is parallel with a $(d-2)$-dimensional base face F of S such that $E \not\subseteq F$? The very definition of a face gives that if $E \cap F \neq \emptyset$ and $E \not\subseteq F$, then E cannot be parallel with F (if it was, then $E \subseteq \langle F \rangle$, so $E \subseteq S \cap \langle F \rangle$, but the latter set is F). Hence, E and F are disjoint and F contains a segment I parallel with E. Let ℓ be the line through 0 which is parallel with E. If E_1, E_2 are the two adjacent apex edges to E, then ℓ lies in the 2-dimensional plane $\langle E_1, E_2 \rangle$, as well as in the $(d-1)$-dimensional hyperplane $L = \langle 0, F \rangle$: indeed, if we translate F by a vector u so that $0 \in I + u$, then clearly $\ell \subseteq \langle F + u \rangle$. Hence, $\ell = \langle E_1, E_2 \rangle \cap L$ (note that since E and F are disjoint, L does not contain $\langle E_1, E_2 \rangle$, but both contain the origin, so their intersection is a line through the origin since $\langle E_1, E_2 \rangle$ is 2-dimensional and L is $(d-1)$-dimensional). Note also that $\Sigma + (b, 0, \ldots, 0)$ contains E, so Σ must contain ℓ. In other words, if property a) is not true, then Σ contains a line ℓ which is the intersection of a 2-dimensional plane $\langle E_1, E_2 \rangle$ determined by two apex edges E_1 and E_2 and of a $(d-1)$-dimensional affine subspace $\langle 0, F \rangle$ generated by a $(d-1)$-dimensional face of K not containing $\langle E_1, E_2 \rangle$. There are only finitely many such lines $\ell_1, \ldots, \ell_m$, and if Σ does not contain any one of them then property a) holds for the cut off S. Now it is clear that if Σ is any supporting hyperplane through 0 to K, then we can change its normal a little so that the changed Σ' will still be a supporting hyperplane and Σ' does not contain any of the lines $\ell_1, \ldots, \ell_m$ (the prohibited normals are on the union of m main hypercircles on the unit ball, so they form a nowhere dense set there). It is also clear that by appropriately choosing the original Σ and change it only little, the changed Σ' will also satisfy property b), i.e. the height lies inside S.

■

11.2 Proof of B)

The proof follows the $\mathbf{R}^3$ case in sections 4–9. First of all, as we have already proved in Section 4, Theorem 1.1 holds for simple polytopes.

Next, let S be a pyramid as in part B), and consider the pieces K_a, $a = a_k = 2^k/n^2$, $k = d^2r, r+1, \ldots, m$ where $1/16 \leq 2^m/n^2 < 1/8$, of S determined by (6.1). These are simple polytopes, and if we can get the local estimate (7.2) on K_a, then the rest is the same as in the $\mathbf{R}^3$-case. In (7.2) the main thing was the bound (6.18), i.e.

$$\overline{\omega}^r_{K_a}\left(F, \frac{1}{n\sqrt{a}}\right) \prec \omega^r_S\left(f, \frac{1}{n}\right) \tag{11.1}$$

for the $\overline{\omega}$-modulus of smoothness on K_a (taken in the direction of edges of K_a), since this allows $\prec \omega^r_S(f, 1/n)$ rate of approximation on K_a by polynomials of degree at most $n\sqrt{a}$ as in (7.2). In the proof of (11.1) the key was the inequality (5.7) in Proposition 5.1 on the rhombus T described there, and this is at our disposal in $\mathbf{R}^d$ (i.e. Proposition 5.1 is still used on plane rhombi). Besides that, the proof of (11.1) depended solely on the inequality

$$\min\{d_S(e_1, x), d_S(e_2, x)\} \geq c_0 d_S(e, x) \tag{11.2}$$

in Proposition 6.1. Therefore, if we can prove the following proposition, which is the complete analogue of Proposition 6.1, then the rest of the proof remains the same.

Proposition 11.1 *If a base edge E is adjacent to the apex edges E_1, E_2 of S and e, e_1, e_2 are their respective directions, then for $x \in S$, $x_1 = 1$ (and hence by similarity also for all $x \in S$, $0 \leq x_1 \leq 1$) we have*

$$\min\{d_S(e_1, x), d_S(e_2, x)\} \geq c_0 d_S(e, x) \tag{11.3}$$

with a c_0 depending only on S.

Proof. For x lying on the 2-dimensional plane $\langle E_1, E_2 \rangle$ (spanned by E_1 and E_2) this is clear, so from now on let $x \in S \setminus \langle E_1, E_2 \rangle$. Consider the 2-dimensional plane H_x that is parallel with $\langle E_1, E_2 \rangle$ and goes through x, and set $V_x = S \cap H_x$.

Let AB be the chord of V_x that contains x and which is parallel with E. Let $S^\infty = \cup_{n>0}(nS)$ be the infinite cone with vertex at 0 generated by the apex edges (more precisely, by their one-way infinite extensions) of S, and let $V_x^\infty = S^\infty \cap H_x$. It is no longer true (as was in the 3-dimensional case) that all these V_x^∞'s (which depend on the location of x) are similar. However, any edge of V_x^∞ lies on the boundary of S^∞, hence it lies on a $(d-1)$-dimensional face of S^∞. In other words, any edge E' of V_x^∞ is the intersection of the 2-dimensional plane H_x with a $(d-1)$-dimensional face F' of S^∞, so the line $\langle E'\rangle$ is the intersection of H_x with $\langle F'\rangle$: $\langle E'\rangle = H_x \cap \langle F'\rangle$. Now changing x means a translation of H_x, which, for its intersection with $\langle F'\rangle$ results in a translation of $H_x \cap \langle F'\rangle$, so the corresponding edge E'' will be parallel with E' (see below). As a consequence, these edges E' can form only finitely many angles $\varphi_1, \ldots, \varphi_k$ with AB (which is parallel with E). Here we have used that if $H_y = H_x + v$ is a translation of H_x, then $H_y \cap \langle F'\rangle$ is a translation of $H_x \cap \langle F'\rangle$. Indeed, the translation vector v can be written as $v = v_0 + v_1$, where v_1 is parallel with $\langle F'\rangle$ and v_0 is parallel with H_x, i.e. with $\langle E_1, E_2\rangle$ (this is due to the fact that H_x and $\langle F'\rangle$ span the whole space $\mathbf{R}^d$). Now

$$\begin{aligned} H_y \cap \langle F'\rangle &= (H_x + v) \cap \langle F'\rangle = (H_x + v_1) \cap \langle F'\rangle = \big(H_x \cap (\langle F'\rangle - v_1)\big) + v_1 \\ &= (H_x \cap \langle F'\rangle) + v_1. \end{aligned}$$

We claim that neither of the angles $\varphi_1, \ldots \varphi_k$ is 0. Indeed, a zero angle would mean that E and E' are parallel. Enlarge E' from 0 so that one of its points becomes a point on the hyperplane $\{u \mid u_1 = 2\}$ (the hyperplane of the base of S). Let this dilation be Φ. Since F' is invariant under Φ, the segment $\Phi(E')$ lies in the intersection $F' \cap \Phi(H_x)$. Let $P \in \Phi(E') \cap \{u \mid u_1 = 2\}$, and $F = F' \cap \{u \mid u_1 = 2\}$. Then F is a $(d-2)$-dimensional base face of S and we claim that $\Phi(E') \subseteq F$. Since $\Phi(E') \subseteq F'$, to this end all we have to show is that $\Phi(E') \subset \{u \mid u_1 = 2\}$. But this follows, since $\Phi(E')$ is parallel with E' so also with E, E lies in $\{u \mid u_1 = 2\}$ and $\Phi(E')$ has a common point (P) with $\{u \mid u_1 = 2\}$. However, $\Phi(E') \subseteq F$ means that the base edge E and the $(d-2)$-dimensional base face F are parallel (and $E \not\subseteq F$, for otherwise $\langle F'\rangle$ would contain $\langle E_1, E_2\rangle$ and then it could not intersect H_x, which is a translation of $\langle E_1, E_2\rangle$, in a line). But this is not possible by property a) of S, and this contradiction shows that all the angles φ_j are different from 0.

Thus, if $\phi_0 > 0$ is their minimum, then the triangle ABC depicted in Figure 11.1 lies inside V_x^∞ and S. Now

$$d_S(e,x) = d_{ABC}(e,x), \qquad d_S(e_j,x) \geq d_{ABC}(e_j,x), \quad j = 1,2, \tag{11.4}$$

and it is clear that

$$d_{ABC}(e,x) \leq \frac{1}{\sin\phi_0} \min_{j=1,2} d_{ABC}(e_j,x). \tag{11.5}$$

Now (11.3) is an immediate consequence of (11.4) and (11.5).

■

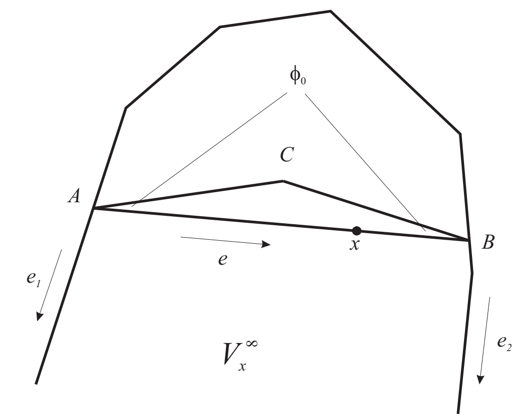

Figure 11.1:

12 A K-functional and the equivalence theorem

K-functionals are important tools in functional analysis and approximation theory. Very often they are equivalent with some kind of moduli of smoothness, and that allows one to prove direct and converse theorems not directly through the moduli of smoothness, but through the K-functionals, see [12] for a systematic treatment, and e.g. [1]–[4], [11] and [17] for various K-functionals and moduli of smoothness in several variables related to polynomial approximation.

As a typical example, consider the φ-modulus of smoothness (1.1) and its equivalence given in (2.9) to the K-functional (2.8). We did not follow that path, since a direct proof of the equivalence of the moduli of smoothness (1.8) with a K-functional seems to be quite hard. Remarkably, however, Theorem 1.1 *does give* this equivalence.

Let K be a polytope in $\mathbf{R}^d$, and consider the relative distances (1.6). The K-functional we need is

$$\mathcal{K}_r(f,t) = \mathcal{K}_r(f,t)_K = \inf_g \left(\|f-g\|_K + t \sup_{e\in S^{d-1}} \left\| \tilde{d}_K(e,\cdot)^r \frac{\partial^r g}{\partial e^r} \right\|_K \right), \tag{12.1}$$

where the infimum is taken for all g that are in $C^r(K)$ (all partial derivatives of order at most r are continuous on K) and the supremum is taken for all directions $e \in S^{d-1}$ in $\mathbf{R}^d$.

Theorem 12.1 *Let K be a convex polytope in $\mathbf{R}^d$. There is a constant M depending only on r and K such that for all $f \in C(K)$ and for all $0 < \delta \le 1$ we have*

$$\frac{1}{M}\mathcal{K}_r(f,\delta^r) \le \omega_K^r(f,\delta) \le M\mathcal{K}_r(f,\delta^r). \tag{12.2}$$

Proof. Let P_n be polynomials of degree at most n $(\geq rd)$ such that

$$\|f - P_n\|_K \prec \omega_K^r(f, n^{-1}),$$

the existence of which is given by Theorem 1.1. Now apply (5.5) on any chord AB of K with $\omega(\delta) = \omega_K^r(f, \delta)$ to conclude that if e is the direction of the chord AB, then for $x \in AB$ we have

$$\left| \tilde{d}_K(e, x)^r \frac{\partial^r P_n(x)}{\partial e^r} \right| \prec n^r \omega_K^r(f, n^{-1}),$$

with $\prec$ independent of the chord AB and the point $x \in AB$. On taking supremum for all $x \in AB$ and for all chords AB of K we can conclude that for all $n \geq rd$

$$\mathcal{K}_r(f, n^{-r}) \leq \|f - P_n\|_K + n^{-r} \sup_{e \in S^{d-1}} \left\| \tilde{d}_K(e, \cdot)^r \frac{\partial^r P_n}{\partial e^r} \right\|_K \prec \omega_K^r(f, n^{-1}).$$

Using simple monotonicity properties of $\mathcal{K}_r$ and ω_K^r (see in particular (2.11)), this is enough to conclude

$$\mathcal{K}_r(f, \delta^r) \prec \omega_K^r(f, \delta)$$

for all $0 < \delta \leq 1$.

The converse inequality is more classical, and it follows from (2.9). Indeed, again if I is a chord of K, then (2.9) (more precisely its transformed form to AB) yields

$$\omega_I^r(f, \delta) \prec \mathcal{K}_r(f, \delta^r)_I \leq \mathcal{K}_r(f, \delta^r),$$

and if we take here the supremum for all chords I of K and apply (1.11), then we get

$$\omega_K^r(f, \delta) \prec \mathcal{K}_r(f, \delta^r).$$

■

Part II

The L^p-case

13 The L^p result

In the second part of the paper we prove a complete analogue of Theorem 1.1 in L^p, $1 \le p < \infty$. We shall be content to do the $\mathbf{R}^3$ case.

Thus, in what follows we shall assume that K is a polytope in $\mathbf{R}^3$ and $f \in L^p(K)$, where the L^p spaces are taken with respect to 3 dimensional Lebesgue measure on K. We shall always assume that $1 \le p < \infty$.

The L^p modulus we are going to use is

$$\omega_K^r(f,\delta)_p := \sup_{e\in S^2}\sup_{h\le\delta}\left(\int_K |\Delta^r_{h\tilde{d}_K(e,x)e} f(x)|^p dx\right)^{1/p} \tag{13.1}$$

with the usual agreement that

$$\Delta^r_{h\tilde{d}_K(e,x)e} f(x) = 0$$

if

$$\left[x - \frac{r}{2}h\tilde{d}_K(e,x)e, x + \frac{r}{2}h\tilde{d}_K(e,x)e\right] \not\subseteq K,$$

i.e. if one of the arguments in Δ^r is outside K. As always in this paper, $\tilde{d}_K(e,x)$ is the normalized distance (1.6), and $\sup_{e\in S^2}$ means that we take the supremum for all directions in $\mathbf{R}^3$.

With this we have the complete analogue of Theorem 1.1.

Theorem 13.1 *Let $K \subset \mathbf{R}^3$ be a 3-dimensional convex polytope and $r = 1,2,\ldots$. Then, for $n \ge 3r$ and $f \in L^p(K)$, we have*

$$E_n(f)_{L^p(K)} \le M\omega_K^r\left(f,\frac{1}{n}\right)_p, \tag{13.2}$$

where M depends only on K, r and p.

Naturally, on the left-hand side $E_n(f)_{L^p(K)}$ is the error of best polynomial approximation of f in $L^p(K)$-norm by polynomials of degree at most n.

The weak converse

$$\omega_K^r\left(f,\frac{1}{n}\right)_p \le \frac{M}{n^r}\sum_{k=0}^{n}(k+1)^{r-1}E_k(f)_{L^p(K)}, \qquad n = 1,2,\ldots,, \tag{13.3}$$

can be proven along standard lines, see [12, Theorem 12.2.3,(12.2.4)], which covers (13.3).

One can easily get from Theorems 13.1 and (13.3), as well as from (16.1) and (16.9) below the following analogue of Corollary 1.2.

Corollary 13.2 *Let K be a convex polytope, $f \in L^p(K)$, $1 \le p < \infty$ and $\alpha > 0$. Assume also that for all directions e we have*

$$\left(\int E_n(f)^p_{L^p(I)} dI \right)^{1/p} \le n^{-\alpha}, \tag{13.4}$$

where the integration is with respect to all chords I of K in the direction of e. Then

$$E_n(f)_{L^p(K)} \le \frac{C}{n^\alpha}, \tag{13.5}$$

where C is independent of f.

"Integration with respect to all chords in the direction of e" means the following: let $e^\perp$ be the hyperplane through the origin which is perpendicular to e. Then, for each point y of $e^\perp$, there is a line $y + \lambda e$, $\lambda \in \mathbf{R}$, through that point which is parallel with e. Now that line intersects K in a (possibly empty) segment $I = I_{e,y}$, and integration with respect to $I = I_{e,y}$ means integration with respect to y on $e^\perp$.

Again, this corollary says that in some sense $n^{-\alpha}$ rate of approximation along chords in a given direction implies for global approximation the rate $n^{-\alpha}$. To prove it consider a direction e and a chord I in that direction. The inequality (1.4) is also true in L^p, see [12, Theorem 7.2.4]. When the L^p version of (1.4) is transformed to I, it takes the form

$$\omega_I^r\left(f, \frac{1}{n}\right)_p \le \frac{C}{n^r} \sum_{k=0}^{n} (k+1)^{r-1} E_k(f)_{L^p(I)}.$$

Now apply Jensen's inequality to get

$$\omega_I^r\left(f, \frac{1}{n}\right)_p^p \le \frac{C_1}{n^{(r-1)p+1}} \sum_{k=0}^{n} (k+1)^{(r-1)p} E_k(f)^p_{L^p(I)}.$$

If we integrate this for all chords I in the direction e and use (13.4), then we obtain

$$\int \omega_I^r\left(f, \frac{1}{n}\right)_p^p dI \le \frac{C_1}{n^{(r-1)p+1}} \sum_{k=0}^{n} (k+1)^{(r-1)p-\alpha p}.$$

If $r - 1 > \alpha$, then the right-hand side is $O(n^{-\alpha p})$, while the supremum of the left-hand side for all directions e is clearly at least as large as $\omega_K^r(f, 1/n)_p^p$, so $\omega_K^r(f, 1/n)_p = O(n^{-\alpha})$ follows. Now (13.5) is a consequence of Theorem 13.1.

In Section 20 we shall prove stronger versions of (13.2) and (13.3). In that stronger version the following fact will play an important role. Let $\mathcal{E}^*$ be a set of directions in $\mathbf{R}^3$, and define the corresponding moduli of smoothness

$$\omega^r_{K,\mathcal{E}^*}(f,\delta)_p := \sup_{e \in \mathcal{E}^*} \sup_{h \le \delta} \left(\int_K |\Delta^r_{h\tilde{d}_K(e,x)e} f(x)|^p dx \right)^{1/p}, \tag{13.6}$$

i.e now we take the supremum of the directional moduli of smoothness only for the directions e that lie in $\mathcal{E}^*$. Besides Theorem 13.1, the proof we give below verifies the following stronger statement.

Theorem 13.3 *Let $K \subset \mathbf{R}^3$ be a 3-dimensional convex polytope and $r = 1, 2, \ldots$. Then there is a finite set $\mathcal{E}^*$ of directions that depends only on K, such that for $n \geq 3r$ and $f \in L^p(K)$ we have*

$$E_n(f)_{L^p(K)} \leq M\omega^r_{K,\mathcal{E}^*}\left(f, \frac{1}{n}\right)_p, \tag{13.7}$$

where M depends only on K, r and p.

In fact, the proof yields an $\mathcal{E}^*$ consisting of at $(r+1)v(v-1)/2$ directions, where v is the number of edges of K.

Note also that if $\mathcal{E}^* = \mathcal{E}$ is the set of the directions of the edges of K, then

$$\omega^r_{K,\mathcal{E}^*}(f, t)_p = \overline{\omega}^r_K(f, t)_p$$

is the modulus of smoothness that will play a significant role in the proof (see (14.1) below), but we do not know if Theorem 13.3 is true with $\mathcal{E}^* = \mathcal{E}$. Nevertheless, the proof starts by claiming that Theorem 13.3 is true with $\mathcal{E}^* = \mathcal{E}$ for simple polytopes, see the next section.

14 Proof of the L^p result

The proof in the L^p case follows the proof given for the continuous case, but at some points there are substantial differences due to the fact that the L^p moduli of smoothness are much more difficult to handle than their continuous cousins. In fact, it would be grossly misleading to state that the L^p proof is the same as the continuous one.

In this section we shall sketch the proof. We shall quickly pass through those steps that are very similar to the continuous case, and elaborate more on parts where some non-trivial change is needed. There is a substantial part of the proof which is totally different in the L^p case, that part will be handled in the sections to follow.

We may always assume that the functions appearing below are Borel-measurable, and then no measurability problems appear when we restrict them to submanifolds.

