Interner Bericht

Information Complexity of Multivariate Fredholm Equations in Sobolev Classes

Karin Frank
Stefan Heinrich
Sergei Pereverzev

263/95

Fachbereich Informatik

Universität Kaiserslautern · Postfach 3049 · D-67653 Kaiserslautern

Information Complexity of Multivariate Fredholm Equations in Sobolev Classes

Karin Frank
Stefan Heinrich
Sergei Pereverzev

263/95

Universität Kaiserslautern
AG Numerische Algorithmen
Postfach 30 49
67653 Kaiserslautern

Jan. 1995

Herausgeber: AG Numerische Algorithmen
Leiter: Professor Dr. S. Heinrich

Information Complexity of Multivariate Fredholm Integral Equations in Sobolev Classes

Karin Frank, Stefan Heinrich, Sergei Pereverzev

Fachbereich Informatik
Universität Kaiserslautern
Postfach 3049, D–67653 Kaiserslautern

11th January 1995

Abstract

In this paper, the complexity of full solution of Fredholm integral equations of the second kind with data from the Sobolev class W_2^r is studied. The exact order of information complexity is derived. The lower bound is proved using a Gelfand number technique. The upper bound is shown by providing a concrete algorithm of optimal order, based on a specific hyperbolic cross approximation of the kernel function. Numerical experiments are included, comparing the optimal algorithm with the standard Galerkin method.

1 Introduction

Information–Based Complexity theory studies the intrinsic difficulty of the approximate solution of numerical problems for which the information is partial, contaminated and priced. Determining the information complexity of basic problems of mathematical physics is a principal goal of Information–Based Complexity theory. One of those basic problems is the solution of Fredholm integral equations of the second kind.

A first estimate of information complexity for the construction of the full solution of integral equations with kernels from Sobolev classes was obtained by Emelyanov and Ilin [EI67]. They restricted themselves to the study of algorithms computing the approximate solution using as information functionals only function values of the kernel and the free term at some points.

The general situation of algorithms using values of arbitrary linear continuous functionals was considered in [Per89]. However, in this paper the lower bound and the upper bound of information complexity differed by a logarithmic factor. In the present paper we determine the exact order of information complexity for Fredholm equations with kernels from Sobolev classes using a Gelfand number technique developed in [Hei93]. This exact order coincides with the order of the upper bound shown in [Per89].

2 Formulation of the problem and some general results

Let V, E and K be normed linear spaces. For the space of all bounded linear operators from V into E we shall write $L(V,E)$ and $L(E) = L(E,E)$. Assume that V is continuously embedded in E and $J_V \in L(V,E)$ is the embedding operator from V into E. Moreover, we assume that there is some linear continuous operator T assigning to each element $k \in K$ an operator $T_k \in L(E)$. Let $V_0 \subset V$ and $K_0 \subset K$ be subsets given in such a way that the operator $(I - T_k)^{-1} \in L(E)$ for any $k \in K_0$, where I is the identity operator. Then setting $X_0 = K_0 \times V_0$ we consider the class of operator equations

$$u - T_k u \;=\; f\,,\; k \in K_0,\; f \in V_0\,. \tag{1}$$

The operator $S : X_0 \to E$ defined by

$$S(k,f) \;=\; (I - T_k)^{-1} f \tag{2}$$

is called the solution operator of equation (1). For the analysis of (1) and (2) we shortly recall the framework of Information–Based Complexity theory. For details the reader is referred to [TWW88].

By a method of specifying information about equations (1) on X_0 we understand any operator $N : X_0 \to \mathbb{R}^{n_1+n_2}$, $N = (N_1, N_2)$ with

$$\begin{aligned} N_1 k &= (\lambda_1(k), \ldots, \lambda_{n_1}(k))\,,\; \lambda_i \in K^*,\; i = 1, \ldots, n_1, \\ N_2 f &= (\sigma_1(f), \ldots, \sigma_{n_2}(f))\,,\; \sigma_j \in V^*,\; j = 1, \ldots, n_2, \end{aligned}$$

where K^* and V^* denote the dual spaces of K and V respectively. $N = (N_1, N_2)$ is also called an information operator. Moreover, we denote by $\mathrm{card}(N)$ the number of all linear functionals taking part in the definition of the information operator N, i.e. $\mathrm{card}(N) = n_1 + n_2$.

By an algorithm φ of approximate solution of equations (1) we mean any operator φ assigning to the information vector $N(k,f) \in \mathbb{R}^{n_1+n_2}$ an element $\varphi(N(k,f)) \in E$ as

an approximate solution of (1). We assume that every algorithm φ is connected with some set F_φ of elements $g(b_1, \ldots, b_l) \in E$ defined by values of numerical parameters $b_1, \ldots, b_l$ and $\varphi(N(k,f)) = g(b_1, \ldots, b_l)$, where $b_1, \ldots, b_l$ depend on $N(k,f)$ and for the calculation of these values it is required to execute only arithmetic operations, including comparisons, on the components of the vector $N(k,f)$. For a concrete algorithm φ we denote by $\mathrm{card}(\varphi)$ the number of arithmetic operations required in order to realize the algorithm. Moreover, for a fixed information operator N we denote by $\Phi(N)$ the set of algorithms φ defined on $N(X_0)$. Considering algorithms of $\Phi(N)$ it is natural to suppose that $\mathrm{card}(N) \leq \mathrm{card}(\varphi) + 1$. Otherwise, the algorithm $\varphi \in \Phi(N)$ cannot use all information represented by the components of the vector $N(k,f)$.

