\textbf{\epsilon-Entropy and moduli of smoothness in $L^p$-spaces}

by

A. KAMONT (Sopot)

\textbf{Abstract.} The asymptotic behaviour of $\epsilon$-entropy of classes of Lipchitz functions in $L^p(\mathbb{R}^d)$ is obtained. Moreover, the asymptotics of $\epsilon$-entropy of classes of Lipchitz functions in $L^p(\mathbb{R}^d)$ whose tail function decreases as $O(\lambda^{-p})$ is obtained. In case $p=1$ the relation between the $\epsilon$-entropy of a given class of probability densities on $\mathbb{R}^d$ and the minimax risk for that class is discussed.

1. Introduction. First we recall the notions of $\epsilon$-entropy and $\epsilon$-capacity (cf. [8]).

Let $(X, \rho)$ be a metric space, and let $A$ be a subset of $X$. A family $\{U_t\}_{t \in T}$ of subsets of $X$ is called an $\epsilon$-covering of $A$ if $A \subseteq \bigcup_{t \in T} U_t$ and the diameter of each $U_t$ does not exceed $2\epsilon$. A subset $U \subseteq X$ is called an $\epsilon$-net of $A$ if for each $a \in A$ there exists $u \in U$ such that $\rho(a, u) \leq \epsilon$. A subset $U \subseteq X$ is called $\epsilon$-distinguishable if $\rho(u_1, u_2) > \epsilon$ for any distinct $u_1, u_2 \in U$.

A subset $A \subseteq X$ is called totally bounded if for each $\epsilon > 0$ there exists a finite $\epsilon$-net of $A$. For a totally bounded $A \subseteq X$ define

\[ \mathcal{N}_\epsilon(A) = \min \{ \#T : \{U_t\}_{t \in T} \text{ is an } \epsilon \text{-covering of } A \}, \]
\[ \mathcal{N}^X_\epsilon(A) = \min \{ \#U : U \text{ is an } \epsilon \text{-net of } A, U \subseteq X \}, \]
\[ \mathcal{M}_\epsilon(A) = \sup \{ \#M : M \text{ is an } \epsilon \text{-distinguishable subset of } A \}, \]

where $\#T$ denotes the cardinality of $T$, and

\[ \mathcal{H}_\epsilon(A) = \ln \mathcal{N}_\epsilon(A), \quad \mathcal{H}^X_\epsilon(A) = \ln \mathcal{N}^X_\epsilon(A), \quad \mathcal{C}_\epsilon(A) = \ln \mathcal{M}_\epsilon(A), \]

where $\ln$ is the natural logarithm.

$\mathcal{H}_\epsilon(A)$ and $\mathcal{C}_\epsilon(A)$ are called the $\epsilon$-entropy and $\epsilon$-capacity of $A$ respectively. Note that if $A \subseteq B$ then $\mathcal{N}_\epsilon(A) \leq \mathcal{N}_\epsilon(B)$, $\mathcal{M}_\epsilon(A) \leq \mathcal{M}_\epsilon(B)$, $\mathcal{H}_\epsilon(A) \leq \mathcal{H}_\epsilon(B)$ and $\mathcal{C}_\epsilon(A) \leq \mathcal{C}_\epsilon(B)$.

The following inequalities are satisfied ([8]):

\[ \mathcal{M}_{2\epsilon}(A) \leq \mathcal{N}_\epsilon(A) \leq \mathcal{N}^X_\epsilon(A) \leq \mathcal{N}^A_\epsilon(A) \leq \mathcal{M}_\epsilon(A), \]
\[ \mathcal{C}_{2\epsilon}(A) \leq \mathcal{H}_\epsilon(A) \leq \mathcal{H}^X_\epsilon(A) \leq \mathcal{H}^A_\epsilon(A) \leq \mathcal{C}_\epsilon(A). \]
To formulate the main result we recall the definitions of moduli of smoothness of order $m$ in $L^p(d^d)$ and $L^p(d^d)$ for $d \in N$, $1 \leq p < \infty$, where $I = [0, 1].$

For $f \in L^p(d^d)$, $m \in N$ and $u \in d^d$ let

$$\Delta^m u f(x) = \sum_{i=0}^{m} (-1)^{m+i} \binom{m}{i} f(x + i u)$$

and for $\delta > 0$ define

$$\omega_m(p, f, \delta) = \sup_{\|u\| \leq \delta} \|\Delta^m u f\|_p,$$

where $\|u\| = (u_1^2 + \ldots + u_d^2)^{1/2}$ for $u = (u_1, \ldots, u_d) \in d^d$.

For $f \in L^p(d^d)$ and $u \in d^d$ let

$$\Pi^d(u) = \{ x \in d^d : x + u \in d^d \}$$

and for $m \in N$, $0 < \delta \leq 1/m$ define

$$\omega_m(p, f, \delta) = \sup_{\|u\| \leq \delta} \left( \int_{\Pi^d(u)} |\Delta^m u f(x)|^p \, dx \right)^{1/p}.$$

The functional $\omega_m(p, f, \delta)$ is the modulus of smoothness of the function $f$ of order $m$ in the $L^p$-norm. In addition we put $\omega_m(p, f, \delta) = \|f\|_p$ for $m = 0$.

For $d, m \in N$, $0 < \alpha < m$, $1 \leq p < \infty$ and $C > 0$ define

$$A(p, d, m, \alpha, C) = \{ f \in L^p(d^d) : \|f\|_p \leq C, \omega_{m,d}(p, f, \delta) \leq C \delta^{\alpha} \text{ for } 0 < \delta \leq 1/m \}.$$

For $f \in L^p(d^d)$ the tail function $\Phi_p(f, \lambda)$ is defined as follows:

$$\Phi_p(f, \lambda) = \left( \int_{d^d} |f(x)|^p \, dx \right)^{1/p}$$

Now for $d, m \in N$, $0 < \alpha < m$, $\gamma > 0$ $0 < \alpha < p$ and $C > 0$ define

$$U(p, d, m, \alpha, \gamma, C) = \{ f \in L^p(d^d) : \|f\|_p \leq C, \omega_{m,d}(f, \delta) \leq \lambda \delta^{\alpha} \text{ for } \lambda > 0 \}.$$

We will discuss the asymptotic behaviour of the $\varepsilon$-entropy of $A(p, d, m, \alpha, C)$ and $U(p, d, m, \alpha, \gamma, C)$ in the metric induced from the spaces $L^p(d^d)$ and $L^p(d^d)$ respectively. It will be shown that there exist constants $\varepsilon_1, \varepsilon_2 > 0$ and $\varepsilon_0 > 0$, independent of $C$ and $\varepsilon$, such that for all $0 < \varepsilon \leq \varepsilon_0$

$$\varepsilon_1 (C/\varepsilon)^{d/\alpha} \leq H_{\varepsilon}(A(p, d, m, \alpha, C)) \leq \varepsilon_2 (C/\varepsilon)^{d/\alpha}.$$

Results of this type are already known ([9], Theorem 10); however, we present the proof, since the dependence of $H_{\varepsilon}(A(p, d, m, \alpha, C))$ on $C$ is needed for estimating $H_{\varepsilon}(U(p, d, m, \alpha, \gamma, C)).$ Moreover, it will be shown that there exist constants $k_1, k_2 > 0$ and $\varepsilon_0 > 0$, independent of $C$ and $\varepsilon$, such that for all $0 < \varepsilon \leq \varepsilon_0$

$$k_1 (C/\varepsilon)^{d/(1/\alpha + 1/\gamma)} \leq H_{\varepsilon}(U(p, d, m, \alpha, \gamma, C)) \leq k_2 (C/\varepsilon)^{d/(1/\alpha + 1/\gamma)}.$$

The proofs rely very much on the following theorem of G. G. Lorentz ([9], Theorem 2):

**Theorem 1.3.** Let $(X, \| \cdot \|)$ be a real separable Banach space, and let $\Phi = \{ \varphi_1, \varphi_2, \ldots \}$ be a sequence of linearly independent elements of $X$ such that $X = \text{span} \Phi$. For a given $f \in X$ set

$$E_0(f) = \|f\|, \quad E_n(f) = \inf \left\{ \|f - \sum_{k=1}^{n} a_k \varphi_k\| : a_1, \ldots, a_n \in \mathbb{R} \right\}.$$

Let $A = \{ \delta_0, \delta_1, \ldots \}$ be a nonincreasing sequence of positive numbers such that $\lim_{n \to \infty} \delta_n = 0$, and define

$$A(A, \Phi) = \{ f \in X : E_n(f) \leq \delta_n, \ n = 0, 1, \ldots \}.$$

For $c \in \mathbb{R}$ with $1 < c < 4$ put

$$N_0 = 0, \quad N_i = \min \{ k : \delta_k \leq c^{-i} \} \quad \text{for } i = 1, 2, \ldots$$

Given $0 < \varepsilon < 1$, let $j \in \mathbb{N}$ be such that $c^{-j-1} < \varepsilon \leq c^{-j-2}$. Then

$$N_{j-3} \leq c \leq C \leq C \left( A(A, \Phi) \right),$$

$$N_{j-3} \leq \sum_{k=1}^{N_j} (1 + \ln c) + N_j \ln \frac{12}{c^{-1} + 1} + N_1 \ln \delta_0.$$

In the sequel the following notation will be used. For a multi-index $\alpha = (a_1, \ldots, a_d) \in (\mathbb{N} \cup \{ 0 \})^d$ define $|\alpha| = a_1 + \ldots + a_d$, $D^\alpha = \partial^{a_1} / \partial x_1^{a_1} \ldots \partial x_d^{a_d}$.

In addition $b = (b_1, \ldots, b_d) \in (\mathbb{N} \cup \{ 0 \})^d$, then we write $a \preceq b$ if $a_j \leq b_j$ for $j = 1, \ldots, d$.

**2. The asymptotics of $H_{\varepsilon}(A(p, d, m, \alpha, C))$.** In this section the asymptotics for $H_{\varepsilon}(A(p, d, m, \alpha, C))$ will be obtained by means of Theorem 1.3. We will find a sequence $\Phi$ in $L^p(d^d)$ and two sequences of positive numbers $\Delta'$ and $\Delta''$ such that

$$A(\Delta', \Phi) \subset A(p, d, m, \alpha, C) \subset A(\Delta'', \Phi).$$

It occurs that the sequence $\Phi$ can be chosen in such a way that its elements are spline functions with dyadic knots. Some necessary definitions are recalled below.

For each $n \in \mathbb{N}$, $n \geq 2$, there exists exactly one pair of integers $\mu, k$ such that $n = 2^\mu + k$, $\mu \geq 0$, $0 \leq k \leq 2^\mu$; put

$$s_{n,j} = \begin{cases} j/2^{\mu+1} & \text{for } j \leq 2^\mu, \\ (j-k)/2^{\mu} & \text{for } j \geq 2k+1. \end{cases}$$
For \( r \in \mathbb{N} \) the B-splines \( N_{n,j}^{(r)} \) are defined by the formula
\[
N_{n,j}^{(r)}(t) = (s_{n,j+r} - s_{n,j})[s_{n,j}, \ldots, s_{n,j+r}; (s-t)_+^{r-1})
\]
(where the square brackets denote the divided difference of \((s-t)_+^{r-1})\), taken in the variable \( s \) at the points \( s_{n,j}, \ldots, s_{n,j+r} \). For the properties of the B-splines we refer e.g. to [1]; some of those properties are listed below:

1. \( N_{n,j}^{(r)} \) is a function of class \( C^{r-2} \) and is a polynomial of degree at most \( r-1 \) on each interval \([s_{n,i}, s_{n,i+1}]\).

2. \( N_{n,j}^{(r)} \geq 0 \), supp \( N_{n,j}^{(r)} = [s_{n,j}, s_{n,j+r}] \).

3. For a given \( n \in \mathbb{N} \) and an interval \((a, b)\) those functions \( N_{n,j}^{(r)} \) which are nontrivial when restricted to \((a, b)\), are linearly independent over that interval.

