Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2010.m9011

Vol. **3**, No. 4, pp. 431-448
November 2010

Fast Solvers of Fredholm Optimal Control Problems

Mario Annunziato[1] and Alfio Borzì[2,3,*]

[1] *Università degli Studi di Salerno, Dipartimento di Matematica e Informatica, Via Ponte Don Melillo, 84084 Fisciano (SA), Italia.*
[2] *Università degli Studi del Sannio, Dipartimento e Facoltà di Ingegneria, Palazzo Dell'Aquila Bosco Lucarelli, Corso Garibaldi 107, 82100 Benevento, Italia.*
[3] *Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstr. 36, 8010 Graz, Austria.*

Received 25 June 2009; Accepted (in revised version) 30 September 2009

Available online 9 September 2010

Abstract. The formulation of optimal control problems governed by Fredholm integral equations of second kind and an efficient computational framework for solving these control problems is presented. Existence and uniqueness of optimal solutions is proved. A collective Gauss-Seidel scheme and a multigrid scheme are discussed. Optimal computational performance of these iterative schemes is proved by local Fourier analysis and demonstrated by results of numerical experiments.

AMS subject classifications: 49K22, 65K10, 65R20

Key words: Optimal control theory, Fredholm integral equations of second kind, iterative methods.

1. Introduction

Fast iterative methods and optimization related to differential and integral equations are two important fields of research in applied mathematics. The purpose of optimization is to define ways of how optimally change or influence real world systems to meet a given target. This requires to realize large-scale optimization strategies with increasing complexity that in turn motivates the development of fast iterative schemes for optimization purposes.

We focus on the optimization framework provided by infinite-dimensional optimal control theory as pioneered in [12] with partial differential equations. In this framework, we consider a governing state equation, a description of the control mechanism, and a criterion defining the objective that models the purpose of the control and describes the cost of its action. An optimal control problem is then formulated as the minimization of the objective under the constraint given by the modeling equations.

*Corresponding author. *Email addresses:* mannunzi@unisa.it (M. Annunziato), alfio.borzi@unisannio.it (A. Borzì)

While optimization with partial differential equations (PDE) has received much attention [3, 9], much less is known on optimization problems with integro-differential equations and with integral equations. The purpose of this paper is to discuss the formulation of an optimal control of a system governed by a Fredholm integral equation of second kind and to present two iterative schemes that solve the corresponding optimality system with optimal computational complexity.

Also from the scientific computing point of view, the numerical solution of Fredholm integral equations is a well established mathematical field [1,7], while much less is known on efficient solution procedures of related optimal control problems. We present onegrid and multigrid iterative schemes that solve linear distributed optimal control problems governed by Fredholm integral equations. We prove mesh-independent convergence of these scheme and robustness with respect to the value of optimization parameters.

Fredholm integral equations of second kind arise naturally in the theory of signal processing [10,14] and in inverse problems [6]. They also play a main role in the modeling of thin wires antennas [16]. Optimal control of systems governed by integral equations are important in applications. In particular, we consider an optimal control problem related to the Ornstein-Uhlenbeck process that arises from statistical communication theory [10].

In the next section, a class of optimal control problems governed by Fredholm integral equations of second kind is formulated and existence and uniqueness of optimal solution is proved. In Section 3, a collective Gauss-Seidel scheme and a linear multigrid scheme are presented. The convergence properties of these iterative schemes are discussed in Section 4 in the framework of one-grid and two-grid local Fourier analysis. We show that the proposed iterative schemes are efficient and robust with respect to changes of the value of the optimization parameters. Specifically, using local Fourier analysis we obtain multigrid convergence factors that are mesh independent and these factors improve as the weight of the cost of the control becomes smaller. In Section 5, results of numerical experiments are reported that demonstrate optimal computational complexity and robustness of the proposed iterative solvers. These results appear in sharp agreement with the estimates obtained by Fourier analysis. A section of conclusion completes this work.

2. Optimal control with Fredholm integral equations of the second kind

We consider Fredholm integral equations of the second kind with linear distributed control mechanism. The purpose of the control is to determine a control function such that the resulting state $y \in L^2(\Omega)$ tracks as close as possible a desired target configuration $z \in L^2(\Omega)$ where Ω is the domain. The corresponding optimal control problem is formulated as the minimization of a cost functional J subject to the constraint given by an integral equation. We have

$$\min_{u \in L^2(\Omega)} J(y,u) := \frac{1}{2}\|y-z\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}, \tag{2.1}$$

$$y = f(y) + u + g \qquad \text{in } \Omega. \tag{2.2}$$

Here $g \in L^2(\Omega)$ is given, $u \in L^2(\Omega)$ is the control function, and the optimization parameter $\nu > 0$ is the weight of the cost of the control. The term $f(y)$ represents the integral operator and is given by

$$f(y)(x) = \int_\Omega K(x,t)\, y(t)\, dt. \tag{2.3}$$

Regarding the governing model, we assume an integral equation of the second kind where the kernel K satisfies the conditions of the Fredholm alternative theorem [11] such that existence and uniqueness of solution for a given u is guaranteed. In particular, we consider a symmetric integral operator $f(\cdot) = f^T(\cdot)$, where $f^T(y)(x) := \int_\Omega K(t,x)\, y(t)\, dt$, and we require that

$$\|K\|_{L^2(\Omega\times\Omega)} = \iint_{\Omega\times\Omega} |K(x,t)|^2\, dx\, dt < 1. \tag{2.4}$$

This condition on itself is sufficient to prove existence and uniqueness of solution [11] and it can be easily verified in application.