The error of the algorithm $\varphi \in \Phi(N)$ on the class X_0 is defined as

$$e(X_0, \varphi) \;=\; \sup_{(k,f) \in X_0} \|S(k,f) - \varphi(N(k,f))\|_E \;.$$

The information complexity of equations (1) on the class X_0 is determined by the quantity

$$e_n(X_0) \;=\; \inf_{N:card(N) \leq n} \; \inf_{\substack{\varphi \in \Phi(N) \\ card(\varphi) \leq n}} e(X_0, \varphi) \;.$$

This is the minimal error which can be reached performing at most n arithmetic operations on the values of at most n information functionals. Moreover, for a fixed information operator $N = (N_1, N_2)$ we introduce the following quantities

$$\begin{aligned} r(X_0, N) \;&=\; \inf_{\varphi \in \Phi(N)} e(X_0, \varphi) \;, \\ d(X_0, N) \;&=\; \sup_{\substack{k,h \in K_0 \\ N_1 k = N_1 h}} \; \sup_{\substack{f,g \in V_0 \\ N_2 f = N_2 g}} \|S(k,f) - S(h,g)\|_E \;, \end{aligned}$$

called the radius and the diameter of information N respectively. It is well–known that

$$\frac{1}{2} d(X_0, N) \;\leq\; r(X_0, N) \;\leq\; d(X_0, N) \;. \tag{3}$$

Note that the so–called n–th minimal radius of information

$$r_n(X_0) \;=\; \inf_{N:card(N) \leq n} r(X_0, N)$$

serves as a lower bound for the information complexity $e_n(X_0)$, i.e.

$$e_n(X_0) \;\geq\; r_n(X_0) \;. \tag{4}$$

Now, as in [Hei93] we establish some relation between the n–th minimal radius $r_n(X_0)$ and Gelfand numbers of certain operators. First of all we impose some assumptions on our class X_0 of equations (1).

Let B_K and B_V be unit balls of the spaces K and V respectively. Fix constants $\rho = (\rho_1, \ldots, \rho_6)$, $\rho_i > 0$ for $i = 1, \ldots, 6$ and $\rho_4, \rho_6 > 1$. We assume that we are given a subset $K_\rho \subset K$ such that

(i) $\rho_1 B_K \subset K_\rho \subset \rho_2 B_K$,

(ii) for any $k \in K_\rho$:

$$\|T_k\|_{E\to E} \le \rho_3\,, \quad \|(I - T_k)^{-1}\|_{E\to E} \le \rho_4\,,$$
$$\|T_k\|_{V\to V} \le \rho_5\,, \quad \|(I - T_k)^{-1}\|_{V\to V} \le \rho_6\,.$$

In the sequel we shall consider equations (1) on the class $X_\rho = K_\rho \times B_V$.

Now we introduce the so–called Gelfand numbers of an operator. Given two Banach spaces Y and Z, for an operator $U \in L(Y, Z)$ the n–th Gelfand number of U is defined by

$$c_n(U : Y \to Z) = \inf_{\lambda_1,\ldots,\lambda_{n-1}\in Y^*} \sup_{\substack{y\in B_Y \\ \lambda_i(y)=0, i=1,\ldots,n-1}} \|Uy\|_Z\,.$$

Moreover, let us define the operator $\Psi : K \to L(V, E)$ by

$$\Psi k = T_k J_V\,,$$

where J_V is the embedding operator from V into E. For a fixed information operator $N = (N_1, N_2)$ we consider the following quantities

$$c(\Psi, N) = \sup_{\substack{k\in B_K, \\ N_1 k = 0}} \|\Psi k\|_{V\to E}\,,$$
$$c(J_V, N) = \sup_{\substack{f\in B_V, \\ N_2 f = 0}} \|J_V f\|_E\,.$$

Theorem 1 *For any information operator $N = (N_1, N_2)$,*

$$d_1 \cdot [c(\Psi, N) + c(J_V, N)] \le r(X_\rho, N) \le d_2 \cdot [c(\Psi, N) + c(J_V, N)]\,,$$

where the constants d_1, d_2 depend only on ρ.

Proof:

Let us fix an arbitrary $N = (N_1, N_2)$ and consider the quantities

$$A = \sup_{\substack{k,h \in K_\rho \\ N_1 k = N_1 h}} \sup_{f \in B_V} \|S(k,f) - S(h,f)\|_E ,$$

$$B = \sup_{k \in K_\rho} \sup_{\substack{f,g \in B_V \\ N_2 f = N_2 g}} \|S(k,f) - S(k,g)\|_E .$$

It is obvious that

$$\max(A,B) \le d(X_\rho, N) \le A + B . \tag{5}$$

Now we shall derive some estimate for A. First of all we note that

$$\begin{aligned} A &= \sup_{\substack{k,h \in K_\rho \\ N_1 k = N_1 h}} \|((I - T_k)^{-1} - (I - T_h)^{-1}) J_V\|_{V \to E} \\ &= \sup_{\substack{k,h \in K_\rho \\ N_1 k = N_1 h}} \|(I - T_k)^{-1}(T_k - T_h)(I - T_h)^{-1} J_V\|_{V \to E} . \end{aligned} \tag{6}$$

