Probabilistic Complexity Analysis for Linear Problems in Bounded Domains

Stefan Heinrich

Karl-Weierstraß-Institute of Mathematics, Academy of Sciences, Mohrenstraße 39, PF 1304, Berlin 1086, German Democratic Republic

Received September 5, 1989

We study the probabilistic setting of information-based complexity for bounded domains and determine the order of the probabilistic cardinality number $m^{\text{prob}}(\epsilon, \delta, q)$ for an approximation problem with Wiener-type measure.

1. Introduction

Many numerical problems and methods for their approximate solution can be brought into the following abstract form: We are given a bounded linear operator $S \in L(X, Y)$ between Banach spaces $X$ and $Y$, the "solution operator." That is, for $x \in X$, $Sx$ is the true solution of the problem. We assume that we have only partial information on the "datum" $x \in X$, which is given by a mapping $N: X \rightarrow \mathbb{R}^n$, the "information operator." Finally, we have a mapping $\varphi: N(X) \rightarrow Y$ that represents the action of the "algorithm" by which we obtain an approximation $\varphi(N(x))$ to $Sx$ using the information $N(x)$. $N$ and $\varphi$ are required to belong to certain classes of not necessarily linear, not necessarily continuous mappings.

This approach is developed in the monograph "Information-Based Complexity," by Traub,Wasilkowski, and Woźniakowski, (1988), later referred to as IBC (see also Traub and Woźniakowski, 1980; Traub, Wasilkowski, and Woźniakowski, 1983). The aim of the theory is to study concrete $N$ and $\varphi$ as well as optimality over $\varphi$ or both $N$ and $\varphi$. The quality of $N$ and $\varphi$ is judged by the behavior of the error $\|Sx - \varphi(N(x))\|$.

In the worst case setting, one takes the supremum of the error over a
bounded set, which usually can be arranged to be a multiple \( qB_x \) of the unit ball \( B_x \) of \( X \), that is, one considers

\[
\sup_{x \in qB_x} \| Sx - \varphi(N(x)) \|
\]

(here and below we admit \(+ \infty\) as a possible value of suprema or integrals).

A major alternative way of estimating the quality of \( N \) and \( \varphi \) proceeds via probability. Given a Borel probability measure \( \mu \) on \( X \) and Borel measurable \( N \) and \( \varphi \), one determines in the average case setting

\[
\int_X \| Sx - \varphi(N(x)) \| d\mu(x),
\]

while in the probabilistic setting more detailed information about the distribution of the error is supplied by the quantity

\[
\inf\{ \varepsilon > 0 : \mu(\{ x \in X : \| Sx - \varphi(N(x)) \| > \varepsilon \} \leq \delta \},
\]

depending on \( \delta > 0 \). Usually the measures in consideration (Wiener measures on function spaces, general Gaussian measures) are defined over the whole space \( X \), which makes it hard to compare the average and probabilistic settings with the worst case setting. To overcome this difficulty, Woźniakowski (1987) considered the average case setting for bounded domains by studying the normalized restriction of the measure \( \mu \) to a ball of radius \( q > 0 \), i.e.,

\[
\mu_q(A) = \mu(A \cap qB_x)/\mu(qB_x)
\]

(for Borel sets \( A \)). Similarly one can proceed in the probabilistic setting, which was carried out in IBC, Section 8.5.5.

In this framework, one studies the probabilistic \((\varepsilon, \delta, q)\)-cardinality \( m^{\text{prob}}(\varepsilon, \delta, q) \), which is the smallest nonnegative integer \( n \) such that there are admissible \( N: X \to \mathbb{R}^n \) and \( \varphi: N(X) \to Y \) with

\[
\mu_q(\{ x \in X : \| Sx - \varphi(N(x)) \| > \varepsilon \} \leq \delta
\]

(precise definitions are given below). The quantity \( m^{\text{prob}}(\varepsilon, \delta, q) \) can be interpreted as the minimal number of information operations needed to solve the problem with the given precision requirements. In the model of computation used in IBC, the complexity of a linear problem is, in general, proportional to this number. Thus, \( m^{\text{prob}}(\varepsilon, \delta, q) \) describes the probabilistic complexity of solving \( S \) on a bounded domain.

Only a few estimates of \( m^{\text{prob}}(\varepsilon, \delta, q) \) are known (see IBC, Sect. 8.5.5).
So far, in no (nontrivial) case of a solution operator $S$ could the order of $m_{\text{prob}}(\varepsilon, \delta, q)$ as a function of all three variables be determined. It is the aim of this paper to solve this problem for the approximation of functions of the periodic Sobolev class in Hilbert space, the measure being of Wiener type and naturally related to the smoothness scale. We provide two-sided estimates of $m_{\text{prob}}(\varepsilon, \delta, q)$ with upper and lower bounds differing only by a constant factor independent of $\varepsilon$, $\delta$, and $q$.

It is intuitively clear that for small $\delta$, $m_{\text{prob}}(\varepsilon, \delta, q)$ must be close to the cardinality for the worst case setting, while for large $q$ the unbounded probabilistic setting will dominate the situation. So we can also provide an answer to the following principal, qualitative aspect of the problem above: How is the "passage from one setting to the other" accomplished and what happens in the intermediate zone?

Finally, our result settles a problem in the limit case $q = +\infty$, as well. For approximation problems the order of $m_{\text{prob}}(\varepsilon, \delta, +\infty)$ was known so far only for $\varepsilon \to 0$, but fixed $\delta$, i.e., the constants depending on $\delta$. Moreover, a further interesting aspect occurs here—the relation to the average case. We discuss this in Section 5.

Our approach uses tools from Banach space theory. The main ingredient is a recent result of Maurey and Pisier on the deviation of Gaussian measures from their mean. We adopt an operator-theoretic point of view on Gaussian measures (see, e.g., Linde and Pietsch, 1974), which enables us to relate distribution estimates to the approximability of the operator generating the measure. In the preliminary Section 2 we expose this approach in some detail. In Section 3 we obtain certain distribution estimates for the general Banach space situation. The main result concerning the approximation problem in Hilbert space is presented in Section 4. The final Section 5 is devoted to the discussion of certain further aspects of the main result. We also mention some open problems.

Our reference for terminology and facts of information-based complexity is IBC. For Banach space theory we refer to Lindenstrauss and Tzafriri (1977, 1979), and for operators in Banach spaces to Pietsch (1978).

2. Notation and Preliminaries on Gaussian Measures

We consider only Banach spaces over the field of reals. Given a Banach space $X$ we let $X^*$ be its dual space and $B_X$ the unit ball of $X$. Subspace means always closed linear subspace. Given another Banach space $Y$, $L(X, Y)$ denotes the space of all bounded linear operators $T$ from $X$ to $Y$, equipped with the operator norm $\|T\|$. The closure of a set $A \subseteq X$ is denoted by $\overline{A}$.

Let $\mathcal{B}(X)$ be the $\sigma$-algebra of Borel subsets of $X$, and let $\mathcal{B}(X)$ be the
algebra of cylindrical subsets of $X$ (see Kuo, 1975, I, Sect. 6, or Pietsch, 1978, 25.2). A Gaussian measure on $X$ is a Radon probability measure $\mu$ on $\mathcal{B}(X)$ such that each $x^* \in X^*$ is a symmetric Gaussian random variable on $(X, \mu)$ (possibly degenerate, i.e., $=0$ almost everywhere). Note that we consider only symmetric, that is mean zero Gaussian, measures. For a Hilbert space $H$ let $\gamma_H$ denote the standard cylindrical Gaussian probability on $H$ (see Kuo, 1975, I, Def. 4.2; Pietsch, 1978, 25.5.1).