Theorem 2.1. *Let K be such that existence and uniqueness of solution for* (2.2) *is guaranteed and define the gradient $\nabla\hat{J}(u) := \nu u - p$, where p is the solution to the integral equation $p = f^T(p) - y + z$ in Ω. Then the control problem* (2.1)-(2.2) *has a unique solution in $L^2(\Omega)$ if and only if $\nabla\hat{J}(u) = 0$. Therefore, the optimal solution is characterized as the solution of the following first-order optimality system*

$$\begin{cases} y - f(y) - u = g, \\ p - f^T(p) + y = z, \\ \nu u - p = 0. \end{cases} \tag{2.5}$$

Proof. Since the integral equation (2.2) has a unique solution y for a given u, we denote this dependence by $y = y(u)$. Therefore, to discuss existence and characterization of the unique solution to (2.1)-(2.2), we can introduce the so-called reduced cost functional $\hat{J}$ [9,12] given by

$$\hat{J}(u) = J(y(u), u). \tag{2.6}$$

In fact, the optimal solution corresponds to the unique minimizer of $\hat{J}(u)$. Notice that the mapping $u \to y(u)$ from L^2 to L^2 is affine and continuous. Let us denote its first derivative at u in the direction δu by $y'(u)\delta u$. It is characterized as the solution to

$$y'(u)\delta u = f(y'(u)\delta u) + \delta u \quad \text{in } \Omega. \tag{2.7}$$

The second derivative of $u \to y(u)$ is zero, due to the linear dependence on u in (2.2). Hence from (2.1), we find for the second derivative of $u \to \hat{J}(u)$

$$\hat{J}''(u)(\delta u, \delta u) = \|y'(u)\delta u\|^2_{L^2(\Omega)} + \nu\, \|\delta u\|^2_{L^2(\Omega)},$$

and thus $u \to \hat{J}(u)$ is uniformly convex. This implies existence of a unique solution u^* to (2.1). Moreover, the minimum is characterized by $\hat{J}'(u^*;\delta u) = 0$ for all δu and consequently

$$\hat{J}'(u^*;\delta u) = (y^* - z, y'(u^*)\delta u)_{L^2(\Omega)} + \nu(u^*,\delta u)_{L^2(\Omega)} = 0, \quad \text{for all } \delta u \in L^2,$$

where $y^* = y(u^*)$. To remove y' from this equation, we introduce $p^* \in L^2$ as the solution to

$$p^* = f^T(p^*) - (y^* - z) \quad \text{in } \Omega, \tag{2.8}$$

that is unique under the same condition on K required for the state equation. Then by (2.7) and (2.8), and using the fact that

$$(f^T(p^*), y'(u^*)\delta u) = (p^*, f(y'(u^*)\delta u)),$$

we have

$$\begin{aligned} \hat{J}'(u^*;\delta u) &= -(p^*,\delta u)_{L^2(\Omega)} + \nu(u^*,\delta u)_{L^2(\Omega)} \\ &= (\nu u^* - p^*, \delta u)_{L^2(\Omega)} = 0, \end{aligned} \tag{2.9}$$

for all $\delta u \in L^2$. Thus, we have obtained the gradient of the reduced cost functional,

$$\nabla \hat{J}(u^*) := \nu u^* - p^*,$$

and formally $\nabla \hat{J}(u^*) = 0$ constitutes the necessary and sufficient (because of convexity) optimality condition for (2.1)-(2.2). Notice that p^* depends on u^* via (2.2) and (2.8). From (2.2), (2.8), and (2.9), we obtain the optimality system (2.5). □

In the present case of optimal control problems without constraints on the control u, we have the scalar equation $\nu u - p = 0$. Thus, we can replace $u = p/\nu$ in the state equation and obtain the following equivalent system

$$\begin{cases} y - f(y) - p/\nu = g, \\ p - f(p) + y = z. \end{cases} \tag{2.10}$$

Notice that system (2.10) corresponds to two coupled integral equations that can be recasted as a unique integral equation system, as follows

$$\begin{pmatrix} 1 & -1/\nu \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y \\ p \end{pmatrix} = \begin{pmatrix} f(y) + g \\ f^T(p) + z \end{pmatrix},$$

that is,

$$\begin{pmatrix} y \\ p \end{pmatrix} = \frac{1}{1+\nu} \begin{pmatrix} \nu & 1 \\ -\nu & \nu \end{pmatrix} \begin{pmatrix} f(y) + g \\ f^T(p) + z \end{pmatrix}. \tag{2.11}$$