In the following the restrictions of \( N_{n,j}^{(r)} \) to \( I \) will be considered. Let
\[
S_n^{(r)} = \text{span}\{N_{n,j}^{(r)} : j = -r + 1, \ldots, n-1\} \quad \text{for } n \geq 2,
\]
and for \( n = -r + 2, \ldots, 1 \) let \( S_n^{(r)} \) be the space of polynomials of degree at most \( n + r - 1 \), restricted to \( I \). Then \( dim S_n^{(r)} = n + r - 1 \) and \( S_n^{(r)} \subset S_{n+1}^{(r)} \). Let the sequence of functions \((f_j^{(r)}, j \geq 2 - r)\) be defined as follows:
\[
f_{2-r}^{(r)} = 1, \quad f_{j+1}^{(r)} = f_j^{(r)} - f_{j+1}^{(r)}, \quad f_2^{(r)} \text{ is orthogonal to } S_{2-r}^{(r)} \text{ with respect to the inner product in } L^2(I); \quad \|f_j^{(r)}\|_2 = 1.
\]
Using the notation \( Hf(t) = f'_t f(x) dx, Df(t) = \frac{df(t)}{dt} \), define for \( k \in \mathbb{Z}, |k| < r,
\[
f_j^{(r,k)} = \begin{cases} f_j^{(r)} & \text{for } k = 0, \\ (H^k f_j^{(r)}) & \text{for } 0 < k < r, \\ (D^k f_j^{(r)}) & \text{for } -r < k < 0. 
\end{cases}
\]

Now we introduce the tensor product spline functions on \( \mathbb{R}^d \).

For \( r = (r_1, \ldots, r_d) \in \mathbb{N}^d, k = (k_1, \ldots, k_d) \in \mathbb{Z}^d, |k| < r \), \( j = (j_1, \ldots, j_d) \), \( j_i \geq 2 - r_i + k_i \), \( t = (t_1, \ldots, t_d) \in \mathbb{R}^d \), define
\[
f_j^{(r,k)}(t) = f_{j_1}^{(r_1,k_1)}(t_1) \cdots f_{j_d}^{(r_d,k_d)}(t_d).
\]

For \( m \in \mathbb{N} \), let \( m = (m + 2, \ldots, m + 2) \in \mathbb{N}^d, r_m = 2 - r \), let
\[
F_j^{(m)} = f_j^{(r_m, m)}, \quad G_j^{(m)} = f_j^{(r_m, -m)}
\]
for \( j = (j_1, \ldots, j_d) \) such that \( j_i \geq -m \) for \( i = 1, \ldots, d \); observe that \( F_j^{(m)} \) are functions of class \( C^m \) and \( (F_j^{(m)}, G_j^{(m)}) = \delta_{j_0} \). Now, for \( f \in L^p(I^d), \)
\[
Q_n^{(m)}(f) = \sum_{-m \leq j \leq n} \langle f, G_j^{(m)} \rangle F_j^{(m)}.
\]
Note that \( Q_n^{(m)} \) is a projection onto the space
\[
V_n^{(m)} = \text{span}\{F_j^{(m)} : -m \leq j \leq n\}.
\]

For \( f \in L^p(I^d) \), let
\[
E_n^{(m)}(f) = \inf \{ \|f - g\|_p : g \in V_n^{(m)} \}
\]

The following lemmas establish some relations between \( E_n^{(m)}(f) \) and \( \omega_n^{(m)}(f, \delta) \). Lemma 2.4 was proved in [5] as Theorem 9.18. The proof of Lemma 2.5 is omitted, since it can be proved similarly to Theorem 10 of [2], with the help of (5.14) of [4] and the inequality
\[
\left( \int_{I^d} |A^{m+k} f(x)|^p dx \right)^{1/p} \leq \|A^{m+k} f\|_p \sum_{|a|=k} \left( \int_{I^d} |A^{n} D^a f(x)|^p dx \right)^{1/p},
\]
which holds for \( k \in \mathbb{N}, f \in C^k_0(I^d) \), \( m \in \mathbb{N} \cup \{0\} \), \( u \in \mathbb{R}^d, \|u\|_1 \leq 1/(m+k) \), \( 1 \leq p < \infty \), where \( A_{k,d} \) is a constant independent of \( f, u, m \) and \( p \).

**Lemma 2.4.** Let \( 1 \leq p < \infty \) and \( m, d \in \mathbb{N} \) be given. There exists a constant \( M_{m,p,d} \), such that
\[
E_n^{(m)}(f) \leq M_{m,p,d} \omega_n^{(m)}(f, 1/n)
\]
for every \( n \geq m \) and \( f \in L^p(I^d) \).

**Lemma 2.5.** Let \( 1 \leq p < \infty \) and \( m, d \in \mathbb{N} \) be given. There exists a constant \( M_{m,d}^{(p)} \), such that
\[
\omega_n^{(m)}(f, 1/n) \leq M_{m,d}^{(p)} n^{-m} \left( \|f\|_p + \sum_{i=m}^{n} i^{m-1} E_i^{(m)}(f) \right)
\]
for every \( n \geq m \) and \( f \in L^p(I^d) \).

Let
\[
B(p, d, m, \alpha, C) = \{ f \in L^p(I^d) : \|f\|_p \leq C, \ E_n^{(m)}(f) \leq C/n^\alpha \text{ for } n \geq m \}.
\]
Corollary 2.6. Let $1 \leq p < \infty$, $m, d \in \mathbb{N}$ and $0 < \alpha < m$ be given. There exist constants $c_1, c_2 > 0$, depending only on $p, m, \alpha$ and $d$, such that for every $C > 0$

$$B(p, d, m, \alpha, c_1 C) \subset A(p, d, m, \alpha, C) \subset B(p, d, m, \alpha, c_2 C).$$

Proof. The existence of $c_2$ follows from Lemma 2.4.

As $0 < \alpha < m$, there exists $a_{m, \alpha}$ such that $1 + \sum_{i=m}^{m-1} i^{m-1-\alpha} \leq a_{m, \alpha} n^{m-\alpha}$ for every $n \geq m$. The existence of $c_1$ now follows from Lemma 2.5 and the inequality $\omega_{n, p}(f, \delta_1) \leq \omega_{n, p}(f, \delta_2)$ for $0 < \delta_1 \leq \delta_2 \leq 1/m$.

Now we will find the asymptotic behaviour of the $\varepsilon$-entropy of $B(p, d, m, \alpha, M)$.

Lemma 2.7. Let $1 \leq p < \infty$, $m, d \in \mathbb{N}$ and $0 < \alpha < m$ be given. There exist constants $a_1, a_2 > 0$ such that the following inequalities hold for each $M > 0$:

$$a_1(M/\varepsilon)^{d/\alpha} \leq \mathcal{H}_\varepsilon(B(p, d, m, \alpha, M)) \quad \text{for } 0 < \varepsilon \leq M/2,$$

$$\mathcal{H}_\varepsilon(B(p, d, m, \alpha, M)) \leq a_2(M/\varepsilon)^{d/\alpha} \quad \text{for } 0 < \varepsilon \leq M.$$

Proof. As $B(p, d, m, \alpha, M) = \left\{ M f : f \in B(p, d, m, \alpha, 1) \right\}$ and $\mathcal{H}_\varepsilon(B(p, d, m, \alpha, M)) = \mathcal{H}_\varepsilon(B(p, d, m, \alpha, 1))$, it is enough to estimate $\mathcal{H}_\varepsilon(B(p, d, m, \alpha, 1))$. We will use Theorem 1.3.

Consider $V_n^{(m)}$ for $n \geq m$ as subspaces of $X = L^p([0, 1]^d)$ and set

$$d_n^{(m)} = \dim V_n^{(m)} = (n + m + 1)^d.$$

The sequence $\Phi = \left\{ \varphi_1, \varphi_2, \ldots \right\}$ is obtained by ordering $\left\{ F_n^{(m)} : j = (j_1, \ldots, j_d) \right\}$ so that $\text{span}\{ \varphi_1, \ldots, \varphi_d^{(m)} \} = V_n^{(m)}$ for every $n \geq m$. The sequence $\Delta$ is defined as follows:

$$d_0 = \ldots = d_m^{(m)} = 1, \quad \delta_k = 1/n^{\alpha} \quad \text{for } d_{k-1}^{(m)} \leq k \leq d_{k-1}^{(m)} + 1.$$

Then $B(p, d, m, \alpha, 1) = A(\Delta, \Phi)$. Let $c = e$ in Theorem 1.3 and write $n_1 = \max\left( \lfloor e^{1/\alpha} \rfloor, m \right)$ (where $[a] = \min\{ k \in \mathbb{Z} : a \leq k \}$); then $N_1 = d_{n_1}^{(m)}$.

First the constant $a_1$ will be found.

Inequality (1.2) implies

$$C_2 B(p, d, m, \alpha, 1) \leq \mathcal{H}_\varepsilon(B(p, d, m, \alpha, 1)).$$

For $\varepsilon \leq 1/2$ choose $j$ such that $e^{-(J-1)} < 2\varepsilon \leq e^{-(J-2)}$; then (1.4) yields

$$C_2 B(p, d, m, \alpha, 1) \geq N_{j-1} = d_{n_{j-2}}^{(m)} = (n_{j-2} + m + 1)^d \geq (e^{(j-3)/\alpha} + m + 1)^d \geq e^{(j-3)d/\alpha} \geq e^{-2d/\alpha} e^{(j-1)d/\alpha} \geq e^{-2d/\alpha} \left( \frac{1}{2\varepsilon} \right)^{d/\alpha}.$$

Setting $a_1 = e^{-2d/\alpha} (1/2)^{d/\alpha}$ we obtain for any $0 < \varepsilon \leq 1/2$

$$a_1(1/\varepsilon)^{d/\alpha} \leq C_\varepsilon(B(p, d, m, \alpha, 1)).$$

Now the constant $a_2$ will be found. For $\varepsilon \leq 1$ choose $j$ such that $e^{-(j-1)} < \varepsilon \leq e^{-(j-2)}$. As $N_i \leq (e^{i/\alpha} + m + 2)^d + (2m + 1)^d$, it follows from (1.5) that

$$\mathcal{H}_\varepsilon(B(p, d, m, \alpha, 1)) \leq 2 \sum_{i=1}^{j} N_i + N_j \ln \frac{12}{\varepsilon} \leq \frac{b}{\varepsilon} \sum_{i=1}^{j} N_i \leq b(2m + 1)^d + b \sum_{i=1}^{j} (e^{i/\alpha} + m + 2)^d \leq \frac{b}{\varepsilon} \sum_{k=1}^{d} \left( \frac{e^{(k+1)/\alpha} - e^{k/\alpha}}{e^{k/\alpha} - 1} \right) + b(2m + 1)^d + (2m + 1)^d \frac{e^{(k+1)/\alpha}}{e^{k/\alpha} - 1},$$

where $b = 2 + \ln(12/(\varepsilon - 1))$. Setting

$$b_1 = \frac{b}{\varepsilon} \sum_{k=1}^{d} \left( \frac{1}{k} \right) \left( \frac{e^{(k+1)/\alpha} - e^{k/\alpha}}{e^{k/\alpha} - 1} \right), \quad b_2 = b(2m + 1)^d + (2m + 1)^d \frac{e^{(k+1)/\alpha}}{e^{k/\alpha} - 1},$$

we get

$$\mathcal{H}_\varepsilon(B(p, d, m, \alpha, 1)) \leq b_1 e^{d/\alpha} + b_2 e^{d/\alpha} \left( \frac{1}{\varepsilon} \right)^{d/\alpha} + b_2 \left( \frac{1}{\varepsilon} \right)^{d/\alpha}.$$

As

$$\ln \frac{1}{\varepsilon} \leq \frac{a}{d_\varepsilon} \left( \frac{1}{\varepsilon} \right)^{d/\alpha},$$

by setting $a_2 = b_1 e^{d/\alpha} + b_2 e^{d/\alpha}$ we obtain

$$\mathcal{H}_\varepsilon(B(p, d, m, \alpha, 1)) \leq a_2 (1/\varepsilon)^{d/\alpha}$$

for any $0 < \varepsilon \leq 1$. Our lemma now follows from (2.8) and (2.9).

Corollary 2.10. Let $1 \leq p < \infty$, $m, d \in \mathbb{N}$ and $0 < \alpha < m$ be given. There exist constants $a_1, a_2 > 0$ and $a_0 > 0$ such that

$$a_1(M/\varepsilon)^{d/\alpha} \leq \mathcal{H}_\varepsilon(A(p, d, m, \alpha, M)) \leq a_2(M/\varepsilon)^{d/\alpha}$$

for all $M > 0$ and $0 < \varepsilon \leq M a_0$.

This follows from Lemma 2.7 and Corollary 2.6.