Gaussian measures are closely related to certain classes of operators. To introduce them, define for $T \in L(H, X)$, $H$ a Hilbert space, $X$ a Banach space

$$E_\gamma(T) = \sup_{\dim F \leq n} \int_F \|Th\| d\gamma_T(h).$$

Let $\Pi_\gamma(H, X)$ denote the set of all $T \in L(H, X)$ with $E_\gamma(T) < \infty$. This class is studied in Linde and Pietsch (1974). It is readily checked that $\Pi_\gamma(H, X)$ endowed with $E_\gamma$ as a norm is a Banach space. Moreover, for $T \in \Pi_\gamma(H, X)$,

$$\|T\| \leq (\pi/2)^{1/2} E_\gamma(T), \quad (1)$$

and for a further Hilbert space $H_0$, a Banach space $X_0$, $S \in L(H_0, H)$, and $U \in L(X, X_0)$,

$$E_\gamma(UTS) \leq \|U\|E_\gamma(T)\|S\|. \quad (2)$$

(Relation (1) is easily verified, while (2) is Lemma 2 of Linde and Pietsch (1974), up to a modification of the norm, which does not affect the proof.) Let $R_\gamma(H, X)$ be the closure of the finite rank operators in $\Pi_\gamma(H, X)$. For $T \in L(H, X)$ let $T \gamma_H$ denote the cylindrical probability measure induced on $X$ by $T$, that is $T \gamma_H(B) = \gamma_H(T^{-1}(B))$ for $B \in \mathcal{B}(X)$. Now $T \in R_\gamma(H, X)$ if and only if $T \gamma_H$ has an extension $T \gamma_H$ to $\mathcal{B}(X)$ which is a Radon measure (such an extension is unique). So $T \in R_\gamma(H, X)$ implies that $T \gamma_H$ is Gaussian. Conversely, if $\mu$ is a Gaussian measure on $X$, there is a separable Hilbert space $H$ and an injection $J \in R_\gamma(H, X)$ with $\mu = J \gamma_H$. $H$ and $J$ are essentially unique (up to isometries). Note that $(J, H, X)$ is then an abstract Wiener space (see Kuo, 1975, I, Sect. 4). Let us also mention that if $\mu = T \gamma_H$, $T \in R_\gamma(H, X)$, then $C_\mu = TT^*$ is the covariance operator of $\mu$, $\text{Im} \; T$ is the support of $\mu$, and

$$E_\gamma(T) = \int_X \|x\| d\mu(x). \quad (3)$$
These facts can be found in Kuo (1975, I, Th. 4.1; III, Th. 1.1; I, Lemma 4.6), Linde and Pietsch (1974, Th. 4), and Vakhania, Tarieladse, and Chobanjan (1985); see also the more detailed guideline to Proposition 1.3 in Heinrich (1990). If \( X = G \) is a Hilbert space, then \( R_\gamma(H, G) \) coincides with the class of Hilbert–Schmidt operators \( S_2(H, G) \), and

\[
(1 + (\pi/2)^{3/2})^{-1} \sigma_2(T) \leq E_\gamma(T) \leq \sigma_2(T),
\]

where \( \sigma_2(T) \) denotes the Hilbert–Schmidt norm. This is a consequence of Pisier (1986, Cor. 2.5; inequality (2.7); compare also Heinrich (1990, Props. 1.2 and 1.4). For further characterizations of \( R_\gamma \) we refer to Chevet, Chobanjan, Linde, and Tarieladse (1977), Chobanjan and Tarieladse (1977), and Kühn (1981). We generally refer to Heinrich (1990, Sect. 1) for a similar, but more detailed, exposition of the facts quoted above, including their versions for the complex case.

Finally, we state two important results which we use. The first result is an estimate of the deviation of a Gaussian measure from its mean, which is due to Maurey and Pisier (see Pisier, 1986, Th. 2.1; Remark on p. 180).

**2.1. Proposition.** Let \( X \) and \( Y \) be Banach spaces, and let \( \mu \) be a Gaussian measure on \( X \), \( \mu = \mathcal{F}_{\gamma_H} \), where \( J \in R_{\gamma}(H, X) \) and \( H \) is a Hilbert space. Let \( T \in L(X, Y) \). Then for all \( t > 0 \) and \( \tau = \pm 1 \)

\[
\mu\{x \in X : \tau(\|Tx\| - E_\gamma(TJ)) > t\} \leq exp(-t^2/(2\|TJ\|^2)).
\]

The second result is the logarithmic concavity of Gaussian measures due to Borell (1974, Cor. 2.1).

**2.2. Proposition.** Let \( \mu \) be a Gaussian measure on a Banach space \( X \), let \( A, B \in \mathcal{B}(X) \), and let \( 0 \leq \alpha \leq 1 \). Then

\[
\mu(\alpha A + (1 - \alpha)B) \geq \mu(A)^\alpha \mu(B)^{1 - \alpha}.
\]

Now we describe the class of information operators and algorithms which we consider. Let \( n \) be a positive integer. A mapping \( N : X \to \mathbb{R}^n \) is called adaptive linear information of cardinality \( n \) if there exist mappings

\[
L_1 : X \to \mathbb{R} \\
L_2 : X \times \mathbb{R} \to \mathbb{R} \\
\vdots \\
L_n : X \times \mathbb{R}^{n-1} \to \mathbb{R}
\]
such that for $1 \leq i \leq n$ and all $(z_1, \ldots, z_{i-1}) \in \mathbb{R}^{i-1}$, $L_i(\cdot, z_1, \ldots, z_{i-1}) \in X^*$, and for $x \in X$

$$N(x) = (L_1(x), L_2(x, a_1), \ldots, L_n(x, a_1, \ldots, a_{n-1})),$$

where

$$a_1 = L_1(x)$$
$$a_2 = L_2(x, a_1)$$
$$\ldots$$
$$a_{n-1} = L_{n-1}(x, a_1, \ldots, a_{n-2})$$

(see IBC, 3.2.2.). The set of mappings $(L_1, \ldots, L_n)$ is called a representation of $N$. $\mathcal{R}^n(X)$ denotes the set of all adaptive linear informations of cardinality $n$. For the probabilistic analysis we must impose certain measurability conditions. Let $\mathcal{R}_0^0(X)$ be the subset of those $N \in \mathcal{R}^n(X)$ which possess a measurable representation $(L_1, \ldots, L_n)$. By this we mean that for $2 \leq i \leq n$,

$$(z_1, \ldots, z_{i-1}) \rightarrow L_i(\cdot, z_1, \ldots, z_{i-1})$$

is a Borel measurable mapping from $\mathbb{R}^{i-1}$ to $X^*$ (cf., e.g., Lee and Wasilkowski, 1986). Given $z \in \mathbb{R}^n$, $z = (z_1, \ldots, z_n)$, it is convenient to use the notation

$$L_{i,z} = L_i(\cdot, z_1, \ldots, z_{i-1})$$

for $1 \leq i \leq n$. Let us note that for $N \in \mathcal{R}_0^0(X)$, $N(X)$ is a Borel subset of $\mathbb{R}^n$.