The advantage of this formulation is the possibility to prove existence and uniqueness of solution to (2.10) using condition (2.4). Notice that the coefficient matrix is never singular for $\nu > 0$. Following [11], we have that, in general, the integral system

$$\phi_i(x) = \lambda \sum_{j=1}^{n} \int_{\Omega} \tilde{K}_{ij}(x,t)\phi_j(t)\,dt + g_i(x), \qquad i = 1, \cdots, n,$$

where $\Omega = (a,b)$, $g_i \in L^2(\Omega)$, and $\tilde{K} \in L^2(\Omega \times \Omega)$, has a unique solution $\phi_i \in L^2(\Omega)$ provided that $|\lambda| < 1/C$, where

$$C^2 = \sum_{i,j=1}^{n} \int\int_{\Omega\times\Omega} |\tilde{K}_{ij}(x,t)|^2\,dx\,dt.$$

In our case, we have

$$C^2 = \frac{1}{(1+\nu)^2} \int\int_{\Omega\times\Omega} (2\nu^2|K(x,t)|^2 + (1+\nu^2)|K(x,t)|^2)\,dx\,dt.$$

Therefore, applying (2.4) to the optimality system, we find that this system admits a unique solution provided that

$$\int\int_{\Omega\times\Omega} |K(x,t)|^2\,dx\,dt < \frac{(1+\nu)^2}{1+3\nu^2}, \tag{2.12}$$

that is less restrictive than (2.4) when ν is sufficiently small. This result shows that controlled solutions may exist under weaker conditions than those required for the uncontrolled problem. A similar result is obtained in [2], in the case of singular elliptic control problems.

3. Discretization and fast iterative schemes

In this section, we discuss the discretization of the Fredholm optimality system by the Nyström method. To solve the resulting discretized problem, we present two iterative schemes which result to be efficient and robust with respect to changes of value of the optimization parameter.

To approximate the optimality system (2.10), we consider the discretization of the problem on a finite difference grid and use direct quadrature (DQ) with the Nyström method [1,7]. We take $x \in \Omega = (-D, D)$ and set the grid points

$$\Omega_h := \{x_i = ih,\ i = -N, -N+1, \cdots, 0, 1, \cdots, N-1, N\},$$

where $h = D/N$. On this grid, we consider the following semi-discrete version of equation (2.10), that is,

$$\begin{cases} y_N(x) - f_N(y_N)(x) - p_N(x)/\nu = g(x), \\ p_N(x) - f_N(p_N)(x) + y_N(x) = z(x), \end{cases} \tag{3.1}$$

where $x \in \Omega$. Here $y_N(x), p_N(x)$ are approximation to the solutions $y(x), p(x)$, for $x \in \Omega$, and $f_N(y_N)$ is the approximation to the integral with a direct quadrature:

$$f(y)(x) \approx f_N(y_N)(x) = h \sum_{j=-N}^{N} v_j K(x, t_j)\, y_N(t_j),$$

where the v_j are the weights of a DQ of order q.

Following the Nyström method, we fully discretize (3.1), by setting $y_i = y_N(x_i)$, $p_i = p_N(x_i)$, having suppressed the evidence of N for an easier notation. Thus, we obtain the following discrete optimality system

$$\begin{cases} y_i - h\sum_{j=-N}^{N} w_{ij} y_j - p_i/\nu = g_i, \\ p_i - h\sum_{j=-N}^{N} w_{ij} p_j + y_i = z_i, \end{cases} \tag{3.2}$$

where $i = -N, \cdots, N$ and $w_{ij} = v_j K(x_i, t_j)$, with v_j given by the quadrature rule. In the following, we denote with $y_h = \{y_{-N}, \cdots, y_N\}$ and $p_h = \{p_{-N}, \cdots, p_N\}$. The solution of (3.2) gives us the approximate solution (y_h, p_h) of (2.10) at the mesh points. Assuming that $K(x,t)\,y(t)$ is q-time continuously differentiable in t, and uniformly differentiable in x [1, 7], and assuming g be continuous, we obtain that the associate solution error is bounded by

$$\|y - y_h\|_\infty + \|p - p_h\|_\infty \leq \mathcal{O}(h^q).$$

In a semi-discrete setting, approximation formulas for $y_N(x)$ and $p_N(x)$, $x \in \Omega$ can be found from (3.1), i.e. by the Nyström interpolation formula

$$\begin{cases} y_N = \dfrac{1}{1+\nu}\left[z + f_N(p_N) + \nu(g + f_N(y_N))\right], \\ p_N = \dfrac{\nu}{1+\nu}\left[z + f_N(p_N) - (g + f_N(y_N))\right]. \end{cases} \tag{3.3}$$

Under the same regularity condition on the kernel as given above, uniqueness of the solution of (3.2) results from the uniqueness of solution to (3.3). For these functions, we obtain a convergence order

$$\|y - y_N\|_\infty + \|p - p_N\|_\infty \leq \mathcal{O}(h^q).$$

For an efficient and robust solution of the discretized optimality system (3.2) we discuss two iterative methods: a onegrid Gauss-Seidel scheme and a multigrid scheme.