The case $m = 1$ will be considered more carefully. Define

$$A^+(p, d, 1, \alpha, M) = \{ f \in A(p, d, 1, \alpha, M) : f \geq 0 \}.$$
For \( f \in L^p(\mathbb{I}^d) \) set \( f_+ = \max(f, 0) \), \( f_- = \max(-f, 0) \); then \( f_+ \geq f_- \). As \( g(a) = \max(a, 0) \) is a Lipschitz function with constant 1, \( f_+ \geq f_- \). \( \mid f(x) - f_+(y) \mid \leq \mid f(x) - f(y) \mid, \mid f(x) - f_-(y) \mid \leq \mid f(x) - f(y) \mid, \) which implies

\[
\omega_{1,p}(f_+, \delta) \leq \omega_{1,p}(f, \delta) \quad \text{and} \quad \omega_{1,p}(f_-, \delta) \leq \omega_{1,p}(f, \delta)
\]

for each \( 0 < \delta \leq 1 \); in addition \( \| f_+ \|_p \leq \| f \|_p \), \( \| f_- \|_p \leq \| f \|_p \). Therefore, if \( f \in A^+(p, \alpha, M) \) then \( f_+, f_- \in A^+(p, d, \alpha, M) \). Let \( U = \{ U_1, \ldots, U_k \} \) be an \( \varepsilon/2 \)-covering of \( A^+(p, d, \alpha, M) \), and \( V_{ij} = \{ f \in A(p, d, \alpha, M) : f_+ \in U_i, f_- \in U_j \} \) for \( i, j = 1, \ldots, k \). Then \( V = \{ V_{ij} : i, j = 1, \ldots, k \} \) is an \( \varepsilon \)-covering of \( A(p, d, \alpha, M) \) with \( \#V \leq k^2 = (\#U)^2 \); this implies

\[
\mathcal{H}_e(A(p, d, 1, \alpha, M)) \leq 2\mathcal{H}_{\varepsilon/2}(A^+(p, d, 1, \alpha, M)).
\]

As \( A^+(p, d, 1, \alpha, M) \subset A(p, d, 1, \alpha, M) \), the last inequality and Corollary 2.10 yield

\[
\text{Corollary 2.11. Let } 1 \leq p < \infty, d \in \mathbb{N} \text{ and } 0 < \alpha < 1 \text{ be given. There exist constants } a_1, a_2 > 0 \text{ and } \varepsilon_0 > 0 \text{ such that}
\]

\[
a_1(M/\varepsilon)^{d/\alpha} \leq \mathcal{H}_e(A^+(p, d, 1, \alpha, M)) \leq a_2(M/\varepsilon)^{d/\alpha}
\]

for every \( M > 0 \) and \( 0 < \varepsilon \leq \varepsilon_0 \).

\[
\text{Lemma 2.12. Let } 1 \leq p < \infty, d \in \mathbb{N}, 0 < \alpha < 1, M_1, M_2 > 0 \text{ be given; let}
\]

\[
M(p, d, \alpha, M_1, M_2) = \{ f \in L^p(\mathbb{I}^d) : \| f \|_p = M_1, \omega_{1,p}(f, \delta) \leq M_2 \delta^{-\alpha} \text{ for } 0 < \delta \leq 1 \}.
\]

Then there are constants \( m_1, m_2 > 0 \) and \( \varepsilon_0 > 0 \) such that

\[
m_1(1/\varepsilon)^{d/\alpha} \leq \mathcal{H}_e(M(p, d, \alpha, M_1, M_2)) \leq m_2(1/\varepsilon)^{d/\alpha}
\]

for \( 0 < \varepsilon \leq \varepsilon_0 \).

\[
\text{Proof. As for } M_0 = \max(M_1, M_2) \text{ we have } M(p, d, \alpha, M_1, M_2) \subset A^+(p, d, 1, \alpha, M_0), \text{ the existence of } m_2 > 0 \text{ and } \varepsilon_0 > 0 \text{ such that}
\]

\[
\mathcal{H}_e(M(p, d, \alpha, M_1, M_2)) \leq m_0(1/\varepsilon)^{d/\alpha}
\]

for every \( 0 < \varepsilon \leq \varepsilon_0 \) follows from Corollary 2.11.

Now we prove the existence of \( m_1 \). Define

\[
\tilde{M}(p, d, \alpha, M_1, M_2) = \{ f \in L^p(\mathbb{I}^d) : \| f \|_p \leq M_1, \omega_{1,p}(f, \delta) \leq M_2 \delta^{-\alpha} \text{ for } 0 < \delta \leq 1 \}.
\]

Set \( C_0 = \min(M_1/2, M_2/2) \); for every constant \( c \) and \( f \in L^p(\mathbb{I}^d) \) we have \( \omega_{1,p}(f, \delta) \leq \omega_{1,p}(f + c\delta) \), and so \( \{ f + C_0 : f \in A^+(p, d, 1, \alpha, C_0) \} \subset \tilde{M}(p, d, \alpha, M_1, M_2) \). Thus \( \tilde{M}(p, d, \alpha, M_1, M_2) \) contains a subset isometric to \( A^+(p, d, 1, \alpha, C_0) \) and it follows from Corollary 2.11 that there are \( m^* > 0 \) and \( \eta > 0 \) such that

\[
m^*(1/\varepsilon)^{d/\alpha} \leq \mathcal{H}_e(\tilde{M}(p, d, \alpha, M_1, M_2))
\]

for each \( 0 < \varepsilon \leq \eta \).

Now, given an \( \varepsilon/2 \)-net of \( \tilde{M}(p, d, \alpha, M_1, M_2) \), we construct an \( \varepsilon \)-net of \( \tilde{M}(p, d, \alpha, M_1, M_2) \).

First notice that, for any \( a > 0 \), \( \| f - g \|_p \leq \varepsilon \Leftrightarrow \| af - ag \| \leq \varepsilon a \) and

\[
\text{(2.14) } M(p, d, \alpha, aM_1, M_2) = \{ af : f \in M(p, d, \alpha, M_1, M_2) \}.
\]

Let \( N^*(\alpha) \) denote the minimal number of elements of an \( \varepsilon \)-net of \( M(p, d, \alpha, aM_1, M_2) \), consisting only of elements of this set. It follows from (2.14) that \( N^*(\alpha) = N^*_{1/\alpha}(1) \). Let \( \varepsilon > 0 \) be given, and \( \tau = [M_1/\varepsilon] + 1 \). For each \( f \in \tilde{M}(p, d, \alpha, M_1, M_2) \) choose an integer \( k, 0 \leq k \leq \tau \), such that

\[
M_1(2k - 1)/2r \leq \| f \|_p \leq M_1(2k + 1)/2r.
\]

For \( \| f \|_p \neq 0 \) set

\[
f^* = \frac{M_1 k}{\tau} \frac{f}{\| f \|_p};
\]

then \( f^* \geq 0 \), \( \| f^* \|_p = (\tau/k) M_1 \),

\[
\omega_{1,p}(f^*, \delta) = \frac{M_1 k}{\tau} \frac{1}{\| f \|_p} \omega_{1,p}(f, \delta) \leq \frac{\tau}{k} M_1 M_2 \delta^{-\alpha}
\]

and \( f^* \in M(p, d, \alpha, (\tau/k) M_1, M_2) \). In addition,

\[
\| f - f^* \|_p \leq \frac{M_1 k}{\tau} \leq \frac{M_1}{2} \leq (M_1/\varepsilon)^{d/\alpha} = \frac{\varepsilon}{2}.
\]

Set \( U = \{ 0 \} \cup \bigcup_{k=1}^r \{ f_{k,1}, \ldots, f_{kn,k} \} \), where \( n_k = N^*_1(\varepsilon/2/k) \), \( k = 1, \ldots, r \), and \( \{ f_{k,1}, \ldots, f_{kn,k} \} \subset M(p, d, \alpha, (k/r) M_1, M_2) \) is an \( \varepsilon/2 \)-net of \( M(p, d, \alpha, (k/r) M_1, M_2) \). Then \( U \) is an \( \varepsilon \)-net of \( \tilde{M}(p, d, \alpha, M_1, M_2) \). This and (1.1) yield

\[
N^*_\varepsilon(\tilde{M}(p, d, \alpha, M_1, M_2)) \leq \#U + 1 \leq 1 + \sum_{k=1}^r N^*_\varepsilon(\frac{\tau}{k}) \leq 1 + \sum_{k=1}^r N^*_\varepsilon(\frac{\epsilon}{2}) \leq 1 + (1 + \tau) \sum_{k=1}^r N^*_\varepsilon(\frac{\epsilon}{4}) \leq (1 + \tau) N^*_\varepsilon(\frac{\epsilon}{4}),
\]

which implies

\[
\mathcal{H}_e(\tilde{M}(p, d, \alpha, M_1, M_2)) \leq \ln(M_1/\varepsilon + 2) + \mathcal{H}_e(M(p, d, \alpha, M_1, M_2)).
\]
As \( \lim_{\varepsilon \to 0} \varepsilon \ln(M_1/\varepsilon + 2) = 0 \) for any \( s > 0 \), it follows from the last inequality and (2.13) that there exist \( \varepsilon_1 > 0 \) and \( m_1 > 0 \) such that
\[
m_1(1/\varepsilon)^{d/\alpha} \leq \mathcal{H}_e(M(p, d, \alpha, M_1, M_2))
\]
for every \( 0 < \varepsilon \leq \varepsilon_1 \).

**3. The asymptotics of** \( \mathcal{H}_e(U(p, d, m, \alpha, \gamma, C)) \).

Recall that for \( m, d \in \mathbb{N}, 1 \leq p < \infty, 0 < \alpha < m \) and \( \gamma > 0 \) and \( C > 0 \),
\[
U(p, d, m, \alpha, \gamma, C) = \{ f \in L^p(\mathbb{R}^d) : \| f \|_p \leq C, \omega_{m,p}(f, \delta) \leq C \delta^\alpha \text{ for } \delta > 0, \Phi_p(f, \lambda) \leq C \lambda^{-\gamma} \text{ for } \lambda > 0 \},
\]
where \( \omega_{m,p}(f, \delta) = \sup_{|u| \leq \delta} \| \Delta_m^\omega f \|_p \) is the modulus of smoothness of \( f \) of order \( m \) in \( L^p(\mathbb{R}^d) \), and \( \Phi_p(f, \lambda) = \left( \int_{|x| \leq Z_\lambda} |f(x)|^p \, dx \right)^{1/p} \) is the tail function (recall that \( Z_\lambda = [-\lambda, \lambda]^d \)).

**Theorem 3.1.** Let \( 1 \leq p < \infty, m, d \in \mathbb{N}, 0 < \alpha < m \) and \( \gamma > 0 \) be given. There exist constants \( k_1, k_2 > 0 \) and \( \varepsilon_0 > 0 \), depending only on \( p, d, m, \alpha \) and \( \gamma \), such that
\[
k_1(C/\varepsilon)^{(d(1/\alpha + 1)/\gamma)} \leq \mathcal{H}_e(U(p, d, m, \alpha, \gamma, C)) \leq k_2(C/\varepsilon)^{(d(1/\alpha + 1)/\gamma)}
\]
for every \( C > 0 \) and \( 0 < \varepsilon \leq C \varepsilon_0 \).

The proof of Theorem 3.1 is split into several lemmas. First the upper estimate for \( \mathcal{H}_e(U(p, d, m, \alpha, \gamma, C)) \) is obtained.

**Lemma 3.2.** Let \( 1 \leq p < \infty, m, d \in \mathbb{N}, 0 < \alpha < m \) and \( \gamma > 0 \) be given. There exist constants \( k > 0 \) and \( \varepsilon_0 > 0 \), depending only on \( p, d, m, \alpha \) and \( \gamma \), such that
\[
\mathcal{H}_e(U(p, d, m, \alpha, \gamma, C)) \leq k(C/\varepsilon)^{(d(1/\alpha + 1)/\gamma)}
\]
for every \( C > 0 \) and \( 0 < \varepsilon \leq C \varepsilon_0 \).