Now let $\mu$ be a Gaussian measure on $X$. Given $N \in \mathcal{R}_0^0(X)$, a measurable representation $(L_1, \ldots, L_n)$ of $N$ is called $\mu$-orthonormal, if for all $z \in \mathbb{R}^n$

$$\langle C_\mu L_{i,z}, L_{j,z} \rangle = \delta_{ij}.$$

Let $\mathcal{R}_\mu^n(X)$ be the subset of all $N \in \mathcal{R}_0^0(X)$ possessing such a representation. As remarked in IBC, p. 221, for complexity analysis it is no loss of a generality to assume $N \in \mathcal{R}_\mu^0(X)$. In fact, for each $n \leq \dim(\text{supp } \mu)$ and $N \in \mathcal{R}_\mu^0(X)$ there is an $N_1 \in \mathcal{R}_\mu^0(X)$ and a measurable $\varphi_1: N_1(X) \rightarrow N(X)$ with

$$N = \varphi_1 \circ N_1.$$
Note that for each \( N \in \mathcal{R}_\mu^n(\mathbb{X}) \), \( N(\mathbb{X}) = \mathbb{R}^n \). It is easily checked that each \( N \in \mathcal{R}_\mu^n(\mathbb{X}) \) has a unique \( \mu \)-orthonormal representation, provided \( \text{supp} \mu = \mathbb{X} \).

Given a subset \( A \subseteq \mathbb{X} \), \( \Phi(A, \mathbb{Y}) \) denotes the set of all (not necessarily linear, not necessarily continuous) mappings from \( A \) to \( \mathbb{Y} \). If \( A \) is a Borel set, \( \Phi_0(A, \mathbb{Y}) \) stands for the set of all Borel measurable \( \varphi \in \Phi(A, \mathbb{Y}) \). Given \( N \in \mathcal{R}_\mu^n(\mathbb{X}) \), the elements of \( \Phi(N(\mathbb{X}), \mathbb{Y}) \) are called algorithms (using information \( N \)).

3. General Estimates

In this section we prove some distribution estimates for the general Banach space situation. Let \( \mathbb{X} \) and \( \mathbb{Y} \) be Banach spaces and let \( S \in L(\mathbb{X}, \mathbb{Y}) \). Let \( \mu \) be a (symmetric) Gaussian measure on \( \mathbb{X} \), let \( n \) be a positive integer, and let \( N \in \mathcal{R}_\mu^n(\mathbb{X}) \). We assume that \( \text{supp} \mu = \mathbb{X} \). For our purposes this is no loss of generality, since we can always restrict \( S \) and \( N \) to \( \text{supp} \mu \), this way neglecting a set of measure zero. We represent \( \mu = \mathcal{F}_\gamma H \), where \( H \) is a separable Hilbert space and \( J \subset R(H, \mathbb{X}) \) is an injection (see Section 2). It follows that \( J(H) \) is dense in \( \mathbb{X} \). Let \( (L_1, \ldots, L_n) \) be the \( \mu \)-orthonormal representation of \( N \). For \( z = (z_1, \ldots, z_n) \in \mathbb{R}^n \) define

\[
m_z = \sum_{i=1}^n z_i C_\mu L_{i,z},
\]

and observe that

\[
N(m_z) = z. \tag{5}
\]

Define

\[
U_z = \{ x \in \mathbb{X} : L_{i,z}(x) = 0, 1 \leq i \leq n \}
\]
\[
F_z = J^{-1}(U_z)
\]
\[
J_z = J|_{F_z}
\]

and

\[
\mu_z = \mathcal{F}_z \gamma F_z.
\]

Then the support of \( \mu_z \) is \( U_z \), and its covariance operator is given by

\[
C_{\mu_z} = J_z J_z^* = JQ_z J_z^*,
\]
where $Q_z$ is the orthogonal projection onto $F_z$. Hence

$$C_{\mu_z} = C_{\mu} - \sum_{i=1}^{n} C_{\mu} L_{i,z} \otimes C_{\mu} L_{i,z}.$$ 

By IBC (Appendix, Lemma 2.9.7), the family of measures $\mu_z(\cdot - m_z)$ is the conditional measure of $\mu$ with respect to $\mathcal{N}$. This means that for each $B \in \mathcal{B}(\mathcal{X})$, $\mu_z(B - m_z)$ is Borel measurable (with respect to $z$), and

$$\mu(B) = \int_{\mathbb{R}^n} \mu_z(B - m_z) d\gamma_n(z).$$

Here $\gamma_n$ is the canonical Gaussian measure on $\mathbb{R}^n$ (i.e., centered, with covariance identity). For our subsequent estimates the following form of this relation is convenient:

$$\mu(B) = \int_{\mathbb{R}^n} \mu_z\{u \in U_z: m_z + u \in B\} d\gamma_n(z). \quad (6)$$

Finally, we define the mean element algorithm $\varphi_0: \mathbb{R}^n \rightarrow Y$ by

$$\varphi_0(z) = Sm_z$$

(see IBC, 6.5.2). Let us finally fix the following constants:

$$\delta_1 = (1 - 2 \exp(-4/\pi))(1 - \exp(-1/(4\pi)) - \exp(-49/\pi))$$

$$c_1 = (1 - 2 \exp(-4/\pi))(2\pi)^{-1/2} \int_1^{+\infty} \exp(-t^2/2) dt.$$ 

Then $\delta_1 > 0.033$, and the simple estimate

$$\int_1^{+\infty} \exp(-t^2/2) dt \geq 1/2 \int_1^{2} t \exp(-t^2/2) dt$$

gives $c_1 \geq 0.041$. (If a constant is denoted by a symbol, this symbol remains reserved for this particular value throughout the paper.) Now we are ready to formulate the estimates for the general situation.

3.1. **PROPOSITION.** Assume the notation as above and suppose that $SJ_z \neq 0$ for some $z \in \mathbb{R}^n$. Then for all $q \geq 14E_r(J)$ and all $\varepsilon > 0$ the following hold:
PROBABILISTIC COMPLEXITY ANALYSIS

(i) If $\varepsilon \geq 2 \sup_z E_\gamma(SJ_z)$, then

$$\mu\{x \in qB_x: \|Sx - \varphi_0(N(x))\| > \varepsilon\} \leq \exp(-\varepsilon^2/(8 \sup \|SJ_z\|^2)).$$

(ii) If $\varepsilon \leq 1/2 \inf_z E_\gamma(SJ_z)$, then for each $\varphi \in \Phi_0(\mathbb{R}^n, Y)$

$$\mu\{x \in qB_x: \|Sx - \varphi(N(x))\| > \varepsilon\} \geq \delta_1.$$

(iii) If $\inf_z (\|SJ_z\|/\|J_z\|) \geq \sqrt{20} \varepsilon/q$, then for each $\varphi \in \Phi_0(\mathbb{R}^n, Y)$

$$\mu\{x \in qB_x: \|Sx - \varphi(N(x))\| > \varepsilon\} \geq c_1 \exp(-2\varepsilon^2/\inf \|SJ_z\|^2).$$

Remark. If $SJ_z = 0$ for all $z \in \mathbb{R}^n$, then (i) is just trivial. Indeed, it is easy to verify that the density of $J(H)$ in $X$ implies the density of $J(F_z)$ in $U_z$. Hence $SJ_z = 0$ gives $Sx = 0$ for all $x \in U_z$. Consequently $N(x) = N(y)$ implies $Sx = Sy$; i.e., we have "full information." Moreover, it follows from (5) that $Sx = \varphi_0(N(x))$ for all $x \in X$.