Let $\mathcal{E}$ be the direction of edges of K, and define the analogue of (2.1) as

$$\overline{\omega}^r_K(f, \delta)_p := \sup_{e\in\mathcal{E}} \sup_{h\leq\delta} \left(\int_K |\Delta^r_{h\tilde{d}_K(e,x)e} f(x)|^p dx\right)^{1/p}. \tag{14.1}$$

Recall also that K is a simple polytope if at every vertex of K there are precisely 3 edges. For cubes

$$E_n(f)_{L^p(K)} \le M\overline{\omega}_K^r\left(f, \frac{1}{n}\right)_p, \qquad n \ge 3r \tag{14.2}$$

was proved in [12, Theorem 12.1.1]. More precisely, [12, Theorem 12.1.1] contains the inequality in (14.2) for all $n \ge r$ if the best approximation $E_n(f)$ is considered for polynomials of degree at most n in each variables. Since we are working with total degree at most n, we wrote in (14.2) $n \ge 3r$. From (14.2) on cubes we get the validity of (14.2) for parallelepipeds by affine transformation. Now a first major step in the proof is to verify

Proposition 14.1 *For all simple polytopes* (14.2) *is true.*

Proof of Proposition 14.1. The proposition can be proven with the method of Section 4, once (14.2) is known for parallelepipeds, as we have just seen. The only difference in the L^p case is that now we have to use the L^p version of Lemma 4.1:

Lemma 14.2 *Let $U \subset K$ be a set, $T \subset K$ a K-parallelepiped with side-lengths in between ε and 2ε such that $U \cap T$ contains a ball B of radius δ. Then there is an l that depends only on ε, δ and K, and there is a C that depends only on p, for which*

$$E_{ln}(f)_{L^p(U\cup T)} \le C\Big(E_n(f)_{L^p(U)} + E_n(f)_{L^p(2T\cap K)}\Big). \tag{14.3}$$

Recall that $2T$ is obtained from T by a dilation about its center by a factor 2.

The original Lemma 4.1 was based on Lemma 4.2, the L^p version of which is

Lemma 14.3 *If B is a ball of radius ρ lying in the unit ball $B_1(0)$, then for any polynomial Q_n of degree at most n*

$$\|Q_n\|_{B_1(0)} \le C\|Q_n\|_{L^p(B)}\left(\frac{17}{\delta}\right)^{n+1}, \tag{14.4}$$

$$\|Q_n\|_{L^p(B_1(0))} \le C\|Q_n\|_{L^p(B)}\left(\frac{17}{\delta}\right)^{n+1} \tag{14.5}$$

with some constant C that depends only on p.

Proof. It is enough to prove (14.4). Let $B = B_\delta(A)$ (δ is the radius and A is the center of B), and set $B' = B_\delta(A) \setminus B_{\delta/2}(A)$. In the proof of Lemma 4.2 we verified

$$|q_n(x)| \le \|q_n\|_{[\alpha-\delta,\alpha+\delta]}(2 \cdot \mathrm{dist}(x, \alpha)/\delta)^n \tag{14.6}$$

for any q_n which is a polynomial of a single variable of degree at most n. Now if X is any point in the unit ball, then let l be the line through A and X. The polynomial Q_n in the lemma when restricted to l, is a polynomial q_n of a single variable of degree at most n, and on this line X lies from A closer than 2. Also, this line intersects B' in two segments of length $\delta/2$, and the center of one of these segments, call it I, is closer to X than 2. Hence, on applying (14.6) to this last segment I we obtain

$$|Q_n(X)| \leq \|Q_n\|_I(2 \cdot \operatorname{dist}(X, A)/(\delta/4))^n \leq \|Q_n\|_I(16/\delta)^n. \tag{14.7}$$

Next, we need Nikolskii's inequality [5, Theorem 4.2.6]: if p_n is a polynomial of degree at most n, then

$$\|p_n\|_{[-1,1]} \leq Cn^{2/p}\|p_n\|_{L^p[-1,1]}. \tag{14.8}$$

When applied on an interval I this changes to

$$\|p_n\|_I \leq C\frac{n^{2/p}}{|I|^{1/p}}\|p_n\|_{L^p(I)}. \tag{14.9}$$

Thus, (14.7) gives

$$|Q_n(X)| \leq C\frac{n^{2/p}}{(\delta/2)^{1/p}}\|Q_n\|_{L^p(I)}(16/\delta)^n \leq C\|Q_n\|_{L^p(I)}(17/\delta)^{n+1} \tag{14.10}$$

with a possibly larger C on the right.

■

Next, we indicate the changes needed in the proof of Lemma 14.2.

Proof of Lemma 14.2. Let P_1 and P_2 be polynomials of degree n such that

$$\|f - P_1\|_{L^p(2T\cap K)} \leq E_n(f)_{L^p(2T\cap K)}, \qquad \|f - P_2\|_{L^p(U)} \leq E_n(f)_{L^p(U)}.$$

On the ball $B \subseteq U \cap T$ we have

$$\begin{aligned} \|P_1 - P_2\|_{L^p(B)} &\leq \|f - P_2\|_{L^p(U)} + \|f - P_1\|_{L^p(2T\cap K)} \\ &\leq E_n(f)_{L^p(U)} + E_n(f)_{L^p(2T\cap K)}, \end{aligned}$$

hence, by Lemma 14.3,

$$\|P_1 - P_2\|_{B_1(0)} \leq C(17/\delta)^{n+1}\big(E_n(f)_{L^p(U)} + E_n(f)_{L^p(2T\cap K)}\big). \tag{14.11}$$

With $\eta = \delta^2/18$ choose the polynomials R_n as in Lemma 4.3, and set $P = R_nP_1 + (1 - R_n)P_2$. This is a polynomial of degree at most $Ln + n$ (with some L which is independent of n), and for it we have on $U \cap 2T = U \cap (2T \cap K)$

$$|f - P| \leq R_n|f - P_1| + (1 - R_n)|f - P_2|, \tag{14.12}$$

so

$$
\begin{aligned}
\int_{U\cap(2T\cap K)} |f-P|^p &\le 2^{p-1}\left(\int_{U\cap(2T\cap K)} |f-P_1|^p + \int_{U\cap(2T\cap K)} |f-P_2|^p\right) \\
&\le E_n(f)^p_{L^p(U)} + E_n(f)^p_{L^p(2T\cap K)}. \qquad (14.13)
\end{aligned}
$$

On T

$$
\begin{aligned}
|f-P| &= |f-P_1+(1-R_n)(P_1-P_2)| \le |f-P_1| + (1-R_n)|P_1-P_2| \\
&\le |f-P_1| + C\eta^n(17/\delta)^{n+1}\big(E_n(f)_{L^p(U)} + E_n(f)_{L^p(2T\cap K)}\big) \\
&\le |f-P_1| + C\big(E_n(f)_{L^p(U)} + E_n(f)_{L^p(2T\cap K)}\big), \qquad (14.14)
\end{aligned}
$$

which gives

$$
\int_T |f-P|^p \le C\left(\int_T |f-P_1|^p + E_n(f)^p_{L^p(U)} + E_n(f)^p_{L^p(2T\cap K)}\right). \qquad (14.15)
$$

Similarly, on $U\setminus 2T$

$$
\begin{aligned}
|f-P| &= |(f-P_2-R_n(P_1-P_2)| \le |f-P_2| + R_n|P_1-P_2| \\
&\le |f-P_2| + C\eta^n(17/\delta)^{n+1}\big(E_n(f)_{L^p(U)} + E_n(f)_{^p(2T\cap K)}\big) \\
&\le |f-P_2| + C\big(E_n(f)_{L^p(U)} + E_n(f)_{^p(2T\cap K)}\big), \qquad (14.16)
\end{aligned}
$$

from which we get

$$
\int_{U\setminus 2T} |f-P|^p \le C\left(\int_{U\setminus 2T} |f-P_2|^p + E_n(f)^p_{L^p(U)} + E_n(f)^p_{L^p(2T\cap K)}\right). \qquad (14.17)
$$

Since $U\cap 2T$, $U\setminus 2T$ and T cover $U\cup T$, (14.13)–(14.17) verify the lemma. ∎

To handle all $n\ge 3r$ not just large n, in Section 4 we used Lemma 4.4, the L^p variant of which is

Lemma 14.4 *Let $U\subset K$ be a set, $T\subset K$ a K-parallelepiped with side-lengths in between ε and 2ε such that $U\cap T$ contains a ball B of radius δ. Then there is a C that depends only on ε,δ,n,p and K for which*

$$
E_n(f)_{L^p(U\cup T)} \le C\Big(E_n(f)_{L^p(U)} + E_n(f)_{L^p(2T\cap K)}\Big). \qquad (14.18)
$$

For the proof just follow the proof of Lemma 4.4, and use the just proven Lemma 14.3 instead of Lemma 4.2.

This completes the discussion on simple polytopes, and with these changes the proof of Section 4 goes over, and it follows that (14.2) is true for all simple polytopes. Thus, Proposition 14.1 holds.

∎

The case of general polytope is reduced to the simple polytope case and that of a pyramid as in Section 10 by cutting off small pyramids around the vertices. That argument remains valid in the present L^p situation (use Lemma 14.3 instead of Lemma 4.2).

Thus, it is enough to consider pyramids S with the properties set forth in Sections 6 and 10, and it is enough to find an appropriate polynomial approximant on some $S_{\eta/4}$ instead of on $S_{1/64}$ as was done in Section 9, (9.2) (recall that S_η is obtained from S by a dilation with a factor η from the apex of S). Instead of (8.3) consider now

$$S_n^* := S \cap \left\{ x \,\middle|\, \frac{\Xi}{n^2} \leq x_1 \leq \eta \right\} \tag{14.19}$$

with some fixed Ξ and $\eta > 0$. In the following sections we are going to show (see (19.10)) that for appropriate Ξ, η and L and for sufficiently large n there are polynomials P_n of degree at most n such that

$$\|P_n - F\|_{L^p(S_n^*)} \prec \omega_S^r \left(f, \frac{1}{n} \right)_p, \tag{14.20}$$

where $F_n(x) = f(x - v_n)$, $v_n = (-L/n^2, 0, 0)$, which is the analogue of (8.9). Once this is done, we can invoke the technique of Section 9 to show that (14.20) automatically implies

$$\|p_n - f\|_{L^p(S_{\eta/4})} \prec \omega_S^r \left(f, \frac{1}{n} \right)_p, \tag{14.21}$$

which will complete the proof of Theorem 13.1. The argument in Section 9 was based on Lemma 9.1, the L^p-version of which is

Proposition 14.5 *Let $g \in L^p[0,1]$ and let $\Lambda > 0$ be fixed. Then, for any polynomial q_n of a single variable and of degree at most n, $n^2 \geq 2\Lambda$, we have*

$$\|g - q_n\|_{L^p[0,1]} \leq C \left(\|g - q_n\|_{L^p[\Lambda/n^2,1]} + \omega_{[0,1]}^r(g, 1/n)_p \right), \tag{14.22}$$

where C depends only on Λ, r and p.

Proof. Just follow the original proof in Lemma 9.1 with the modification, that instead of (9.6) we use its L^p-version

$$\|\varphi^r(q_n^*)^{(r)}\|_{L^p[0,1]} \leq M_r n^r \omega_\varphi^r(q_n^*, n^{-1})_p \prec n^r \theta_n, \tag{14.23}$$

see [12, Theorem 7.3.1] or (17.4)–(17.5) below. The only step from that proof that needs to be explained for L^p is the use of Remez' inequality. The L^∞ version of Remez' inequality, that was used in Lemma 9.1 is this: for every λ there is a C_λ such that if h_n is a polynomial of a single variable of degree at most n, then

$$\|h_n\|_{[-1-\lambda/n^2, 1+\lambda/n^2]} \leq C_\lambda \|h_n\|_{[-1,1]}. \tag{14.24}$$

This is a very simple form of Remez's inequality [16], and actually it is a simple consequence of the inequality (4.14). Instead of it, in Proposition 14.5 we need to use its L^p-variant:

$$\|h_n\|_{L^p[-1-\lambda/n^2, 1+\lambda/n^2]} \leq C_\lambda \|h_n\|_{L^p[-1,1]}. \tag{14.25}$$

To prove this note first of all that, by Nikolskii's inequality (14.8), and by the L^∞ Remez inequality (14.24), we have

$$\|h_n\|_{L^\infty[-1-\lambda/n^2, 1+\lambda/n^2]} \leq C_\lambda n^{2/p} \|h_n\|_{L^p[-1,1]}.$$

Hence,

$$\left(\int_{-1-\lambda/n^2}^{-1} + \int_1^{1+\lambda/n^2}\right) |h_n|^p \leq 2\left(\frac{\lambda}{n^2}\right) C_\lambda^p n^2 \|h_n\|_{L^p[-1,1]}^p,$$

and if we add to both sides to the integral of $|h_n|^p$ over $[-1,1]$, then we obtain (14.25).

■

With these modifications the proof of Theorem 13.1 reduces to the verification of (14.20), which will be done in Sections 15–19. Indeed, this part of the proof is much more difficult than in the L^∞ case, and it proceeds along quite a different path.

15 The dyadic decomposition

Consider a pyramid S as in Section 6 with apex at 0 and with base in the hyperplane $\{x = (x_1, x_2, x_3) \mid x_1 = 2\}$, and let S^∞ be the infinite cone with apex at 0 determined by S: $S^\infty = \cup_{n=1}^\infty nS$. Recall that about S we assumed

a) no two base edges of S are parallel,

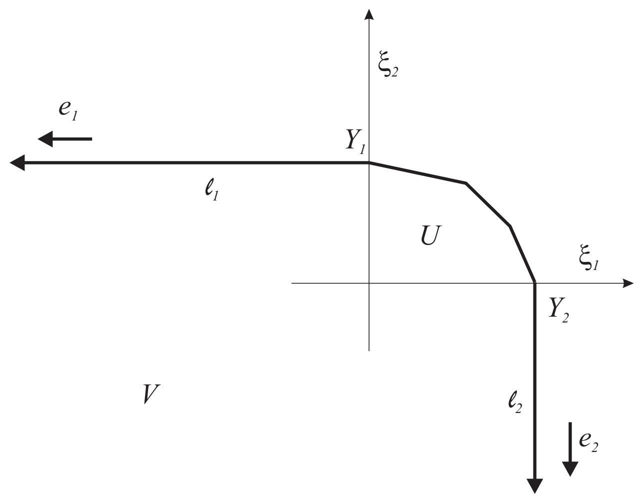

Figure 15.1: The (ξ_1, ξ_2) coordinate system

b) the height of S lies in the interior of S.

We keep these properties here, as well.

Consider a base edge E and the two apex edges E_1 and E_2 such that E lies in the plane $\langle E_1, E_2 \rangle$ spanned by E_1 and E_2. Let e_1 and e_2 be the directions of the edges E_1, E_2. Without loss of generality we may assume that E_1 and E_2 are orthogonal to each other. Indeed, this can always be achieved by an affine transformation (an alternative to handle the non-orthogonality case would be to use dyadic rhombi instead of dyadic squares below).

Let H be a translate of the plane $\langle E_1, E_2 \rangle$ so that $V = V(H) := H \cap S \neq \emptyset$, and let τ be the distance of H and $\langle E_1, E_2 \rangle$. We can parametrize these H's by τ. If $V^\infty = V^\infty(H) = H \cap S^\infty$, then all these $V^\infty(H)$ are similar to one another, and on the boundary of V^∞ there are two infinite edges (half-lines) ℓ_1 and ℓ_2 parallel with E_1 and E_2, respectively. Let Y_1 and Y_2 be the endpoints of ℓ_1 and ℓ_2, and let us place a (ξ_1, ξ_2) coordinate system on H in the following way (see Figure 15.1):

- the positive direction of ξ_j is $-e_j$, $j = 1, 2$,
- Y_1 lies on the ξ_2-axis and Y_2 lies on the ξ_1-axis.

In what follows all reference (like positive quadrant) is made to this coordinate system, but remember, that we also have the (x_1, x_2, x_3) coordinate system in $\mathbf{R}^3$. It is clear that the origin, as well as the whole negative quadrant lies in V^∞, and for small τ, say for $0 \leq \tau \leq \tau_0$, the origin also lies in $S_{1/2}$ (which is S dilated from its apex by a factor $1/2$), so it lies within $V = H \cap S$. Let U be the intersection of V with the positive quadrant (see Figure 15.1). Then the boundary of V^∞ consists of three parts: ℓ_1, ℓ_2 and the part of the boundary of

U that does not lie on the (ξ_1, ξ_2) coordinate axis. Note that, by the assumption that no two base edges are parallel to one another, this boundary does not have an edge parallel with the base edge E we started with (see the discussion in Section 6). Note also that, by the similarity of the infinite polygons V^∞ and by the way we placed the (ξ_1, ξ_2) coordinate system on H, all these U (for different H's) are similar to one another.

For an integer j consider the vertical lines $\xi_1 = k/2^j$, $k = 0, \pm 1, \pm 2, \ldots$ and the horizontal lines $\xi_2 = l/2^j$, $l = 0, \pm 1, \pm 2, \ldots$, which give the j-th level dyadic division of H. A j-th level dyadic square T is

$$T = \left\{ (\xi_1, \xi_2) \,\middle|\, \frac{k}{2^j} \le \xi_1 < \frac{k+1}{2^j}, \quad \frac{l}{2^j} \le \xi_2 < \frac{l+1}{2^j} \right\}.$$

We call the lower left corner $(k/2^j, l/2^j)$ the main vertex of T, and we denote by $2d_T$ the side-length of T, i.e. $d_T = 1/2^{j+1}$. Clearly, two dyadic squares (from any levels) are either disjoint or one of them contains the other.

In this section we make a dyadic decomposition of part of $V = S \cap H$ which is similar to dyadic decompositions used in harmonic analysis, and discuss some geometric properties of them.

For a $\lambda > 0$ let λT be the dilation of T from its center by a factor λ. Fix a small $\beta > 0$ so that

$$S_\beta \subset \bigcup_{0 \le \tau \le \tau_0} V(H_\tau). \tag{15.1}$$

For $x \in V^\infty$ let $T(x)$ by the largest dyadic square T containing x for which $4T \subseteq V^\infty$. If β is sufficiently small, then for $x \in S_\beta$ this is also the largest dyadic square T containing x for which $4T \subseteq V$. The set

$$\mathcal{T} = \mathcal{T}(H) = \{T(x) \,|\, x \in V \cap S_\beta\} \tag{15.2}$$

is the dyadic decomposition of V^∞ we are going to use, see Figure 15.2. Note that $4T \subset V$ for $T \in \mathcal{T}$. Simple geometry shows that if T is a dyadic square at level j, then the square $9T$ contains a $4T'$ where T' is a dyadic square at level $(j-1)$, therefore we get from the definition of $T(x)$ that $9T(x) \not\subseteq V^\infty$, so $9T(x)$ must contain a point on the boundary of V^∞. Furthermore, if x lies sufficiently close to the apex of S, then this means that $9T(x)$ must contain a point on the boundary of V^∞ that lies in $V := H \cap S$, i.e. a point of $\partial V^\infty \cap \partial V$. It is also clear that the union

$$\bigcup_{T \in \mathcal{T}} T$$

of the dyadic squares in this decomposition covers (the interior of) $V \cap S_\beta$. Thus, when we take the union of all these unions (with respect to H_τ with $0 \le \tau \le \tau_0$), then they cover S_β.