Moreover, using the properties of K_ρ, for $h, k \in K_\rho$ we have

$$\|(I - T_k)^{-1}(T_k - T_h)(I - T_h)^{-1} J_V\|_{V \to E} \le \rho_4 \rho_6 \cdot \|T_k - T_h\|_{V \to E} , \tag{7}$$

and on the other hand

$$\begin{aligned} &\|(I - T_k)^{-1}(T_k - T_h)(I - T_h)^{-1} J_V\|_{V \to E} \\ &\quad \ge \|I - T_k\|_{E \to E}^{-1} \|T_k - T_h\|_{V \to E} \|I - T_h\|_{V \to V}^{-1} \\ &\quad \ge (1 + \rho_3)^{-1} (1 + \rho_5)^{-1} \|T_k - T_h\|_{V \to E} . \end{aligned} \tag{8}$$

Further, from the definition of $c(\Psi, N)$ and assumption (i) we obtain

$$\begin{aligned} \rho_1 \cdot c(\Psi, N) &= \sup_{\substack{k \in \rho_1 B_K \\ N_1 k = 0}} \|T_k\|_{V \to E} \\ &\le \frac{1}{2} \sup_{\substack{k,h \in K_\rho \\ N_1 k = N_1 h}} \|T_k - T_h\|_{V \to E} \\ &\le \frac{1}{2} \sup_{\substack{k \in 2\rho_2 B_K \\ N_1 k = 0}} \|T_k\|_{V \to E} \\ &= \rho_2 \cdot c(\Psi, N) . \end{aligned} \tag{9}$$

In such a way it follows from (6) – (9) that

$$2\rho_1 (1 + \rho_3)^{-1} (1 + \rho_5)^{-1} \cdot c(\Psi, N) \le A \le 2\rho_2 \rho_4 \rho_6 \cdot c(\Psi, N) . \tag{10}$$

In a similar manner we can prove that

$$2(1+\rho_3)^{-1} \cdot c(J_V, N) \;\leq\; B \;\leq\; 2\rho_4 \cdot c(J_V, N)\,. \tag{11}$$

Now the statement of the theorem follows from (3), (5), (10) and (11).

□

In the sequel we shall use the following notation: If $P(b)$ and $Q(b)$ are functions defined on some set B, we write

$$P(b) \prec Q(b)\,,$$

if there is a constant $c > 0$ such that for all $b \in B$: $P(b) \leq c \cdot Q(b)$. We write

$$P(b) \asymp Q(b)\,,$$

if $P(b) \prec Q(b)$ and $Q(b) \prec P(b)$.

Corollary 1

$$r_n(X_\rho) \;\asymp\; \inf_{n_1+n_2\leq n} \left(c_{n_1+1}(\Psi : K \to L(V,E)) + c_{n_2+1}(J_V : V \to E)\right)\,.$$

This relation follows immediately from Theorem 1 and the definition of Gelfand numbers.

3 The main result

Let $d \in \mathbb{N}$, $G = [0, 2\pi]^d$, and let $L_2(G)$ be the space of square–summable functions on G with the usual norm $\|\cdot\|$ and the usual inner product $(\cdot,\cdot)$. We set

$$\begin{aligned} e_0(s) &= \frac{1}{\sqrt{2\pi}}\,, \\ e_n(s) &= \frac{1}{\sqrt{\pi}} \cos(ns)\,, \\ e_{-n}(s) &= \frac{1}{\sqrt{\pi}} \sin(ns) \end{aligned}$$

for $n \in \mathbb{N}$ and $s \in [0, 2\pi]$. Then for a given multiindex $i = (i_1, \ldots, i_d) \in \mathbb{Z}^d$ the basis function $e_i \in L_2(G)$ is defined by

$$e_i(t) \;=\; e_{i_1}(t_1) \cdot \ldots \cdot e_{i_d}(t_d)\,,$$

where $t = (t_1, \ldots, t_d) \in G$. The Fourier coefficients of $f \in L_2(G)$ are given by $\hat{f}(i) = (f, e_i)$, $i \in \mathbb{Z}^d$. Moreover, for $i \in \mathbb{Z}^d$ we set $|i| = \sqrt{i_1^2 + \ldots + i_d^2}$.

Let $r \in \mathbb{R}^+$. Then the Sobolev space $H^r(G)$ is defined as

$$H^r(G) = \left\{ f : f \in L_2(G), \|f\|_r := \left(\sum_{i \in Z^d} (1 + |i|^2)^r \hat{f}(i)^2 \right)^{\frac{1}{2}} < \infty \right\} .$$

Note that for $r \in \mathbb{N}$, the space $H^r(G)$ consists of all periodic functions on G which have square-summable generalized partial derivatives up to the order r. We shall use the following abbreviations: $L_2 = L_2(G)$, $H^r = H^r(G)$, $\mathcal{H}^r = H^r(G^2)$.