Proof of 3.1. To show the upper estimate in (i), we use (5) and (6), and get

$$\mu\{x \in qB_x: \|Sx - \varphi_0(N(x))\| > \varepsilon\} = \int_{\mathbb{R}^n} \mu_z\{u \in U_z: \|S(m_z + u) - \varphi_0(N(m_z + u))\| > \varepsilon\} d\gamma_n(z)$$

$$= \int_{\mathbb{R}^n} \mu_z\{u \in U_z: \|Su\| > \varepsilon\} d\gamma_n(z).$$

By 2.1, for each $z \in \mathbb{R}^n$ with $SJ_z \neq 0$,

$$\mu_z\{u \in U_z: \|Su\| > \varepsilon\} \leq \exp((-\varepsilon - E_\gamma(SJ_z))^2/(2\|SJ_z\|^2))$$

$$\leq \exp(-\varepsilon^2/(8 \sup \|SJ_z\|^2)).$$

If $SJ_z = 0$, then for all $u \in U_z$, $Su = 0$, so

$$\mu_z\{u \in U_z: \|Su\| > \varepsilon\} = 0.$$

This proves (i). To verify the lower estimates (ii) and (iii), let $\lambda = 6E_\gamma(J)$. 
Then $0 < \lambda < q$. Denote

$$M_\lambda = \{z \in \mathbb{R}^n : \|m_z\| \leq \lambda\}.$$  

By (5) and (6) we have for arbitrary $\varphi \in \Phi_0(\mathbb{R}^n, Y)$,

$$\mu\{x \in qB_X : \|Sx - \varphi(N(x))\| > \varepsilon\}$$

$$= \int_{\mathbb{R}^n} \mu_x\{u \in U_z : m_z + u \in qB_X$$

$$\text{and } \|S(m_z + u) - \varphi(N(m_z + u))\| > \varepsilon\} dy_n(z)$$

$$- \int_{\mathbb{R}^n} \mu_x\{u \in U_z : m_z + u \in qB_X \text{ and } \|Sm_z - \varphi(z) + Su\| > \varepsilon\} dy_n(z)$$

$$\geq \int_{M_h} \mu_x\{u \in U_z : \|u\| \leq q - \lambda \text{ and } \|Sm_z - \varphi(z) + Su\| > \varepsilon\} dy_n(z)$$

$$\geq \mu_x\{u \in U_z : \|Su\| > \varepsilon\}$$

(7)

Next we show that for any $z \in \mathbb{R}^n$

$$\mu_x\{u \in U_z : \|Sm_z - \varphi(z) + Su\| > \varepsilon\} \geq \mu_x\{u \in U_z : \|Su\| > \varepsilon\}. \quad (8)$$

Indeed, fix $z$ and denote $Sm_z - \varphi(z) = y$

$$A_+ = \{u \in U_z : \|y + Su\| \leq \varepsilon\}$$

$$A_- = \{u \in U_z : \|-y + Su\| \leq \varepsilon\}.$$  

Clearly, $A_- = -A_+$, hence $\mu_z(A_+) = \mu_z(A_-)$. Moreover it is easily checked that

$$\frac{1}{2}A_+ + \frac{1}{2}A_- \subseteq \{u \in U_z : \|Su\| \leq \varepsilon\}.$$  

Now we get from 2.2

$$\mu_z(\frac{1}{2}A_+ + \frac{1}{2}A_-) \geq \mu_z(A_+)^{1/2} \mu_z(A_-)^{1/2} = \mu_z(A_+).$$

Combining the last two relations and passing to complements, we obtain (8). This, in turn, together with (7) gives
Next we estimate \( \gamma_n(M_\lambda) \). For this observe first that \( \|m_z\| > \lambda \) implies

\[
\mu_z\{u \in U_z : \|m_z + u\| \leq \lambda/2\} \\
\leq \mu_z\{u \in U_z : \|u\| > \lambda/2\} \\
\leq \exp(-\left(\lambda/2 - E_\gamma(J_z)\right)^2/(2\|J_z\|^2)) \\
\leq \exp(-\left(\lambda/2 - E_\gamma(J)\right)^2/(2\|J\|^2)).
\]

Here we used 2.1. Together with (6) we get

\[
\mu(\lambda/2B_X) = \mu(\lambda/2B_X) - \exp(-\left(\lambda/2 - E_\gamma(J_z)\right)^2/(2\|J_z\|^2)) \\
\geq 1 - 2 \exp(-\left(\lambda/2 - E_\gamma(J)\right)^2/(2\|J\|^2)).
\]

By 2.1, the choice of \( \lambda = 6E_\gamma(J) \), and (1),

\[
\gamma_n(M_\lambda) \geq \mu(\lambda/2B_X) - \exp(-\left(\lambda/2 - E_\gamma(J)\right)^2/(2\|J\|^2)) \\
\geq 1 - 2 \exp(-\left(\lambda/2 - E_\gamma(J)\right)^2/(2\|J\|^2)) \\
= 1 - 2 \exp(-2E_\gamma(J)^2/\|J\|^2) \geq 1 - 2 \exp(-4/\pi). \tag{10}
\]

From 2.1 we get for all \( z \in \mathbb{R}^n \),

\[
\mu_z\{u \in U_z : \|u\| > q - \lambda\} \leq \exp(-\left(q - \lambda - E_\gamma(J_z)\right)^2/(2\|J_z\|^2)).
\]

Since \( q \geq 14E_\gamma(J) \), we have

\[
q - \lambda - E_\gamma(J_z) \geq q - 7E_\gamma(J) \geq q/2.
\]

Thus, for all \( z \in \mathbb{R}^n \)

\[
\mu_z\{u \in U_z : \|u\| > q - \lambda\} \leq \exp(-q^2/(8\|J_z\|^2)). \tag{11}
\]

We also need a slightly different form of this estimate, which we derive from (11), using (1) again.
\( \mu_\varepsilon \{ u \in U_\varepsilon : ||u|| > q - \lambda \} \leq \exp(-14E_\gamma(J)^2/(8||J_\varepsilon||^2)) \leq \exp(-49E_\gamma(J)^2/(2||J||^2)) \leq \exp(-49/\pi) \). 

(12)

From now on we must distinguish between the two cases (ii) and (iii). We start with (ii). Using 2.1 (with \( \tau = -1 \)) and the assumption \( \varepsilon \leq 1/2 \inf_\varepsilon E_\gamma(SJ_\varepsilon) \), we get

\[
\mu_\varepsilon \{ u \in U_\varepsilon : ||Su|| \leq \varepsilon \} \leq \exp(-E_\gamma(SJ_\varepsilon) - \varepsilon)^2/(2||SJ_\varepsilon||^2) \leq \exp(-1/(4\pi)).
\]

Combining this with (9), (10), and (12), we get

\[
\mu_\varepsilon \{ x \in qB_X : ||Sx - \varphi(N(x))|| > \varepsilon \}
\geq (1 - 2 \exp(-4/\pi))(1 - \exp(-1/(4\pi)) - \exp(-49/\pi)) = \delta_\varepsilon.
\]

To prove (iii), let \( z \in \mathbb{R}^n \). Observe first that \( J \) as an element of \( R_\varepsilon(H, X) \) is the norm limit of finite rank operators, hence \( J \) is compact. Then so are \( SJ_\varepsilon = SJ_\varepsilon|_{F_z} \) and its dual \( (SJ_\varepsilon)^* \in L(Y^*, F_z) \) (we identify \( F_z^* \) with \( F_z \)). This together with the weak*-compactness of the ball in \( Y^* \) implies that \( (SJ_\varepsilon)^* \) attains its norm; i.e., there is a \( y^* \in Y^* \) with \( ||y^*|| = 1 \) and

\[
||(SJ_\varepsilon)^*y^*|| = ||(SJ_\varepsilon)^*|| = ||SJ_\varepsilon||.
\]