Fix a large number M, and for a large integer n consider

$$\mathcal{T}_n = \mathcal{T}_n(H) = \left\{ T \,\middle|\, T \in \mathcal{T},\ \mathrm{dist}(T, \partial S) \ge \frac{M}{n^2} \right\}. \tag{15.3}$$

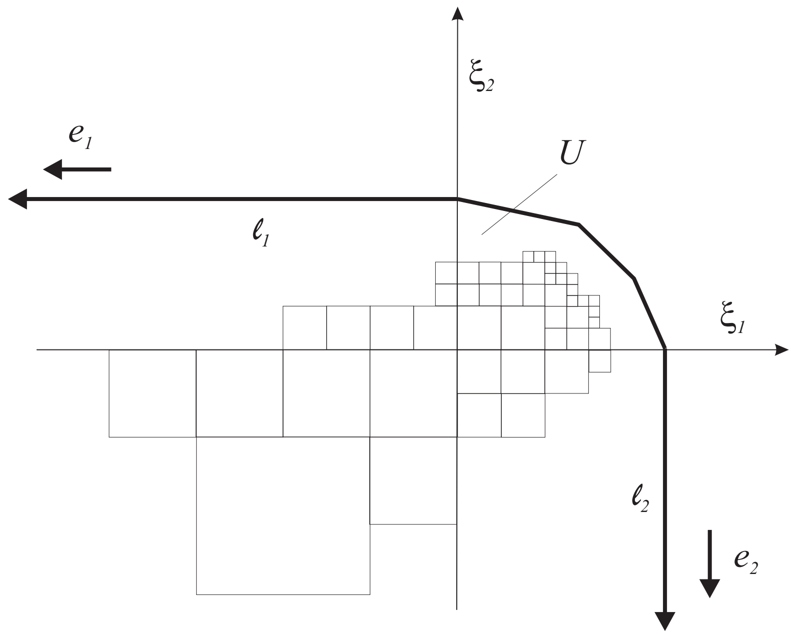

Figure 15.2: Part of the dyadic decomposition

Then, there is an M_1 such that for sufficiently large n

$$\left\{x \in V \cap S_\beta \,|\, \mathrm{dist}(x, \partial S) \geq \frac{M_1}{n^2}\right\} \subseteq \bigcup_{T \in \mathcal{T}_n} T, \tag{15.4}$$

i.e. the union of the squares in $\mathcal{T}_n$ covers that part of $V \cap S_\beta$ that lies of distance $\geq M_1/n^2$ from the boundary ∂S of S.

Since for all $T = T(x)$ we must have $\mathrm{dist}(T, \partial V^\infty \cap \partial V) \leq 8\sqrt{2} d_T \leq 12 d_T$ (otherwise $9T \subset V$), and since $\mathrm{dist}(T, \partial V) \geq \mathrm{dist}(T, \partial S)$, we must have (by the definition of $\mathcal{T}_n$ in (15.3))

$$d_T \geq \frac{M}{12n^2}, \qquad T \in \mathcal{T}_n. \tag{15.5}$$

The next lemmas summarize the most important properties of the dyadic decomposition. In them we write $A \sim B$ for $A \prec B$ and $B \prec A$.

Recall from (1.5) that $d_S(e, x)$ is the distance from x to the boundary of S in the direction of e.

Lemma 15.1 *If $T \in \mathcal{T}$ and $y \in T$, then either $d_S(e_1, y) \sim d_T$ or $d_S(e_2, y) \sim d_T$. Furthermore, if L is some number and if $\mathrm{dist}(T, U) \leq L d_T$, then $d_S(e_j, y) \sim_L d_T$, $j = 1, 2$.*

Of course, here $\sim_L$ means that the constants in $\sim$ depend on L.

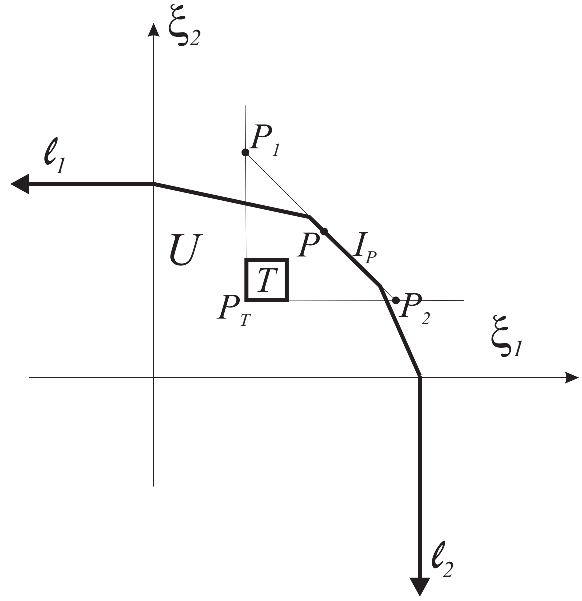

Figure 15.3: The segment I_P and the points P_1, P_2

Proof. We have $d_S(e_j, y) \geq d_T$ because $4T \subseteq V \subset S$. We know that $9T$ has a point P on $\partial V^\infty \cap \partial V$. If P lies on ℓ_1, then $d_T \leq d_S(e_2, y) \leq 10d_T$. Similarly, if P lies on ℓ_2, then $d_T \leq d_S(e_1, y) \leq 10d_T$. Finally, consider the case when P lies on $\partial U \cap \partial V$. Then it lies on a boundary segment I_P of U. I_P has negative slope which can only take finitely many values (here we use that all the U's for different H are similar). Let P_1 and P_2 be as in Figure 15.3. If P_T is the main (lower left) vertex of T, then $\mathrm{dist}(P_T, P_j) \leq Ad_T$ with some A that depends only on the finitely many slopes in question, and so

$$d_T \leq d_S(e_j, y) \leq Ad_T, \qquad j = 1, 2. \tag{15.6}$$

If $\mathrm{dist}(T, U) \leq Ld_T$, then $(L+1)T$ contains a point $Y \in U$, and

$$d_{T(Y)} \leq \mathrm{dist}(Y, \partial V^\infty).$$

If z is the center of T, then

$$\begin{aligned} \mathrm{dist}(Y, \partial V^\infty) \leq (L+1)\sqrt{2}d_T + \mathrm{dist}(z, \partial V^\infty) &\leq (L+1)\sqrt{2}d_T + 9\sqrt{2}d_T \\ &= (L+10)\sqrt{2}d_T. \end{aligned}$$

Therefore,

$$d_{T(Y)} \leq (L+10)\sqrt{2}d_T.$$

Since $T(Y)$ lies in the positive quadrant, we have (in view of $9T(Y) \cap (\partial U \cap \partial V) \neq \emptyset$)

$$\mathrm{dist}(Y, \partial U \cap \partial V) \leq 10\sqrt{2}d_{T(Y)},$$

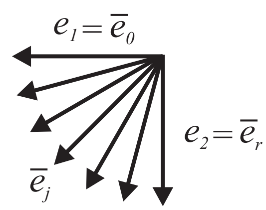

Figure 15.4: The directions $\overline{e}_j$

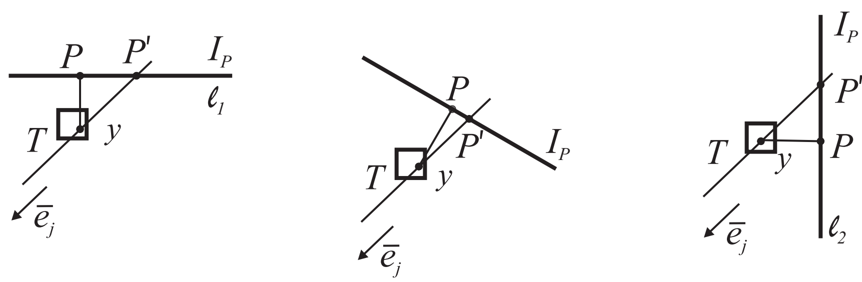

Figure 15.5: The various possibilities for P and I_P in case of the directions $\overline{e}_j$

so there is a point P on $\partial U \cap \partial V$ of distance

$$\leq \big((L+2)\sqrt{2} + 10\sqrt{2}(L+10)\sqrt{2}\big)d_T$$

from y. Now repeat the proof of (15.6) with this P and T.

■

We shall need to use the directions $\overline{e}_j$, $j = 0, 1, \ldots, r$, ($\overline{e}_0 = e_1$, $\overline{e}_r = e_2$) that cut the angle in between e_1 and e_2 into r equal parts (see Figure 15.4 and the proof of Proposition 5.1).

Lemma 15.2 *If $T \in \mathcal{T}$ and $y \in T$, then for all $1 \leq j \leq r-1$ we have $d_S(\overline{e}_j, y) \sim d_T$.*

The lemma is not true for $j = 0$ and $j = r$—these cases were discussed in the preceding lemma.

Proof. Note first of all, that the line through the origin (in the (ξ_1, ξ_2) coordinate system) in the direction $\overline{e}_j$ has strictly positive and finite slope. Let $P \in \partial V^\infty$ be the closest point on the boundary of V^∞ to y, see Figure 15.5. Then P belongs to a segment I_P of the boundary ∂V and the segment yP is perpendicular to I_P. $\overline{e}_j$ cannot be parallel with I_P (the latter has either negative, 0 or ∞ slope) and $\overline{e}_j$ and I_P can form only finitely many angles. Hence, if P' is the intersection of the line of I_P with the line through y in the direction of $\overline{e}_j$, then we have for some constant A

$$d_T \le d_S(\overline{e}_j, y) \le \mathrm{dist}(y, P') \le A \cdot \mathrm{dist}(y, P) \le A \cdot 10\sqrt{2} d_T.$$

■

Recall now that we have started this section with a base edge E with direction e, which was the linear combination of e_1 and e_2. The next lemma is about this direction.

Lemma 15.3 *If $T \in \mathcal{T}$ and $y \in T$, then $d_S(e, y) \sim d_T$.*

This lemma is true because of the basic assumption that no two base edges are parallel (if we had parallel base edges the lemma would be false for them).

Proof. The line ℓ through y in the direction e has negative slope (it must intersect the horizontal and vertical sides ℓ_1 and ℓ_2 of V). Now follow the preceding proof, see also Figure 15.6. If I_P is horizontal or vertical, then we are done as before. If I_P lies on the boundary of U, and so it has a negative slope, then, using that no two base edges are parallel, we get that ℓ is not parallel with I_P, and they can form only finitely many angles, so

$$d_T \le d_S(e, y) \le \mathrm{dist}(y, P') \le A \cdot \mathrm{dist}(y, P) \le A \cdot 10\sqrt{2} d_T$$

for some A that depends only on S.

■

Lemma 15.4 **(a)** *If $T_j, T_k \in \mathcal{T}$ and $T_k \subseteq \frac{7}{2} T_j$, then $d_{T_k} \ge d_{T_j}/16$.*

(b) *If $T_j, T_k \in \mathcal{T}$ and $3T_j \cap 3T_k \ne \emptyset$, then $d_{T_k} \ge d_{T_j}/16$.*

(c) *No $y \in V$ can belong to more than 80^2 of the $3T_k$'s, $T_k \in \mathcal{T}$.*

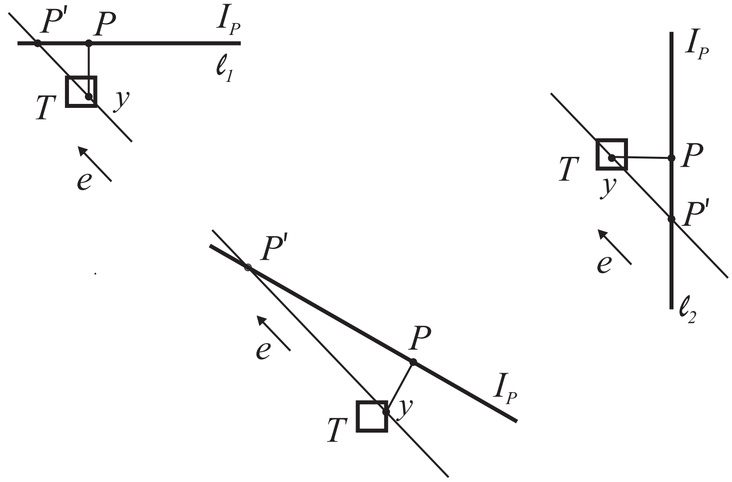

Figure 15.6: The various possibilities for P and I_P in case the direction e

Proof. **(a)** If we had $d_{T_k} < d_{T_j}/16$ then $d_{T_k} \leq d_{T_j}/32$, in which case $8T_k \subseteq 4T_j$ would be true, since the side-length of $8T_k$ would be $\leq d_{T_j}/2$ (and recall that $T_k \subseteq \frac{7}{2}T_j$). But, by the definition of the dyadic decomposition, $8T_k \not\subseteq V^\infty$, while $4T_j \subseteq V^\infty$, and this is a contradiction.

(b) If $T_k \subseteq \frac{7}{2}T_j$, then we can use part (a). If $T_k \not\subseteq \frac{7}{2}T_j$ and $d_{T_k} \leq d_{T_j}/32$, then $T_k \subseteq 4T_j \setminus \frac{7}{2}T_j$ (note that $\frac{7}{2}T_j$ is the union of dyadic squares of side-length $d_{T_j}/2$), hence $3T_k$ cannot intersect $3T_j$ (the distance from $3T_j$ to $4T_j \setminus \frac{7}{2}T_j$ is $\geq d_{T_j}/2 > 3d_{T_k}$). So this last assumption is impossible.

(c) Suppose $y \in 3T_k, 3T_j$, and assume that $d_{T_j} \geq d_{T_k}$. According to (b) we have $d_{T_j}/16 \leq d_{T_k} \leq d_{T_j}$, and $T_k \subseteq 5T_j$ (otherwise $T_k \cap 5T_j = \emptyset$ and then $3T_k \cap 3T_j \neq \emptyset$). Finally, there are at most 80^2 such T_k, since the area of $5T_j$ is $10^2 d_{T_j}^2$, while the area of each such T_k is $4d_{T_k}^2 \geq 4(d_{T_j}/16)^2$.

∎

Recall now the sets $K_a = S_a \setminus S_{a/4}$ from (6.1).

Lemma 15.5 *If $T \in \mathcal{T}$ and $T \cap K_a \neq \emptyset$, then $3T \subseteq \tilde{K}_a$, where*

$$\tilde{K}_a = K_{a/8} \cup K_{a/4} \cup K_{a/2} \cup K_a \cup K_{2a} \cup K_{4a}. \tag{15.7}$$

Proof. What we need to show is that if there is a $y = (y_1, \cdot, \cdot) \in T$ for which $a/2 \leq y_1 \leq 2a$ and $w = (w_1, \cdot, \cdot) \in 3T$, then $a/16 \leq w_1 \leq 8a$. Let $v_1 \in 4T$ have smallest x_1-coordinate. Then $v_1 \geq 0$ (because $4T \subset V \subset S$ and S lies in the half-space $\{x \,|\, x_1 \geq 0\}$).

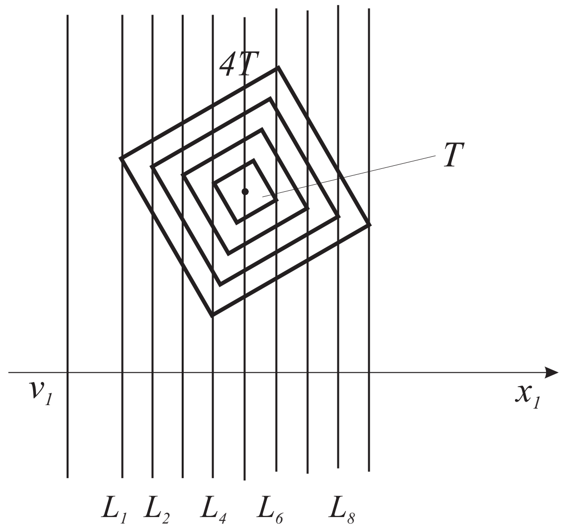

Figure 15.7:

Through each endpoint of the appropriate diagonal (which has longer projection onto the x_1-axis) of the squares $T, 2T, 3T$ and $4T$ draw a line parallel with e (these are the lines in H for which the x_1-coordinate is fixed since e is parallel with the base which lies in the hyperplane $x_1 = 2$), see Figure 15.7. Since w is a point in $3T$, simple consideration based on parallel lines (see Figure 15.7, and note that the two extremal cases are when $y \in L_4$ and $w \in L_8$, resp. when $y \in L_6$ and $w \in L_2$) gives that

$$w_1 \geq y_1 - \frac{y_1 - v_1}{5} 4 = \frac{y_1}{5} + \frac{4}{5} v_1 \geq \frac{y_1}{5} \geq \frac{a}{10},$$

and similarly

$$w_1 \leq y_1 + \frac{4}{3}(y_1 - v_1) \leq \frac{7}{3} y_1 \leq \frac{14}{3} a.$$

■

16 Some properties of L^p moduli of smoothness

One of the most important properties of L^p moduli of smoothness ω_K^r (see (13.1)) that will be frequently used below is

$$\omega_K^r(f, \delta)_p^p \sim \sup_{e \in S^2} \frac{1}{\delta} \int_0^\delta \left(\int_K |\Delta^r_{u\tilde{d}_K(e,x)e} f(x)|^p dx \right) du, \tag{16.1}$$

and (see (14.1) for the definition of $\overline{\omega}_K^r$)

$$\overline{\omega}_K^r(f,\delta)_p^p \sim \max_{e\in\mathcal{E}} \frac{1}{\delta}\int_0^\delta \left(\int_K |\Delta^r_{u\tilde{d}_K(e,x)e} f(x)|^p dx\right) du. \tag{16.2}$$

As before, $\mathcal{E}$ is the direction of the edges of K.

Recall the agreement that in the integral

$$\int_K |\Delta^r f|^p$$

the integrand is considered to be zero if one of the arguments of $\Delta^r f$ lies outside the set K. Sometimes, however, we shall have integrals of the form

$$\int_V |\Delta^r f|^p$$

in which the integrand may have meaning even if some of the arguments in $\Delta^r f$ lie outside the set V of integration. Therefore, for clearer notation and for further emphasis we introduce the

Agreement. The notation

$$\int_U^* |\Delta^r f|^p \tag{16.3}$$

means that the integrand is considered to be zero if one of the arguments of $\Delta^r f$ *lies outside the set* U.

Thus, in this sense, the integrals in (16.1) and (16.2) (as well as the integrals in the definition of the modulus of smoothness) are actually $\int^*$.

Define the directional modulus of smoothness in the direction e as

$$\omega_{K,e}^r(f,\delta)^p := \sup_{u\le\delta} \int_K^* |\Delta^r_{u\tilde{d}_K(e,x)e} f(x)|^p dx. \tag{16.4}$$

Instead of (16.1) and (16.2) we only need to show that

$$\omega_{K,e}^r(f,\delta)^p \sim \frac{1}{\delta}\int_0^\delta \int_K^* |\Delta^r_{u\tilde{d}_K(e,x)e} f(x)|^p dx du. \tag{16.5}$$

Proof of (16.5). (16.5) follows from its variant on segments: if I is a segment and e is its direction, then for $\delta \le 1$

$$\sup_{u\le\delta}\int_I^* |\Delta^r_{u\tilde{d}_I(e,x)e} f(x)|^p dx \sim \frac{1}{\delta}\int_0^\delta \left(\int_I^* |\Delta^r_{u\tilde{d}_I(e,x)e} f(x)|^p dx\right) du, \tag{16.6}$$

where $\sim$ is universal, it does not depend on $\delta \le 1$ or I. Clearly, it is enough to do this for $I = [a,b]$, in which case it takes the form

$$\sup_{u\le\delta}\int_I^* |\Delta^r_{u\sqrt{(x-a)(b-x)}} f(x)|^p dx \sim \frac{1}{\delta}\int_0^\delta \left(\int_I^* |\Delta^r_{u\sqrt{(x-a)(b-x)}} f(x)|^p dx\right) du, \tag{16.7}$$

For $I=[0,1]$ this was proven in [12, Sec. 2.3], see particularly formulae (2.3.2)–(2.3.3). The general case is obtained by a linear transformation: assume that Φ is a linear transformation mapping $I=[a,b]$ into $\Phi(I)=[\Phi(a),\Phi(b)]$. Since, with $y=\Phi(x)$, $F(y)=f(x)$, we have

$$\Delta^r_{u\sqrt{(x-a)(b-x)}}f(x)=\Delta^r_{u\sqrt{(y-\Phi(a))(\Phi(b)-y)}}F(y),$$

which implies with $\beta=|\Phi'|$

$$\int_I^* |\Delta^r_{u\sqrt{(x-a)(b-x)}}f(x)|^p dx=\frac{1}{\beta}\int_{\Phi(I)}^* |\Delta^r_{u\sqrt{(y-\Phi(a))(\Phi(b)-y)}}F(y)|^p dy,$$

it follows that (16.7) is true for an I precisely if it is true for $\Phi(I)$. Since, as we have just mentioned, (16.7) is true for $[0,1]$, its validity for all segments I follows.