Now, referring to the notation of the previous section, we set $E = L_2$, $V = H^r$ and $K = \mathcal{H}^r$. Moreover, for each $k \in \mathcal{H}^r$ the operator $T : \mathcal{H}^r \to L(L_2)$ is defined by

$$T_k g(t) = \int_G k(t,s)\, g(s)\, ds . \tag{12}$$

Let $\alpha_1 > 0$, $\alpha_2 > 1$, $\alpha = (\alpha_1, \alpha_2)$ and consider the set

$$\mathcal{H}^r_\alpha = \{ k \in \mathcal{H}^r : \|k\|_r \le \alpha_1, \ \|(I - T_k)^{-1}\|_{L_2 \to L_2} \le \alpha_2 \} .$$

Denote by X^r_α the class of equations (1) with free terms $f \in B_{H^r} = B_V$ and operators (12) with kernels $k(t,s)$ from $\mathcal{H}^r_\alpha$, i.e. $X^r_\alpha = \mathcal{H}^r_\alpha \times B_{H^r}$. It is easy to see that for some $\rho_1, \ldots, \rho_6$ depending on α the set $\mathcal{H}^r_\alpha$ may be considered as K_ρ treated in the previous section, thus

$$\mathcal{H}^r_\alpha = K_\rho , \ X^r_\alpha = X_\rho , \ \rho = \rho(\alpha) . \tag{13}$$

Now we can state the main result.

Theorem 2 *For the class of equations (1) on X^r_α defined above*

$$e_n(X^r_\alpha) \asymp n^{-\frac{r}{d}} \log^{\frac{r}{d}} n .$$

Proof:

The required upper estimate for the quantity $e_n(X^r_\alpha)$ follows from [Per89]. In this paper an algorithm for the approximate solution of Fredholm integral equations with kernels from Sobolev spaces was constructed which provides the upper complexity bound. This algorithm uses as information the values of Fourier coefficients of the

kernels with numbers from a so–called hyperbolic cross. Some modification of this algorithm will be presented in the next section.

Now we shall prove the lower estimate for $e_n(X_\alpha^r)$. Using (4) and Corollary 1, we have

$$e_n(X_\alpha^r) \;\geq\; r_n(X_\alpha^r) \;\succ\; c_n(\Psi : \mathcal{H}^r \to L(H^r, L_2)) \,. \tag{14}$$

To show the lower bound, we shall estimate the Gelfand numbers of Ψ. Let $\{b_{ij}\}_{i,j\in Z^d}$ be the unit vector basis of $l_2(\mathbb{Z}^{2d})$ and define the operator $W : l_2(\mathbb{Z}^{2d}) \to \mathcal{H}^r$ as

$$W\, b_{ij} \;=\; (1 + |i|^2 + |j|^2)^{-\frac{r}{2}} \cdot e_i(t)\, e_j(s) \,.$$

Moreover, we define the operator $U \;:\; L(H^r, L_2) \;\to\; l_\infty(\mathbb{Z}^{2d})$ assigning to each operator $A \in L(H^r, L_2)$ an element

$$U\, A \;=\; \left((1 + |m|^2)^{-\frac{r}{2}} \cdot (Ae_m, e_l)\right)_{m,l\in Z^d} \,.$$

It is obvious that the operator W is an isometry, so $\|W\| = 1$, and the operator U is an injection with $\|U\| \leq 1$. Now we compose the operators W and U with Ψ and obtain an operator

$$\begin{aligned} D \;&=\; U\, \Psi\, W \,, \\ D \;&:\; l_2(\mathbb{Z}^{2d}) \to l_\infty(\mathbb{Z}^{2d}) \,. \end{aligned}$$

It is easily verified that D is a diagonal operator which acts in the following manner:

$$D\, b_{ij} \;=\; \xi_{ij}\, b_{ij} \,,\; \xi_{ij} \;=\; [(1 + |i|^2 + |j|^2)(1 + |j|^2)]^{-\frac{r}{2}} \,.$$

Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq \ldots$ be the elements of the sequence $\{\xi_{ij}\}$ arranged in nonincreasing order. Namely,

$$\begin{aligned} \lambda_n \;&=\; \inf\{\varepsilon : \mathrm{card}\{(i,j) : \xi_{ij} > \varepsilon\} < n\} \\ &=\; \max_{\substack{Q\subseteq Z^{2d} \\ card(Q)=n}} \min\{\xi_{ij} : (i,j) \in Q\} \,. \end{aligned}$$

If

$$Q_n \;=\; \{(i,j) : \xi_{ij} \geq n^{-r}\} \,,$$

then it follows from the definition of $\{\lambda_n\}_{n\in\mathbb{N}}$ that

$$\lambda_{card(Q_n)} \;\geq\; n^{-r} \;>\; \lambda_{card(Q_n)+1} \,. \tag{15}$$

Let us estimate $\mathrm{card}(Q_n)$. For this end we consider the set

$$B_n \;=\; \{(i,j) : \max\{|i_1|,\dots,|i_d|,|j_1|,\dots,|j_d|,1\}\cdot\max\{|j_1|,\dots,|j_d|,1\} \le n\}\ .$$

Keeping in mind that

$$Q_n \;=\; \left\{(i,j) : (1+|i|^2+|j|^2)^{\frac{1}{2}}(1+|j|^2)^{\frac{1}{2}} \le n\right\}\ ,$$

we have

$$Q_n \;\subset\; B_n \;\subset\; Q_{\lceil p\cdot n\rceil}\,,\ p=\sqrt{(2d+1)(d+1)}\,, \tag{16}$$

where $\lceil pn\rceil$ denotes the smallest integer greater than or equal to pn. Moreover,

$$\begin{aligned}
&\mathrm{card}(B_n) \;= \\
&\quad= \mathrm{card}\left\{(i,j) : \max\{|i_1|,\dots,|i_d|\} \le \sqrt{n},\ \max\{|j_1|,\dots,|j_d|\} \le \sqrt{n}\right\} \;+ \\
&\qquad+ \sum_{\sqrt{n}<l\le n} \mathrm{card}\left\{(i,j) : \max\{|i_1|,\dots,|i_d|\} = l,\ \max\{|j_1|,\dots,|j_d|\} \le \frac{n}{l}\right\} \\
&\quad\asymp n^d + \sum_{\sqrt{n}<l\le n} l^{d-1}\left(\frac{n}{l}\right)^d \\
&\quad\asymp n^d\cdot\sum_{\sqrt{n}<l\le n}\frac{1}{l} \\
&\quad\asymp n^d\log n\ .
\end{aligned}$$