**Proof.** For \( f \in L^p(Z_\lambda) \) and \( 0 < \delta \leq 2\lambda/m \) define
\[
\omega_{m,p}^{(\lambda)}(f, \delta) = \sup_{|u| \leq \delta} \left( \int_{Z_{\lambda}(mu)} |\Delta_m^\omega f(x)|^p \, dx \right)^{1/p},
\]
where \( Z_{\lambda}(mu) = \{ x \in Z_\lambda : x + mu \in Z_\lambda \} \). Set
\[
U_\lambda(p, d, m, \alpha, C) = \{ f \in L^p(Z_\lambda) : \| f \|_{L^p(Z_\lambda)} \leq C, \omega_{m,p}^{(\lambda)}(f, \delta) \leq C \delta^\alpha \text{ for } 0 < \delta \leq 2\lambda/m \}.
\]
Then for \( f \in U(p, d, m, \alpha, \gamma, C) \) and \( \lambda > 0 \) we have \( f_\lambda = f|_{Z_\lambda} \in U_\lambda(p, d, m, \alpha, C) \). For \( \varepsilon > 0 \) write \( \lambda_\varepsilon = (2C/\varepsilon)^{1/\gamma} \); then for \( f, g \in U(p, d, m, \alpha, \gamma, C) \)
\[
\| f - g \|_p \leq \left( \int_{Z_{\lambda_\varepsilon}} |f(x) - g(x)|^p \, dx \right)^{1/p} + \Phi_p(f - g, \lambda_\varepsilon)
\]
\[
\leq \| f - g \|_{L^p(Z_{\lambda_\varepsilon})} + \varepsilon,
\]
so if \( \| f_\lambda - g_\lambda \|_{L^p(Z_{\lambda_\varepsilon})} \leq \varepsilon \) then \( \| f - g \|_p \leq 2\varepsilon \); therefore
\[
\mathcal{H}_e(U(p, d, m, \alpha, \gamma, C)) \leq \mathcal{H}_{e,2}(U_\lambda(p, d, m, \alpha, C)).
\]

Now \( \mathcal{H}_{e,2}(U_\lambda(p, d, m, \alpha, C)) \) will be estimated from above; we will find \( \eta_\varepsilon > 0 \) and a subset \( A_\varepsilon \subset L^p(\mathbb{R}^d) \) such that \( \mathcal{H}_{e,2}(U_\lambda(p, d, m, \alpha, C)) \leq \mathcal{H}_{e,2}(A_\varepsilon) \).

Define
\[
\psi_\varepsilon : [0, 1] \to [-\lambda_\varepsilon, \lambda_\varepsilon], \quad \psi_\varepsilon(t) = \lambda_\varepsilon(2t - 1),
\]
\[
\psi_\varepsilon : [0, 1]^d \to Z_{\lambda_\varepsilon}, \quad \psi_\varepsilon(z_1, \ldots, z_d) = (\psi_\varepsilon(z_1), \ldots, \psi_\varepsilon(z_d)).
\]
Then for any \( f \in L^p(Z_{\lambda_\varepsilon}) \)
\[
\| f \|_{L^p(\psi_\varepsilon)} = (2\lambda_\varepsilon)^{d/p} \omega_{m,p}(f \circ \psi_\varepsilon, \delta/Z_{\lambda_\varepsilon}),
\]
\[
\Delta_m^\omega f(z) = \Delta_m^\omega (f \circ \psi_\varepsilon)(\psi_\varepsilon^{-1}(z)),
\]
\[
Z_{\lambda_\varepsilon}(mu) = \psi_\varepsilon \left( \left( \frac{m - \mu}{2\lambda_\varepsilon} \right) \right),
\]
which implies
\[
\omega_{m,p}^{(\lambda_\varepsilon)}(f, \delta) = (2\lambda_\varepsilon)^{d/p} \omega_{m,p}(f \circ \psi_\varepsilon, \delta/Z_{\lambda_\varepsilon})
\]
and for \( f \in U_\lambda(p, d, m, \alpha, C) \)
\[
\| f \circ \psi_\varepsilon \|_{L^p(\psi_\varepsilon)} \leq (2\lambda_\varepsilon)^{-d/p} C,
\]
\[
\omega_{m,p}(f \circ \psi_\varepsilon, \delta/Z_{\lambda_\varepsilon}) \leq (2\lambda_\varepsilon)^{-d/p} C \delta^\alpha.
\]

Set \( \eta_\varepsilon = \varepsilon (2\lambda_\varepsilon)^{-d/p}, A_\varepsilon = A(p, d, m, \alpha, (2\lambda_\varepsilon)^{-d/p} C \eta_\varepsilon) \). Notice that if \( f, g \in U_\lambda(p, d, m, \alpha, C) \) are such that \( \| f \circ \psi_\varepsilon - g \circ \psi_\varepsilon \|_{L^p(\psi_\varepsilon)} \leq \eta_\varepsilon \) then \( \| f - g \|_{L^p(Z_{\lambda_\varepsilon})} \leq \varepsilon \). In addition, \( \lambda_\varepsilon \geq 1 \) for \( \varepsilon \leq 2C \), and if \( f \in U_\lambda(p, d, m, \alpha, C) \) then \( f \circ \psi_\varepsilon \in A_\varepsilon \). Therefore
\[
\mathcal{H}_{e,2}(U_\lambda(p, d, m, \alpha, C)) \leq \mathcal{H}_{e,2}(A_\varepsilon).
\]

Corollary 2.10 implies the existence of constants \( \xi > 0 \) and \( \alpha > 0 \), independent of \( C \) and \( \varepsilon \), such that if \( \eta_\varepsilon \leq (2\lambda_\varepsilon)^{-d/p} C \xi \) then
\[
\mathcal{H}_{e,2}(A_\varepsilon) \leq a((2\lambda_\varepsilon)^{-d/p} C \eta_\varepsilon)^{d/\alpha} = k(C/\varepsilon)^{d(1/\alpha + 1/\gamma)}.
\]
where \( k = 2d(d+1)/\alpha \).
\( \varepsilon_0 \) must be chosen so that the inequalities \( \varepsilon \leq 2C \) and \( \eta_C \leq (2\lambda_C)^{\alpha-d/p} C \xi \) hold for any \( 0 < \varepsilon \leq C \varepsilon_0 \). It is enough to take \( \varepsilon_0 = \min(2, 2\xi/(\alpha+\gamma)) \). Now our lemma follows from (3.3)-(3.6). \( \blacksquare \)

To estimate \( H_n(U(p, d, m, \alpha, \gamma, C)) \) from below it is convenient to use Theorem 1.3. Some spaces of spline functions are also needed, and the classical relation between the distances from those subspaces and the moduli of smoothness will be recalled. The following notation will be used:

\[
N^{(s)}(x) = s(0, \ldots, s; \cdots (-x)_{-1}^{-1}) \quad \text{for } s \in \mathbb{N}, \quad x \in \mathbb{R},
\]

\[
N^{(s)}_{i,h}(x) = N^{(s)} \left( \frac{x - ih}{h} \right) \quad \text{for } h > 0, \quad i \in \mathbb{Z}, \quad x \in \mathbb{R},
\]

\[
N^{(s)}_{j,h}(t_1, \ldots, t_d) = N^{(s)}_{j_1,h_1}(t_1) \cdots N^{(s)}_{j_d,h_d}(t_d)
\]

for \( z = (z_1, \ldots, z_d) \in \mathbb{N}^d, \quad j = (j_1, \ldots, j_d) \in \mathbb{Z}^d, \quad h = (h_1, \ldots, h_d) \in \mathbb{R}^d, \quad h_i > 0, \quad t = (t_1, \ldots, t_d) \in \mathbb{R}^d. \)

Some useful properties of spline functions \( N^{(s)}_{j,h} \) are mentioned below. They are multivariant analogues of Schoenberg’s result (10).

(3.6) There exists a constant \( c_\varepsilon > 0 \) such that

\[
c_\varepsilon^{-1} (h_1 \cdots h_d)^{1/p} \left( \sum_{j \in \mathbb{Z}^d} |a_j|^p \right)^{1/p} \leq \left\| \sum_{j \in \mathbb{Z}^d} a_j N^{(s)}_{j,h} \right\|_p \leq c_\varepsilon (h_1 \cdots h_d)^{1/p} \left( \sum_{j \in \mathbb{Z}^d} |a_j|^p \right)^{1/p}
\]

for every sequence \((a_j)_{j \in \mathbb{Z}^d}\). (3.7) For \( a \in (\mathbb{N} \cup \{0\})^d \) and \( z \in \mathbb{N}^d \) with \((0, \ldots, 0) \leq a < z \) and \( g = \sum_{j \in \mathbb{Z}^d} g_j N^{(s)}_{j,h} \) we have

\[
h_1^{a_1} \cdots h_d^{a_d} D^a g = \sum_{j \in \mathbb{Z}^d} (D^a g_j) N^{(s-z)}_{j,a-h}
\]

where

\[
D^a g_j = \sum_{0 \leq b \leq a} \binom{a_1}{b_1} \cdots \binom{a_d}{b_d} (-1)^{|b|} g_{b_1+1} \cdots g_{b_d+1}
\]

(3.8) For each \( m, d \in \mathbb{N} \) there exists a constant \( M_{m, d} > 0 \) such that

\[
\| D^a f \|_p \leq M_{m, d} \| u \|_m \sum_{|a| = m} \| D^a f \|_p
\]

for any \( 1 \leq p < \infty, \ f \in C^m(\mathbb{R}^d) \cap L^p(\mathbb{R}^d) \) and \( u \in \mathbb{R}^d. \)

Define

\[
r_m = (m + 2, \ldots, m + 2) \in \mathbb{N}^d \quad \text{for } m \in \mathbb{N},
\]

\[
h_n = (h_{n,1}, \ldots, h_{n,m}) \in \mathbb{R}^d, \quad h_n = 1/2^n \quad \text{for } n \in \mathbb{N} \cup \{0\},
\]

\[
V_{n,p} \left( m \right) = \text{span} \{ N_{j,h_n}^{(m)} : j \in \mathbb{Z}^d \} \cap L^p(\mathbb{R}^d).
\]

Notice that \( V_{n,p} \left( m \right) \subset V_{n+1,p} \left( m \right). \)

For \( f \in L^p(\mathbb{R}^d) \) define

\[
E_{n,p}^{(m)}(f) = \inf \{ \| f - g \|_p : g \in V_{n,p} \left( m \right) \}.
\]

The following lemma recalls the well-known relation between \( E_{n,p}^{(m)}(f) \) and \( \omega_{m,p}(f, \delta) \) (cf. [3]).

**Lemma 3.9.** Let \( m, d \in \mathbb{N} \) be given. There exists a constant \( C_{m, d} > 0 \) such that

\[
\omega_{m,p}(f, 1/2k) \leq \frac{C_{m, d}}{2k m} \left( \| f \|_p + \sum_{i=0}^{k-1} 2^{im} E_{n,p}^{(m)}(f) \right)
\]

for every \( 1 \leq p < \infty, \ f \in L^p(\mathbb{R}^d) \) and \( k \in \mathbb{N} \cup \{0\}. \)

For \( m, d \in \mathbb{N}, \ 0 < \alpha < m \) and \( \gamma > 0 \) put

\[
k_n = k_n(m, \alpha, \gamma) = (m + 3)2^{m+n}/\gamma \quad \text{for } n \in \mathbb{N} \cup \{0\},
\]

\[
U_n^{(m)} = \text{span} \{ N_{j,h_n}^{(m)} : j \in \mathbb{Z}^d, \text{ supp } N_{j,h_n}^{(m)} \subset [-k_n, k_n]^d \},
\]

and for \( 1 \leq p < \infty, \ f \in L^p(\mathbb{R}^d) \) set

\[
D_{n,p}^{(m)}(f) = \inf \{ \| f - g \|_p : g \in U_n^{(m)} \}.
\]

It follows from the definitions of \( U_n^{(m)} \) and \( V_{n,p} \left( m \right) \) that \( U_n^{(m)} \subset V_{n,p} \left( m \right) \), which implies \( E_{n,p}^{(m)} \leq D_{n,p}^{(m)}(f) \). The properties of the functions \( N_{j,h_n}^{(m)} \) imply that \( U_n^{(m)} \subset U_{n+1}^{(m)}. \)

Define

\[
W(p, d, m, \alpha, \gamma, C) = \{ f \in L^p(\mathbb{R}^d) : \| f \|_p \leq C, \| D^a f \|_p \leq C 2^{m \alpha} \text{ for } n \in \mathbb{N} \cup \{0\} \}
\]