First, we define an iterative procedure on Ω_h, that belongs to the class of collective Gauss-Seidel schemes. It results from a sequential update of the optimization variables at each grid point. The update of the variables (y_i, p_i) at grid point x_i is obtained by solving exactly the discrete optimality system with respect to (y_i, p_i) and considering the remaining variables as constant. Our iterative method is given by the following algorithm. Let an initial approximation $(y^{(0)}, p^{(0)})$ be given. Here, tol is the required tolerance on the

L^2-norm of the residual of the constraint equation. We have the following iterative scheme denoted by S_h.

Algorithm 3.1: Collective Gauss-Seidel (CGS) Iteration Scheme

1. For $m = 0, 1, 2, \cdots, \ell$ do
2. If $\|g_h - y_h^m + f_h(y_h^m) - p_h^m/\nu\|_{L_h^2} < tol$ then stop.
3. For $i = -N, -N+1, \cdots, 0, 1, \cdots, N-1, N$ (lexicographic order) do

$$\begin{pmatrix} y_i \\ p_i \end{pmatrix}^{(m+1)} = \begin{pmatrix} 1 - hw_{ii} & -1/\nu \\ 1 & 1 - hw_{ii} \end{pmatrix}^{-1} \times \left[\begin{pmatrix} g_i \\ z_i \end{pmatrix} + h \sum_{j<i} w_{ij} \begin{pmatrix} y_j \\ p_j \end{pmatrix}^{(m+1)} + h \sum_{j>i} w_{ij} \begin{pmatrix} y_j \\ p_j \end{pmatrix}^{(m)} \right]. \tag{3.4}$$

4. End.

Next, in view of an extension of this work to integro-differential control problems and thus envisioning the need to accelerate the CGS scheme, we discuss the embedding of the CGS scheme in a multigrid framework [15]. For the multigrid formulation, we consider multiple nested grids. Let us index the operators and variables defined on the grid Ω_k with mesh size $h = h_k = h_0/2^k$, $k = 1, \cdots, L$. We now illustrate a linear multigrid scheme [5, 8].

In general, an initial approximation to the solution of the optimality system will differ from the exact solution because of errors involving high-frequency as well as low-frequency components. In order to solve for all frequency components of the error, the multigrid strategy combines two complementary schemes [5, 8]. The high-frequency components of the error are reduced by smoothing iterations while the low-frequency error components are effectively reduced by a coarse-grid correction method.

On the grid of level k with $h = h_k$, the smoothing procedure is denoted by S_k, and S_k^m is the smoothing operator applied m times on the pair $v_k = (y_k, p_k)$. We choose S_k to be the CGS iteration on Ω_k, given by Algorithm 3.1 applied at level k. Later we prove that this iteration has good smoothing properties. To correct for the smooth component of the error, a coarse grid correction (CGC) is defined. To illustrate the CGC scheme, we exploit the fact that the optimality system is linear.

For simplicity of exposition, consider the case of two levels with a fine grid level with mesh size $h = h_k$ and the coarse grid problem is constructed on the grid with mesh size $H = h_{k-1}$. Recall the optimality system

$$y_h - f_h(y_h) - p_h/\nu = g_h, \tag{3.5}$$

$$p_h - f_h^T(p_h) + y_h = z_h. \tag{3.6}$$

Let us denote (3.5)-(3.6) with $A_h v_h = F_h$, where $v_h = (y_h, p_h)$ and $F_h = (g_h, z_h)$. Denote with $v_h^{(\ell)}$ an approximate solution to the discrete problem, obtained after ℓ sweeps of an

iterative scheme. In correspondence, we define the solution error $e_h^{(\ell)} = v_h - v_h^{(\ell)}$ and the corresponding residual $r_h^{(\ell)} = F_h - A_h v_h^{(\ell)}$. Thus the solution of the problem $A_h v_h = F_h$ is equivalent to solving $A_h e_h^{(\ell)} = r_h^{(\ell)}$ and then $v_h = v_h^{(\ell)} + e_h^{(\ell)}$. This equivalence is used to define the coarse problem in a linear multigrid scheme as we discuss next.