We are going to integrate (16.6) for all chords of K in the direction of e. What we are doing precisely is the following: let $e^\perp$ be the hyperplane through the origin which is perpendicular to e. Then for each point y of $e^\perp$, there is a line $y+\lambda e$, $\lambda\in\mathbf{R}$, through that point which is parallel with e, and that line intersects K in a (possibly empty) chord $I_{e,y}$. Now we apply (16.6) on this chord, and integrate the resulting inequality on $e^\perp$ for all y:

$$\int_{y\in e^\perp}\left(\sup_{u\le\delta}\int_{I_{e,y}}^* |\Delta^r_{u\tilde{d}_K(e,y+\lambda e)e}f(y+\lambda e)|^p d\lambda\right)dy\sim$$

$$\sim\int_{y\in e^\perp}\left(\frac{1}{\delta}\int_0^\delta\int_{I_{e,y}}^* |\Delta^r_{u\tilde{d}_K(e,y+\lambda e)e}f(y+\lambda e)|^p d\lambda du\right)dy.$$

This implies

$$\sup_{u\le\delta}\left(\int_{y\in e^\perp}\int_{I_{e,y}}^* |\Delta^r_{u\tilde{d}_K(e,y+\lambda e)e}f(x+\lambda e)|^p d\lambda dy\right)\prec$$

$$\prec\frac{1}{\delta}\int_0^\delta\left(\int_{y\in e^\perp}\int_{I_{e,y}}^* |\Delta^r_{u\tilde{d}_K(e,y+\lambda e)e}f(y+\lambda e)|^p d\lambda dy\right)du.$$

On the other hand, the right-hand side is clearly smaller than the left-hand side, so we can write $\sim$ instead of $\prec$:

$$\sup_{u\le\delta}\left(\int_{y\in e^\perp}\int_{I_{e,y}}^* |\Delta^r_{u\tilde{d}_K(e,y+\lambda e)e}f(y+\lambda e)|^p d\lambda dy\right)\sim$$

$$\sim\frac{1}{\delta}\int_0^\delta\left(\int_{y\in e^\perp}\int_{I_{e,y}}^* |\Delta^r_{u\tilde{d}_K(e,y+\lambda e)e}f(y+\lambda e)|^p d\lambda dy\right)du, \quad (16.8)$$

and this is precisely (16.5).

∎

In the preceding proof we have also verified the following. For a direction e and for $y \in e^{\perp}$ let

$$\omega^r_{e,y}(f,\delta)^p_p = \sup_{u\le\delta} \int^*_{I_{e,y}} |\Delta^r_{u\tilde{d}_K(e,y+\lambda e)e} f(y+\lambda e)|^p d\lambda$$

be the L^p modulus of smoothness of f on the segment that the line through y in the direction of e cuts out of K. Then

$$\omega^r_{K,e}(f,\delta)^p_p \sim \int_{e^{\perp}} \omega^r_{e,y}(f,\delta)^p_p dy, \tag{16.9}$$

i.e. the (p-th power of the) L^p modulus of smoothness in the direction of e is equivalent to the integral of the (p-th power of the) moduli of smoothness on all segments of K in the direction of e.

We shall also need that

$$\omega^r_K(f,\delta)_p \le C\|f\|_{L^p(K)} \tag{16.10}$$

with a C that depends only on r and p. Indeed, for $[-1,1]$, i.e. when $K=[-1,1]$, this follows from the inequality (see [12, (1.2.1)])

$$\int^*_{(a,b)} |g(x \pm h\sqrt{1-x^2})| \le C\int_a^b |g|, \qquad (a,b)\subset[-1,1], \tag{16.11}$$

which is an immediate consequence of the fact that for $x \pm h\sqrt{1-x^2} \in [-1,1]$ we have

$$\left(x \pm h\sqrt{1-x^2}\right)' \ge \frac{\sqrt{2}-1}{\sqrt{2}} \tag{16.12}$$

see [12, (1.2.1)]. Now, (16.11) clearly gives for $I=[-1,1]$, and then by a linear transformation for all segments I, the inequality

$$\omega^r_I(f,\delta)^p_p \le C\|f\|^p_{L^p(I)}. \tag{16.13}$$

From here we get

$$\omega^r_{K,e}(f,\delta)^p \le C\|f\|^p_{L^p(K)} \tag{16.14}$$

by integration with respect to all chords of K in the given direction e. Finally, to get (16.10) take the supremum of both sides for all directions.

Actually, this proof gives a little more that we shall also need: if $K_1 \subset K_2$ are convex sets and e is a direction, then

$$\int^*_{K_1} |\Delta_{u\tilde{d}_{K_2}(e,x)e} f(x)|^p dx \le C\int_{K_1} |f(x)|^p dx. \tag{16.15}$$

Next, we note that, by [12, Theorem 2.1.1], the equivalence (2.9) is true in all L^p spaces, and then the argument leading to (2.11) shows that the L^p-version of (2.11) is also true:

$$\omega^r_{[-1,1]}(f,\lambda t)_p \le M\lambda^r \omega^r_{[-1,1]}(f,t)_p, \qquad \lambda > 1,\ t \le 1. \tag{16.16}$$

This now gives the same estimate on all segments, and then by the (by now) usual argument (via integrating for all chords of K in a given direction) we obtain

$$\omega^r_{K,e}(f,\lambda t)_p \le M\lambda^r \omega^r_{K,e}(f,t)_p, \qquad \lambda > 1,\ t \le 1; \tag{16.17}$$

and

$$\omega^r_{K}(f,\lambda t)_p \le M\lambda^r \omega^r_{K}(f,t)_p, \qquad \lambda > 1,\ t \le 1; \tag{16.18}$$

just use (16.16) and (16.9).

Next, we discuss a few results which allow us to replace a function φ_1 in $\Delta^r_{u\varphi_1} f$ with another one. Most of these are based on the equivalence of a φ-modulus of smoothness with the appropriate K-functional, see [12, Theorem 2.1.1].

Lemma 16.1 *Suppose I is an interval, and for $\xi \in I$ we have*

$$0 < \frac{1}{A}\varphi_2(\xi) \le \varphi_1(\xi) \le \varphi_2(\xi)$$

with some constant A. Then, for all $\delta > 0$,

$$\int_0^\delta \int_I |\Delta^r_{u\varphi_1(\xi)} g(\xi)|^p d\xi du \le A \int_0^\delta \int_I |\Delta^r_{u\varphi_2(\xi)} g(\xi)|^p d\xi du. \tag{16.19}$$

This is [12, Lemma 2.2.1], but for completeness we present the simple proof.

Proof. Substitution gives

$$\begin{aligned}\int_0^\delta |\Delta^r_{u\varphi_1(\xi)} g(\xi)|^p du &= \frac{1}{\varphi_1(\xi)} \int_0^{\delta\varphi_1(\xi)} |\Delta^r_v g(\xi)|^p dv \\ &\le \frac{A}{\varphi_2(\xi)} \int_0^{\delta\varphi_2(\xi)} |\Delta^r_v g(\xi)|^p dv = A\int_0^\delta |\Delta^r_{u\varphi_2(\xi)} g(\xi)|^p du,\end{aligned}$$

and we obtain (16.19) by Fubini's theorem.

■

Often we shall be using Corollary 1.2 in the following form.

Corollary 16.2 *Let $K_0 \subseteq K$ be convex sets, e a direction, and assume that for all $x \in K_0$ we have with some constants τ and A the inequality*

$$\tau \le \tilde{d}_K(e,x) \le A\tau.$$

Then, for all $\delta > 0$,

$$\int_0^\delta \int_{K_0} |\Delta^r_{u\tau e} g(x)|^p dx du \le A \int_0^\delta \int_{K_0} |\Delta^r_{u\tilde{d}_K(e,x)e} g(x)|^p dx du. \tag{16.20}$$

Indeed, just apply Lemma 16.1 on every chord of K_0 in the direction of e and then integrate the so obtained inequalities on the plane perpendicular to e.

Lemma 16.3 *Suppose that $I = [a,b] \subset \mathbf{R}$ is an interval of length $\le D$. Then, for $D\delta \le |I|$,*

$$\frac{1}{\delta}\int_0^\delta \int_I^* |\Delta^r_{u\sqrt{(\xi-a)(b-\xi)}} g(\xi)|^p d\xi du \le C\frac{1}{\delta}\int_0^\delta \int_I^* |\Delta^r_{uD} g(\xi)|^p d\xi du, \tag{16.21}$$

where C is a constant that depends only on r and p.

Proof. Write the right-hand side as

$$\frac{1}{\delta D/|I|}\int_0^{\delta D/|I|} \int_I^* |\Delta^r_{v|I|} g(\xi)|^p d\xi dv.$$

According to [12, Theorem 2.1.1] and (16.1), for any $F \in L^p[0,1]$ and $\sigma \le 1$ we have

$$\frac{1}{C}K(F,\sigma^r) \le \frac{1}{\sigma}\int_0^\sigma \int_{[0,1]}^* |\Delta^r_v F(\xi)|^p d\xi dv \le CK(F,\sigma^r), \tag{16.22}$$

with some constant C (that may depend on r and p), where

$$K(F,\sigma^r) = \inf_Q \left(\int_{[0,1]} |F-Q|^p + \sigma^{rp}\int_{[0,1]} |Q^{(r)}|^p \right).$$

With $\psi(\xi) = \sqrt{\xi(1-\xi)}$ we have, again by [12, Theorem 2.1.1] and (16.1),

$$\frac{1}{C}\widetilde{K}(F,\delta^r) \le \frac{1}{\delta}\int_0^\delta \int_{[0,1]}^* |\Delta^r_{u\psi(\xi)} F(\xi)|^p d\xi du \le C\widetilde{K}(F,\delta^r), \tag{16.23}$$

with some constant C, where

$$\widetilde{K}(F,\delta^r) = \inf_Q \left(\int_{[0,1]} |F-Q|^p + \delta^{rp}\int_{[0,1]} \psi^{rp}|Q^{(r)}|^p \right). \tag{16.24}$$

We set in (16.22) $\sigma = \delta D/|I|$. Since $\delta D/|I| \geq \delta$ and $\psi(\xi) \leq 1$, this gives $\widetilde{K}(F, \delta^r) \leq K(F, (\delta D/|I|)^r)$, and the $I = [0, 1]$ case of the lemma is an immediate consequence.

The general case follows from this by a linear transformation. Indeed, let Φ be a linear transformation mapping I into $[0, 1]$. Then $|\Phi'| = 1/|I| =: \beta$, and with $F = g(\Phi^{-1})$, $\Phi(\xi) = \zeta$ we have

$$\int_I^* |\Delta_{uD}^r g(\xi)|^p d\xi = \frac{1}{\beta} \int_{[0,1]}^* |\Delta_{uD\beta}^r F(\zeta)|^p d\zeta,$$

while

$$\int_I^* |\Delta_{u\sqrt{(\xi-a)(b-\xi)}}^r g(\xi)|^p d\xi = \frac{1}{\beta} \int_{[0,1]}^* |\Delta_{u\sqrt{\zeta(1-\zeta)}}^r F(\zeta)|^p d\zeta.$$

Thus, under this linear transformation $|I|$ changes to 1 and D changes to $D\beta = D/|I|$, so the condition $\delta D \leq |I|$ is unchanged. Therefore, we can use the $I = [0, 1]$ case and the lemma follows.

■

Lemma 16.4 *Suppose that $J \subset I = [a, b] \subset \mathbf{R}$ are intervals, and for $\xi \in J$ we have for some constant τ the inequality $\tau|J| \leq \sqrt{(\xi - a)(b - \xi)}$. Then, for $\delta \leq 1$,*

$$\frac{1}{\delta} \int_0^\delta \int_J^* |\Delta_{u\tau|J|}^r g(\xi)|^p d\xi du \leq C \frac{1}{\delta} \int_0^\delta \int_I^* |\Delta_{u\sqrt{(\xi-a)(b-\xi)}}^r g(\xi)|^p d\xi du, \quad (16.25)$$

where C depends only on r and p.

Proof. The inner integral on the left is invariant under linear transformation, i.e. if Φ is a linear transformation, then

$$\int_J^* |\Delta_{u\tau|J|}^r g(\xi)|^p d\xi = \frac{1}{\beta} \int_{\Phi(J)}^* |\Delta_{u\tau|\Phi(J)|}^r g(\Phi^{-1}(\zeta))|^p d\zeta,$$

where $\beta = |\Phi'|$ is the Jacobian of Φ. The same is true on the right-hand side because

$$\int_I |\Delta_{u\sqrt{(\xi-a)(b-\xi)}}^r g(\xi)|^p d\xi = \frac{1}{\beta} \int_{\Phi(I)} |\Delta_{u\sqrt{(\zeta-\Phi(a))(\Phi(b)-\zeta)}}^r g(\Phi^{-1}(\zeta))|^p d\zeta.$$

Therefore, by choosing Φ so that $\Phi(I) = [0, 1]$, we may assume that $I = [0, 1]$, and in this case $\sqrt{(\xi - a)(b - \xi)} = \sqrt{\xi(1 - \xi)} =: \psi(\xi)$.

In view of (16.23) we can choose an r-times continuously differentiable function G such that

$$\int_{[0,1]} |g - G|^p + \delta^{rp} \int_{[0,1]}^* \psi^{rp} |G^{(r)}|^p \leq C \frac{1}{\delta} \int_0^\delta \int_{[0,1]} |\Delta_{u\psi(\xi)}^r g(\xi)|^p d\xi du. \quad (16.26)$$

By assumption

$$\int_J |g-G|^p + \delta^{rp}\int_J (\tau|J|)^{rp}|G^{(r)}|^p \le \int_{[0,1]} |f-G|^p + \delta^{rp}\int_{[0,1]} \psi^{rp}|G^{(r)}|^p,$$

hence

$$\inf_q \left(\int_J |g-q|^p + (\delta\tau|J|)^{rp}\int_J |q^{(r)}|^p\right) \le C\frac{1}{\delta}\int_0^\delta \int_{[0,1]} |\Delta^r_{u\psi(\xi)} g(\xi)|^p d\xi du. \tag{16.27}$$

Next, let Ψ be the linear transformation that maps $[0,1]$ into J. (16.22) under Ψ changes to

$$\frac{1}{C}K^J(g,(\sigma|J|)^r) \le \frac{1}{\sigma}\int_0^\sigma \int_J^* |\Delta^r_{u|J|} g(\xi)|^p d\xi du \le CK^J(g,(\sigma|J|)^r), \quad \sigma \le 1, \tag{16.28}$$

(with the same constant C as in (16.22)), where

$$K^J(g,(\sigma|J|)^r) = \inf_q \left(\int_J |g-q|^p + (\sigma|J|)^{rp}\int_J |q^{(r)}|^p\right).$$

With $\sigma = \delta\tau$ the expression in the middle of (16.28) is

$$\frac{1}{\delta\tau}\int_0^{\delta\tau} \int_J^* |\Delta^r_{u|J|} g(\xi)|^p d\xi du = \frac{1}{\delta}\int_0^\delta \int_J^* |\Delta^r_{u\tau|J|} g(\xi)|^p d\xi du. \tag{16.29}$$

Now if we put (16.27)–(16.29) together, we obtain (16.25).

■

Lemma 16.5 *Suppose that $J = [\alpha,\beta] \subset I = [a,b]$ are intervals. Then, for $\delta \le 1$,*

$$\frac{1}{\delta}\int_0^\delta \int_J^* |\Delta^r_{u\sqrt{(\xi-\alpha)(\beta-\xi)}} g(\xi)|^p d\xi du \le C\frac{1}{\delta}\int_0^\delta \int_I^* |\Delta^r_{u\sqrt{(\zeta-a)(b-\zeta)}} g(\zeta)|^p d\zeta du, \tag{16.30}$$

where C depends only on r and p. Furthermore, if

$$(\zeta-\alpha)(\beta-\zeta) \le \tau(\zeta-a)(b-\zeta), \qquad \zeta \in (\alpha,\beta)$$

with some constant τ, then for $\delta \le \tau$

$$\frac{\sqrt{\tau}}{\delta}\int_0^{\delta/\sqrt{\tau}} \int_J^* |\Delta^r_{u\sqrt{(\xi-\alpha)(\beta-\xi)}} g(\xi)|^p d\xi du \le C\frac{1}{\delta}\int_0^\delta \int_I^* |\Delta^r_{u\sqrt{(\zeta-a)(b-\zeta)}} g(\zeta)|^p d\zeta du, \tag{16.31}$$

where C depends only on r and p.

Precisely as in Corollary 16.2, we get from this

Corollary 16.6 *Let $K_0 \subset K_1$ be convex sets, and let e be any direction. Then, for $\delta \le 1$,*

$$\frac{1}{\delta}\int_0^\delta \int_{K_0}^* |\Delta^r_{u\tilde{d}_{K_0}(e,x)} g(x)|^p dx du \le C\frac{1}{\delta}\int_0^\delta \int_{K_1}^* |\Delta^r_{u\tilde{d}_{K_1}(e,x)} g(x)|^p dx du \quad (16.32)$$

and

$$\omega^r_{K_0,e}(f,\delta)_p \le C\omega^r_{K,e}(f,\delta)_p, \quad (16.33)$$

where C depends only on r and p.

Note that (16.32) and (16.33) are equivalent in view of (16.5).

This corollary says that the (directional) L^p moduli of smoothness are essentially monotone functions of the underlying sets. Note that, while this is absolutely trivial in the L^∞ case, in the L^p-case it needs verification.

Proof. Just follow the proof of the preceding lemma. The inequality (16.23) is transformed by a linear transformation into

$$\frac{1}{C}\widetilde{K}^I(g,\delta^r) \le \frac{1}{\delta}\int_0^\delta \int_I^* |\Delta^r_{u\sqrt{(\zeta-a)(b-\zeta)}} g(\zeta)|^p d\zeta du \le C\widetilde{K}^I(g,\delta^r), \quad (16.34)$$

with the same constant C as in (16.23), where

$$\widetilde{K}^I(g,\delta^r) = \inf_Q \left(\int_I |g-Q|^p + \delta^{rp}\int_I ((\zeta-a)(b-\zeta))^{rp/2}|Q^{(r)}(\zeta)|^p d\zeta\right). \quad (16.35)$$

Now write this up also for J instead of I and use that $J \subset I$ and $(\zeta-\alpha)(\beta-\zeta) \le (\zeta-a)(b-\zeta)$ for $\zeta \in J$ to conclude first $\tilde{K}^J(g,\delta^r) \le \tilde{K}^I(g,\delta^r)$, and then (16.30).

The proof of (16.31) is identical, only in this case first conclude with the above argument $\tilde{K}^J(g,(\delta/\sqrt{\tau})^r) \le \tilde{K}^I(g,\delta^r)$, and then from this (16.31).

■

17 Local L^p moduli of smoothness

In this section we verify the analogue of (6.18), but due to the fact that L^p moduli are integrals, we shall need to prove a more complicated form.

We use the notations and setup from Section 15. Recall that there we worked with a pyramid S with base on the hyperplane $\{x \mid x_1 = 2\}$ and with apex at the origin such that the height of S is the segment $0 \le x_1 \le 2$. We were also working in Section 15 with a base edge direction e and two apex edges E_1, E_2 with directions e_1, e_2 such that e is their linear combination. Exactly as in Section 15, we may assume e_1 and e_2 to be perpendicular to each other.

Besides these, we shall also need the directions $\overline{e}_j$, $j = 0, 1, \ldots, r$, $\overline{e}_0 = e_1$, $\overline{e}_r = e_2$ that divide the angle in between e_1 and e_2 into r equal parts, see Figure 15.4.

Let H be a translate of the plane $\langle E_1, E_2 \rangle$ with the (ξ_1, ξ_2) coordinate system as discussed in Section 15, $V = V(H) = H \cap S$, and consider the dyadic decomposition $\mathcal{T} = \mathcal{T}(H)$ of V from (15.2). Let n be a fixed large integer and $T \in \mathcal{T}(H)$ be a dyadic square from the decomposition lying from the boundary of S of distance $\geq M/n^2$ with some large and fixed M. These formed the set $\mathcal{T}_n = \mathcal{T}_n(H)$ in Section 15. Recall that $2d_T$ is the side-length of a square T in this decomposition. Then $d_T \geq M/12n^2$ by (15.5). Furthermore, the union

$$\bigcup_{\mathrm{dist}(T, \partial S) \geq M/n^2} T$$

of these squares cover the set

$$\{x \in V \,|\, \mathrm{dist}(x, \partial S) \geq M_1/n^2, \quad x_1 \leq 2\beta\} \tag{17.1}$$

for some M_1, see (15.4).