Now from (16) we find

$$\mathrm{card}(Q_n) \;\asymp\; n^d\log n\ . \tag{17}$$

Combining (15) and (17) we obtain

$$\lambda_{[n^d\log n]} \;\asymp\; n^{-r}\ ,$$

or what is the same

$$\lambda_n \;\asymp\; n^{-\frac{r}{d}}\log^{\frac{r}{d}} n\ .$$

Then it follows from Theorem 11.11.7 in [Pie78] that

$$c_n\left(D : l_2(\mathbb{Z}^{2d}) \to l_\infty(\mathbb{Z}^{2d})\right) \;\asymp\; \lambda_n \;\asymp\; n^{-\frac{r}{d}}\log^{\frac{r}{d}} n\ . \tag{18}$$

Furthermore, by basic properties of Gelfand numbers

$$\begin{aligned} c_n\left(D : l_2(\mathbb{Z}^{2d}) \to l_\infty(\mathbb{Z}^{2d})\right) &\leq \|U\| \cdot c_n(\Psi : \mathcal{H}^r \to L(H^r, L_2)) \cdot \|W\| \\ &\leq c_n(\Psi : \mathcal{H}^r \to L(H^r, L_2)) \,. \end{aligned}$$

Finally, using this inequality and relations (14) and (18) we get

$$e_n(X_\alpha^r) \succ n^{-\frac{r}{d}} \log^{\frac{r}{d}} n \,.$$

The theorem is proved. □

4 The algorithm

In this section, we shall describe the optimal algorithm used in the proof of the upper bound of Theorem 2. The algorithm is based on Fourier coefficients and a hyperbolic cross approximation of the kernel function k. The approximate solution is computed in two steps: first we compute a primary approximation v_0 as the solution of an integral equation with an approximated kernel function, then we perform some kind of iterative refinement.

The algorithm can be used not only for periodic functions on $[0, 2\pi]$, but also for nonperiodic functions, carrying out a periodization of the kernel and the right–hand side by some transformation of variables, in a preprocessing step. We shall explain this transformation, which allows us even to simplify the algorithm by reducing the basis.

Let $(k, f) \in X_\alpha^r$ be given. For any $m \in \mathbb{N}$, let C_m be the following set of multiindices $i = (i_1, \ldots, i_d) \in \mathbb{Z}^d$:

$$C_m \;=\; \{i \in \mathbb{Z}^d : \max\{|i_1|, \ldots, |i_d|\} \leq m\} \,.$$

Then the orthogonal projection operator $P_m : L_2(G) \to \operatorname{span}\{e_i : i \in C_m\}$ is defined as

$$(P_m f)(t) \;=\; \sum_{i \in C_m} (f, e_i)\, e_i(t) \,.$$

Now, fix $n \in \mathbb{N}$ and recall, that the set B_n is described by

$$B_n \;=\; \{(i,j) : \max\{|i_1|, \ldots, |i_d|, |j_1|, \ldots, |j_d|, 1\} \cdot \max\{|j_1|, \ldots, |j_d|, 1\} \leq n\} \,.$$

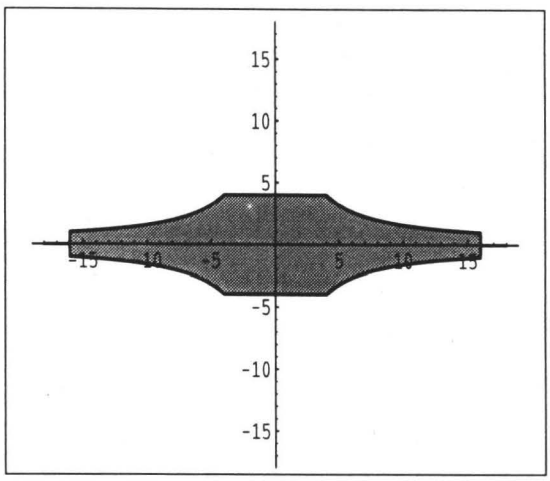

Figure 1: Shape of the hyperbolic cross B_n for $d = 1$, $n = 16$.

For $d = 1$, the shape of this set is shown in figure (1). Then the projections h, f_n of the kernel and the right–hand side, respectively, are defined by its Fourier coefficients

$$\begin{aligned} \hat{h}(i,j) &= \begin{cases} \hat{k}(i,j) & (i,j) \in B_n \\ 0 & \text{otherwise}, \end{cases} \\ \hat{f}_n(i) &= \begin{cases} \hat{f}(i) & i \in C_n \\ 0 & \text{otherwise}. \end{cases} \end{aligned} \tag{19}$$

The algorithm replaces equation (1) by the approximate equation

$$w - T_h w = f_n, \tag{20}$$

and computes an approximation v to w in an iterative way. As can be seen easily, for the construction of equation (20) we need the information $\bar{N} = (\bar{N}_1, \bar{N}_2)$ about the initial equation (1), where

$$\begin{aligned} \bar{N}_1 k &= \left\{ \hat{k}(i,j) \right\}_{(i,j) \in B_n}, \\ \bar{N}_2 f &= \left\{ \hat{f}(i) \right\}_{i \in C_n}. \end{aligned}$$

From relations (16), (17) it follows that

$$\text{card}(\bar{N}) \asymp n^d \cdot \log n.$$

Unfortunately, the direct solution of equation (20) by means of some exact solution method for systems of linear equations would take too many arithmetic operations.