We formulate the coarse-grid problem, $A_H v_H = F_H$, where $v_H = (y_H, p_H)$ should denote the solution errors of the approximation $v_h = (y_h, p_h)$ represented on the coarse grid. Because of the equivalence between the linear problem and its formulation in terms of solution errors and residuals, we set F_H as the residual of the optimality system represented on the coarse grid. Therefore, the coarse-grid problem is given by

$$y_H - f_H(y_H) - p_H/\nu = I_h^H r_h^y, \tag{3.7}$$
$$p_H - f_H^T(p_H) + y_H = I_h^H r_h^p. \tag{3.8}$$

The operator $I_h^H : L_h^2 \to L_H^2$ denotes a restriction operator. We chose full weighting such that

$$r_H(x_j) = (r_h(x_{j-1}) + 2\, r_h(x_j) + r_h(x_{j+1}))/4, \qquad j = 2, 4, \cdots, n-1.$$

In stencil form we have

$$I_h^H = \frac{1}{4} \begin{pmatrix} 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 1 \end{pmatrix}.$$

The right-hand side terms r_h^y and r_h^p are the residuals defined by

$$r_h^y = g_h - (y_h - f_h(y_h) - p_h/\nu), \tag{3.9}$$
$$r_h^p = z_h - (p_h - f_h^T(p_h) + y_h). \tag{3.10}$$

Once the coarse grid problem is solved, the coarse grid correction follows

$$y_h^{new} = y_h + I_H^h\, y_H, \tag{3.11}$$
$$p_h^{new} = p_h + I_H^h\, p_H, \tag{3.12}$$

where $I_H^h : L_H^2 \to L_h^2$ represents an interpolation operator. We take I_H^h to be the piecewise linear interpolation given by [8]

$$I_H^h = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

If the high frequency components of the error on the finer grid are indeed well damped, then the grid Ω_H should provide enough resolution for the error of the solution and hence v_H should be a good approximation to this error. This idea of transferring the problem to be solved to a coarser grid can be applied along the set of nested meshes. One starts at level L with a zero approximation and applies the smoothing iteration m_1 times. Then the problem is transferred down to a coarser grid and so on. Once the coarsest grid is reached, one solves the coarsest problem to convergence by applying, as we do, a few steps of the CGS smoothing iteration. The solution obtained on each grid is then used to correct the approximation on the next finer grid. The coarse grid correction followed by m_2 post-smoothing steps is applied from one grid to the next, up to the finest grid with level L. This entire process represents one V-(m_1, m_2) multigrid cycle.

The multigrid V-(m_1, m_2)-cycle algorithm, expressed in terms of the multigrid iteration operator B_k in recursive form applied to $A_k\, v_k = F_k$ is uniquely defined as follows [4].

Algorithm 3.2: Multigrid V-(m_1, m_2)-Cycle

Set $B_1 = A_1^{-1}$. For $k = 2, \cdots, L$ define B_k in terms of B_{k-1} as follows. Let $q^0 = 0$.

1. Set the starting approximation $v_k^{(0)}$.

2. Pre-smoothing (Algorithm 3.1). Define $v_k^{(l)}$ for $l = 1, \cdots, m_1$, by

$$v_k^{(l)} = S_k(v_k^{(l-1)}, F_k).$$

3. Coarse grid correction. Set $v_k^{(m_1+1)} = v_k^{(m_1)} + I_{k-1}^k\, q^m$ where q^i for $i = 1, \cdots, m$ is defined by

$$q^i = q^{i-1} + B_{k-1}\left[I_k^{k-1}(F_k - A_k(v_k^{(m_1)})) - A_{k-1} q^{i-1}\right].$$

4. Post-smoothing (Algorithm 3.1). Define $v_k^{(l)}$ for $l = m_1 + 2, \cdots, m_1 + m_2 + 1$, by

$$v_k^{(l)} = S_k(v_k^{(l-1)}, F_k).$$

5. Set $B_k\, F_k = v_k^{(m_1+m_2+1)}$.

Notice that we can perform m two-grid iterations at each working level. For $m = 1$ we have a V-cycle and for $m = 2$ we have a W-cycle; m is called the cycle index [15]. A stopping criteria is implemented as in Algorithm 3.1, at the end of each V-cycle on the finest level.

In the next section, we investigate the convergence properties of the CGS scheme and of the V-cycle multigrid scheme.

4. Local Fourier analysis

In this section, we investigate the convergence properties of the CGS scheme and of the V-cycle scheme using Fourier analysis [5, 15]. This is an effective tool for analyzing iterative processes where we assume the problem defined on the infinite grid $G^h = \{jh,\, j \in \mathbb{Z}\}$. Notice that on G^h, only the components (called harmonics) $\phi(x,\theta) = e^{i\theta x/h}$ with $\theta \in (-\pi, \pi]$ are visible, i.e. there is no other component (no aliasing) with frequency $\theta_0 \in (-\pi, \pi]$ with $|\theta_0| < \theta$ such that $e^{i\theta_0 x/h} = e^{i\theta x/h}$, $x \in G^h$.

In local Fourier analysis the notion of low- and high-frequency components on the grid G^h is related to a coarser grid denoted by G^H. In this way $e^{i\theta x/h}$ on G^h is said to be an high-frequency component, with respect to the coarse grid G^H, if its restriction (projection) to G^H is not visible there. If $H = 2h$ then the high frequencies are those with $\frac{\pi}{2} \le |\theta| \le \pi$. We have $e^{i\theta x/h} = e^{i(2\theta)x/H}$.