Let F be a Borel function on S. With an integer $m \geq 5r$ to be specified later and with $3T$ (the 3 times enlarged T from its center) we apply Proposition 14.1 and (16.1), according to which there are polynomials Q_m of degree at most m such that

$$\iint_{3T} |F - Q_m|^p \prec \sum_{j=1}^{2} m \int_0^{1/m} \iint_{3T}^{*} |\Delta^r_{u\tilde{d}_{3T}(e_j, x)e_j} F(x)|^p dx du \tag{17.2}$$

(recall that here $\iint_{3T}^{*} |\Delta^r|^p$ means that the integrand is 0 if any of the arguments in Δ^r is outside $3T$). Since the side-length of $3T$ is $6d_T$ and $10d_T/m \leq 6d_T$, we can employ Lemma 16.3 on every horizontal and every vertical segment of $3T$ in replacing the integrand on the right-hand side by

$$|\Delta^r_{u10d_T e_j} F(x)|^p,$$

and then the substitution $u10d_T^{1/2} = v$ gives that

$$\begin{aligned} \iint_{3T} |F - Q_m|^p \quad &\prec \quad \sum_{j=1}^{2} \frac{m}{10d_T^{1/2}} \int_0^{10d_T^{1/2}/m} \iint_{3T}^{*} |\Delta^r_{vd_T^{1/2} e_j} F(x)|^p dx dv \\ &=: \quad \sum_{j=1}^{2} \mathcal{J}(T, m, e_j, F). \end{aligned} \tag{17.3}$$

Next, we need the analogue of (5.3), namely

$$\|\varphi^r H_m^{(r)}\|_{L^p[-1,1]} \leq M_r m^r \omega_\varphi^r (H_m, m^{-1})_p \tag{17.4}$$

for any polynomial H_m of a single variable and of degree at most m (here $\omega_\varphi^r(f,t)_p = \omega_{[-1,1]}^r(f,t)_p$ is the modulus of smoothness from (1.1) and (5.2), with the supremum norms there replaced by the $L^p[-1,1]$-norm). Indeed, just follow the proof of (5.3), and use that, by [12, Theorem 7.3.1], (5.1) is true in L^p spaces, as well. By the usual linear transformations we get from (17.4) on any segment I

$$\left\| \tilde{d}_I(\overline{e},x)^r \frac{\partial^r H_m(x)}{\partial \overline{e}^r} \right\|_{L^p(I)} \le M_r m^r \omega_I^r(H_m, m^{-1})_p, \tag{17.5}$$

where $\overline{e}$ is the direction of I.

Consider now one of the directions $\overline{e}_j$ from the beginning of this section, and all chords of $3T$ in that direction that intersect $2T$. Let the union of these chords be T_0. Then $2T \subset T_0 \subset 3T$. If, for $H_m = Q_m$, we apply on each chord of T_0 in the direction of $\overline{e}_j$ the inequality (17.5), then we can conclude (use also (16.1))

$$\iint_{T_0} \tilde{d}_{3T}(\overline{e}_j,x)^{rp} \left| \frac{\partial^r Q_m(x)}{\partial \overline{e}_j^r} \right|^p dx$$

$$\prec m^{rp} m \int_0^{1/m} \iint_{T_0}^* |\Delta^r_{u\tilde{d}_{3T}(\overline{e}_j,x)\overline{e}_j} Q_m(x)|^p dx du, \tag{17.6}$$

and call here the right-hand side $m^{rp}\mathcal{J}(Q_m)$. Now

$$\mathcal{J}(Q_m) \le 2^{p-1}(\mathcal{J}(F) + \mathcal{J}(F-Q_m)),$$

and here

$$\mathcal{J}(F-Q_m) \le m \int_0^{1/m} \iint_{3T}^* |\Delta^r_{u\tilde{d}_{3T}(\overline{e}_j,x)\overline{e}_j}(F-Q_m)(x)|^p dx du.$$

In view of (16.10) the right-hand side increases if we replace the inner integrals by the corresponding integrals of $F - Q_m$ over $3T$, and so

$$\mathcal{J}(F-Q_m) \prec m \int_0^{1/m} \left(\iint_{3T} |F-Q_m|^p \right) du = \iint_{3T} |F-Q_m|^p, \tag{17.7}$$

for which (17.3) can be applied.

Next, consider

$$\mathcal{J}(F) = m \int_0^{1/m} \iint_{T_0}^* |\Delta^r_{u\tilde{d}_{3T}(\overline{e}_j,x)\overline{e}_j} F(x)|^p dx du.$$

Note now that each segment of T_0 in the direction of $\overline{e}_j$ is of length $\le 6\sqrt{2}d_T \le 10d_T$ and $\ge 2d_T \ge 10d_T/m$, so we can apply Lemma 16.3 (with $D = 10d_T$) on every such segment to replace on the right

$$|\Delta^r_{u\tilde{d}_{3T}(\overline{e}_j,x)\overline{e}_j} F(x)|^p$$

by

$$|\Delta^r_{u10d_T\overline{e}_j}F(x)|^p,$$

and then the substitution $u10d_T^{1/2}=v$ gives, as in (17.3),

$$\mathcal{J}(F)\prec\mathcal{J}(T,m,\overline{e}_j,F). \tag{17.8}$$

Note also that,

$$ru\tilde{d}_{3T}(\overline{e}_j,x)/2\leq\frac{r}{2m}5d_T\leq d_T,$$

so the integral $\int^*_{T_0}$ on the left of (17.6) is at least as large as the integral $\int_{2T}$ (there is no * here!) with the same integrand, and on $2T$ we have

$$\tilde{d}_{3T}(\overline{e}_j,x)\sim d_T.$$

So we obtain from (17.6)–(17.8)

$$\iint_{2T}\left|\frac{\partial^rQ_m}{\partial\overline{e}_j^r}\right|^p\prec\frac{m^{rp}}{d_T^{rp}}\sum_{j=0}^r\mathcal{J}(T,m,\overline{e}_j,F),$$

where

$$\mathcal{J}(T,m,\overline{e}_j,F)=\frac{m}{10d_T^{1/2}}\int_0^{10d_T^{1/2}/m}\iint^*_{3T}|\Delta^r_{ud_T^{1/2}\overline{e}_j}F(x)|^pdxdu, \tag{17.9}$$

and where we have also used that $e_1=\overline{e}_0$ and $e_2=\overline{e}_r$.

Now, exactly as in the proof of Proposition 5.1, the r-th directional derivative in the direction e (the base edge direction we are interested in) can be bound by the r-th directional derivatives in the direction of $\overline{e}_j$, $j=0,1,\ldots,r$, hence

$$\iint_{2T}\left|\frac{\partial^rQ_m(x)}{\partial e^r}\right|^p\prec\frac{m^{rp}}{d_T^{rp}}\sum_{j=0}^r\mathcal{J}(T,m,\overline{e}_j,F) \tag{17.10}$$

also follows.

Next, recall the sets $K_a=S_a\setminus S_{a/4}$ from (6.1), and let us assume that $x\in T\cap K_a$ for some $a\leq\beta/2$ with the β from (15.4) or (17.1). To estimate the r-th symmetric difference in the direction of e we use

$$\left|\Delta^r_{h\tilde{d}_S(e,x)e}Q_m(x)\right|\prec(h\tilde{d}_S(e,x))^{r-1}\int_{-rh\tilde{d}_S(e,x)/2}^{rh\tilde{d}_S(e,x)/2}\left|\frac{\partial^rQ_m(x+te)}{\partial e^r}\right|dt \tag{17.11}$$

(see [12, (2.4.5)]). Lemma 15.3 gives that $d_S(e,x)\prec d_T$, so in the limits of integration we get

$$r\tilde{d}_S(e,x)/2\leq M_2(d_Ta)^{1/2} \tag{17.12}$$

with some fixed M_2. Recall also that $d_T \geq M/12n^2$ by (15.5), so for $M \geq 12M_2^2$ and for $h \leq 1/n\sqrt{a}$ we have

$$rh\tilde{d}_S(e,x)/2 \leq M_2 h(d_T a)^{1/2} \leq \frac{M_2 d_T^{1/2}}{n} \leq d_T.$$

For this choice we get for $x \in T \cap K_a$

$$\left|\Delta^r_{h\tilde{d}_S(e,x)e} Q_m(x)\right|^p \prec (hM_2(d_T a)^{1/2})^{(r-1)p} \times$$

$$\times \left(\int_{-hM_2(d_T a)^{1/2}}^{hM_2(d_T a)^{1/2}} \left|\frac{\partial^r Q_m(x+te)}{\partial e^r}\right| dt\right)^p.$$

Now apply Hölder's inequality on the right-hand side and integrate the obtained inequality for $x \in T \cap K_a$. Noting that

$$t \in [-hM_2(d_T a)^{1/2}, hM_2(d_T a)^{1/2}] \subset [-d_T, d_T],$$

and so for $x \in T$

$$x + te \in 2T, \tag{17.13}$$

we obtain this way

$$\iint_{T\cap K_a} \left|\Delta^r_{h\tilde{d}_S(e,x)e} Q_m(x)\right|^p dx$$

$$\prec (hM_2(d_T a)^{1/2})^{rp-1} \int_{-hM_2(d_T a)^{1/2}}^{hM_2(d_T a)^{1/2}} \iint_{2T} \left|\frac{\partial^r Q_m(y)}{\partial e^r}\right|^p dy dt$$

$$\prec (h(d_T a)^{1/2})^{rp} \iint_{2T} \left|\frac{\partial^r Q_m}{\partial e^r}\right|^p.$$

To the right-hand side we can apply (17.10) (see also (17.9)) to get

$$\iint_{T\cap K_a} \left|\Delta^r_{h\tilde{d}_S(e,x)e} Q_m(x)\right|^p dx \prec (h(d_T a)^{1/2})^{rp} \frac{m^{rp}}{d_T^{rp}} \sum_{j=0}^{r} \mathcal{J}(T, m, \overline{e}_j, F). \tag{17.14}$$

Now if we replace in a similar integral for $|\Delta^r F|$ the difference $\Delta^r F$ by $\Delta^r(F - Q_m) + \Delta^r(Q_m)$ and apply

$$\iint_{T\cap K_a} \left|\Delta^r_{h\tilde{d}_S(e,x)e}(F - Q_m(x))\right|^p dx \prec \iint_{2T} |F - Q_m|^p \tag{17.15}$$

(see (16.10) with $K_1 = 2T$ and $K_2 = S$, and use also (17.13)), then we can conclude from (17.3) and (17.14)

$$\begin{aligned}\iint_{T\cap K_a} \left|\Delta^r_{h\tilde{d}_S(e,x)e}F(x)\right|^p dx \quad &\prec \quad \sum_{j=1}^{2} \mathcal{J}(T,m,e_j,F) \\ &+ \quad (h(d_T a)^{1/2})^{rp}\frac{m^{rp}}{d_T^{rp}}\sum_{j=0}^{r}\mathcal{J}(T,m,\overline{e}_j,F).\end{aligned}$$

Up to this point m was arbitrary, and $h \le 1/n\sqrt{a}$. Now if $m = \lceil n10d_T^{1/2}\rceil$ and $h \le 1/n\sqrt{a}$, then

$$h(d_T a)^{1/2}\frac{m}{d_T} \prec 1,$$

and we obtain

$$\iint_{T\cap K_a} \left|\Delta^r_{h\tilde{d}_S(e,x)e}F(x)\right|^p dx \prec \sum_{j=0}^{r}\mathcal{J}(T,m,\overline{e}_j,F) \tag{17.16}$$

(recall that $e_1 = \overline{e}_0$ and $e_2 = \overline{e}_r$).

The following observation will be important. The argument from (17.11) to (17.16) shows that (17.16) is true when, on the left-hand side, $\tilde{d}_S(e,x)$ is replaced by some smaller quantity $\delta(x) = \tilde{d}_{S_0}(e,x) \le \tilde{d}_S(e,x)$, $S_0 \subset S$:

$$\iint_{T\cap K_a} \left|\Delta^r_{h\delta(x)e}F(x)\right|^p dx \prec \sum_{j=0}^{r}\mathcal{J}(T,m,\overline{e}_j,F), \tag{17.17}$$

and here $\prec$ does not depend on $\delta(x) = \tilde{d}_{S_0}(e,x) \le \tilde{d}_S(e,x)$. Indeed, this follows from two facts:

- the right-hand side of (17.11) is decreasing in $\tilde{d}_S(e,x)$ in the sense that if we replace it with $\delta(x) = \tilde{d}_{S_0}(e,x) \le \tilde{d}_S(e,x)$, then the right-hand side decreases,
- by (16.15), if $\delta(x) = \tilde{d}_{S_0}(e,x)$, $S_0 \subset S$, then

$$\int_{2T}^{*} |\Delta^r_{u\delta(x)e}g(x)|^p dx \le C\int_{2T}^{*} |g(x)|^p dx.$$

If, instead of (17.15), we use the latter fact with $g = F - Q_m$ and follow the proof from (17.11) to (17.16), then we obtain (17.17).

On the right-hand side in (17.16)–(17.17) we have the sum of

$$\mathcal{J}(T,m,\overline{e}_j,F) \prec n\int_0^{1/n}\iint_{3T} |\Delta^r_{ud_T^{1/2}\overline{e}_j}F(x)|^p dx du \tag{17.18}$$

(no * in the integral!) for $j = 0, 1 \dots, r$. In what follows we are going to estimate these terms for $j = 0, 1, \dots, r$, and to this end we mention first of all the following. Consider the line through a point x lying sufficiently close to the apex, say x is lying in S_β with the small β from (17.1). As in (1.5)–(1.6), this line intersects S in a segment $A_{\overline{e}_j,x}B_{\overline{e}_j,x}$. The minimum of the distances between x and $A_{\overline{e}_j,x}, B_{\overline{e}_j,x}$ is $d_S(\overline{e}_j, x)$, for which Lemmas 15.1, 15.2 can be applied. However, the other distance between x and $A_{\overline{e}_j,x}, B_{\overline{e}_j,x}$, say $\mathrm{dist}(x, B_{\overline{e}_j,x})$, is bounded by a number depending only on S, and (more importantly), it is at least 1, since x lies in the half-space $\{x \,|\, x_1 \le 1\}$, while $B_{\overline{e}_j,x}$ lies on the base of S, which is in the hyperplane $\{y \,|\, y_1 = 2\}$. Therefore, if $j = 1, 2, \dots, r-1$, then Lemma 15.2 gives that for $x \in 3T$ we have

$$d_T^{1/2} \le \tilde{d}_S(\overline{e}_j, x) \prec d_T^{1/2}, \tag{17.19}$$

hence, for such j, Lemma 16.1 (cf. also its Corollary 16.2) shows that

$$n \int_0^{1/n} \iint_{3T} |\Delta^r_{u d_T^{1/2} \overline{e}_j} F(x)|^p dx du \prec n \int_0^{1/n} \iint_{3T} |\Delta^r_{u \tilde{d}_S(\overline{e}_j,x)\overline{e}_j} F(x)|^p dx du. \tag{17.20}$$

When $j = 0$ or $j = 1$, then this is not necessarily true, since then d_T can be much smaller than $d_S(\overline{e}_j, x)$. Consider e.g. the case $j = 0$, the $j = r$ case is completely parallel. Then $\overline{e}_0 = e_1$, and if $T \in \mathcal{T}$ is a small square close to the side ℓ_1 of V and far from U (see Figure 17.1), then d_T is much smaller than $d_S(\overline{e}_0, x)$, and care should be exercised when we want to replace $d_T^{1/2}$ by $\tilde{d}_S(\overline{e}_0, x)$ like in (17.20). Indeed, (17.20) may not be true in such cases. Note however, that (17.20) is still true if $\mathrm{dist}(T, U) \le 20 d_T$ or $\mathrm{dist}(T, \ell_1) \ne \mathrm{dist}(T, \partial V)$, since then, by Lemma 15.1, we have (17.19).

So let $\mathcal{T}^1 = \mathcal{T}_n^1(H)$ be the collection of all $T \in \mathcal{T}_n(H)$ for which $\mathrm{dist}(T, U) > 20 d_T$ and $\mathrm{dist}(T, \ell_1) = \mathrm{dist}(T, \partial V)$. These are the squares in $\mathcal{T}_n$ that lie much closer to ℓ_1 than to the rest of the boundary of V (see Figure 17.1). In a similar manner, let $\mathcal{T}^2 = \mathcal{T}_n^2(H)$ be the collection of all $T \in \mathcal{T}_n(H)$ for which $\mathrm{dist}(T, U) > 20 d_T$ and $\mathrm{dist}(T, \ell_2) = \mathrm{dist}(T, \partial V)$. These are the squares in $\mathcal{T}_n$ that lie much closer to ℓ_2 than to the rest of the boundary of V.

Let now $a \le \beta/2$ with the β from (15.1) or (17.1). We take the sum of both sides in (17.16) for all T with $T \cap K_a \ne \emptyset$, apply (17.18) and (17.20), and use that, by Lemma 15.5, for all T in question the relation $3T \subset \tilde{K}_a$ is true with $\tilde{K}_a$ in that lemma. If we recall that, by (15.4) (see also (17.1)) the union of all such $T \cap K_a$ cover the set

$$K_{a,n} := \{x \in K_a \,|\, \mathrm{dist}(x, \partial S) \ge M_1/n^2\}, \tag{17.21}$$

then we obtain this way the estimate

$$\sup_{h \le 1/n\sqrt{a}} \iint_{H \cap K_{a,n}} \left|\Delta^r_{h \tilde{d}_S(e,x) e} F(x)\right|^p dx \prec \Theta_0(a, H) + \Theta_1(a, H) + \Theta_2(a, H), \tag{17.22}$$

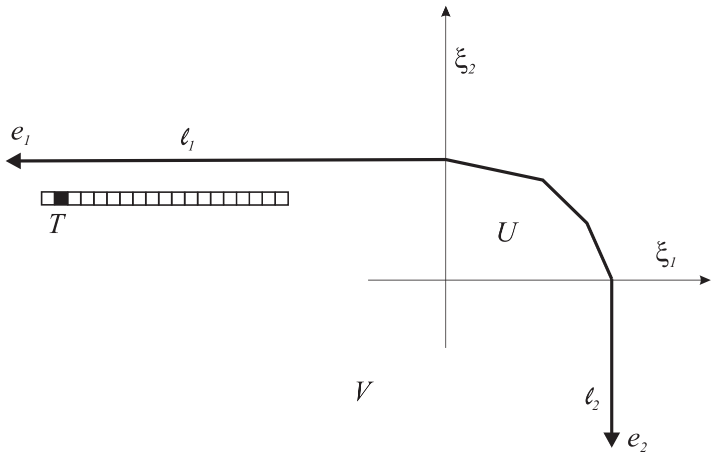

Figure 17.1: Squares in the dyadic decomposition lying closer to ℓ_1 than to the rest of the boundary of V

where

$$\begin{aligned}\Theta_0(a,H) \quad &+ \quad \Theta_1(a,H)+\Theta_2(a,H) := \qquad (17.23)\\ &\prec \quad \sum_{j=0}^{r} n\int_0^{1/n}\iint_{H\cap\tilde{K}_a} |\Delta^r_{u\tilde{d}_S(\overline{e}_j,x)\overline{e}_j}F(x)|^p dxdu\\ &+ \quad \sum_{T\in\mathcal{T}^1(H),\ T\cap K_a\neq\emptyset} n\int_0^{1/n}\iint_{3T} |\Delta^r_{ud_T^{1/2}e_1}F(x)|^p dxdu\\ &+ \quad \sum_{T\in\mathcal{T}^2(H),\ T\cap K_a\neq\emptyset} n\int_0^{1/n}\iint_{3T} |\Delta^r_{ud_T^{1/2}e_2}F(x)|^p dxdu.\end{aligned}$$

This estimate depends on $H = H_\tau$, and recall that H_τ was a translate of the plane $\langle E_1, E_2\rangle$ spanned by apex edges E_1 and E_2 by $\tau\mathbf{n}$, where $\mathbf{n}$ is a normal vector to that plane. In what follows we shall need to take the integrals of these estimates for all $H = H_\tau$ by which we mean integration with respect to the parameter τ from $\tau = 0$ to $\tau = \tau_0$ (here τ_0 is from (15.1)).