So we set $m = \lceil n^{1/3} \rceil$, and compute the approximate solution v in the following two steps. First, we calculate the solution v_0 of the equation

$$v_0 - T_h P_m v_0 = f_n , \tag{21}$$

and then the final approximation as

$$\bar{\varphi}(\bar{N}; k, f) = v = v_0 + (I - T_h P_m)^{-1} (T_h v_0 - v_0 + f_n) . \tag{22}$$

The algorithm (21), (22) was analyzed in [Per89] for the case $d = 1$. This analysis can be extended easily to the d–dimensional case ($d > 1$), and we get

$$\mathrm{card}(\bar{\varphi}) \asymp n^d \log n . \tag{23}$$

Related algorithms were discussed in [FH94] and [Fra94], where the analysis for arbitrary $d \in \mathbb{N}$ was carried out in greater detail. In a similar way as there, one can prove that the algorithm $\bar{\varphi}$ satisfies the error estimate

$$e(X_\alpha^r, \bar{\varphi}) \asymp n^{-r} . \tag{24}$$

Consequently, from Theorem 2 and relations (21) - (24) it follows that the algorithm $\bar{\varphi}$ is error–optimal for the class X_α^r in the sense of information complexity.

A modification of the algorithm $\bar{\varphi}$ described by equation (22), which proved to be useful in implementations, is the use of a so–called dyadic hyperbolic cross instead of the hyperbolic cross B_n. For this end we replace the operator T_h in (20) by the operator

$$T_{h_1} = \sum_{l=1}^{M} P_{2^{2M-l}} T_k (P_{2^l} - P_{2^{l-1}}) + P_{2^{2M}} T_k P_1 , \tag{25}$$

where $n = 2^N$ and $M = \lceil \frac{N}{2} \rceil$. Then T_{h_1} is also an integral operator, whose kernel function is defined by

$$\hat{h}_1(i,j) = \begin{cases} \hat{k}(i,j) & (i,j) \in \tilde{B}_n \\ 0 & \text{otherwise} , \end{cases}$$

$$\tilde{B}_n = \bigcup \{ C_{2^l} \times C_{2^m} : l, m \geq 0, \max(l,m) + m \leq 2M \} .$$

The shape of the dyadic set $\tilde{B}_n$ is shown in figure 2.

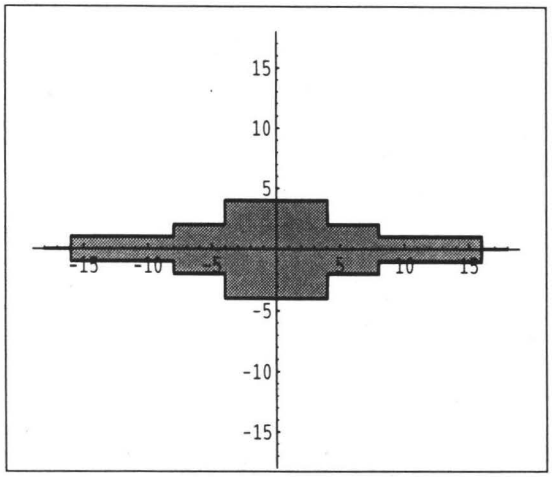

Figure 2: Shape of the hyperbolic cross $\tilde{B}_n$ for $d = 1$, $n = 16$.

In the presented form, the algorithm works only for periodic on G functions. It can be applied also to Fredholm integral equations with nonperiodic kernels and free terms

$$z(x) - \int_{[0,1]^d} k_1(x,y)\, z(y)\, dy \;=\; g(x)\,,$$

where $x, y \in \mathbb{R}^d$. For this end, we make e.g. the following change of variables

$$\begin{aligned} x_j &= \frac{1}{2}(1 - \cos t_j)\,, \\ y_j &= \frac{1}{2}(1 - \cos s_j) \end{aligned} \tag{26}$$

for $t_j, s_j \in [0, \pi]$, $j = 1, .., d$, and consider the transformed equation

$$u - T_k u \;\equiv\; u(t) - \int_{[0,\pi]^d} k(t,s)\, f(s)\, ds \;=\; f(t)\,, \tag{27}$$

where

$$\begin{aligned} k(t_1, .., t_d, s_1, .., s_d) &= k_1\left(\frac{1-\cos t_1}{2}, .., \frac{1-\cos t_d}{2}, \frac{1-\cos s_1}{2}, .., \frac{1-\cos s_d}{2}\right) * \\ &\quad * 2^{-d} \prod_{i=1}^{d} \sin s_i\,, \end{aligned} \tag{28}$$

$$u(t_1, .., t_d) = z\left(\frac{1-\cos t_1}{2}, .., \frac{1-\cos t_d}{2}\right), \tag{29}$$

$$f(t_1,..,t_d) \;=\; g\left(\frac{1-\cos t_1}{2},..,\frac{1-\cos t_d}{2}\right). \tag{30}$$

From (28) – (30) it follows that the integral operator T_k of equation (27) can be considered as acting in the space of even 2π–periodic functions. Hence, for the approximate solution of equation (27) we can reduce our basis to the cosine–basis in $L_2(G)$.