Then

\[
\mu_\varepsilon \{ u \in U_\varepsilon : ||Su|| > \varepsilon \} \geq \mu_\varepsilon \{ u \in U_\varepsilon : |\langle Su, y^* \rangle| > \varepsilon \}
= \gamma_{F_z} \{ f \in F_z : |\langle SJ_\varepsilon f, y^* \rangle| > \varepsilon \}
= \gamma_{F_z} \{ f \in F_z : |(f, (SJ_\varepsilon)^*y^*)/(||(SJ_\varepsilon)^*y^*||)| > \varepsilon/||SJ_\varepsilon|| \}
= \psi(\varepsilon/||SJ_\varepsilon||),
\]

(13)

where

\[
\psi(a) = (2/\pi)^{1/2} \int_{a}^{+\infty} \exp(-t^2/2)dt.
\]

(Note that the assumption of (iii) implies \( SJ_\varepsilon \neq 0 \).) We make use of the following simple relations

\[
\psi(a) > (2/\pi)^{1/2} \int_{a}^{a+1} \exp(-t^2/2)dt > (2/\pi)^{1/2} \exp(-(a + 1)^2/2)
\]

(14)
and

$$\psi(1) < (2/\pi)^{1/2} \int_1^{+\infty} t \exp(-t^2/2) dt = (2/\pi)^{1/2} e^{-1/2}. \quad (15)$$

If $\varepsilon \geq ||SJ_z||,$ we get from (14)

$$\psi(\varepsilon/||SJ_z||) \geq (2/\pi)^{1/2} \exp(-\varepsilon^2/||SJ_z||^2).$$

If $\varepsilon < ||SJ_z||,$ then

$$\psi(\varepsilon/||SJ_z||) < \psi(1) = (2/\pi)^{1/2} \exp(-\varepsilon^2/||SJ_z||^2).$$

Since by (15), $\psi(1) < (2/\pi)^{1/2}$, we get in both cases from (13)

$$\mu_{\varepsilon}\{u \in U_{J^z}: ||Su|| > \varepsilon\} \geq \psi(1) \exp(-\varepsilon^2/||SJ_z||^2). \quad (16)$$

By (9), (10), (11), and (16),

$$\mu_{\varepsilon}\{x \in qB_x: ||Sx - \varphi(N(x))|| > \varepsilon\} \geq (1 - 2 \exp(-4/\pi)) \times \inf_{z} (\psi(1) \exp(-\varepsilon^2/||SJ_z||^2) - \exp(-q^2/(8||J_z||^2))). \quad (17)$$

We show that for all $z \in \mathbb{R}^n$

$$\psi(1) \exp(-\varepsilon^2/||SJ_z||^2) \geq 2 \exp(-q^2/(8||J_z||^2)). \quad (18)$$

Clearly (17) and (18) imply (iii). Inequality (18) is equivalent to

$$2\varepsilon^2/||SJ_z||^2 \leq q^2/(8||J_z||^2) - \log(2/\psi(1)) \quad (19)$$

(log always means the natural logarithm). To verify (19) note that by (1),

$$q^2/(8||J_z||^2) \geq 49E_{\gamma}(J)^2/(2||J||^2) \geq 49/\pi > 15.$$

On the other hand, by (14), $\psi(1) > (2/\pi)^{1/2} e^{-2}$, so

$$\log(2/\psi(1)) < \log((2\pi)^{1/2} e^{-2}) < 3.$$

Hence

$$\log(2/\psi(1)) < q^2/(40||J_z||^2),$$

and consequently
The assumption of (iii) obviously implies
\[ 2\alpha^2/\|\mathcal{J}z\|^2 \leq q^2/(10\|J_z\|^2) - \log(2/\psi(1)). \] (20)

Relations (20) and (21) give (19), which concludes the proof.

4. An Approximation Problem

We study the approximation of functions in periodic Sobolev spaces. To define these spaces, let \( \Gamma = \{e^{it}: 0 \leq t \leq 2\pi\} \) be the unit circle and \( \lambda \) the Lebesgue measure on \( \Gamma \). Let \( L_2 = L_2(\Gamma, \lambda) \) and let \( (e_n)_{n=-\infty}^{\infty} \) be the normalized in \( L_2 \) trigonometric basis, i.e.,
\[
e_0(t) = (2\pi)^{-1/2}, \quad e_n(t) = \pi^{-1/2} \cos nt, \quad e_{-n}(t) = \pi^{-1/2} \sin nt
\]
\((n \in \mathbb{N})\). For any real \( r \geq 0 \) the periodic Sobolev space \( H' = H'(\Gamma, \lambda) \) is defined as
\[
H' = \left\{ f \in L_2: \|f\|_r^2 = \sum_{j \in \mathbb{Z}} (1 + j^2)^r(f, e_j)^2 < \infty \right\},
\]
where \(( , , )\) denotes the scalar product of \( L_2 \). For \( r \geq s \) let \( I_r: H' \to H^s \) be the identical embedding. We write \( I_r \) for \( I_{r_0}: H' \to L_2 \). Put
\[
f_j = (1 + j^2)^{-r/2}e_j, \quad g_j = (1 + j^2)^{-s/2}e_j \quad (j \in \mathbb{Z}).
\]
Then \((f_j)\) and \((g_j)\) are orthonormal bases in \( H' \) and \( H^s \), respectively, and \( I_{rs} \) is of diagonal type
\[
I_{rs}f_j = (1 + j^2)^{-(r-s)/2}g_j. \quad (22)
\]
Therefore it is quite simple to verify the following (well-known) approximation properties. Denote
\[
E_k = \text{span}\{e_j: |j| < k\}
\]
and let \( P_k \) be the orthogonal projection onto \( E_k \). Then for all \( k \geq 1 \)
\[
\|h - P_k h\|_r \leq k^{-(r-s)}\|h\|_r \quad (h \in H') \quad (23)
\]
\[
\|h\|_s \geq k^{-s}\|h\|_s \quad (h \in E_k). \quad (24)
\]
Moreover, for all \( n \geq 1 \)

\[
c_n(I_n) \geq n^{-(r-s)},
\]

where for an operator \( T \in L(X, Y) \), \( c_n(T) \) denotes the \( n \)th Gelfand number

\[
c_n(T) = \inf_{F \subseteq Y} \| T|_F \|
\]

(see, e.g., Pietsch (1978, Chap. 11) for this notion and for the approximation of diagonal operators). In order to use (iii) of Proposition 3.1 to obtain lower estimates we must be able to handle the quotient \( \| S J_s \| / \| J_s \| \). This is achieved on the basis of the following

4.1. PROPOSITION. Let \( r > s \geq 0 \), let \( J = I_{I_s}; H^r \to H^s \), \( S = I_s; H^s \to L_2 \), and let \( n \) be a positive integer. Then for each subspace \( F \subset H^r \) with codim \( F < n \),

\[
\| S J_s \| \geq c_2 n^{-s} \| J_s \|,
\]

where \( c_2 = (2^{1/(r-s)} + 1)^{-s}/2 \).