We shall apply (17.22) with

$$a = b_k = \frac{2^k + L}{n^2}$$

for k satisfying

$$\frac{L}{n^2} \le \frac{2^k}{n^2} \le \frac{\beta}{4},$$

where L is some large number to be specified later. If we integrate the left-hand side of (17.22) with respect to $H = H_\tau$, $0 \le \tau \le \tau_0$ and sum for all such k (which is the same as summing for all such k and integrating with respect to H), then we get a quantity that is at least as large as

$$\sum_k \sup_{h \le 1/n\sqrt{b_k}} \iint_{K_{b_k,n}} \left|\Delta^r_{h\tilde{d}_S(e,x)e} F(x)\right|^p dx. \tag{17.24}$$

At the same time, the same operation (i.e integration with respect to H and summation with respect to k) when applied to $\Theta_0(H, b_k)$ gives a quantity that is bounded by the sum of the directional moduli of smoothness:

$$\sum_k \int \Theta_0(H, b_k) dH \tag{17.25}$$

$$\begin{aligned}
&= \sum_k \int \left(\sum_{j=0}^r n \int_0^{1/n} \iint_{H \cap \tilde{K}_{b_k}} |\Delta^r_{u\tilde{d}_S(\overline{e}_j,x)\overline{e}_j} F(x)|^p dx du \right) dH \\
&\prec \sum_{j=0}^r n \int_0^{1/n} \int \left(\iint_{S \cap H} |\Delta^r_{u\tilde{d}_S(\overline{e}_j,x)\overline{e}_j} F(x)|^p dx \right) dH du \\
&\prec \sum_{j=0}^r n \int_0^{1/n} \int_S |\Delta^r_{u\tilde{d}_S(\overline{e}_j,x)\overline{e}_j} F(x)|^p dx du \prec \omega^r_S(F, 1/n)^p_p,
\end{aligned}$$

where we used that, for $2^k \ge L$, no point can lie in more than 8 of the sets $\tilde{K}_{b_k}$. Indeed, if x belongs to $\tilde{K}_{b_k}$ (see (15.7) for the definition of the set $\tilde{K}_{b_k}$), then for its first coordinate (in the (x_1, x_2, x_3) coordinate system of $\mathbf{R}^3$) we have

$$\frac{2^k + L}{16n^2} \le x_1 \le 8\frac{2^k + L}{n^2} < 16\frac{2^k}{n^2},$$

and for an x_1 there are at most 8 such k.

It is clear that

$$\sum_k \int \Theta_1(H, b_k) dH \tag{17.26}$$

$$\prec \int \left(n \int_0^{1/n} \sum_{T \in \mathcal{T}^1(H)} \iint_{3T} |\Delta^r_{ud_T^{1/2} e_1} f(x)|^p dx du \right) dH$$

where we used that no T can intersect more than 8 of the sets K_{b_k}. Indeed, by Lemma 15.5 if T intersects K_{b_k}, then $3T \subset \tilde{K}_{b_k}$, and we have just seen that this can only happen for at most 8 of the k's. In what follows we are going to show that

$$n\int_0^{1/n}\sum_{T\in\mathcal{T}^1(H)}\iint_{3T}|\Delta^r_{ud_T^{1/2}e_1}F(x)|^p dxdu$$

$$\prec n\int_0^{1/n}\int_{S\cap H}|\Delta^r_{u\tilde{d}_S(e_1,x)e_1}F(x)|^p dxdu. \tag{17.27}$$

Then this and (17.26) imply (via integration with respect to dH)

$$\sum_k\int\Theta_1(H,b_k)dH\prec n\int_0^{1/n}\int_S|\Delta^r_{u\tilde{d}_S(e_1,x)e_1}F(x)|^p dxdu\prec\omega_S^r(F,1/n)_p^p. \tag{17.28}$$

A similar procedure yields

$$\begin{aligned}&\sum_k\int\Theta_2(H,b_k)dH \\ &\prec\quad n\int_0^{1/n}\int_S|\Delta^r_{u\tilde{d}_S(e_2,x)e_2}F(x)|^p dxdu\prec\omega_S^r(F,1/n)_p^p,\end{aligned} \tag{17.29}$$

which, together with the preceding estimates imply

$$\sum_k\sup_{h\le 1/n\sqrt{b_k}}\iint_{K_{b_k,n}}\left|\Delta^r_{h\tilde{d}_S(e,x)e}F(x)\right|^p dx\prec\omega_S^r(F,1/n)_p^p. \tag{17.30}$$

Recall that here the summation is for all k with

$$\frac{L}{n^2}\le\frac{2^k}{n^2}\le\frac{\beta}{4}. \tag{17.31}$$

If we start out with (17.17) instead of (17.16), then we get instead of (17.30) the estimate

$$\sum_k\sup_{h\le 1/n\sqrt{b_k}}\iint_{K_{b_k,n}}\left|\Delta^r_{h\delta(x)e}F(x)\right|^p dx\prec\omega_S^r(F,1/n)_p^p, \tag{17.32}$$

for any $\delta(x)=\tilde{d}_{S_0}(e,x)\le\tilde{d}_S(e,x)$, $S_0\subset S$ (recall also (17.21) for the definition of the sets $K_{b_k,n}$).

To complete the proof of (17.30)–(17.32), we still need to prove (17.27). Let $T\in\mathcal{T}^1$, and note that if (in the (ξ,ξ_2) coordinate system) we shift T to the left into a dyadic square, then the shifted square also belongs to $\mathcal{T}^1$ unless it gets outside the set S_β (see (15.2) and Figure 17.1). Thus, if we introduce the equivalence relation $T_1\sim T_2$ on $\mathcal{T}^1$ meaning that T_1 and T_2 can be obtained from each other by a horizontal shift, then the equivalence classes are consisting of continuous chains of dyadic squares of equal side-length lying in a horizontal strip. Let $\mathcal{A}$ be an equivalence class and

$$A = A(\mathcal{A}) = \bigcup_{T\in\mathcal{A}} 3T.$$

Then A is a rectangle with horizontal and vertical sides and with vertical side-length equal to $6d_T$, where $T\in\mathcal{A}$ is an arbitrary element; say $A=[\alpha,\beta]\times[\gamma,\delta]$ where $\delta-\gamma=6d_T$. Note that even though the different equivalence classes are disjoint, the sets $A(\mathcal{A})$ need not be disjoint, but it follows from Lemma 15.4,(c) that any point can belong to at most 80^2 such $A(\mathcal{A})$. By the same lemma every $x\in A(\mathcal{A})$ can belong to at most 80^2 squares $3T$ with $T\in A(\mathcal{A})$ (actually only to 6 such $3T$, but that is indifferent), therefore it follows that

$$\begin{aligned}\sum_{T\in\mathcal{A}}\iint_{3T}|\Delta^r_{ud_T^{1/2}e_1}F(x)|^p dx \quad &\prec \quad \iint_A |\Delta^r_{ud_T^{1/2}e_1}F(x)|^p dx \\ &= \quad \int_\gamma^\delta\int_\alpha^\beta |\Delta^r_{ud_T^{1/2}e_1}F(\xi_1,\xi_2)|^p d\xi_1\xi_2,\end{aligned} \tag{17.33}$$

where, on the plane H, we identified a point x by its (ξ_1,ξ_2)-coordinates. For $\xi_2\in(\gamma,\delta)$ let $I_{\xi_2}=\{x=(\xi_1,\xi_2)\in V\}$ be the chord of V consisting of those points for which the second coordinate is ξ_2. We are going to show that

$$n\int_0^{1/n}\int_\alpha^\beta |\Delta^r_{ud_T^{1/2}e_1}F(\xi_1,\xi_2)|^p d\xi_1 du \prec n\int_0^{1/n}\int^*_{I_{\xi_2}} |\Delta^r_{u\tilde{d}_S(e_1,x)e_1}F(\xi_1,\xi_2)|^p d\xi_1 du. \tag{17.34}$$

If we integrate this inequality with respect to ξ_2 on $[\gamma,\delta]$, then we obtain (see also (17.33))

$$\begin{aligned} &n\int_0^{1/n}\left(\sum_{T\in\mathcal{A}}\iint_{3T}|\Delta^r_{ud_T^{1/2}e_1}F(x)|^p dx\right)du \\ &\prec n\int_0^{1/n}\iint^*_{V\cap\,\mathrm{Strip}(\mathcal{A})}|\Delta^r_{u\tilde{d}_S(e_1,x)e_1}F(x)|^p dx du,\end{aligned}$$

where $\mathrm{Strip}(\mathcal{A})$ is the strip $\{x=(\xi_1,\xi_2)\mid \xi_2\in[\gamma,\delta]\}$. Now if we sum this up for all equivalence classes $\mathcal{A}$ and take into account that no point can belong to more than 80^2 such strips $\mathrm{Strip}(\mathcal{A})$ (see Lemma 15.4,(c)), we obtain (17.27), pending the proof of (17.34).

To prove (17.34), note first of all that $[\alpha-d_T,\beta+d_T]\subset I_{\xi_2}$, and for $u\le 1/n$ we have $rud_T^{1/2}/2\le d_T$ (cf. (15.5)), so

$$n\int_0^{1/n}\int^*_{[\alpha-d_T,\beta+d_T]}|\Delta^r_{ud_T^{1/2}e_1}F(\xi_1,\xi_2)|^p d\xi_1 du$$

$$\prec n\int_0^{1/n}\int_{I_{\xi_2}}^* |\Delta^r_{u\tilde{d}_S(e_1,x)e_1}F(\xi_1,\xi_2)|^p d\xi_1 du. \tag{17.35}$$

is a stronger inequality than (17.34). Finally, since $\tilde{d}_S(e_1,x)=\tilde{d}_{I_{\xi_2}}(e_1,x)$, and on $[\alpha-d_T,\beta+d_T]$ we have $d_T^{1/2}\le \tilde{d}_{I_{\xi_2}}(e_1,x)$, (17.35) follows from Lemma 16.4 with the choice $I=I_{\xi_2}$, $J=[\alpha-d_T,\beta+d_T]$, $\tau=d_T^{1/2}/|J|$, $\delta=1/n$ and $g(\xi)=F(\xi,\xi_2)$.

This completes the proof of (17.30). Note that since e is a base edge direction, in (17.30) we have $\tilde{d}_S(e,x)=\tilde{d}_{K_{b_k}}(e,x)$ for $x\in K_{b_k}$. So in (17.32) we can set $S_0=K_{b_k,n}$, $\delta(x)=\tilde{d}_{K_{b_k,n}}(e,x)$ (see (17.21) for the definition of the sets $K_{b_k,n}$), and we get from (17.32)

$$\sum_k \sup_{h\le 1/n\sqrt{b_k}} \iint_{K_{b_k,n}} \left|\Delta^r_{h\tilde{d}_{K_{b_k,n}}(e,x)e}F(x)\right|^p dx \prec \omega_S^r(F,1/n)_p^p. \tag{17.36}$$

Since e was any base edge direction, this completes the discussion on base edge directions.

We still need to prove an analogue for apex edge directions, for example for the direction e_1 we need to prove

$$\sum_k \sup_{h\le 1/n\sqrt{b_k}} \int_{K_{b_k,n}}^* \left|\Delta^r_{h\tilde{d}_{K_{b_k,n}}(e_1,x)e_1}F(x)\right|^p dx \prec \omega_S^r(F,1/n)_p^p. \tag{17.37}$$

Once this is done, we get from (17.36) and (17.37) (as well as from the analogue of (17.37) for all other apex edge directions) the estimate

$$\sum_k \max_{e\in\mathcal{E}} \sup_{h\le 1/n\sqrt{b_k}} \int_{K_{b_k,n}}^* \left|\Delta^r_{h\tilde{d}_{K_{b_k,n}}(e,x)e}F(x)\right|^p dx \prec \omega_S^r(F,1/n)_p^p, \tag{17.38}$$

where $\mathcal{E}$ is the direction of all edges of S, which is the same as the edge directions of each K_{b_k} (see the definition (6.1) of the sets K_η). This is the analogue of (6.18) we have been seeking.

In view of (16.1), it is enough to prove the version of (17.37) when, on the left-hand side, the supremum is replaced by integrals, i.e. to prove

$$\sum_k n\sqrt{b_k}\int_0^{1/n\sqrt{b_k}} \int_{K_{b_k,n}}^* \left|\Delta^r_{u\tilde{d}_{K_{b_k,n}}(e_1,x)e_1}F(x)\right|^p dxdu \prec \omega_S^r(F,1/n)_p^p. \tag{17.39}$$

The proof of this uses the technique employed so far. Since for $x\in K_{b_k,n}$ we have

$$\tilde{d}_{K_{b_k,n}}(x,e_1)\le 2\sqrt{b_k}\tilde{d}_S(x,e_1),$$

Lemma 16.5, (16.31) easily gives (see the argument below) that each individual term on the left of (17.39) is bounded by the right-hand side. However, we need the stronger statement that the sum of these terms is also bounded by the right-hand side.

Consider again the segment I_{ξ_2} from the preceding proof. Its intersection with the set of points that lie of distance $\geq M_1/n^2$ from the boundary of S is some interval, say

$$I_{\xi_2} \cap \left\{ z \,\middle|\, \operatorname{dist}(z, S) \geq \frac{M_1}{n^2} \right\} =: [Z, W],$$

where (if this intersection is not empty) the point W lies on the base of S, so it has 2 as its first coordinate (in the (x_1, x_2, x_3) coordinate system of $\mathbf{R}^3$), and where M_1 is from (17.1) (see also (17.21)). Let Z_1 be the first coordinate of Z. Then the projection of the intersection $I_{\xi_2} \cap K_{b_k,n}$ for the definition of the sets $K_{a,n}$) onto the x_1-axis is

$$[Z_1, 2] \cap \left[\frac{b_k}{2}, 2b_k\right],$$

and typically this has length $3b_k/2$. However, if $b_k/2 < Z_1 < 2b_k (\leq \beta)$, i.e. when

$$\frac{1}{2}\frac{2^k + L}{n^2} \leq Z_1 \leq 2\frac{2^k + L}{n^2},$$

then this intersection is of length smaller than $3b_k/2$. In view of $2^k \geq L$, the preceding inequality implies

$$\frac{1}{4} Z_1 n^2 \leq 2^k < 2 Z_1 n^2,$$

and there are at most 3 such k. In a similar fashion, unless $b_k/2$ lies closer to Z_1 than $b_k/4$, the distance from any point of $[b_k/2, 2b_k]$ to Z_1 is at least $b_k/4$, and there are at most four k's for which this is not the case (i.e. when $b_k/2 < Z_1 + b_k/4$).

Thus, the x_1-axis (in the (x_1, x_2, x_3) coordinate system of $\mathbf{R}^3$) intersects the set $K_{b_k,n}$ in a segment of length $3b_k/2$ or less, hence the sets $K_{b_k,n}$ divide the segment I_{ξ_2} into overlapping subsegments J_k of length $\alpha(3/2)b_k$ or less (only the first three of these segments lying closest to the ξ_2 axis may be shorter), where $\alpha = 1/\sin\theta$, with θ the angle between the direction e_1 and the base $x_1 = 2$. Furthermore, according to what has just been said, except maybe for the first four of these segments (for which the following argument is valid in view of Lemma 16.5, (16.31)), we have on J_k the relations

$$\tilde{d}_{K_{b_k,n}}(e_1, x) \leq \alpha b_k, \qquad \sqrt{\alpha b_k/4} \leq \tilde{d}_S(e_1, x) \leq C\sqrt{\alpha b_k},$$

hence, first

$$\int_0^{1/n\sqrt{b_k}} \int_{J_k}^{*} \left|\Delta^r_{u\tilde{d}_{K_{b_k,n}}(e_1,x)e_1} F(x)\right|^p dx du \prec \int_0^{1/n\sqrt{b_k}} \int_{J_k}^{*} \left|\Delta^r_{u2\alpha b_k e_1} F(x)\right|^p dx du$$

by Lemma 16.3, and then by Lemma 16.1

$$\int_0^{1/n\sqrt{b_k}} \int_{J_k}^{*} \left|\Delta^r_{u2\alpha b_k e_1} F(x)\right|^p dxdu \prec \int_0^{1/n\sqrt{b_k}} \int_{J_k} \left|\Delta^r_{u4\sqrt{\alpha b_k}\tilde{d}_S(e_1,x)e_1} F(x)\right|^p dxdu$$

(no * in the integral on the right!). Therefore,

$$\begin{aligned} & n\sqrt{b_k} \int_0^{1/n\sqrt{b_k}} \int_{J_k}^{*} \left|\Delta^r_{u\tilde{d}_{K_{b_k,n}}(e_1,x)e_1} F(x)\right|^p dxdu \\ \prec \quad & n\sqrt{b_k} \int_0^{1/n\sqrt{b_k}} \int_{J_k} \left|\Delta^r_{u4\sqrt{\alpha b_k}\tilde{d}_S(e_1,x)e_1} F(x)\right|^p dxdu \\ = \quad & \frac{n}{4\sqrt{\alpha}} \int_0^{4\sqrt{\alpha}/n} \int_{J_k} \left|\Delta^r_{v\tilde{d}_S(e_1,x)e_1} F(x)\right|^p dxdv. \end{aligned}$$

If we integrate for ξ_2 (recall that $\cup J_k = I_{\xi_2}$) and then for $dH = dH_\tau$ (i.e. for $d\tau$), then we get

$$\begin{aligned} & n\sqrt{b_k} \int_0^{1/n\sqrt{b_k}} \int_{K_{b_k,n}}^{*} \left|\Delta^r_{h\tilde{d}_{K_{b_k,n}}(e_1,x)e_1} F(x)\right|^p dx \\ & \qquad \prec \frac{n}{4\sqrt{\alpha}} \int_0^{4\sqrt{\alpha}/n} \int_{K_{b_k}} \left|\Delta^r_{v\tilde{d}_S(e_1,x)e_1} F(x)\right|^p dxdv. \end{aligned}$$

Now sum this up for all k with the property (17.31) to obtain

$$\sum_k n\sqrt{b_k} \int_0^{1/n\sqrt{b_k}} \int_{K_{b_k,n}}^{*} \left|\Delta^r_{h\tilde{d}_{K_{b_k,n}}(e_1,x)e_1} F(x)\right|^p dx \prec \omega_S^r(F, 4\sqrt{\alpha}/n)_p^p,$$

and, in view of (16.18), on the right-hand side we can replace $\omega_S^r(F, 4\sqrt{\alpha}/n)_p^p$ by $\omega_S^r(F, 1/n)_p^p$.

This completes the proof of (17.39), and with it also the proof of (17.38).