In this framework, in order to analyze a given iteration we represent solution errors in terms of their θ components $e^{(m)}(x) = \sum_\theta E_\theta^{(m)} e^{i\theta x/h}$ and $e^{(m+1)}(x) = \sum_\theta E_\theta^{(m+1)} e^{i\theta x/h}$ (with formal summation on θ), where $E_\theta^{(m)}$ and $E_\theta^{(m+1)}$ denote the error amplitudes of the θ component, before and after smoothing, respectively. The action of an iteration operator M is $e^{(m+1)} = M\, e^{(m)}$. In the Fourier space this action is represented by $E_\theta^{(m+1)} = \hat{M}(\theta)\, E_\theta^{(m)}$, and $\hat{M}(\theta)$ is the Fourier symbol of M.

Let $(e_y(j), e_p(j)) = \sum_\theta W_\theta\, \phi(j,\theta)$ denotes the errors for the state and the adjoint variables on the grid points $x_j = jh$. Here, $W_\theta = (Y_\theta, P_\theta)$ are the corresponding Fourier coefficients. The action of one smoothing step on the errors can be expressed by $W_\theta^{(1)} = \hat{S}(\theta)\, W_\theta^{(0)}$.

Now, consider applying the CGS step for solving our distributed control problem. We assume that the kernel is symmetric, i.e. $w_{i-j} = w_{|i-j|}$ and that it is decaying sufficiently fast such that we can truncate the sum approximating the integral,

$$\sum_{k=-\ell}^{\ell} \phi(k) \approx \sum_{k=-N}^{N} \phi(k).$$

Substituting $(e_y(j), e_p(j))$ in (3.2) and applying the CGS Algorithm 3.1, we obtain

$$\begin{aligned}
&\begin{pmatrix} \left(1 - h\sum_{k=-\ell}^{0} w_{|k|} e^{i\theta k}\right) & -\dfrac{1}{\nu} \\ 1 & \left(1 - h\sum_{k=-\ell}^{0} w_{|k|} e^{i\theta k}\right) \end{pmatrix} \begin{pmatrix} Y_\theta^{(1)} \\ P_\theta^{(1)} \end{pmatrix} \\
&= \begin{pmatrix} h\sum_{k=1}^{\ell} w_{|k|} e^{i\theta k} & 0 \\ 0 & h\sum_{k=1}^{\ell} w_{|k|} e^{i\theta k} \end{pmatrix} \begin{pmatrix} Y_\theta^{(0)} \\ P_\theta^{(0)} \end{pmatrix}.
\end{aligned}$$

Hence

$$\hat{S}(\theta)=\begin{pmatrix}\left(1-h\sum_{k=-\ell}^{0}w_{|k|}e^{i\theta k}\right) & -\frac{1}{\nu}\\ 1 & \left(1-h\sum_{k=-\ell}^{0}w_{|k|}e^{i\theta k}\right)\end{pmatrix}^{-1}\times\begin{pmatrix}h\sum_{k=1}^{\ell}w_{|k|}e^{i\theta k} & 0\\ 0 & h\sum_{k=1}^{\ell}w_{|k|}e^{i\theta k}\end{pmatrix}. \tag{4.1}$$

We consider the entire frequency domain spanned by the two sets of frequencies $\theta\in[-\pi/2,\pi/2)$ and define

$$\overline{\theta}:=\theta-sign(\theta)\pi.$$

Here θ represents low frequencies components while $\overline{\theta}$ contains the high frequencies components. This choice results in a basis of the two harmonics with low- and high-frequencies, $e^{i\theta x/h}$ and $e^{i\overline{\theta}x/h}$ respectively. In this framework, a way to characterize the smoothing property of the smoothing operator S is to consider its action on both sets of frequencies as follows

$$\widehat{S}(\theta)=\begin{bmatrix}\hat{S}(\theta) & 0\\ 0 & \hat{S}(\overline{\theta})\end{bmatrix}$$

and to assume an ideal coarse grid correction which annihilates the low frequency error components and leaves the high frequency error components unchanged. That is, one defines the projection operator on the high-frequency harmonics as follows

$$\widehat{Q}(\theta)=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$$

In this framework, the smoothing property μ of S is defined as follows

$$\mu=\max\left\{r(\widehat{Q}(\theta)\widehat{S}(\theta)):\ \theta\in[-\pi/2,\pi/2)\right\}, \tag{4.2}$$

where $r(\cdot)$ is the spectral radius. This number provides a measure of the ability of the iterative scheme to damp the high-frequency components of the solution error.