Proof. Choose \( h \in F \) with \( \| h \|_r = 1 \) and

\[
\| h \|_s = \| J_f h \|_s = \| J_s f \| \geq n^{-(r-s)},
\]

where the last relation follows from (25). For arbitrary \( k \geq 1 \) we have by (23) and (24)

\[
\| h \|_s \geq \| P_k h \|_s \geq k^{-s} \| P_k h \|_s \geq k^{-s} (\| h \|_s - \| h - P_k h \|_s)
\]

\[
\geq k^{-s} (\| h \|_s - k^{-(r-s)}) = k^{-s} \| h \|_s - k^{-r}.
\]

Consequently,

\[
\| S J_s \| / \| J_s \| \geq \| h \|_s / \| h \|_s \geq k^{-s} - k^{-r} \| h \|_s
\]

\[
\geq k^{-s} - k^{-r} n^{-s} = k^{-s} (1 - (n/k)^{-s}).
\]

Now choose any \( k \) satisfying

\[
2^{1/(r-s)} n \leq k \leq (2^{1/(r-s)} + 1)n.
\]

Then

\[
(n/k)^{-s} \leq 1/2,
\]
and we get
\[ \|SJ_F\|/\|J_F\| \geq k^{-s}/2 \geq 1/2(2^{1/(r-s)} + 1)^{-s}n^{-s} = c_2 n^{-s}. \]

Before we state the next result we need two more approximation properties, this time with respect to the norm \( E_\gamma \). Since in Hilbert spaces \( R_\gamma \) coincides with the Hilbert–Schmidt class, it follows that \( I_r \in R_\gamma(H', L_2) \) iff \( r > 1/2 \). If this is satisfied, we get on the basis of (4) and (22) for \( k \geq 1 \),

\[
\begin{align*}
E_\gamma(I_r - P_k I_r)^2 &\leq \sigma_2(I_r - P_k I_r)^2 \\
&= 2 \sum_{j=k}^{+\infty} (1 + j^2)^{-r} \leq 2(k^2 + 1)^{-r} + 2 \int_k^{+\infty} x^{-2r} dx \\
&\leq 2(k^2 + 1)^{-r} + 2/(2r - 1)k^{-2r+1}.
\end{align*}
\]

Consequently,

\[ E_\gamma(I_r - P_k I_r) \leq c_3 k^{-r+1/2} \] (26)

with \( c_3 = (4r/(2r - 1))^{1/2} \). In a similar way, for \( r - s > 1/2 \),

\[ E_\gamma(I_{rs}) \leq c_4 = (1 + 2^{1-(r-s)} + 2/(2(r - s) - 1))^{1/2}. \] (27)

We also need a lower bound with respect to the norm \( E_\gamma \). For this recall that for an operator \( T \in S_2(G, H) \) between Hilbert spaces \( G \) and \( H \) and for positive integer \( n \) we have

\[ \sigma_2(T)^2 = \sum_{k=1}^{\infty} c_k(T)^2 \geq n c_n(T)^2 \] (28)

(see Pietsch, 1978, 11.3). Now let \( F \subset H' \) be any subspace of codimension \( <n \). Then (28) and (25) give

\[
\begin{align*}
\sigma_2(I_r|_F) &\geq n^{1/2} c_n(I_r|_F) \geq n^{1/2} c_{2n-1}(I_r) \\
&\geq n^{1/2}(2n - 1)^{-r} \geq 2^{-r} n^{-r+1/2}.
\end{align*}
\]

Combining this with (4) we get

\[
\inf_{F \subset H'} \left\{ \begin{array}{l}
\text{codim } F < n \\
E_\gamma(I_r|_F) \geq c_5 n^{-r+1/2}
\end{array} \right. \] (29)
with \( c_5 = (1 + (\pi/2)^{3/2})^{-1}2^{-r} \). (Clearly, the quantity on the left-hand side of (29) is the analogue of the Gelfand number for the norm \( E_r \).)

Now we come to the main result of this paper. We let \( X = H^s \), \( Y = L_2 \), and \( S = I_\delta \); \( H^s \to L_2 \). We assume \( r - 1/2 > s > 0 \). Then \( I_\delta \in R_\gamma(H', H^s) \) and we define the measure \( \mu \) on \( H^s \) by

\[
\mu = \tilde{I}_\delta r_{H'}.
\]

Thus, in the notation of Section 2, \( H = H', J = I_\delta \). The measure \( \mu \) is of Wiener type in the sense that it is generated by embedding spaces of smooth functions, i.e., that it represents a certain smoothness. It is closely related to the measure studied in IBC, 7.3.1 (put \( d = 1, v = 2(r - s), r_1 = s \) there—for \( s \) an integer). For \( \varepsilon > 0, \delta > 0, q > 0 \) we denote by \( m_{\text{prob}}(\varepsilon, \delta, q) \) the smallest positive integer \( n \) such that there are \( N \in \mathcal{N}_0(X) \) and \( \varphi \in \Phi_0(N(X), Y) \) with

\[
\mu_q \{ x \in X : \| Sx - \varphi(N(x)) \| > \varepsilon \} \leq \delta,
\]

where \( \mu_q(A) = \mu(A \setminus qB_X)/\mu(qB_X) \) for \( A \in \mathcal{B}(X) \). So far we consider only finite \( q \); the limit case \( q = +\infty \) is discussed in Corollary 4.4. By the remarks in Section 2 concerning \( \mu \)-orthonormal representations the definition of \( m_{\text{prob}}(\varepsilon, \delta, q) \) does not change if we replace \( \mathcal{N}_0(X) \) by \( \mathcal{N}_{\text{prob}}(X) \).

Finally, let us introduce the following constants. Let \( \delta_2 \) be any positive real satisfying

\[
\delta_2 \leq \min(\delta_1, 1/\varepsilon), \quad \delta_2 < c_1,
\]

let

\[
c_6 = (1 - \exp(-169/\pi))^{-1} = 1 + 4.3 \times 10^{-24}
\]

\[
d_1 = \min((c_2/\sqrt{20})^{1/5}, (1 - \log c_1/\log \delta_2)/2) \big/ (2^{1/2}), \quad (c_5/2) \big/ (r - 1/2),
\]

\[
d_2 = 2 \max((2c_3)^{1/2}, (8c_6)^{1/2})
\]

and

\[
q_1 = 14c_4.
\]

4.2. Theorem. Let \( \varepsilon > 0, 0 < \delta < \delta_2, \) and \( q \geq q_1 \). Then

\[
d_1 \eta(\varepsilon, \delta, q) - 1 \leq m_{\text{prob}}(\varepsilon, \delta, q) \leq d_2 \eta(\varepsilon, \delta, q) + 1,
\]
where

\[ \eta(\varepsilon, \delta, q) = \min((q/\varepsilon)^{1/5}, \max(((\log(1/\delta))^{2/5}/\varepsilon)^{1/r}, (1/\varepsilon)^{(1/r-1/2)})). \]

**Proof.** First observe that by assumption and (27) \( q \geq 14c_4 \geq 14E_\varepsilon(J) \).

With 2.1 and (1) this gives

\[ \mu\{x \in X: \|x\| > q\} \leq \exp(-(q - E_\varepsilon(J))^2/(2\|J\|^2)) \leq \exp(-169/\pi). \]

Hence for \( A \in \mathcal{B}(X) \),

\[ \mu(A \cap qB_X) \leq \mu_q(A) \leq c_6 \mu(A \cap qB_X). \quad (30) \]

First we prove the upper bound stated in the theorem. Let

\[ k = \lceil (q/\varepsilon)^{1/5} \rceil. \]