■

18 Local approximation

Here we use the notations and setup from Section 7. So let f be an L^p function on the pyramid S, $v_n = (-L/n^2, 0, 0)$ with a large but fixed L, and

$$F(x) = F_n(x) = f(x - v_n) \quad \text{on} \quad \mathbf{S} = \mathbf{S}^{(n)} = S + v_n. \tag{18.1}$$

In what follows we are going to use again the notations K_1, S_η, K_η from (6.1), but when we apply them to $\mathbf{S}$, then we denote them by $\mathbf{K}_1$, $\mathbf{S}_\eta$, $\mathbf{K}_\eta$. Note that the edge directions for the polytopes $\mathbf{S}$, $\mathbf{K}_1$, $\mathbf{S}_\eta$, $\mathbf{K}_\eta$ is the same as that for S, i.e. it is $\mathcal{E}$. We apply (17.38) on the pyramid $\mathbf{S} = \mathbf{S}^{(n)}$ to the function $F = F_n$, for which it takes the form

$$\sum_k \max_{e\in\mathcal{E}} \sup_{h\le 1/n\sqrt{b_k}} \int_{\mathbf{K}_{b_k,n}(M_1/n^2)} \left|\Delta^r_{h\tilde{d}_{\mathbf{K}_{b_k,n}}(e,x)e} F(x)\right|^p dx \prec \omega^r_{\mathbf{S}}(F,1/n)^p_p, \quad (18.2)$$

where the summation is taken for all k for which (17.31) is true. Recall that here the polytope $\mathbf{K}_{b_k}$ was cut out of $\mathbf{S}$ by two hyperplanes L_1 and L_2 that are parallel with the base and are of distance $b_k/2$ and $2b_k$, $b_k = (2^k + L)/n^2$, resp. from the apex v_n of $\mathbf{S}$, and $\mathbf{K}_{b_k,n}$ is the part of $\mathbf{K}_{b_k}$ that lies of distance $\ge M_1/n^2$ from the boundary of $\mathbf{S}$. If L is sufficiently large (and this is the only requirement on L), then this set includes the portion of S (the original pyramid) that are cut out of S by L_1 and L_2. Now the distance of L_1 resp. L_2 from the origin (which is the ape of S) is

$$\frac{b_k}{2} - \frac{L}{n^2} = \frac{2^k + L}{2n^2} - \frac{L}{n^2} \le \frac{2^{k-1}}{n^2} := \frac{a_k}{2}$$

resp.

$$2b_k - \frac{L}{n^2} = 2\frac{2^k + L}{n^2} - \frac{L}{n^2} \ge \frac{2^{k+1}}{n^2} := 2a_k,$$

so the polytope $\mathbf{K}_{b_k,n}$ includes K_{a_k} with $a_k = 2^k/n^2$. But then we can invoke Corollary 16.6 and (16.5) to conclude from (18.2)

$$\sum_k \max_{e\in\mathcal{E}} \sup_{h\le 1/n\sqrt{b_k}} \int_{K_{a_k}} \left|\Delta^r_{h\tilde{d}_{K_{a_k}}(e,x)e} F(x)\right|^p dx \prec \omega^r_{\mathbf{S}}(F,1/n)^p_p, \quad (18.3)$$

i.e.

$$\sum_k \overline{\omega}^r_{K_{a_k}}\left(F, \frac{1}{n\sqrt{b_k}}\right)^p_p \prec \omega^r_{\mathbf{S}}(F,1/n)^p_p, \quad (18.4)$$

where $\overline{\omega}$ is the modulus of smoothness (14.1). Since $2^k \ge L$, here $b_k = (2^k + L)/n^2 \le 2\cdot 2^k/n^2 = 2a_k$, so an application of (16.18) gives

$$\sum_{L\le 2^k\le \beta n^2/4} \overline{\omega}^r_{K_{a_k}}\left(F, \frac{1}{n\sqrt{a_k}}\right)^p_p \prec \omega^r_{\mathbf{S}}(F,1/n)^p_p. \quad (18.5)$$

For $a > 0$ set now $F^*(x^*) = F(ax^*)$. Then F^* is an L^p function on K_1 such that

$$\Delta^r_{h\tilde{d}_{K_1}(e,x^*)e} F^*(x^*) = \Delta^r_{h\tilde{d}_{K_a}(e,x)e} F(x), \qquad x \in K_a.$$

Since K_1 is a simple polytope (at each vertex there are 3 edges), we can apply Proposition 14.1 from Section 14 to conclude that for $n\sqrt{a} \ge 3r$, i.e. for $a \ge$

$9r^2/n^2$, there are polynomials $P^*_{n\sqrt{a}} = P^*_{a,n\sqrt{a}}$ of 3 variables of degree at most $n\sqrt{a}$ such that

$$\|F^* - P^*_{n\sqrt{a}}\|_{L^p(K_1)} \prec \overline{\omega}^r_{K_1}\left(F^*, \frac{1}{n\sqrt{a}}\right)_p = \overline{\omega}^r_{K_a}\left(F, \frac{1}{n\sqrt{a}}\right)_p.$$

With

$$p_{n\sqrt{a}}(x) = p_{a,n\sqrt{a}}(x) = P^*_{n\sqrt{a}}(x/a)$$

this is the same as

$$\|F - p_{n\sqrt{a}}\|_{L^p(K_a)} \prec \overline{\omega}^r_{K_a}\left(F, \frac{1}{n\sqrt{a}}\right)_p. \tag{18.6}$$

For the polynomials

$$q = p_{n\sqrt{a}} - p_{n\sqrt{2a}} = p_{a,n\sqrt{a}} - p_{2a,n\sqrt{2a}} \tag{18.7}$$

of degree at most $n\sqrt{2a}$ this yields

$$\|q\|_{L^p(K_a\cap K_{2a})} \le \|F - p_{n\sqrt{a}}\|_{L^p(K_a)} + \|F - p_{n\sqrt{2a}}\|_{L^p(K_{2a})},$$

and here

$$K_a \cap K_{2a} = S \cap \{x \,|\, a \le x_1 \le 2a\}.$$

Consider the polar coordinates (r, φ, ψ) in $\mathbf{R}^3$, and let $\ell = \ell_{\varphi,\psi}$ be the half-line $\{(r, \varphi, \psi) \,|\, r \ge 0\}$. We have

$$\int_{K_a} |q|^p \sim a^2 \iint \left(\int_{\ell_{\varphi,\psi}\cap K_a} |q|^p\right) d\varphi d\psi, \tag{18.8}$$

and a similar formula holds when K_a is replaced by $K_a \cap K_{2a}$. We have seen in (7.3) that for $x \in \ell$

$$|q(x)| = |\tilde{q}(x_1)| \le \left(\frac{|x_1 - 3a/2|}{a/2}\right)^{n\sqrt{2a}} \|\tilde{q}\|_{[a,2a]},$$

where x_1 is the first coordinate of $x \in \ell$ and $\tilde{q}(t_1) = q(t)$, $t \in \ell$, is a polynomial of degree at most $n\sqrt{2a}$ in the variable t_1. By Nikolskii's inequality (14.9), on the right

$$\|\tilde{q}\|_{[a,2a]} \prec \frac{(n\sqrt{2a})^{2/p}}{a^{1/p}} \|\tilde{q}\|_{L^p[a,2a]}.$$

For $0 \le x_1 \le a$ this yields (cf. (7.4))

$$|q(x)| \prec 3^{n\sqrt{2a}} \frac{(n\sqrt{2a})^{2/p}}{a^{1/p}} \|\tilde{q}\|_{L^p[a,2a]} \prec \frac{e^{4n\sqrt{a}}}{a^{1/p}} \|\tilde{q}\|_{L^p[a,2a]}, \tag{18.9}$$

while for $2a \le x_1 \le 2$ we obtain (cf. (7.5))

$$|q(x)| \prec \left(\frac{8x_1}{a}\right)^{n\sqrt{2a}} \frac{(n\sqrt{2a})^{2/p}}{a^{1/p}} \|\tilde{q}\|_{L^p[a,2a]} \prec \frac{e^{3n\sqrt{a}\log(8x_1/a)}}{a^{1/p}} \|\tilde{q}\|_{L^p[a,2a]}. \tag{18.10}$$

On the right we have

$$\|\tilde{q}\|_{L^p[a,2a]} \sim \|q\|_{L^p(K_a \cap K_{2a} \cap \ell)},$$

and since a similar formula holds when a is replaced by $2^l a$ with an arbitrary integer l, we obtain for $l \le 0$ from (18.9)

$$\|q\|_{L^p(K_{2^l a} \cap \ell)} \prec 2^{l/p} e^{4n\sqrt{a}} \|q\|_{L^p(K_a \cap K_{2a} \cap \ell)},$$

while for $l \ge 0$ from (18.10)

$$\|q\|_{L^p(K_{2^l a} \cap \ell)} \prec 2^{l/p} e^{3n\sqrt{a}\log(16\cdot 2^l)} \|q\|_{L^p(K_a \cap K_{2a} \cap \ell)}.$$

Raising these to the p-th power and integrating with respect to $d\varphi d\psi$ we get from (18.8)

$$\|q\|_{L^p(K_{2^l a})} \prec 2^{3l/p} e^{4n\sqrt{a}} \|q\|_{L^p(K_a \cap K_{2a})},$$

when $l \le 0$ and (using $a \ge 9r^2/n^2$)

$$\|q\|_{L^p(K_{2^l a})} \prec 2^{3l/p} e^{3n\sqrt{a}\log(16\cdot 2^l)} \|q\|_{L^p(K_a \cap K_{2a})} \prec e^{4nl\sqrt{a}} \|q\|_{L^p(K_a \cap K_{2a})}$$

when $l \ge 0$ (and n is sufficiently large). Recall that here $q = p_{n\sqrt{a}} - p_{n\sqrt{2a}}$ is the polynomial from (18.7), and it has degree at most $n\sqrt{2a}$. All this is for $a \ge 9r^2/n^2$.

We shall use these with $a = a_k$ and $2^l a = 2^l a_k = a_{k_0}$, i.e. with $l = k_0 - k$, for which they take the form

$$\|p_{n\sqrt{a_k}} - p_{n\sqrt{2a_k}}\|_{L^p(K_{a_{k_0}})} \prec e^{3(k_0-k)/p} e^{4n\sqrt{a}_k} \|p_{n\sqrt{a_k}} - p_{n\sqrt{2a_k}}\|_{L^p(K_{a_k} \cap K_{2a_k})}, \tag{18.11}$$

when $k_0 \le k$ and

$$\|p_{n\sqrt{a_k}} - p_{n\sqrt{2a_k}}\|_{L^p(K_{a_{k_0}})} \prec e^{4n(k_0-k)\sqrt{a_k}} \|p_{n\sqrt{a_k}} - p_{n\sqrt{2a_k}}\|_{L^p(K_{a_k} \cap K_{2a_k})} \tag{18.12}$$

when $k_0 \ge k$. Here the restriction on k is that $2^k/n^2 = a_k \ge 9r^2/n^2$, i.e. $2^k \ge 9r^2$, and on k_0 is that $a_{k_0} \le 1/2$, i.e. $2^{k_0} \le n^2/2$.

Moreover, (18.5) and (18.6) give

$$\sum_{L \le 2^k \le \beta n^2/4} \|F - p_{n\sqrt{a_k}}\|^p_{L^p(K_{a_k})} \prec \omega^r_{\mathbf{S}}(F, 1/n)^p_p. \tag{18.13}$$

19 Global L^p approximation excluding a neighborhood of the apex

We shall see in this section how the argument of Section 8 changes in the L^p case.

We use the preceding estimates (18.11) and (18.12) with $k = L, \dots, m$, where m is chosen so that $\beta/8 \le 2^m/n^2 < \beta/4$ (see (15.1) for the definition of β). We also assume that $2^{L-1} > 9r^2$. We combine the polynomials $p_{n\sqrt{a}} = p_{a,n\sqrt{a}}$ with the fast decreasing polynomials

$$R_{n,a}(x) := R_{n,a}^{(4)}(x_1), \tag{19.1}$$

where $R_{n,a}^{(4)}(x_1)$ is the polynomial of the single variable x_1 (the first coordinate of x) from (3.11) with some large A, and set, as in (8.2)

$$P_n = \sum_{k=L}^{m} \left(R_{n,a_k} - R_{n,a_{k-1}}\right) p_{n\sqrt{a}_k} + R_{n,a_{L-1}} p_{n\sqrt{a_L}} + (1 - R_{n,a_m}) p_{n\sqrt{a_m}}. \tag{19.2}$$

This is a polynomial of degree at most Cn with some constant C that depends only on A. We shall estimate the L^p distance of this polynomial from $F(x) = F_n(x) = f(x - v_n)$ (see (18.1)) on the set (cf. (8.3))

$$S_n^* := S \cap \left\{ x \,\middle|\, \frac{2^{L+1}}{n^2} \le x_1 \le \frac{\beta}{8} \right\}. \tag{19.3}$$

Exactly as in Section 8

$$\begin{aligned} P_n - F &= \sum_{k=L}^{m} \left(R_{n,a_k} - R_{n,a_{k-1}}\right)(p_{n\sqrt{a_k}} - F) \\ &+ R_{n,a_{L-1}}(p_{n\sqrt{a_L}} - F) + (1 - R_{n,a_m})(p_{n\sqrt{a_m}} - F), \end{aligned} \tag{19.4}$$

and for $x \in \{x \,|\, a_{k_0} \le x_1 \le a_{k_0+1}\}$ with $L + 1 \le k_0 \le m - 1$ the first sum on the right-hand side can be written in the form

$$\begin{aligned} &\sum_{k=L}^{k_0-1} R_{n,a_k}(p_{n\sqrt{a_k}} - p_{n\sqrt{a_{k+1}}}) + R_{n,a_{k_0}}(p_{n\sqrt{a_{k_0}}} - F) - R_{n,a_{L-1}}(p_{n\sqrt{a_L}} - F) \\ &\quad + \sum_{k=k_0+1}^{m-1} (R_{n,a_k} - 1)(p_{n\sqrt{a_k}} - p_{n\sqrt{a_{k+1}}}) + (R_{n,a_m} - 1)(p_{n\sqrt{a_m}} - F) \\ &\quad -(R_{n,a_{k_0}} - 1)(p_{n\sqrt{a_{k_0+1}}} - F) =: A_1 + A_2 - A_3 + A_4 + A_5 - A_6. \end{aligned} \tag{19.5}$$

Here $-A_3$ cancels the second term, while A_5 cancels the third term on the right of (19.4). Since

$$\{x \,|\, a_{k_0} \le x_1 \le a_{k_0+1}\} \cap S \subseteq K_{a_{k_0+1}},\ K_{a_{k_0}},$$

we obtain from (18.6)

$$\begin{aligned}\|A_2\|_{L^p(K_{a_{k_0}})} &+ \|A_6\|_{L^p(K_{a_{k_0}})} \prec \|p_{n\sqrt{a_{k_0}}} - F\|_{L^p(K_{a_{k_0}})} \\ &+ \|p_{n\sqrt{a_{k_0+1}}} - F\|_{L^p(K_{a_{k_0+1}})}. \end{aligned} \tag{19.6}$$

In A_1 we have $\sum_{k<k_0} R_{n,a_k}(x) \prec 1$ on the set $x \in \{x \mid a_{k_0} \le x_1 \le a_{k_0+1}\}$ (see (19.8) below), so by Jensen's (or Hölder's) inequality

$$A_1^p \prec \sum_{k=L}^{k_0-1} R_{n,a_k} |p_{n\sqrt{a_k}} - p_{n\sqrt{a_{k+1}}}|^p. \tag{19.7}$$

For $L+1 \le k \le k_0 - 1$ and $a_{k_0} \le x_1 \le a_{k_0+1}$ we get from (3.12),

$$\begin{aligned} R_{n,a_k}(x) = R_n^{(4)}(x_1) \le e^{-An\sqrt{a_k}\log(16x_1/a_k)} &\le e^{-An\sqrt{a_k}\log(8a_{k_0}/a_k)} \\ &\le e^{-An\sqrt{a_k}(k_0-k)/2}. \end{aligned} \tag{19.8}$$

So, in view of (18.12) (the summations $\sum_{k_0}$ below are for $L+1 \le k_0 \le m-1$)

$$\begin{aligned}\sum_{k_0} \int_{K_{a_{k_0}}} |A_1|^p &\prec \sum_{k_0} \sum_{k=L}^{k_0-1} e^{-An\sqrt{a_k}(k_0-k)/2} e^{4np(k_0-k)\sqrt{a}_k} \times \\ &\quad \times \|p_{n\sqrt{a_k}} - p_{n\sqrt{2a_k}}\|^p_{L^p(K_{a_k}\cap K_{2a_k})}\end{aligned}$$

Here, for large A,

$$e^{-An\sqrt{a_k}(k_0-k)/2} e^{4np(k_0-k)\sqrt{a}_k} \le e^{-n\sqrt{a_k}(k_0-k)} \le e^{-\sqrt{2^k}(k_0-k)},$$

therefore

$$\sum_{k_0} \int_{K_{a_{k_0}}} |A_1|^p \prec \sum_k \|p_{n\sqrt{a_k}} - p_{n\sqrt{2a_k}}\|^p_{L^p(K_{a_k}\cap K_{2a_k})} \prec \sum_k \|F - p_{n\sqrt{a_k}}\|^p_{L^p(K_{a_k})}.$$

In a similar manner, for $k_0 + 1 \le k \le m-1$ and $a_{k_0} \le x_1 \le a_{k_0+1}$

$$0 \le 1 - R_{n,a_k}(x) = 1 - R_n^{(4)}(x_1) \le e^{-An\sqrt{a_k}},$$

so we obtain from (18.11) for A_4 via another use of Jensen's inequality as in (19.7) (note that in A_4 we have $\sum_{k>k_0}(1 - R_{n,a_k}) \prec 1$ for $x \in \{x \mid a_{k_0} \le x_1 \le a_{k_0+1}\}$)

$$\begin{aligned}\sum_{k_0} \int_{K_{a_{k_0}}} |A_4|^p &\prec \sum_{k_0} \sum_{k=k_0+1}^{m-1} e^{3(k_0-k)p} \|p_{n\sqrt{a_k}} - p_{n\sqrt{2a_k}}\|^p_{L^p(K_{a_k})} \\ &\prec \sum_k \|p_{n\sqrt{a_k}} - p_{n\sqrt{2a_k}}\|^p_{L^p(K_{a_k}\cap K_{2a_k})} \\ &\prec \sum_k \|F - p_{n\sqrt{a_k}}\|^p_{L^p(K_{a_k})}.\end{aligned}$$

Collecting the estimates in this section we can conclude

$$\sum_{L+1\le k_0\le m-1} \|P_n - F\|^p_{K_{k_0}} \prec \sum_{L\le k\le m} \|F - p_{n\sqrt{a_k}}\|^p_{L^p(K_{a_k})}. \tag{19.9}$$

Since the union of the sets K_{k_0}, $L+1 \le k_0 \le m-1$, include the set S_n^* from (19.3), it follows from here and from (19.9) and (18.13) that

$$\|P_n - F\|_{L^p(S_n^*)} \prec \omega^r_{\mathbf{S}}(F, 1/n)_p.$$

But

$$\omega^r_{\mathbf{S}}(F, 1/n)_p = \omega^r_S(f, 1/n)_p,$$

so we get

$$\|P_n - F\|_{L^p(S_n^*)} \prec \omega^r_S(f, 1/n)_p. \tag{19.10}$$

This is the inequality that we wanted to prove and that was used in Section 14 in the proof of Theorem 13.1, see (14.20).

■

20 Strong direct and converse inequalities

When $1 < p < \infty$, the inequalities (13.2) and (13.3) have a sharper form.

Theorem 20.1 *Let $K \subset \mathbf{R}^3$ be a 3-dimensional convex polytope, $1 < p < \infty$, $f \in L^p(K)$, and $r = 1, 2, \ldots$. With $s = \max(2, p)$ we have for $n \ge 3r$*

$$\frac{1}{n^r}\left(\sum_{k=3r}^{n} k^{sr-1} E_k(f)_p^s\right)^{1/s} \le M\omega^r_K\left(f, \frac{1}{n}\right)_p, \tag{20.1}$$

where M depends only on K, r and p.

Here we used the notation $E_n(f)_p = E_n(f)_{L^p(K)}$.

Theorem 20.2 *Let $K \subset \mathbf{R}^3$ be a 3-dimensional convex polytope, $1 < p < \infty$, $f \in L^p(K)$, and $r = 1, 2, \ldots$. With $q = \min(2, p)$ we have for $n = 1, 2, \ldots$*

$$\omega^r_K\left(f, \frac{1}{n}\right)_p \le M\frac{1}{n^r}\left(\sum_{k=0}^{n} (k+1)^{qr-1} E_k(f)_p^q\right)^{1/q} \tag{20.2}$$

where M depends only on K, r and p.

(20.1) is clearly stronger than (13.2), and one can easily see that (20.2) is stronger than (13.3), see the discussion in the papers [9], [10].

The history of such strong inequalities is briefly as follows. It was M. F. Timan [18], [19] who found them for trigonometric approximation. The case for approximation on an interval by algebraic polynomials was done in [2] and [20]. The papers [9] and [10] provide a very general framework for such strong inequalities, and they also established an important connection to Banach space geometry.

We shall obtain (20.1) and (20.2) from the analogous results for $[-1,1]$ using their equivalence to strong inequalities (so called strong Marchaud inequalities) connecting different orders of moduli of smoothness, see [9] and [10].

Proof of Theorem 20.1. Without loss of generality we may assume f to be a Borel-function, so no measurability problems will appear below.

For simpler notation we shall drop the subscript p in $\omega_K^r(f,\delta)_p$, since all moduli of smoothness are L^p moduli in this section.

We are going to show that

$$t^r\left(\int_t^1 \frac{\omega_K^{r+1}(f,u)^s}{u^{sr+1}}du\right)^{1/s} \le C\omega_K^r(f,t), \qquad t\le 1. \tag{20.3}$$

Using the monotonicity of ω^r this is easily seen to be equivalent to

$$\frac{1}{n^r}\left(\sum_{k=1}^{n} k^{sr-1}\omega_K^{r+1}(f,1/k)^s\right)^{1/s} \le C\omega_K^r(f,1/n), \qquad n=1,2,\ldots, \tag{20.4}$$

and then an application of

$$E_k(f)_p \prec \omega_K^{r+1}(f,1/k), \qquad k\ge 3r$$

(see Theorem 13.1) gives (20.1).