In Fig. 1, we depict the smoothing factor of the CGS scheme as a function of ν and h. It appears that μ is almost independent of the value of the discretization parameter h and increases by increasing the value of the weight ν. The CGS scheme is not only a good smoother, it provides also good convergence properties for all frequencies. In Table 1, we report estimates of $\rho(S_h)$ resulting from the local Fourier convergence analysis. We see that the convergence of the CGS scheme is robust with respect to ν and is almost mesh independent.

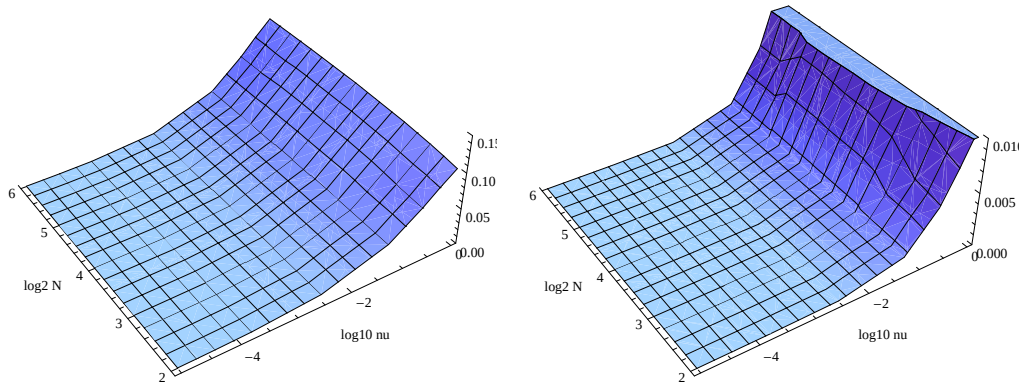

Figure 1: Smoothing (left) and twogrid convergence factors as functions of ν and N.

Table 1: Estimates for $\rho(S_h)$ with $w_{|i-j|} = -\frac{1}{2}\exp(-|i-j|h)$.

ν	$N=64$	$N=128$	$N=256$	$N=512$
1.0e-02	1.93e-02	1.94e-02	1.94e-02	1.95e-02
1.0e-03	6.15e-03	6.18e-03	6.20e-03	6.20e-03
1.0e-04	1.94e-03	1.95e-03	1.96e-03	1.96e-03
1.0e-05	6.15e-04	6.19e-04	6.20e-04	6.21e-04
1.0e-06	1.94e-04	1.95e-04	1.96e-04	1.96e-04

Next, we discuss twolevel Fourier analysis [5] to estimate the convergence of the multigrid iteration. For this purpose, we need to construct the Fourier symbol of the twolevel coarse-grid correction operator corresponding to (3.7)-(3.12). We have

$$CG_h^H = \left[I_h - I_H^h (A_H)^{-1} I_h^H A_h \right].$$

Notice that this operator refers to the entire optimality system. Therefore, the intergrid transfer operators are two copies of the transfer operators I_H^h and I_h^H, respectively.

We denote the Fourier symbol of CG_h^H as follows

$$\widehat{CG}_h^H(\theta) = \left[\hat{I}_h - \hat{I}_H^h(\theta)(\widehat{A}_H(2\theta))^{-1} \hat{I}_h^H(\theta)\widehat{A}_h(\theta) \right].$$

(Recall that $e^{i\theta x/h} = e^{i(2\theta)x/H}$.) The symbol of the full-weighting restriction operator for all frequency components is

$$\hat{I}_h^H(\theta) = \frac{1}{2}\begin{bmatrix} (1+\cos(\theta)) & 0 & (1-\cos(\theta)) & 0 \\ 0 & (1+\cos(\theta)) & 0 & (1-\cos(\theta)) \end{bmatrix}.$$

For the linear prolongation operator we have $\hat{I}_H^h(\theta) = \hat{I}_h^H(\theta)^T$.

Consider the optimality system (3.5)-(3.6). We obtain that the symbol of the fine grid

operator is

$$\widehat{A}_h(\theta) = \begin{bmatrix} a_y^h(\theta) & -1/\nu & 0 & 0 \\ 1 & a_p^h(\theta) & 0 & 0 \\ 0 & 0 & a_y^h(\bar{\theta}) & -1/\nu \\ 0 & 0 & 1 & a_p^h(\bar{\theta}) \end{bmatrix},$$

where

$$a_y^h(\theta) = 1 - h\sum_{k=-\ell}^{\ell} w_{|k|} e^{i\theta k}, \qquad a_p^h(\theta) = a_y^h(\theta).$$

Similarly, for the frequency represented on the coarse grid, the symbol of the coarse grid operator follows

$$\widehat{A}_H(\theta) = \begin{bmatrix} a_y^H(2\theta) & -1/\nu \\ 1 & a_p^H(2\theta) \end{bmatrix}.$$

The symbol of the twolevel method is given by

$$\widehat{TG}_h^H(\theta) = \hat{S}_h(\theta)^{m_2}\, \widehat{CG}_h^H(\theta)\, \hat{S}_h(\theta)^{m_1}.$$