**Lemma 3.10.** Let \( d, m \in \mathbb{N}, \ 0 < \alpha < m, \ \gamma > 0 \) and \( 1 \leq p < \infty \) be given. There exists a constant \( a > 0 \) such that for any \( C > 0 \)

\[
W(p, d, m, \alpha, \gamma, aC) \subset U(p, d, m, \alpha, \gamma, C).
\]

**Proof.** First notice that for \( f \in W(p, d, m, \alpha, \gamma, aC) \)

\[
\| f \|_p \leq aC.
\]

Now we estimate \( \omega_{m,p}(f, \delta) \). For \( \delta > 1 \) we have

\[
\omega_{m,p}(f, \delta) \leq 2^{m} \| f \|_p \leq 2^{m} aC \delta^{\alpha}.
\]
For $0 < \delta \leq 1$ choose $k \in \mathbb{N} \cup \{0\}$ such that $1/2^{k+1} \leq \delta \leq 1/2^k$; then it follows from Lemma 3.9 that

$$\omega_{m,p}(f,\delta) \leq \omega_{m,p}(f,\frac{1}{2^k}) \leq \frac{C_{m,d}}{2^{km}} \left( \|f\|_p + \sum_{i=0}^{k} 2^{in} D_{i,p}^{(m)}(f) \right) \leq \frac{C_{m,d}}{2^{km}} \left( \|f\|_p + \sum_{i=0}^{k} 2^{in} D_{i,p}^{(m)}(f) \right) \leq 2^{m+1} C_{m,d} a C^a.$$

We also have to estimate $\Phi_p(f,\lambda)$. As $f \in W(p, d, m, \alpha, \gamma, aC)$, the definition of $U_n^{(m)}$ implies

$$\Phi_p(f, \lambda_n) \leq \left( \|f\|_p + \sum_{i=0}^{k} 2^{in} D_{i,p}^{(m)}(f) \right) \leq \frac{(m+3)^\gamma}{\lambda_n} a C.$$

For $0 < \lambda < \lambda_0$ we have

$$\Phi_p(f, \lambda) \leq \|f\|_p + a C \leq \left( \frac{(m+3)^\gamma}{\lambda} \right) a C.$$

For $\lambda \geq \lambda_0$ choose $n \in \mathbb{N} \cup \{0\}$ such that $\lambda_n \leq \lambda \leq \lambda_{n+1}$; then

$$\Phi_p(f, \lambda) \leq \Phi_p(f, \lambda_{n+1}) \leq \frac{(m+3)^\gamma}{\lambda_{n+1}} a C \leq \frac{(m+3)^\gamma}{\lambda_n} a C.$$

It follows from (3.11)-(3.15) that it is enough to take

$$a = \min \{ (2^{-m_2}, 2^{-(m_1+1)} C_{m,d}, 2^{-\gamma} (m+3)^\gamma) \}.$$

Now the $\varepsilon$-entropy of $W(p, d, m, \alpha, \gamma, C)$ will be estimated from below.

**Lemma 3.16.** Let $d, m \in \mathbb{N}_0$, $0 < \alpha < m$, $\gamma > 0$ and $1 \leq \rho < \infty$ be given. There exist constants $k_0 > 0$ and $\varepsilon_0 > 0$ such that

$$k_0(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \leq \mathcal{H}_\varepsilon(W(p, d, m, \alpha, \gamma, C))$$

for every $C > 0$ and $0 < \varepsilon \leq C_0$.

**Proof.** We will use Theorem 1.3.

First let $C = 1$. Write $a_n^{(m)} = \dim U_n^{(m)}$; then $2^n k_n^d \leq a_n^{(m)} < \infty$ for $n \in \mathbb{N} \cup \{0\}$. Choose a sequence $\Phi = \{ \varphi_1, \varphi_2, \ldots \}$ in $L^p(\mathbb{R}^d)$ so that $U_n^{(m)} = \text{span}\{ \varphi_1, \ldots, \varphi_{a_n^{(m)}}^{(m)} \}$ for each $n \in \mathbb{N} \cup \{0\}$. The sequence $\Delta = \{ \delta_0, \delta_1, \ldots \}$ is defined as follows:

$$\delta_0 = \ldots = \delta_{d_{(m)-1}^{(m)}} = 1, \quad \delta_k = \frac{1}{2^{n\alpha}} \quad \text{for} \quad a_n^{(m)} \leq k < a_{n+1}^{(m)}, \ n \geq 0.$$

Then $W = W(p, d, m, \alpha, \gamma, 1) = A(\Delta, \Phi)$.

Let $c = 2$ in Theorem 1.3. For $0 < \varepsilon \leq 1/2$ choose $j \in \mathbb{N}$ such that $2^{-j+1} < \varepsilon \leq 2^{-j+1}$; then $N_{j-3} = a_{n_{j-1}}^{(m)}$, where $n_i = \max\{n_i, i/\alpha\}$.

(1.2) and (1.4) now imply

$$\mathcal{H}_\varepsilon(W) \geq C_{2\varepsilon}(W) \geq N_{j-3} \ln 2 = a_{n_{j-1}}^{(m)} \ln 2 \geq \left( \frac{2^{n_{j-1}} \ln 2}{\alpha} \right)^{d/(1/\alpha+1/\gamma)} (m+3)^d (1/\varepsilon)^{d(1/\alpha+1/\gamma)} \ln 2.$$

Take $k = 2^{-d(1/\alpha+1/\gamma)} (m+3)^d \ln 2, \ v_0 = 1/2$. As $W(p, d, m, \alpha, \gamma, C) = \{ f : f \in W(p, d, m, \alpha, \gamma, 1) \}$, it follows that

$$\mathcal{H}_\varepsilon(W(p, d, m, \alpha, \gamma, C)) = \mathcal{H}_\varepsilon(W(p, d, m, \alpha, \gamma, 1)) = k_0(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \leq \mathcal{H}_\varepsilon(W(p, d, m, \alpha, \gamma, C)).$$

Theorem 3.1 now follows from Lemmas 3.2, 3.10 and 3.16.

Now let $U^+(p, d, m, \alpha, \gamma, C) = \{ f \in U(p, d, m, \alpha, \gamma, C) : f \geq 0 \}$ for $m = 1, 2$. The following result can be proved similarly to Corollary 2.11.

**Corollary 3.17.** Let $d \in \mathbb{N}_0$, $0 < \alpha < 1$, $\gamma > 0$ and $1 \leq p < \infty$ be given. There exist constants $k_1, k_2 > 0$ and $\varepsilon_0 > 0$ such that

$$k_1(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \leq \mathcal{H}_\varepsilon(U^+(p, d, m, \alpha, \gamma, C)) \leq k_2(C/\varepsilon)^{d(1/\alpha+1/\gamma)}$$

for every $C > 0$ and $0 < \varepsilon \leq C_0$.

Similar inequalities will be proved for $m = 2$ and $0 < \alpha < 2$. First some more notation will be introduced. Define

$$M_{j,k}^{(s)} = \{ (i_1, \ldots, i_k)^{-1} N_{j,k}^{(s)} \}$$

for $j \in \mathbb{Z}^d$, $k \in \mathbb{N}^d$, $1 = (i_1, \ldots, i_k) \in \mathbb{R}^d$, $k_0 > 0$, $Q_j^{(s)}(x) = \sum_{j \in \mathbb{Z}^d} \int_{\mathbb{R}^d} M_j^{(s)}(t) f(t) dt$. For $f \in L^p(\mathbb{R}^d)$,

$$\omega_{k_0, e_i, p}(f, \delta) = \sup_{|\xi| \leq \delta} \left| \Delta_{k_0}^{(s)} f(x) \right|$$

for $f \in L^p(\mathbb{R}^d)$.

where $e_i = (\delta_{i_1}, \ldots, \delta_{i_k})$. It was proved in [3] that for any $f \in L^p(\mathbb{R}^d)$

$$\| f - Q_j^{(s)} f \|_p \leq 16 \sum_{i=1}^{d} \omega_{k_0, e_i, p}(f, \varepsilon_i).$$

For $m, n \in \mathbb{N} \cup \{0\}$, $f \in C^{k_0}(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ and $\delta > 0$ the following inequality holds:

$$\omega_{k_0, e_i, p}(f, \delta) \leq k_0 \omega_{k_0, e_i, p}(f, \delta).$$

**Lemma 3.20.** Let $d \in \mathbb{N}_0$, $0 < \alpha < 2$, $\gamma > 0$ and $1 \leq p < \infty$ be given. There exist constants $k > 0$ and $\varepsilon_0 > 0$ such that for every $C > 0$ and $0 < \varepsilon \leq C_0$

$$k_0(C/\varepsilon)^{d(1/\alpha+1/\gamma)} \leq \mathcal{H}_\varepsilon(U^+(p, d, 2, \alpha, \gamma, C)).$$
Proof. Let \( z = (4, \ldots, 4) \in \mathbb{N}^d \) and recall that
\[
U_n^{(2)} = \text{span} \left\{ N_{j_n, h_n}^{(2)} : h_n = \left( \frac{1}{2^m}, \ldots, \frac{1}{2^n} \right) \in \mathbb{R}^d, \supp N_{j_n, h_n}^{(2)} \subset [-k_n, k_n]^d \right\}.
\]
For \( f \in L^p(\mathbb{R}^d) \) set
\[
D_n^{(+)}(f) = \inf \{ ||f - g||_p : g \in U_n^{(2)}, \quad g \geq 0 \}
\]
and define
\[
W^+(p, d, \alpha, \gamma, C) = \{ f \in L^p(\mathbb{R}^d) : ||f||_p \leq C, \quad D_n^{(+)}(f) \leq C/2n^\alpha \text{ for } n \in \mathbb{N} \cup \{0\} \}.
\]
Notice that if \( f \in W^+(p, d, \alpha, \gamma, C) \) then \( f \geq 0 \). As in the proof of Lemma 3.10, it is possible to find a constant \( c > 0 \) such that for every \( C > 0 \)
\begin{equation}
W^+(p, d, \alpha, \gamma, aC) \subset U^+(p, d, 2, \alpha, \gamma, C).
\end{equation}

Now the \( \varepsilon \)-capacity of \( W^+(p, d, \alpha, \gamma, C) \) will be estimated from below. It follows from (3.6), (3.7), (3.18) and (3.19) that there exists a constant \( A > 0 \) such that for any \( s \in \mathbb{N} \cup \{0\} \), \( g \in U_s^{(2)} \) and \( i = (i_1, \ldots, i_d) \) with \( i > 0 \)
\[
||g - Q^{(2)}_s g||_p \leq A ||g||_p \frac{\sum_{i=1}^d \frac{1}{2^{i}}} {h_s^2},
\]
(recall that \( h_s = 1/2^s \)). Let \( s \in \mathbb{N} \) be chosen so that \( 2 \cdot 2^{(1+\alpha/\gamma)2^{\alpha/\gamma} - 1} \geq 1 \); then for \( g \in U_s^{(2)} \) and \( n > s \) we have \( \supp Q^{(2)}_n g \subset [-k_n, k_n]^d \). As \( Q^{(2)}_s g \geq 0 \) for \( g \geq 0 \), the above inequality gives for \( n > s \) and \( g \in U_s^{(2)} \)
\[
||g - Q^{(2)}_n g||_p \leq A ||g||_p \frac{\sum_{i=1}^d \frac{1}{2^{i}}} {h_n^2}.
\]
As \( D_n^{(+)}(g) \leq ||g||_p \) for \( n \leq s \), putting \( C_0 = \min(1, 1/(4A)) \) we obtain for \( g \in U_s^{(2)}, \quad g \geq 0, \quad ||g||_p \leq C_0Ch_n^\alpha \)
\[
D_n^{(+)}(g) \leq C_0Ch_n^\alpha \quad \text{for } n \in \mathbb{N} \cup \{0\}.
\]
Set \( E_s = \{ j \in \mathbb{Z}^d : \supp N_{j, h_s}^{(2)} \subset [-k_s, k_s]^d \} \); the last inequality implies
\[
W_s, C = \left\{ g = \sum_{j \in E_s} a_j N_{j, h_s}^{(2)} : a_j \geq 0, \quad ||g||_p \leq C_0Ch_s^\alpha \right\} \subset W^+(p, d, \alpha, \gamma, C),
\]
so for any \( \varepsilon > 0 \)
\begin{equation}
\mathcal{C}_\varepsilon(W_s, C) \leq \mathcal{C}_\varepsilon(W^+(p, d, \alpha, \gamma, C)).
\end{equation}

Now the \( \varepsilon \)-capacity of \( W_s, C \) will be estimated from below. It follows from (3.6) that for some \( c_\varepsilon > 0 \)
\[
c_\varepsilon \cdot h_s^d/p \left( \sum_{j \in E_s} |a_j|_p \right)^{1/p} \leq \left( \sum_{j \in E_s} a_j N_{j, h_s}^{(2)} \right)**_p \leq c_\varepsilon h_s^d/p \left( \sum_{j \in E_s} |a_j|_p \right)^{1/p}.
\]

Let
\[
G = \left\{ a = (a_j)_{j \in E_s} : ||a||_p = \left( \sum_{j \in E_s} |a_j|_p \right)^{1/p} \leq c_\varepsilonCh_s^{-d/p}, \quad a_j \geq 0 \right\}.
\]

Then for any \( a = (a_j)_{j \in E_s} \in G \) we have \( g_a = \sum_{j \in E_s} a_j N_{j, h_s}^{(2)} \in W_s, C \), and for any \( a, b \in G \), if \( ||a - b||_p \geq \eta_s = c_\varepsilonCh_s^{-d/p} \) then \( ||g_a - g_b||_p \geq \varepsilon \), so \( \mathcal{C}_\varepsilon(W_s, C) \geq \mathcal{C}_\varepsilon(G) \). \( G \) is a subset of \( R^{\frac{d^2}{2}} \), where \( d_s^2 = \dim U_s^{(2)} = \#E_s \).