Thus, it is enough to prove (20.3). The $K=[-1,1]$ case of this is [2, Theorem 2.1,(2.2)], and from it we get the validity of (20.3) for all intervals/segments (with a constant independent of the segment in question).

Let e be a direction, and I a chord of K in the direction of e. According to what we have just said,

$$t^{rs}\int_t^1 \frac{\omega_I^{r+1}(f,u)^s}{u^{sr+1}}du \le C\omega_I^r(f,t)^s, \qquad t\le 1. \tag{20.5}$$

First let $p\ge 2$. Then $s=p$, and we have

$$t^{rp}\int_t^1 \frac{\omega_I^{r+1}(f,u)^p}{u^{pr+1}}du \le C\omega_I^r(f,t)^p, \qquad t\le 1.$$

Now integrate this inequality for all chords of K in the direction of e, and use that, by (16.9),

$$\omega_{K,e}^r(f,\delta)^p := \sup_{u\le\delta}\int_K^* |\Delta^r_{u\tilde{d}_K(e,x)e}f(x)|^p \sim \int \omega_I^r(f,t)^p dI, \tag{20.6}$$

where dI indicates integration with respect to all chords of K in the direction of e (in the notations of (16.9) dI is actually dy over $e^{\perp}$). Thus,

$$t^{rp}\int_t^1 \frac{\omega_{K,e}^{r+1}(f,u)^p}{u^{pr+1}}du \le C\omega_{K,e}^r(f,t)^p \le C\omega_K^r(f,t)^p. \tag{20.7}$$

Recall now that, by Theorem 13.3, there is a finite set $\mathcal{E}^*$ of directions (depending only on K) such that

$$\omega_K^{r+1}(f,u)^p \prec \max_{e\in\mathcal{E}^*}\omega_{K,e}^{r+1}(f,u)^p \le \sum_{e\in\mathcal{E}^*}\omega_{K,e}^{r+1}(f,u)^p. \tag{20.8}$$

Now if we take the sum of (20.7) for all $e\in\mathcal{E}^*$ and use the preceding inequality, then we obtain (20.3).

Next, let $1<p<2$, in which case $s=2$, and (20.5) takes the form

$$t^{2r}\int_t^1 \frac{\omega_I^{r+1}(f,u)^2}{u^{2r+1}}du \le C\omega_I^r(f,t)^2, \tag{20.9}$$

while we want to prove

$$t^{2r}\int_t^1 \frac{\omega_{K,e}^{r+1}(f,u)^2}{u^{2r+1}}du \le C\omega_{K,e}^r(f,t)^2. \tag{20.10}$$

Indeed, if we can show (20.10), and we take the sum of (20.10) for all $e\in\mathcal{E}^*$ and use (20.8), then we obtain (20.3). But

$$\left(t^{2r}\int_t^1 \frac{\omega_{K,e}^{r+1}(f,u)^2}{u^{2r+1}}du\right)^{p/2} = \left(t^{2r}\int_t^1\left(\int\frac{\omega_I^{r+1}(f,u)^p}{u^{(2r+1)(p/2)}}dI\right)^{2/p}du\right)^{p/2}.$$

On the right we use Minkowskii's inequality (*"the norm of an integral is at most as large as the integral of the norms"*; here the integral is taken with respect to dI, and the $L^{2/p}$-norm with respect to du) to continue the preceding displayed formula as

$$\le \int\left(t^{2r}\int_t^1\left(\frac{\omega_I^{r+1}(f,u)^p}{u^{(2r+1)(p/2)}}\right)^{2/p}du\right)^{p/2}dI = \int\left(t^{2r}\int_t^1\frac{\omega_I^{r+1}(f,u)^2}{u^{2r+1}}du\right)^{p/2}dI$$

For the inner integral on the right we can apply (20.9) and we can continue as

$$\le C\int \omega_I^r(f,t)^p dI = C\omega_{K,e}^r(f,t)^p,$$

and this proves (20.10).

■

Proof of Theorem 20.2. First we show that

$$\omega_K^r(f,t) \le Ct^r \left(\int_t^1 \frac{\omega_K^{r+1}(f,u)^q}{u^{qr+1}} du + \|f\|_{L^p(K)}^q \right)^{1/q} \tag{20.11}$$

with a C that is independent of f and $t \le 1$. The $K = [-1,1]$ case was proved in [20]. Indeed, there, in Theorem 1, the $K = [-1,1]$ case of (20.2), namely

$$\omega_{[-1,1]}^r \left(f, \frac{1}{n} \right)_p \le C \frac{1}{n^r} \left(\sum_{k=0}^{n} (k+1)^{qr-1} E_k(f)_{L^p[-1,1]}^q \right)^{1/q} \tag{20.12}$$

was verified. For $k \ge r+1$ we know from (1.3) (apply it with $r+1$ instead of r)

$$E_k(f)_{L^p[-1,1]} \le C\omega_{[-1,1]}^{r+1} \left(f, \frac{1}{k} \right)_{L^p[-1,1]}.$$

For $k < r$ we can only write

$$E_k(f)_{L^p[-1,1]} \le C\|f\|_{L^p[-1,1]}.$$

If put these into (20.12), then we obtain

$$\omega_{[-1,1]}^r \left(f, \frac{1}{n} \right)_p \le C \frac{1}{n^r} \left(\sum_{k=1}^{n} k^{qr-1} \omega_{[-1,1]}^{r+1} \left(f, \frac{1}{k} \right)_{L^p[-1,1]}^q + \|f\|_{L^p[-1,1]}^q \right)^{1/q},$$

which is equivalent to (20.11) for $K = [-1,1]$.

By the usual linear transformation we obtain from the $K = [-1,1]$ case of (20.11) the inequality (20.11) for all segments $K = I$.

Let e be a direction, and let first $p \le 2$. Then $q = p$ and (20.11) for a chord I of K takes the form

$$\omega_I^r(f,t)^p \le Ct^{rp} \int_t^1 \frac{\omega_I^{r+1}(f,u)^p}{u^{qr+1}} du + \|f\|_{L^p(I)}^p.$$

If we integrate this for all chords I of K in the direction of e we get (with the notation (20.6))

$$\begin{aligned} \omega_{K,e}^r(f,t)^p &\le Ct^{rp} \left(\int_t^1 \frac{\omega_{K,e}^{r+1}(f,u)^p}{u^{qr+1}} du + \|f\|_{L^p(K)}^p \right) \\ &\le Ct^{rp} \left(\int_t^1 \frac{\omega_K^{r+1}(f,u)^p}{u^{qr+1}} du + \|f\|_{L^p(K)}^p \right). \end{aligned}$$

Now if we take the supremum of the left-hand side for all directions e we obtain (20.11).

Next, let $p > 2$, in which case $q = 2$, and (20.11) for a chord I of K (that has been verified above) takes the form

$$\omega_I^r(f,t)^2 \le Ct^{2r}\left(\int_t^1 \frac{\omega_I^{r+1}(f,u)^2}{u^{qr+1}}du + \|f\|_{L^p(I)}^2\right). \tag{20.13}$$

Now

$$\omega_{K,e}^r(f,t)^2 = \left(\int \omega_I^r(f,t)^p dI\right)^{2/p},$$

where dI indicates integration with respect to all chords I of K that is in the direction of e. On the right-hand side we can apply (20.13) to get

$$\begin{aligned}
\omega_{K,e}^r(f,t)^2 \quad &\prec \quad \left(\int \left(t^{2r}\int_t^1 \frac{\omega_I^{r+1}(f,u)^2}{u^{qr+1}}du + t^{2r}\|f\|_{L^p(I)}^2\right)^{p/2} dI\right)^{2/p} \\
&\le \quad \left(\int \left(t^{2r}\int_t^1 \frac{\omega_I^{r+1}(f,u)^2}{u^{qr+1}}du\right)^{p/2} dI\right)^{2/p} \\
&\qquad + \left(\int \left(t^{2r}\|f\|_{L^p(I)}^2\right)^{p/2} dI\right)^{2/p}
\end{aligned}$$

The second term on the right is

$$t^{2r}\|f\|_{L^p(K)}^2.$$

For the first term on the right we can apply again Minkowskii's inequality (*"the norm of an integral is at most as large as the integral of norm"*, where now the norm is the $L^{p/2}(dI)$ norm, and the integral is with respect to du), from which we get the bound

$$\begin{aligned}
t^{2r}\int_t^1 \frac{\left(\int \omega_I^{r+1}(f,u)^{2(p/2)}dI\right)^{2/p}}{u^{qr+1}}du \quad &= \quad t^{2r}\int_t^1 \frac{\omega_{K,e}^{r+1}(f,u)^2}{u^{qr+1}}du \\
&\le \quad t^{2r}\int_t^1 \frac{\omega_K^{r+1}(f,u)^2}{u^{qr+1}}du
\end{aligned}$$

for that first term. Thus,

$$\omega_{K,e}^{r+1}(f,t)^2 \prec t^{2r}\int_t^1 \frac{\omega_K^{r+1}(f,u)^2}{u^{qr+1}}du + t^{2r}\|f\|_{L^p(K)}^2,$$

and all we have to do to get (20.11) is to take the supremum of the left-hand side for all directions e. Thus, (20.11) has been fully verified.

Now let us get back to (20.2). We can rewrite (20.11) in the form

$$\omega_K^r\left(f,\frac{1}{n}\right)^q \le C\frac{1}{n^{rq}}\left(\sum_{k=1}^{n} k^{qr-1}\omega_K^{r+1}(f,1/k)^q + \|f\|^q_{L^p(K)}\right), \qquad n=1,2,\dots, \tag{20.14}$$

and here we apply (13.3), according to which

$$\omega_K^{r+1}\left(f,\frac{1}{k}\right) \le \frac{M}{k^{r+1}}\sum_{j=0}^{k}(j+1)^r E_j(f)_p.$$

Write this into (20.14) and apply Jensen's inequality:

$$\begin{aligned}\omega_K^r\left(f,\frac{1}{n}\right)^q \quad &\prec \quad \frac{1}{n^{rq}}\sum_{k=1}^{n} k^{qr-1}\left(\frac{1}{k^{r+1}}\sum_{j=0}^{k}(j+1)^r E_j(f)_p\right)^q \\ &\qquad\qquad\qquad\qquad + \frac{1}{n^{rq}}\|f\|^q_{L^p(K)} \\ &\prec \quad \frac{1}{n^{rq}}\sum_{k=1}^{n} k^{qr-1}\frac{1}{k^{rq+1}}\sum_{j=0}^{k}(j+1)^{rq}E_j(f)_p^q + \frac{1}{n^{rq}}\|f\|^q_{L^p(K)}.\end{aligned}$$

If we interchange here the order of the two summations we obtain

$$\omega_K^r\left(f,\frac{1}{n}\right) \le M\frac{1}{n^r}\left(\sum_{j=0}^{n}(j+1)^{qr-1}E_j(f)_p^q + \|f\|^q_{L^p(K)}\right)^{1/q}. \tag{20.15}$$

Finally, replace here f by $f - P_0$, where P_0 is the constant which minimizes the L^p-norm of $f - P_n$:

$$E_0(f)_p = \|f - P_n\|_{L^p(K)}.$$

In this substitution the left-hand side in (20.15) does not change, and on the right of (20.15) we can write $E_0(f)_p$ instead of $\|f\|$, and we obtain (20.2).

■

21 The K-functional in L^p and the equivalence theorem

Just as in Section 12, Theorem 13.1 allows us to prove the equivalence of the moduli of smoothness ω_K^r with a K-functionals.

In L^p the form of the K-functional is

$$\mathcal{K}_{r,p}(f,t)_K = \inf_g\left(\|f-g\|_{L^p(K)} + t\sup_{e\in S^2}\left\|\tilde{d}_K(e,\cdot)^r\frac{\partial^r g}{\partial e^r}\right\|_{L^p(K)}\right), \tag{21.1}$$

where the infimum is taken for all g that are in $C^r(K)$ (all partial derivatives of order at most r are continuous on K) and the supremum is taken for all directions $e \in S^2$ in $\mathbf{R}^3$.

Theorem 21.1 *Let K be a 3-dimensional convex polytope in $\mathbf{R}^3$. There is a constant M depending only on r, K and $1 \le p < \infty$ such that for all $f \in L^p(K)$ and for all $0 < \delta \le 1$ we have*

$$\frac{1}{M}\mathcal{K}_{r,p}(f, \delta^r)_K \le \omega_K^r(f, \delta)_p \le M\mathcal{K}_{r,p}(f, \delta^r)_K. \tag{21.2}$$

Proof. Let P_n be polynomials of degree at most n ($\ge 3r$) such that

$$\|f - P_n\|_{L^p(K)} \prec \omega_K^r(f, n^{-1})_p, \tag{21.3}$$

the existence of which is given by Theorem 13.1. Let e be a direction, and apply to $H_m = P_n$ the inequality (17.5) on every chord I of K in the direction e:

$$\int_I \tilde{d}_I(\overline{e}, x)^{rp} \left| \frac{\partial^r P_n}{\partial e^r}(x) \right|^p dx \prec n^{rp} \omega_I^r(P_n, n^{-1})_p^p$$

with $\prec$ independent of the chord I. On integrating this for all chords I in the direction e and using (16.9) we obtain $n \ge 3r$

$$\left(\frac{1}{n}\right)^r \left\| \tilde{d}_K(e, \cdot)^r \frac{\partial^r P_n}{\partial e^r} \right\|_{L^p(K)} \prec \omega_{K,e}^r(P_n, n^{-1})_p, \tag{21.4}$$

where $\omega_{K,e}^r(P_n, \delta)$ is the directional modulus of smoothness from (16.4) and (16.5). But, by (16.14), here

$$\omega_{K,e}^r(P_n, n^{-1})_p \prec \|f - P_n\|_{L^p(K)} + \omega_{K,e}^r(f, n^{-1})_p \prec \omega_K^r(f, n^{-1})_p,$$

and so

$$\|f - P_n\|_{L^p(K)} + \left(\frac{1}{n}\right)^r \left\| \tilde{d}_K(e, \cdot)^r \frac{\partial^r P_n}{\partial e^r} \right\|_{L^p(K)} \prec \omega_K^r(f, n^{-1})_p, \tag{21.5}$$

which implies the first inequality in (21.2) for $t = 1/n$. Using simple monotonicity properties of $\mathcal{K}_{r,p}$ and ω_K^r (see in particular (16.17)), this is enough to conclude

$$\mathcal{K}_{r,p}(f, \delta^r)_K \prec \omega_K^r(f, \delta)_p$$

for all $0 < \delta \le 1$.

The converse inequality

$$\omega_K^r(f, \delta)_p \prec \mathcal{K}_{r,p}(f, \delta^r)_K \tag{21.6}$$

is an easy consequence of (16.34)–(16.35). Indeed, let h be an r-times differentiable function such that

$$\|f-h\|_{L^p(K)} + \delta^r \sup_{e\in S^2} \left\| \tilde{d}_K(e,\cdot)^r \frac{\partial^r h}{\partial e^r} \right\|_{L^p(K)} \le 2\mathcal{K}_{r,p}(f,\delta^r). \tag{21.7}$$

On applying (16.34)–(16.35) to $g=f$ and $Q=h$ on a chord I of K in a given direction e we obtain

$$\frac{1}{\delta}\int_0^\delta \int_I^* |\Delta^r_{u\tilde{d}_K(e,x)e} f(x)|^p dx du \prec \int_I |f-h|^p + \delta^{rp} \int_I \tilde{d}_K(e,x)^{rp} \left| \frac{\partial^r h(x)}{\partial e^r} \right|^p dx.$$

If we integrate this for all chords I in the direction e, apply (16.5) and (21.7), then we obtain

$$\omega^r_{K,e}(f,\delta)_p \prec \mathcal{K}_{r,p}(f,\delta^r)_K,$$

where $\prec$ is independent of e. Now (21.6) is a consequence of we take the supremum on the left hand-side for all directions.

■

The following corollary may be of interest.

Corollary 21.2 *There is a finite set $\mathcal{E}^*$ of directions such that*

$$\omega^r_K(f,\delta)_p \le M\omega^r_{K,\mathcal{E}^*}(f,\delta)_p, \qquad \delta>0, \tag{21.8}$$

where

$$\omega^r_{K,\mathcal{E}^*}(f,\delta)_p := \max_{e\in\mathcal{E}^*} \omega^r_{K,e}(f,\delta)_p$$

is the restricted modulus of smoothness from (13.6)*, and where M depends only on r,p and K.*

Indeed, by Theorem 13.3 there is a finite set $\mathcal{E}^*$ of directions such that for $n\ge 3r$ there are polynomials P_n of degree at most n for which

$$\|f-P_n\|_{L^p(K)} \prec \omega^r_{K,\mathcal{E}^*}(f,n^{-1})_p. \tag{21.9}$$

Now instead of (21.3) use this in the preceding proof to conclude

$$\mathcal{K}_{r,p}(f,\delta^r)_K \prec \omega^r_{K,\mathcal{E}^*}(f,\delta)_p,$$

and then the corollary follows from the second inequality in (21.2).

Exactly as in Theorem 13.3, there is such an $\mathcal{E}^*$ consisting of at $(r+1)v(v-1)/2$ directions, where v is the number of edges of K.

References

[1] F. Dai, Z. Ditzian and H. Huang, Equivalence of measures of smoothness in $L_p(S^{d-1})$, $1 < p < \infty$, *Studia Math.*, **196**(2010), 179–205.

[2] F. Dai, Z. Ditzian, S. Tikhonov, Sharp Jackson inequalities, *J. Approx. Theory*, **151**(2008), 86-112.

[3] F. Dai and K. Wang, A survey of polynomial approximation on the sphere, *Int. J. Wavelets Multiresolut. Inf. Process.*, **7**(2009), 749-771.

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

[5] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Grundlehren de mathematischen Wissenschaften, **303**, Springer-Verlag, Berlin, Heidelberg, New York 1993.

[6] Z. Ditzian, Polynomial approximation in $L_p(S)$ for $p > 0$, *Constr. Approx.*, **12**(1996), 241-269.

[7] Z. Ditzian, On best polynomial approximation in $L_w^2(S)$, *J. Math. Anal. Appl.*, **194**(1995), 548-559.

[8] Z. Ditzian, Multivariate Bernstein and Markov inequalities, *J. Approx. Theory*, **70**(1992), 273-283.

[9] Z. Ditzian and A. Prymak, Sharp Marchaud and converse inequalities in Orlicz spaces, *Proc. Amer. Math. Soc.*, **135**(2007), 1115-1121.

[10] Z. Ditzian and A. Prymak, Convexity, moduli of smoothness, and a Jackson type inequality, *Acta Math. Hungar.,* **130**(2011), 254-285.

[11] Z. Ditzian, K. Runovskii, Averages and K-functionals related to the Laplacian, *J. Approx. Theory*, **97**(1999), 113-139.

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

[13] M. Dubiner, The theory of multi-dimensional polynomial approximation, *J. Anal. Math.*, **67**(1995), 39-116.

[14] K. G. Ivanov and V. Totik, Fast decreasing polynomials, *Constr. Approx.*, **6**(1990), 1–20.

[15] E. Levin and D. S. Lubinsky, *Orthogonal polynomials for exponential weights*, CMS Books in Mathematics/Ouvrages de Mathmatiques de la SMC, 4. Springer-Verlag, New York, 2001.

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

[17] K.V. Rustamov, On equivalence of different moduli of smoothness on the sphere, *Proc. Stekelov Inst. Math.*, **204**(1994), 235-260.

[18] M. F. Timan, Converse theorem of the constructive theory of functions in the space L_p, *Math. Sb.*, **44**(1958), 125-132. (Russian)

[19] M. F. Timan, On Jackson's theorem in L^p spaces, *Ukrain. Mat. Zh.*, **18**(1966), no. 1, 134-137. (Russian)

[20] V. Totik, Sharp converse theorem for L^p polynomial approximation, *Constructive Approx.*, **4**(1988), 419-433.

[21] Y. Xu, Analysis on the unit ball and on the simplex, *Electron. Trans. Numer. Anal.*, **25**(2006), 284-301

[22] Y. Xu, Generalized translation operator and approximation in several variables, *J. Comput. Appl. Math.*, **178**(2005), 489-512.

[23] Y. Xu, Weighted approximation of functions on the unit sphere, *Construcive Approx.*, **21**(2005), 1–28.

Analysis and Stochastics Research Group of the Hungarian Academy of Sciences
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