Then \( k^5 \geq q/\varepsilon \), hence \( qk^{-5} \leq \varepsilon \). By (23) (replacing \( r \) by \( s \) and \( s \) by \( 0 \)) we get

\[ \|h - P_k h\|_0 \leq \varepsilon \quad (h \in H_s, \|h\|_s \leq q). \]

Therefore the information \( N_k: H_s \rightarrow \mathbb{R}^{2k-1}, \)

\[ N_k h = ((1 + j^2)^{s/2}(h, e_j))_{j=-k}^{k-1} \]

(scalar product in \( L_2 \)) and the algorithm \( \varphi_k \in L(\mathbb{R}^{2k-1}, L_2) \)

\[ \varphi_k((\alpha_j)_{j=-k}^{k-1}) = \sum_{j=-k}^{k-1} \alpha_j(1 + j^2)^{-n^2}e_j \]

are as desired and give

\[ m_{\text{prob}}(\varepsilon, \delta, q) \leq 2k - 1 \leq 2(q/\varepsilon)^{1/5} + 1. \quad (31) \]

Next we put

\[ k = \lceil \max((2c_3/\varepsilon)^{1/(r-1/2)}, (8c_6 \log(1/\delta)/\varepsilon^2)^{(1/(2r))}) \rceil \quad (32) \]

and let \( N_k \) and \( \varphi_k \) be as above. Note that \( N_k \in \mathcal{R}_n^*(H^s) \) and that \( \varphi_k \) is the mean element algorithm corresponding to \( N_k \) (and to the choice of \( S \) and \( \mu \) in this section). Denote

\[ J_k = J|_{\text{Ker}(N_k)} \]
It follows from (32) and (26) that

$$\varepsilon \geq 2c_3 k^{-r+1/2} \geq 2E_\gamma (I_r - P_k I_r) \equiv 2E_\gamma (SJ_k).$$

(33)

Since $\delta < \delta_2 \leq 1/\varepsilon$, $\log(1/\delta) \geq 1$, and hence

$$c_6 \log(1/\delta) \geq (1 + \log c_6) \log(1/\delta) \geq \log(c_6/\delta).$$

We get from this and from (32),

$$k^{2r} \geq 8 \log(c_6/\delta)/\varepsilon^2,$$

and therefore

$$\exp(-\varepsilon^2 k^{2r}/8) \leq \delta/c_6.$$

Inequality (23) gives

$$\|SJ_k\| \leq \|I_r - P_k I_r\| \leq k^{-r}.$$

Combining the last two relations, we get

$$\exp(-\varepsilon^2/(8\|SJ_k\|^2)) \leq \delta/c_6.$$

Because of (33) we can apply 3.1(i) and obtain, also using (30),

$$\mu_q\{x \in X: \|Sx - \varphi(N_k x)\| > \varepsilon\} \leq c_6 \exp(-\varepsilon^2/(8\|SJ_k\|^2)) \leq \delta.$$

This gives

$$m_{\text{prob}}(\varepsilon, \delta, q) \leq 2k - 1 \leq 2 \max((2c_3/\varepsilon)^{1/(r-1/2)}, (8c_6 \log(1/\delta)/\varepsilon^2)^{1/(2r)}) + 1.$$

Since $c_6 > 1$, it follows that $d_2 > 2$, and the last estimate together with (31) proves the upper bounds.

To verify the lower bounds, let $n \geq 1$, $N \in \mathcal{R}_\mu^n(X)$, and $\varphi \in \Phi_0(\mathbb{R}^n, Y)$ be such that

$$\mu_q\{x \in X: \|Sx - \varphi(N(x))\| > \varepsilon\} \leq \delta.$$

Then by (3) we also have

$$\mu\{x \in qB_x: \|Sx - \varphi(N(x))\| > \varepsilon\} \leq \delta.$$
We distinguish between two cases. First we assume that there is a \( z \in \mathbb{R}^n \) with

\[ \|SJ_z\| < (\sqrt{20} \varepsilon / q) \|J_z\|. \]

Then we can use Proposition 4.1 and get

\[ c_2(n + 1) \varepsilon \leq \sqrt{20} \varepsilon / q. \]

thus

\[ n \geq (c_2 / \sqrt{20})^{1/2} (q / \varepsilon)^{1/8} - 1. \tag{34} \]

This proves the lower bound in the first case. Now we treat the second case and assume that for all \( z \in \mathbb{R}^n \),

\[ \|SJ_z\| \geq (\sqrt{20} \varepsilon / q) \|J_z\|. \]

Then Proposition 3.1(iii) is applicable and gives

\[ \delta \geq c_1 \exp(-2 \varepsilon^2 / \inf_z \|SJ_z\|^2). \]

In view of (25),

\[ \|SJ_z\| \geq c_{n+1}(SJ) \geq (n + 1)^{-r}, \]

which yields

\[ \delta \geq c_1 \exp(-2 \varepsilon^2 (n + 1)^{2r}), \]

and therefore

\[ (n + 1)^{2r} \geq (-\log \delta + \log c_1) / 2 \varepsilon^2 \]

\[ \geq (1 - \log c_1 / \log \delta_2) \log(1/\delta)/(2 \varepsilon^2). \]

Hence

\[ n \geq ((1 - \log c_1 / \log \delta_2) / 2)^{1/(2r)}((\log(1/\delta))^{1/2} / \varepsilon)^{1/r} - 1. \tag{35} \]

Since \( \delta < \delta_1 \), we conclude from Proposition 3.1(ii) that there is a \( z \in \mathbb{R}^n \) with

\[ \varepsilon > E_\nu(SJ_z)/2. \]
Moreover, (29) yields

\[ E_\gamma(SJ_z) = E_\gamma(SJ|F_\gamma) \geq c_5(n + 1)^{-\alpha/2}. \]

It follows that

\[ n \geq (c_5/2)^{1/(\alpha/2)} (1/\varepsilon)^{1/(\alpha/2)} - 1. \]

This together with (35) implies the lower bound also in the second case and thus concludes the proof of the theorem.

**Remark.** The proof shows that the information \( N_k \) given essentially by the Fourier coefficients and the mean element algorithm \( \varphi_k \) is quasi-optimal in the following sense: There is a constant \( d > 0 \) such that for all \( \varepsilon, \delta, q \) satisfying the assumptions of 4.2 there is a positive integer \( k \) with

\[ \mu_q \{ x \in X : \| Sx - \varphi_k(N_k x) \| > \varepsilon \} \leq \delta \]

and \( 2k - 1 \leq d\text{prob}(\varepsilon, \delta, q) \). Note that \( N_k \) is nonadaptive (i.e., the \( L_{i,z} \) actually do not depend on \( z \)) and that \( \varphi_k \) is linear.

Let us briefly recall the model of computation used in IBC, 3.2.3 and 3.3. First, one can carry out information operations to get the components of \( N(x) \), that means we can obtain the value of arbitrary continuous linear functionals over \( X \) at \( x \). Assume that each information operation has the same fixed cost \( c \). Second, to compute \( \varphi(N(x)) \), one can perform combinatorial operations, which are arithmetic operations, comparison of real numbers, the evaluation of a certain finite set of elementary functions (which has to be fixed), vector addition in \( Y \), and multiplication of an element of \( Y \) by a scalar. We assume that the combinatorial operations are carried out with infinite precision and that all of them are of unit cost. Then the probabilistic \((\varepsilon, \delta, q)\)-complexity \( \text{comp}_{\text{prob}}(\varepsilon, \delta, q) \) is the minimal cost of an \( N \in \mathcal{M}_0(X) \) and a \( \varphi \in \Phi_0(N(X), Y) \) with

\[ \mu_q \{ x \in X : \| Sx - \varphi(N(x)) \| > \varepsilon \} \leq \delta \]

(see IBC for details). The following is a direct consequence of 4.2 and the above remark.