It can be checked (using the method of the proof of Lemma 1 of [9]) that
\[
\mathcal{M}_{\eta}(G) \geq \left( \frac{1}{2\eta} \right)^{d_s^2} \left( \frac{C_0Ch_s^{-d/p}}{c_\varepsilonCh_s^{-d/p}} \right)^{d_s^2},
\]
which implies
\[
\mathcal{C}_\varepsilon(G) \geq d_s^2 \ln \left( \frac{C_0Ch_s^{-d/p}}{2\eta c_\varepsilon} \right).
\]

Therefore
\[
\mathcal{C}_\varepsilon(W_s, C) \geq \mathcal{C}_\varepsilon(G) \geq d_s^2 \ln \left( \frac{C_0Ch_s^{-d/p}}{2\eta c_\varepsilon} \right).
\]
Let
\[
\varepsilon_{s, C} = \left[ \frac{1}{2\eta} \log \left( \frac{C_0Ch_s^{-d/p}}{2\eta c_\varepsilon} \right) \right] \quad \text{and then}
\]
\[
\frac{C_0Ch_s^{-d/p}}{2\eta c_\varepsilon} \geq \varepsilon,'
\]
which implies
\begin{equation}
\mathcal{C}_\varepsilon(W_s, \alpha, \gamma, C) \geq d_s^2 \geq \left( \frac{2^{(1+\alpha/\gamma)/(2\alpha/\gamma)} - 1} {2\alpha/\gamma} \right)^{d} \geq k(C/s)^{d(1/\alpha+1/\gamma)}.
\end{equation}
Choose \( \varepsilon_0 > 0 \) in such a way that
\[
2 \cdot 2^{(1+\alpha/\gamma)/(2\alpha/\gamma)} - 1 \geq 1
\]
for every \( 0 < \varepsilon \leq \varepsilon_0 \). The lemma now follows from (3.21)–(3.23). \( \blacksquare \)

Lemmas 3.2, 3.20 and inequalities (1.2) imply

Lemmas 3.24. Let \( d \in \mathbb{N}, 0 < \alpha < 2, \gamma > 0 \) and \( 1 < p < \infty \) be given. There exist constants \( k_1, k_2 > 0 \) and \( \varepsilon_0 > 0 \) such that
\[
k_1(C/s)^{d(1/\alpha+1/\gamma)} \geq \varepsilon_0(U^+(p, d, 2, \alpha, \gamma, C)) \leq k_2(C/s)^{d(1/\alpha+1/\gamma)}
\]
for every \( C > 0 \) and \( 0 < \varepsilon \leq \varepsilon_0 \).
Lemma 3.25. Let \( d \in \mathbb{N}, m \in \{1, 2\}, 0 < \alpha < m, \gamma > 0, 1 \leq p < \infty \) and 
\( M_1, M_2, M_3 > 0 \) be given. Set 
\[
M(p, d, m, \alpha, \gamma, M_1, M_2, M_3) = \{ f \in L^p(\mathbb{R}^d) : f \geq 0, \|f\|_p = M_1, \omega_{m, p}(f, \delta) \leq M_2 \delta^\alpha \text{ for } \delta > 0, \Phi_p(f, \lambda) \leq M_3 \lambda^{-\gamma} \text{ for } \lambda > 0 \}.
\]
Assume that there exists \( \varphi \in L^p(\mathbb{R}^d) \) such that \( \varphi \geq 0, \|\varphi\|_p > 0 \),
\[
\omega_{m, p}(\varphi, \delta) \leq a_1 \|\varphi\|_p \delta^\alpha \text{ for } \delta > 0, \\
\Phi_p(\varphi, \lambda) \leq a_2 \|\varphi\|_p \lambda^{-\gamma} \text{ for } \lambda > 0
\]
for some \( 0 < a_1 < M_2, 0 < a_2 < M_3 \). Then there exist \( k_1, k_2, e_0 > 0 \) such that
\[
k_1(1/e)^{d(1/\alpha + 1/\gamma)} \leq \mathcal{H}_e(M(p, d, m, \alpha, \gamma, M_1, M_2, M_3)) \leq k_2(1/e)^{d(1/\alpha + 1/\gamma)}.
\]

Proof. Define
\[
\tilde{M}(p, d, m, \alpha, \gamma, M_1, M_2, M_3) = \{ f \in L^p(\mathbb{R}^d) : f \geq 0, \omega_{m, p}(f, \delta) \leq M_2 \|f\|_p \delta^\alpha \text{ for } \delta > 0, \Phi_p(f, \lambda) \leq M_3 \|f\|_p \lambda^{-\gamma} \text{ for } \lambda > 0 \},
\]
and set \( \varphi_{M_1} = M_1 \varphi/(2\|\varphi\|_p) \) and 
\[
b_1 = \frac{M_1}{2}, \quad b_2 = \frac{M_1(M_2 - a_1)}{2}, \quad b_3 = \frac{M_1(M_2 - a_2)}{2}, \quad b = \min(b_1, b_2, b_3).
\]
Then \( \{ \varphi_{M_1} + f : f \in U^+(p, d, m, \alpha, \gamma, b) \} \subset \tilde{M}(p, d, m, \alpha, \gamma, M_1, M_2, M_3) \) and the rest of the proof is similar to the proof of Corollary 2.12. \( \blacksquare \)

4. \( \varepsilon \)-Entropy and nonparametric density estimation. A measurable function \( F_n : \mathbb{R}^{nd} \times \mathbb{R}^d \to \mathbb{R} \) is called a density estimator if for any \( x_1, \ldots, x_n \in \mathbb{R}^d, F_n(x_1, \ldots, x_n) \) \( : \mathbb{R}^d \to \mathbb{R} \) is a probability density on \( \mathbb{R}^d \).

For a given density \( f : \mathbb{R}^d \to \mathbb{R} \) and density estimator \( F_n : \mathbb{R}^{nd} \times \mathbb{R}^d \to \mathbb{R} \),
\[
E_fD(f, F_n) = \int_{\mathbb{R}^{nd}} D(f, F_n(x_1, \ldots, x_n)) f(x_1) \ldots f(x_n) \, dx_1 \ldots dx_n,
\]
where
\[
D(f, g) = \int_{\mathbb{R}^d} |f(x) - g(x)| \, dx \quad \text{for } f, g \in L^1(\mathbb{R}^d).
\]

Let \( \mathcal{F} \) be a family of densities on \( \mathbb{R}^d \). \( \mathcal{F} \) is considered as a metric space with metric induced from \( L^1(\mathbb{R}^d) \). Define
\[
\mathcal{R}_\mathcal{F}(n) = \inf \sup_{F_n \in \mathcal{F}} E_fD(f, F_n),
\]
where the infimum is taken over all density estimators \( F_n \) (based on a sample of size \( n \)) such that \( F_n(x_1, \ldots, x_n) \in \mathcal{F} \) for any \( x_1, \ldots, x_n \in \mathbb{R}^d \). \( \mathcal{R}_\mathcal{F}(n) \) is called the minimax risk for \( \mathcal{F} \), corresponding to samples of size \( n \) and loss function \( D \). Our aim is to establish some relations between \( \mathcal{R}_\mathcal{F}(n) \) and \( \mathcal{H}_e(\mathcal{F}) \).

Theorem 4.1. Let \( \mathcal{F} \) be a family of densities on \( \mathbb{R}^d \) which is totally bounded in the \( L^1(\mathbb{R}^d) \) metric.

(4.2) If there exist \( \varepsilon_0 > 0, C > 0 \) and \( \eta > 0 \) such that
\[
\mathcal{H}_e(\mathcal{F}) \leq C(1/e)^\eta \quad \text{for } 0 < e \leq e_0,
\]
then there exist \( M > 0 \) and \( n_0 \in \mathbb{N} \) such that
\[
\mathcal{R}_\mathcal{F}(n) \leq M n^{-1/(2+\eta)} \quad \text{for } n \geq n_0.
\]

(4.3) If there exist \( M > 0, \eta > 0 \) and \( n_0 \in \mathbb{N} \) such that
\[
M n^{-1/(2+\eta)} \leq \mathcal{R}_\mathcal{F}(n) \quad \text{for } n \geq n_0,
\]
then there exist \( C > 0 \) and \( e_0 > 0 \) such that
\[
C(1/e)^\eta \leq \mathcal{H}_e(\mathcal{F}) \quad \text{for } 0 < e \leq e_0.
\]

The result analogous to (4.2), but under the additional assumption that the supports of all elements of \( \mathcal{F} \) are contained in a compact set \( S \subset \mathbb{R}^d \), was proved earlier by P. Groeneboom ([7]). The present proof is based on the main idea of the proof in [7]. The following lemma will be used ([7]):

Lemma 4.4. Let \( S \) be a compact subset of \( \mathbb{R}^d \), and \( f_1, f_2 \) be two densities on \( \mathbb{R}^d \) with supports contained in \( S \). For \( \xi > 0 \) define
\[
B(f_1, \xi) = \{ h : \mathbb{R}^d \to \mathbb{R} : h \geq 0, \int_S h(x) \, dx = 1, D(f_1, h) \leq \xi \}.
\]

For any \( n \in \mathbb{N} \) there exists a measurable function \( \varphi_{f_1, f_2} : S^n \to \{0, 1\} \) such that
\[
\sup_{h \in B(f_1, \xi)} \int_{S^n} \varphi_{f_1, f_2}(x_1, \ldots, x_n) \prod_{i=1}^n h(x_i) \, dx_1 \ldots dx_n
\]
\[
+ \sup_{h \in B(f_1, \xi)} \int_{S^n} (1 - \varphi_{f_1, f_2}(x_1, \ldots, x_n)) \prod_{i=1}^n h(x_i) \, dx_1 \ldots dx_n
\]
\[
\leq \exp\left(-\frac{\eta}{8} n(D(f_1, f_2) - 2\xi)^2 \right).
\]

Proof of Theorem 4.1. The following notation will be used:
\[
\mathcal{R}_\lambda = \{ x \in \mathbb{R}^d : \|x\| \leq \lambda \} \quad \text{for } \lambda > 0,
\]
\[
\Phi(f, \lambda) = \int_{\mathbb{R}^d} |f(x)| \, dx \quad \text{for } \lambda > 0, f \in L^1(\mathbb{R}^d),
\]
and $g(\lambda) = \sup_{f \in \mathcal{F}} \Phi(f, \lambda)$. Notice that $g$ is nonincreasing and nonnegative; we now show that $\lim_{\lambda \to \infty} g(\lambda) = 0$. It is enough to check that for every $\xi > 0$ there is a $\lambda_\xi > 0$ such that $g(\lambda_\xi) \leq \xi$. Let $\{f_1, \ldots, f_k\}$ be a finite $\xi/2$-net of $\mathcal{F}$ (which exists since $\mathcal{F}$ is totally bounded); then for each $i$, $1 \leq i \leq k$, there is a $\lambda_i$ such that $\Phi(f_i, \lambda_i) \leq \xi/2$. It is enough to put $\lambda = \max\{\lambda_1, \ldots, \lambda_k\}$.