The cardinality of required information as well as the number of arithmetic operations performed by the algorithm have changed only by a constant factor. If the functions (k, f) in (28), (30) are in $\mathcal{H}^r_\alpha$, then the order of accuracy is preserved.

5 Numerical examples

To present a numerical example we restrict ourselves to the case $d = 1$ and consider the following two integral equations.

$$u_1(s) \;-\; \int_0^{2\pi} k_1(s,t)\,u_1(t)\,dt \;=\; f_1(s) \tag{31}$$

with

$$\begin{aligned} k_1(s,t) &= \frac{1}{\pi}\sum_{i,j>0}\frac{\sin(is)\cdot\sin(jt)}{(1+i^2+j^2)^2}\,, \\ u_1(s) &= \frac{1}{\sqrt{\pi}}\sum_{i>0}\frac{\sin(is)}{i^{\frac{7}{2}}}\,. \\ f_1(s) &= u_1(s) - \frac{1}{\sqrt{\pi}}\sum_{i>0}\sin(is)\cdot\sum_{j>0}(1+i^2+j^2)^{-2}\cdot j^{-\frac{7}{2}} \end{aligned}$$

Note that k_1 and f_1 are chosen in such a way that they (roughly) represent the same degree of smoothness: Both belong to the Sobolev space $H^{3-\alpha}$ for all $\alpha > 0$, but not to H^3. Figure 3 shows the graphs of k_1 and u_1.

The second example equation will be

$$u_2(s) \;-\; \int_0^{2\pi} k_2(s,t)\,u_2(t)\,dt \;=\; f_2(s) \tag{32}$$

with

$$k_2(s,t) \;=\; \frac{1}{2\gamma}\sum_{i,j>0}\frac{\sin(js)\sin(it)}{i^3\cdot j^3} \;=\; \frac{(s^3-3\pi s^2+2\pi^2 s)(t^3-3\pi t^2+2\pi^2 t)}{288\gamma}\,,$$

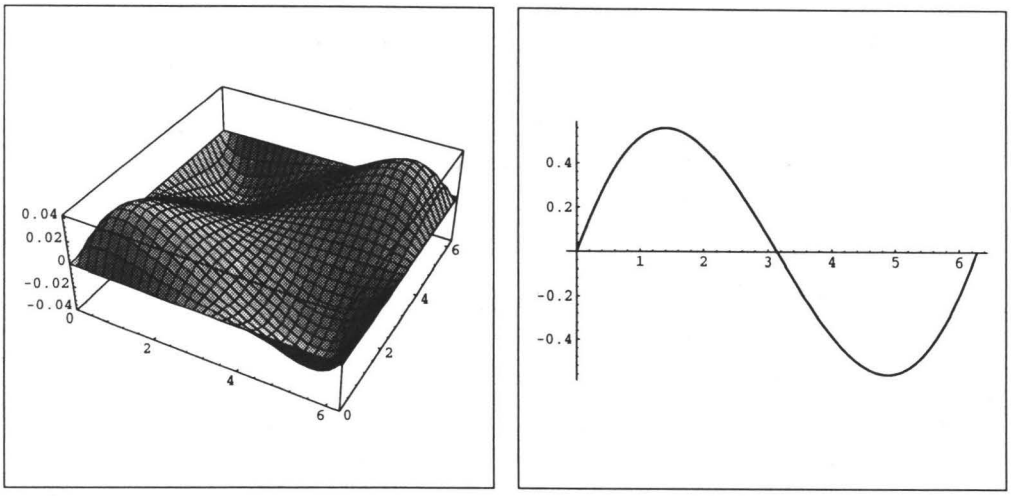

Figure 3: Kernel function and solution of example equation (31).

$$u_2(s) = \sum_{j>0} \frac{\sin(js)}{j^3} = \frac{s^3 - 3\pi s^2 + 2\pi^2 s}{12},$$
$$f_2(s) = \frac{1}{2} u_2(s),$$

where

$$\gamma = \int_0^{2\pi} \left(\frac{x^3 - 3\pi x^2 + 2\pi^2 x}{12} \right)^2 dx \approx 3.19608 .$$

The kernel of this integral equation belongs to the Sobolev space with dominating mixed derivative $H^{\frac{5}{2}-\alpha,\frac{5}{2}-\alpha}([0,2\pi])^2$ for each $\alpha > 0$, but not to $H^{\frac{5}{2},\frac{5}{2}}([0,2\pi])^2$ (see e.g. [Fra94] for the definition). Both solution and right–hand side belong to the Sobolev space $H^{\frac{5}{2}-\alpha}([0,2\pi])$, $\alpha > 0$.

We applied two algorithms to these equations: the order–optimal algorithm $\bar{\varphi}$ with the dyadic hyperbolic cross $\tilde{B}_n$ and the standard Galerkin method. Both algorithms use Fourier coefficients of the kernel function and the free term. They were implemented in C^{++} on a Workstation HP 9000/712/60. In all calculations double precision was used. The required Fourier coefficients of the right–hand sides were computed beforehand, the measured CPU–time does not include this preprocessing step. This procedure is justified by the subject of our attention: we assume that we are given the needed information or at least can compute each functional with constant cost, in order to look at the cost and the accuracy of the algorithm working with this information.