4.3. **Corollary.** With the notation as above, let \( \varepsilon > 0, 0 < \delta < \delta_2, q \geq q_1 \). Then

\[ c(d_1 \eta(\varepsilon, \delta, q) - 1) \leq \text{comp}_{\text{prob}}(\varepsilon, \delta, q) \approx (c + 2)(d_2 \eta(\varepsilon, \delta, q) + 1) - 1. \]

This gives a partial answer to a question posed in IBC (Sect. 8.5.5, p. 346).
In 4.2, \( q \) was assumed to be finite. The case \( q = +\infty \) can easily be derived from it by a simple limit argument. Denote

\[
m_{\text{prob}}(\varepsilon, \delta) = m_{\text{prob}}(\varepsilon, \delta, +\infty)
\]

(we put \( \mu_{+\infty} = \mu \) in the original definition). It is readily seen that for each \( \theta > 0 \) there is a \( q_0 > 0 \) such that for all \( q > q_0 \) and all \( A \in \mathfrak{B}(X) \)

\[
\mu_q(A) - \theta \leq \mu(A) \leq \mu_q(A) + \theta.
\]

Consequently, for all \( \varepsilon > 0, \delta > \theta, q > q_0 \)

\[
m_{\text{prob}}(\varepsilon, \delta + \theta, q) \leq m_{\text{prob}}(\varepsilon, \delta) \leq m_{\text{prob}}(\varepsilon, \delta - \theta, q).
\]

Combining this with 4.2 gives

4.4. Corollary. Let \( \varepsilon > 0 \) and \( 0 < \delta < \delta_2 \). Then

\[
d_1 \zeta(\varepsilon, \delta) - 1 \leq m_{\text{prob}}(\varepsilon, \delta) \leq d_2 \zeta(\varepsilon, \delta) + 1,
\]

where

\[
\zeta(\varepsilon, \delta) = \max(\left((\log(1/\delta))^{1/2}/\varepsilon\right)^{1/r}, (1/\varepsilon)^{1/(r-1/2)}).
\]

Note that the lower and upper estimates given in IBC (Sect. 8.7) differ from each other by some power of \((\log(1/\delta))^{1/2}/\varepsilon\). So also in the unbounded probabilistic setting this is the first sharp order estimate for an approximation problem.

5. Remarks and Open Problems

First we discuss the passages from one setting to another. Let again \( X = H^s, Y = L_2, S = L_s; H^s \rightarrow L_2 \). Denote by \( m_{\text{wor}}(\varepsilon, q) \) the worst case cardinality, that is the smallest integer \( n \geq 1 \) such that there are \( N \in \mathfrak{N}(X) \) and \( \varphi \in \Phi(N(X), Y) \) with

\[
\sup_{x \in \mathfrak{B}_X} \|x - \varphi(N(x))\| \leq \varepsilon.
\]

It follows from IBC (Chap. 4, Th. 5.3.2; cf. also Sect. 5.3.1) that there exist constants \( d_3, d_4 > 0 \) such that for all \( \varepsilon > 0, q > 0 \)

\[
d_3(q/\varepsilon)^{1/s} \leq m_{\text{wor}}(\varepsilon, q) \leq d_4(q/\varepsilon)^{1/s} + 1.
\]
In the average case setting we let $\mu$ be as in Section 4 and define $m_{\text{avg}}(\epsilon)$ to be the smallest $n \geq 1$ such that there is an $N \in \mathbb{N}(X)$ and a $\varphi \in \Phi_0(N(X), Y)$ with

$$\int_X \|Sx - \varphi(N(x))\|d\mu(x) \leq \epsilon.$$ 

By IBC (Chap. 6, Th. 5.5.1; see also Sect. 7.3.1) there are constants $d_5, d_6 > 0$ such that for all $\epsilon > 0$

$$d_5(1/\epsilon)^{1/(r-1/2)} \leq m_{\text{avg}}(\epsilon) \leq d_6(1/\epsilon)^{1/(r-1/2)} + 1.$$

Now we can give some interpretations of Theorem 4.2 and Corollaries 4.3 and 4.4. We first consider the case $q = +\infty$ (the unbounded probabilistic case). By 4.4 we see that for large $\delta$, precisely for

$$\delta \geq \exp(-(1/\epsilon)^{1/(r-1/2)}),$$

probabilistic and average cardinality are of the same order, while for smaller $\delta$, they behave differently. This gives a quantitative foundation to the intuitive understanding that the average estimate also holds with sufficiently large probability.

In the general bounded probabilistic case $q < \infty$ the cardinality $m_{\text{prob}}(\epsilon, \delta, q)$ is of the same order as the unbounded probabilistic cardinality as long as

$$(q/\epsilon)^{1/s} \geq \max(((\log(1/\delta))^{1/2}/\epsilon)^{1/r}, (1/\epsilon)^{1/(r-1/2)}),$$

otherwise $m_{\text{prob}}(\epsilon, \delta, q)$ is of the same order as the worst case cardinality. Roughly speaking, the unbounded probabilistic case turns into the worst case, and nothing else occurs. Let us have a slightly closer look at this when $\epsilon$ and $q$ are fixed and $\delta$ gets small. Suppose that

$$(q/\epsilon)^{1/s} \geq (1/\epsilon)^{1/(r-1/2)},$$

that is,

$$q \geq \epsilon^{(r-s-1/2)/(r-1/2)}$$

(since we have a bound on $q$, $q \geq q_1$, and $r - s - 1/2 > 0$, this just means that $\epsilon$ is not too large). Then, as $\delta \to 0$, first the average case cardinality occurs (as long as (36) holds), then comes the unbounded probabilistic cardinality, as long as
\[ \delta \geq \exp\left(-\frac{1}{\varepsilon}e^{2(1-s)/q}q^{2r/s}\right), \]

and finally, below this value, the worst case dominates.

Let us finally mention some questions which arise from the results of this paper. It follows from 4.2 and the remark after the proof that for this concrete approximation problem nonadaptive information provides the same order as adaptive information. For the worst case, average case, and unbounded probabilistic setting there are general results of this kind (see IBC, 4.5.2, 6.5.6, and 8.5.3). It would be interesting to have counterparts for the bounded probabilistic setting. Furthermore, in Section 4 we restricted our considerations to the Hilbert space case and relied on the identity of certain types of widths in this case. The Banach space case, that is, the approximation of functions from \( W^s_p \) in \( L_q \), is open. In particular, it would be interesting to understand if the order of the cardinality of the bounded probabilistic case always reduces to either the worst case or the unbounded probabilistic case. Finally, other numerical problems like, e.g., integration (with natural restrictions on the information available) should be analyzed from this point of view.

**Acknowledgment**

I thank H. Woźniakowski for some motivating discussions, for the pressure to replace the nonadaptive first version by an adaptive one, and for finally convincing me by providing the estimate (10).

**References**

Köhn, T. (1981), \( \gamma \)-Summing operators in Banach spaces of type \( p \) (\( 1 < p \leq 2 \)) and cotype \( q \) (\( 2 \leq q < \infty \)), *Teor. Veroyatnost. Primenen.* 26, 121–131.  