Now let $\lambda \geq \lambda_1/2$; then $\Phi(f, \lambda) \leq g(\lambda) \leq 1/2$ for $f \in \mathcal{F}$. Moreover, define

$$f_\lambda(x) = \begin{cases} 0 & \text{for } ||x|| > \lambda, \\ \frac{1}{1 - \Phi(f, \lambda)} f(x) & \text{for } ||x|| \leq \lambda. \end{cases}$$

Clearly $f_\lambda$ is a probability density on $\mathbb{R}^d$ with support contained in $\mathbb{R}^d_\lambda$. Note that for $f, h \in \mathcal{F}$ we have

$$D(f, h) \leq 4g(\lambda) + D(f_\lambda, h_\lambda),$$

(4.6)

$$D(f_\lambda, h_\lambda) \leq 2g(\lambda) + D(f, h).$$

(4.7)

Set $u(\varepsilon) = \mathcal{H}_\mathcal{F}^d(\varepsilon)$; $u(\varepsilon)$ is a nonincreasing, left-continuous function and $u(\varepsilon) \geq \mathcal{H}^d_{\mathcal{F}}$ for $\varepsilon > 0$.

Now let $0 < \varepsilon < 1$ and $\lambda > 0$ such that $g(\lambda) \leq \varepsilon/4$ be given. We are going to construct a suitable density estimator.

Let $\{f_1, \ldots, f_k\}$ be an $\varepsilon$-net of $\mathcal{F}$, consisting of elements of $\mathcal{F}$, such that $\ln k = \mathcal{H}^d_{\mathcal{F}}(\varepsilon) \leq u(\varepsilon)$. Write $h_\lambda = (f_\lambda)$, and put $\xi = \varepsilon + 2g(\lambda)$.

For $i \leq j$ choose $\varphi_{h_\lambda, h_{\lambda}} : (\mathbb{R}^d_\lambda)^n \to \{0, 1\}$ as in Lemma 4.4. Moreover, let $\varphi_{h_\lambda, h_{\lambda}} = 1 - \varphi_{h_\lambda, h_{\lambda}}$ for $i < j$. Clearly, in the latter case inequality (4.5) holds as well.

For $(x_1, \ldots, x_n) \in (\mathbb{R}^d_\lambda)^n$ put

$$J_{h_\lambda}(x_1, \ldots, x_n) = \{ h_\lambda : \varphi_{h_\lambda, h_{\lambda}}(x_1, \ldots, x_n) = 1 \}$$

and

$$L_{h_\lambda}(x_1, \ldots, x_n) = \max_{h_\lambda \in J_{h_\lambda}(x_1, \ldots, x_n)} D(h_\lambda, h_{\lambda}) \quad \text{if} \quad J_{h_\lambda}(x_1, \ldots, x_n) \neq \emptyset,$$

$$0 \quad \text{if} \quad J_{h_\lambda}(x_1, \ldots, x_n) = \emptyset.$$

Now define

$$\Theta(x_1, \ldots, x_n) = \begin{cases} f_1 & \text{if } ||x|| > \lambda \text{ for some } 1 \leq l \leq n, \\ f_i \text{ where } L_{h_\lambda}(x_1, \ldots, x_n) = \min_{h_\lambda \in J_{h_\lambda}(x_1, \ldots, x_n)} L_{h_\lambda}(x_1, \ldots, x_n) & \text{if } ||x|| \leq \lambda \text{ for all } 1 \leq l \leq n. \end{cases}$$

For $f \in \mathcal{F}$ and $i \in \mathbb{N}$ we have

$$P^n_{\mathcal{F}}(x_1, \ldots, x_n) \in (\mathbb{R}^d_\lambda)^n : D(f, \Theta(x_1, \ldots, x_n)) \geq (5 + i)\varepsilon$$

$$\leq P^n_{\mathcal{F}}(x_1, \ldots, x_n) \in (\mathbb{R}^d_\lambda)^n : D(f, \Theta(x_1, \ldots, x_n)) \geq (5 + i)\varepsilon,$$

(4.8)

where $P^n_{\mathcal{F}}$ denotes the probability measure on $\mathbb{R}^d$ with density $\prod_{i=1}^n f(x_i)$.

Choose $r, 1 \leq r \leq k$, such that $D(f, f_r) \leq \varepsilon$. Then

$$D(f, \Theta(x_1, \ldots, x_n)) \leq D(f, \Theta(x_1, \ldots, x_n)) + \varepsilon$$

and (by (4.6) and the choice of $\lambda$)

$$D(f, \Theta(x_1, \ldots, x_n)) \leq \varepsilon + D(h_r, \Theta(x_1, \ldots, x_n)),$$

so (4.8) implies

$$P^n_{\mathcal{F}}(x_1, \ldots, x_n) \in (\mathbb{R}^d_\lambda)^n : D(f, \Theta(x_1, \ldots, x_n)) \geq (5 + i)\varepsilon$$

$$\leq P^n_{\mathcal{F}}(x_1, \ldots, x_n) \in (\mathbb{R}^d_\lambda)^n : D(h_r, \Theta(x_1, \ldots, x_n)) \geq (3 + i)\varepsilon.$$ (4.9)

It follows from the definition of $\Theta$ that if $D(h_r, \Theta(x_1, \ldots, x_n)) \geq a$ then there is $j, 1 \leq j \leq k$, such that $\varphi_{h_r, h_j}(x_1, \ldots, x_n) = 1$ and $D(h_r, h_j) \geq a$. Therefore we obtain from (4.9)

$$P^n_{\mathcal{F}}(x_1, \ldots, x_n) \in (\mathbb{R}^d_\lambda)^n : D(f, \Theta(x_1, \ldots, x_n)) \geq (5 + i)\varepsilon$$

$$\leq P^n_{\mathcal{F}}(x_1, \ldots, x_n) \in (\mathbb{R}^d_\lambda)^n : \exists 1 \leq i \leq k \text{ s.t. } D(h_r, h_i) \geq (3 + i)\varepsilon$$

and $\varphi_{h_r, h_j}(x_1, \ldots, x_n) = 1$.

Let

$$k_j = \# \{l : (3 + f_j)\varepsilon < D(h_r, h_j) \leq (4 + f_j)\varepsilon \};$$

then

$$\sum_{j \geq 1} k_j \leq k.$$ (4.10)

As $D(f, f_r) \leq \varepsilon$, (4.7) implies $D(f_r, h_r) \leq \varepsilon + 2g(\lambda) = \xi$. Thus, (4.5) and (4.10) give

$$P^n_{\mathcal{F}}(x_1, \ldots, x_n) \in (\mathbb{R}^d_\lambda)^n : D(f, \Theta(x_1, \ldots, x_n)) \geq (5 + i)\varepsilon$$

$$\leq \sum_{j \geq 1} \exp\left(-\frac{1}{2n}D(h_r, h_j) - 2\xi\varepsilon\right) \leq \sum_{j \geq 1} k_j \exp\left(-\frac{1}{2n}j^2\varepsilon^2\right).$$

Now

$$E_D(f, \Theta) = \int_{\mathbb{R}^d} D(f, \Theta(x_1, \ldots, x_n)) \prod_{i=1}^n f(x_i) \, dx_1 \cdots dx_n$$

$$\leq \left( \int_{\mathbb{R}^d} \left( \int_{(\mathbb{R}^d_\lambda)^n} D(f, \Theta(x_1, \ldots, x_n)) \prod_{i=1}^n f(x_i) \, dx_1 \cdots dx_n \right) dx \right)^{1/n}$$

$$\leq 2ng(\lambda) + \varepsilon \left( \sum_{j \geq 1} k_j \exp\left(-\frac{1}{2n}j^2\varepsilon^2\right) \right).$$

For $\varepsilon \geq 2/\sqrt{n}$ we have $j \exp\left(-\frac{1}{2n}j^2\varepsilon^2\right) \leq \exp\left(-\frac{1}{2n}j^2\varepsilon^2\right)$, which together with the last inequality gives

$$E_D(f, \Theta) \leq 2ng(\lambda) + \varepsilon \left( \sum_{j \geq 1} k_j \exp\left(-\frac{1}{2n}j^2\varepsilon^2\right) \right)$$

$$\leq 2ng(\lambda) + \varepsilon \left( \sum_{j \geq 1} k_j \exp\left(-\frac{1}{2n}j^2\varepsilon^2\right) \right),$$

(4.11)
and this implies
\[(4.12) \quad R_F(n) \leq 2ng(\lambda) + e(6 + \exp\{u(\varepsilon) - \frac{1}{3}n^2\}).\]

Since (4.12) holds for any \(\lambda > 0\) such that \(g(\lambda) \leq 1/4\), and \(\lim_{\lambda \to \infty} g(\lambda) = 0\), we get
\[(4.13) \quad R_F(n) \leq e(6 + \exp\{u(\varepsilon) - \frac{1}{3}n^2\}).\]

We recall that the inequalities (4.11)–(4.13) only hold for \(2/\sqrt{n} \leq \varepsilon \leq 1\).

Now, let the assumption of (4.2) hold. It follows from (1.2) and the definition of \(u\) that
\[(4.14) \quad u(\varepsilon) \leq \bar{C}(1/\varepsilon)^{1/2} \quad \text{for} \quad 0 < \varepsilon \leq \bar{\varepsilon}_0 = 2e_0, \quad \bar{C} = 2^{n}C.\]

If we put \(\varepsilon_n = (8\bar{C}/n)^{1/(2+\gamma)}\), then we find that there is an \(n_0 \in \mathbb{N}\) such that \(2/\sqrt{n} \leq \varepsilon_n \leq \min(1, 2\bar{e}_0)\) for \(n > n_0\). Thus, from (4.13) and (4.14) we obtain
\[R_F(n) \leq 7(8\bar{C})^{1/(2+\gamma)}(1/n)^{1/(2+\gamma)} \quad \text{for} \quad n > n_0.\]

Now, let the assumption of (4.3) hold. This assumption and (4.13) imply that \(F\) contains infinitely many distinct elements, so that \(\lim_{x \to 0} u(x) = \infty\).

Let \(\psi(\varepsilon) = 8\bar{e}^{-2}u(\varepsilon)\). Then \(\psi\) is a left-continuous, strictly decreasing function on some interval \((0, \xi)\) and \(\lim_{\varepsilon \to 0} \psi(\varepsilon) = \infty\). For \(m \in \mathbb{N}, m \geq \psi(\xi)\), let
\[\varepsilon_m = \sup\{\varepsilon : \psi(\varepsilon) \geq m\}.\]

Then \(\psi(\varepsilon_m + 0) \leq m \leq \psi(\varepsilon_m) = \psi(\varepsilon_m - 0).\) For convenience assume that \(\varepsilon_m\) is defined for all \(m \in \mathbb{N}\). Note that \(\varepsilon_{m+1} \leq \varepsilon_m\) and \(\lim_{m \to \infty} \varepsilon_m = 0\).

Now, define
\[\xi_1 = \varepsilon_1 = \ldots = \varepsilon_{k_1-1} > \varepsilon_{k_1}, \]
\[\xi_2 = \varepsilon_{k_1} = \ldots = \varepsilon_{k_2-1} > \varepsilon_{k_2}, \ldots \]
\[\xi_{n+1} = \varepsilon_{k_n} = \ldots = \varepsilon_{k_{n+1}-1} > \varepsilon_{k_{n+1}}.\]

Note that \(\lim_{m \to \infty} k_m = \infty, \lim_{m \to \infty} \xi_n = 0\) and \(k_{n+1} - 1 \leq \psi(\xi_{n+1}) = \psi(\xi_n) < k_n\) (if \(k_n \leq \psi(\xi_n)\), then \(\xi_n \leq \xi_n^{n+1}\), but the choice of \(k_n\) implies \(\xi_n > \xi_n^{n+1}\)).

For \(n\) large enough we have \(u(2/\sqrt{k_n}) \geq 1/2\) and \(0 < 2/\sqrt{k_n} < \xi\), hence
\[\psi\left(\frac{2}{\sqrt{k_n}}\right) = 8\bar{e}^{-2}u\left(\frac{2}{\sqrt{k_n}}\right) \geq k_n > \psi(\xi_n),\]

which implies \(\xi_n > 2/\sqrt{k_n}\). Note that \(u(\varepsilon)^{-1/2}k_n\varepsilon^2 < 0\) for \(\varepsilon \geq \xi_n\). Therefore (4.13) implies for \(n\) large enough
\[(4.15) \quad R_F(k_n) \leq 7\xi_n.\]

Our assumption implies that \(M\left(1/n\right)^{(1/(2+\gamma))} \leq R_F(n)\) for \(n > n_0\). Choose \(s \in \mathbb{N}\) such that for any \(n \geq s\) we have \(k_n > n_0, \psi(\xi_n) \geq 1\) and (4.15) holds.