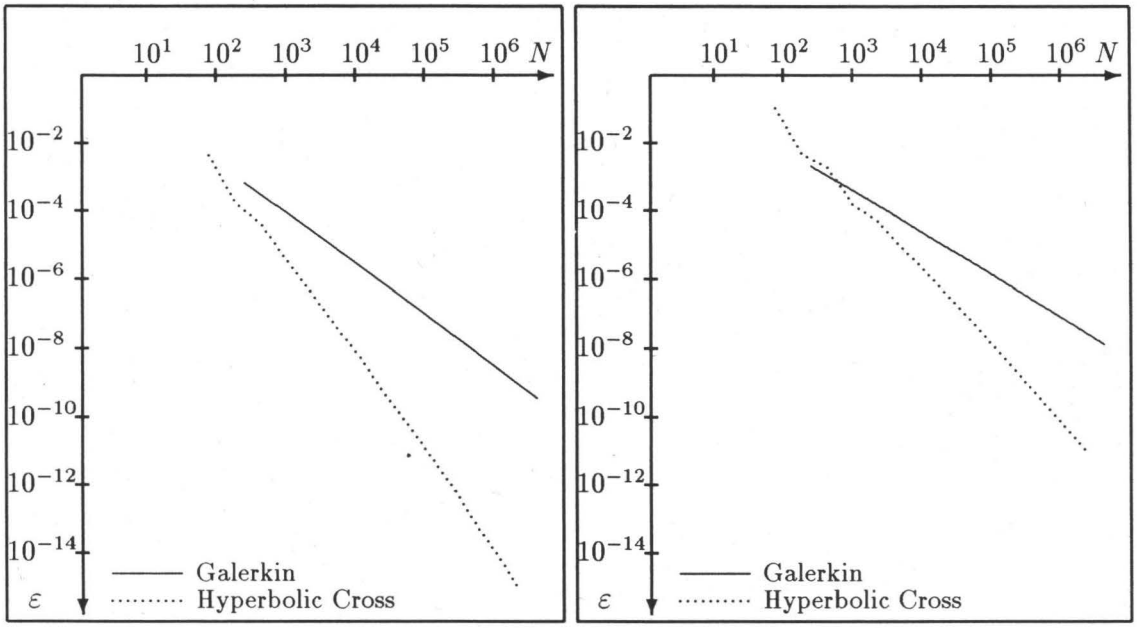

Figure 4: Accuracy of the standard Galerkin method and the Hyperbolic cross method $\bar{\varphi}$ for example equations (31) (left) and (32) (right).

In figure 4, the accuracy of both methods applied to the example equations is shown. There, the abscissa constitutes the full number of Fourier coefficients used by the algorithm. The ordinate is the L_2–error of the approximate solution. Both axes are logarithmically scaled.

The example equation (31) satisfies the conditions of Theorem 2. Consequently, the algorithm $\bar{\varphi}$ reaches the optimal convergence rate of $O(N^{-r}\log^r N)$ for all $r < 3$, whereas the Galerkin method has the well–known convergence rate of $O(N^{-\frac{r}{2}})$ for all $r < 3$, but not $O(N^{-\frac{3}{2}})$.

Looking at the accuracy in example (32), we must recall that the kernel function of this equation is a function with dominating mixed derivative, so the algorithm $\bar{\varphi}$ is not optimal for this class of functions. However, it performs clearly better than the standard Galerkin method also on this example.

We compared both methods on the basis of the number of information functionals (in this case Fourier coefficients) required. We could have used the number of arithmetic operations or the run–time as well, but they would not change the outcome of the comparison. The Galerkin method includes the solution of a large $(O(N^{\frac{1}{2}}) \times O(N^{\frac{1}{2}}))$ linear system, and even iterative variants could at best reach $O(N)$, which is just the run–time of $\bar{\varphi}$.

References

[EI67] K. V. Emelyanov, A. M. Ilin: On the number of arithmetic operations, necessary for the approximate solution of Fredholm integral equations of the second kind (in Russian). Zh. Vychisl. Mat. i Mat. Fiz. **7**(1967), 905–910.

[FH94] K. Frank, S. Heinrich: Complexity of local solution of integral equations. Beiträge zur Angewandten Analysis und Informatik (ed. by E. Schock), 1994.

[Fra94] K. Frank: Complexity of local solution of multivariate integral equations. to appear in J. Complexity, 1994.

[Hei93] S. Heinrich: Complexity of integral equations and relations to s–numbers. J. Complexity **9**(1993), 141–153.

[Per88] S. V. Pereverzev: On the complexity of the problem of finding solutions of Fredholm equations of the second kind with differentiable kernels I (in Russian). Ukrain. Mat. Zh. **40**(1988), 84–91.

[Per89] S. V. Pereverzev: On the complexity of the problem of finding solutions of Fredholm equations of the second kind with differentiable kernels II (in Russian). Ukrain. Mat. Zh. **41**(1989), 189–193.

[Pie78] A. Pietsch: Operator Ideals. Deutscher Verlag der Wissenschaften, Berlin 1978.

[TWW88] J. F. Traub, G. W. Wasilkowski, H. Woźniakowski: Information–Based Complexity. Academic Press, New York 1988.