Now \(k_n \leq 2\psi(\xi_n)\) for \(n \geq s\) and
\[M\left(\frac{1}{2\psi(\xi_n)}\right)^{1/(2+\gamma)} \leq M\left(\frac{1}{k_n}\right)^{1/(2+\gamma)} \leq R_F(k_n) \leq 7\xi_n.\]

The definition of \(\psi\) now implies
\[(4.16) \quad \tilde{\psi}^{-1}(x) = \begin{cases} \frac{x}{\varepsilon} & \text{if} \ \psi(\varepsilon) = x, \\ \frac{x}{\varepsilon} & \text{if} \ \psi(\varepsilon + 0) < x < \psi(\varepsilon). \end{cases}\]

Then \(\tilde{\psi}^{-1}\) is nonincreasing and \(\tilde{\psi}^{-1}(k_n) = \xi_{n+1}\). As \(\tilde{u}\) is nonincreasing, we have
\[2\psi(\varepsilon) = 16\bar{e}^{-2}u(\varepsilon) \leq 8\left(\frac{\varepsilon}{\sqrt{2}}\right)^{-2}u\left(\frac{\varepsilon}{\sqrt{2}}\right) = \psi\left(\frac{\varepsilon}{\sqrt{2}}\right),\]

and the inequality \(k_n - 1 \leq \psi(\xi_n) < k_n \leq \psi(\xi_{n+1})\) implies
\[\xi_{n+1} = \tilde{\psi}^{-1}(k_n) \geq \tilde{\psi}^{-1}(2(k_n - 1)) \geq \tilde{\psi}^{-1}(2\psi(\xi_n)) \geq \tilde{\psi}^{-1}(\psi(\xi_n/\sqrt{2})) = \xi_n/\sqrt{2}.\]

Now, let \(0 < \varepsilon \leq \xi_n\); choose \(n \geq s\) such that \(\xi_n + 1 < \varepsilon \leq \xi_n\); then the last inequality and (4.16) imply
\[(4.17) \quad u(\varepsilon) \geq \tilde{M}2^{-n/2}(1/\varepsilon)^{n}.\]

As \(u(\varepsilon) = H_{\varepsilon}(\mathcal{F})\) and \(H_{\varepsilon}(\mathcal{F})\) is nonincreasing (so that the set of its discontinuity points is countable) it follows from (4.17) that
\[H_{\varepsilon}(\mathcal{F}) \geq \tilde{M}2^{-n/2}(1/\varepsilon)^{n} \quad \text{for} \quad 0 < \varepsilon \leq \xi_n.\]

The last inequality and (1.2) give (4.3). ||

**Examples.**

(4.18) Let \(d, m \in \mathbb{N}, 0 < \alpha < m, \gamma > 0, C_1, C_2 > 0\) be given parameters, and
\[\mathcal{F} = \{f \in L^1(\mathbb{R}^d) : f \geq 0, \|f\|_1 = 1, \omega_{m,1}(f, \delta) \leq C_1\delta^{\alpha} \text{ for } \delta > 0, \omega_k(f, \lambda) \leq C_2\lambda^{-\gamma} \text{ for } \lambda > 0\}.\]

It follows from Lemma 3.2 that there exist \(C > 0\) and \(\varepsilon_0 > 0\) such that
\[H_{\varepsilon}(\mathcal{F}) \leq C(1/e)^{(1/\alpha + 1/\gamma)} \quad \text{for} \quad 0 < \varepsilon \leq \varepsilon_0.\]

(4.2) now implies the existence of \(M > 0\) and \(n_0 \in \mathbb{N}\) such that
\[R_F(n) \leq M(1/n)^{\alpha\gamma/(2\alpha\gamma + 4(\alpha + \gamma))} \quad \text{for} \quad n \geq n_0.\]
(4.19) Let $d, m \in \mathbb{N}, 0 < \alpha < m, C > 0$ be given parameters, and
\[ \mathcal{F} = \{ f \in L^1(\mathbb{I}^d) : f \geq 0, \| f \|_1 = 1, \omega_{m,1}(f, \delta) \leq C\delta^\alpha \text{ for } 0 < \delta \leq 1/m \} \]
where the $L^1$-norm and the modulus of smoothness are taken in $L^1(\mathbb{I}^d)$. $\mathcal{F}$ can also be considered as a family of densities on $\mathbb{R}^d$ whose supports are contained in $\mathbb{I}^d$ (while the modulus of smoothness is still taken in $L^1(\mathbb{I}^d)$). It will be proved that there exist $A_1, A_2, M_1, M_2 > 0, c_0 > 0$ and $n_0 \in \mathbb{N}$ such that
\begin{align*}
(4.20) & \quad A_1(1/\varepsilon)^d/\alpha \leq H_\varepsilon(\mathcal{F}) \leq A_2(1/\varepsilon)^d/\alpha \quad \text{for } 0 < \varepsilon \leq \varepsilon_0, \\
(4.21) & \quad M_1(1/n)^{\alpha/(2\alpha + d)} \leq R_\varepsilon(n) \leq M_2(1/n)^{\alpha/(2\alpha + d)} \quad \text{for } n \geq n_0.
\end{align*}
(No note that (4.20) resembles Lemma 2.12, which was proved for all $1 \leq p < \infty$, but for $0 < \alpha < 1$ only.)

The existence of $A_3$ follows from Corollary 2.10, and the existence of $M_2$ is a consequence of (4.2). Once the existence of $M_1$ has been proved, the existence of $A_1$ will follow from (4.3). So it remains to prove the left-hand side inequality in (4.21). In order to do this, we use Assouad’s Lemma (its proof can be found for example in [6]):

**Theorem 4.22 (Assouad’s Lemma).** For $r \in \mathbb{N}$ set $Z_r = \{-1, 1\}^r$; for $z \in Z_r$ and $1 \leq i \leq r$ define
\[ z_{i+} = (z_1, \ldots, z_{i-1}, 1, z_{i+1}, \ldots, z_r), \]
\[ z_{i-} = (z_1, \ldots, z_{i-1}, -1, z_{i+1}, \ldots, z_r). \]
Let $\{f_z : z \in Z_r\}$ be a family of $2^r$ distinct densities on $\mathbb{R}^d$. Assume that there exists a partition of $\mathbb{R}^d$ into measurable sets $A_1, \ldots, A_r$ such that for any $z \in Z_r$ and $1 \leq i \leq r$
\[ \int_{A_i} |f_{z_{i+}}(x) - f_{z_{i-}}(x)| dx \geq p > 0. \]
Moreover, let
\[ \int_{\mathbb{R}^d} \sqrt{f_{z_{i+}}(x)f_{z_{i-}}(x)} dx \geq q > 0. \]
Then for any density estimator $F_n$
\[ \sup_{x \in \mathbb{R}^d} E_{f_i} D(f, F_n) \geq \frac{p^r}{2} \max \left(1 - \sqrt{1 - 2\Delta n}, \frac{1}{2} \Delta^2 n \right). \]

Now we prove the left-hand side inequality in (4.21). For a given $m \in \mathbb{N}$ take $k \in \mathbb{N}, k \geq m + 2$; recall that $N^{(k)}(x) = k[0, x, \ldots, k_{-1}].$ Choose $h > 0$ and $r \in \mathbb{N}$ such that $2kh = 1$ and put
\[ g_{r, k}(t_1, \ldots, t_d) = N^{(k)}(t_1/h) \ldots N^{(k)}(t_d/h). \]
Then \(g_{r, k}(t_1, \ldots, t_d) = [0, kh]^d; \) for $i_1, \ldots, i_d \in \{0, 1, \ldots, r - 1\}$ define
\[ A_{i_1, \ldots, i_d} = [2kh(i_1, 2kh(i_1 + 1/2)] \times \ldots \times [2kh(i_d, 2kh(i_d + 1/2)) \]
\[ g_{i_1, \ldots, i_d} = g_{r, k}(t_1 - 2kh(i_1), \ldots, t_d - 2kh(i_d)) \]
\[ - g_{r, k}(t_1 - 2kh(i_1 + 1/2), \ldots, t_d - 2kh(i_d + 1/2)). \]
For each $i \in \{1, 2, \ldots, r^d\}$ choose a multi-index $\nu_i \in \{0, 1, \ldots, r - 1\}$ so that $u_i \neq u_{i'}$ for $i \neq i'.$

Let $a = \{\mathbb{R}, 0 < a < 1; \}$ for each $z \in Z_r, z = (z_1, \ldots, z_d)$, set
\[ f_z(t) = \begin{cases} 1 + a \sum_{i=1}^{r^d} z_i g_{u_i}(t) & \text{if } t \in \mathbb{I}^d, \\
0 & \text{if } t \in \mathbb{R}^d \setminus \mathbb{I}^d. \end{cases} \]
First note that $\int_{\mathbb{I}^d} f_z(x) dx = \int_{\mathbb{R}^d} f_z(x) dx = 1$ and $f_z \geq 0$ for any $z \in Z_r.$
The parameters $a$ and $r$ will be chosen in such a way that $f_z$ restricted to $\mathbb{I}^d$ is an element of $\mathcal{F}$.

The properties of the functions $N^{(k)}(x)$ mentioned in Section 3 imply that there exists a constant $C_{m,1} > 0$ such that for every $z \in Z_r, a, r$ and $h$
\[ \omega_{m,1}(f_z, \delta) \leq a \sum_{i=1}^{r^d} \omega_{m,1}(g_{u_i}, \delta) \leq C_{m,1} a \frac{1}{(2kh)^d} \min((\delta/h)^m, 1). \]
Therefore, for given $0 < \alpha < m$ and $k \geq m + 2$ it is possible to find a constant $C_{m,1, r, \alpha}$ independent of $a$ and $r$, such that $\omega_{m,1}(f_z, \delta) \leq C_{m,1, r, \alpha} a^{\alpha/\alpha} a^{\alpha/\alpha}$ for any $z \in Z_r.$
Note that
\[ \int_{\mathbb{R}^d} |f_{z_{i+}}(x) - f_{z_{i-}}(x)| dx = 4ah^d = 4a \frac{1}{(2kh)^d} = p, \]
\[ \int_{\mathbb{R}^d} \sqrt{f_{z_{i+}}(x)f_{z_{i-}}(x)} dx = \int_{\mathbb{R}^d} \sqrt{f_{z_{i+}}(x)f_{z_{i-}}(x)} dx \]
\[ \geq 1 - 2kh^d + 2kh^d(1 - a^2) \geq 1 - 2kh^d a^2 = q. \]
Take $r > (C_{m,1, r, \alpha})^{1/\alpha}, a = (C_{m,1, r, \alpha})^{-\alpha};$ then $f_z \in \mathcal{F}$ for $z \in Z_r.$ It follows from Theorem 4.22 that for any density estimator $F_n$
\[ \sup_{x \in Z_r} E_{f_i} D(f, F_n) \geq \sup_{x \in Z_r} E_{f_i} D(f, F_n) \geq \frac{1}{2} pr^d (1 - \sqrt{2(1 - q^2) \alpha}) \]
\[ \geq \frac{1}{2} pr^d (1 - \sqrt{2^{n}(1 - q^2)}) = s_1 r^{\alpha} (1 - \sqrt{2^{n} r^{-(2\alpha + d)}}) \]
where $s_1, s_2$ are some positive constants, independent of $n$ and $r.$ Take
\[ r_n = (4s_2 n^{1/(2\alpha + d)}) + 1. \]
Then
\[ \sqrt{2^{n} r_n^{-(2\alpha + d)}} \leq \frac{1}{2}. \]
also \( r_n > (C/C_{n,k}(\alpha))^{1/\alpha} \) and \( r_n \leq 2(4\pi n)^{1/(2\alpha+1)} \) for \( n \) large enough. This and (4.23) imply the left-hand side inequality in (4.21).

References


INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
ABRAHAM 13
81-826 POPOW, POLAND

Received February 26, 1998 (2005)