This is an 4×4 matrix corresponding to the two frequency components. The estimate convergence factor is obtained as follows

$$\rho(TG_h^H) = \sup\left\{ r(\widehat{TG}_h^H(\theta)) : \theta \in [-\pi/2, \pi/2) \right\}.$$

In Table 2, we report theoretical estimates of $\rho(TG_h^H)$ resulting from the twogrid convergence analysis. Comparison with results of numerical experiments show that these estimates are sharp; see Table 5 for results of numerical experiments. In Fig. 1, we also plot the twogrid convergence factor as function of ν and N. We can see that the local Fourier analysis predicts mesh-independent smoothing factors and convergence factors and these factors improve as ν becomes smaller. Therefore our multigrid approach provides an iterative scheme with optimal complexity and robustness with respect to the weight of the cost of the control. These results are in agreement with results of numerical experiments.

Table 2: Estimates for $\rho(TG_h^H)$ for the case of $m_1 = m_2 = 1$ smoothing steps; $w_{|i-j|} = -\frac{1}{2}\exp(-|i-j|h)$.

ν	$N = 64$	$N = 128$	$N = 256$	$N = 512$
1.0e-01	3.31e-03	3.35e-03	3.36e-03	3.37e-03
1.0e-02	3.74e-04	3.77e-04	3.79e-04	3.80e-04
1.0e-03	3.78e-05	3.82e-05	3.84e-05	3.85e-05
1.0e-04	3.79e-06	3.83e-06	3.85e-06	3.86e-06
1.0e-05	3.79e-07	3.83e-07	3.85e-07	3.86e-07
1.0e-06	3.80e-08	3.83e-08	3.85e-08	3.86e-08

5. Numerical experiments

In order to validate the optimal control formulation and to test the proposed numerical solution procedures, we present results of numerical experiments with two test problems by assigning the kernel $K(x,t)$, the desired target function $z(x)$, and the known term $g(x)$. For the numerical results, we report solution errors, norms of residuals, and observed convergence rates for different mesh sizes N and different values of the weight ν. The residual norm we use is the discrete L^2-norm denoted with $\|\cdot\|_2$. With res we mean the residual of the integral state equation. For convergence, we set a tolerance $tol = 10^{-12}$ on $\|res\|_2$. The observed rate of convergence ρ is the mean ratio of reduction of the norm of two consecutive residual.

To compare the computational cost between CGS and MG, we use the Work Unit (WU) as the time to execute a single CGS iteration, i.e. a single matrix-vector product. So that the WU cost for CGS is just the number of iteration N_{iter}. For MG we have one WU of cost for one smoothing sweep at the finest level. To count the total WU cost of a MG cycle, we have $m_1 + m_2$ WUs on the finest level and this value reduces by a factor of 4 in the next coarse grid. In total for L levels we have

$$\text{Total MG cost } = (m_1+m_2)\sum_{k=0}^{L} 4^{-k} = (m_1+m_2)(4-4^{-L})/3 \text{ WU},$$

where $L = \log_2(2N)$ is the number of levels.

Test case 1

Our first numerical experiment is to validate the approximation property of the discretization scheme. For this purpose, we consider an operator

$$K(x,t) = \cos(\pi x)\cos(\pi t), \quad g(x) = -\cos(\pi x), \quad z(x) = \cos(\pi x),$$

such that we can compute the exact optimal solution. This is given by

$$y(x) = \cos(\pi x), \quad p(x) = \nu\cos(\pi x), \quad \text{for } \; x \in \Omega = (-1,1).$$

The norm of the kernel for this case is $\|K\|_{L^2(\Omega\times\Omega)} = 1$ however it satisfies condition (2.12) for $\nu < 1$.

In Table 3, we report results corresponding to a required residual tolerance of 10^{-12}. The solution is in agreement with y_{exact}. In Fig. 2, we see the solution for $\nu = 10^{-3}$, $N = 8$, plotted by using the Nyström formula on the interpolation points; the vector values y_h are pointed out with circles. On the same figure, the error between computed and exact solution is plotted versus the parameter ν for constant mesh and iterations. We see that the error reaches the machine precision quickly as ν decrease.

Table 3: Case 1. Results of numerical experiments with exact solution: $y_{exact}(x) = \cos(\pi x)$ and $u_{exact}(x) = \cos(\pi x)$.

ν	N	$\|y - y_{exact}\|$	$\|u - u_{exact}\|$	$\|\|res\|\|_2$
1e-01	64	8.17e-014	1.54e-013	3.37e-014
1e-01	128	8.93e-014	1.99e-013	4.38e-014
1e-01	256	9.29e-014	2.25e-013	5.05e-014
1e-01	512	9.47e-014	2.39e-013	5.43e-014
1e-03	64	3.77e-015	1.55e-013	4.07e-014
1e-03	128	4.11e-015	1.80e-013	4.64e-014
1e-03	256	4.33e-015	1.95e-013	4.96e-014
1e-03	512	4.33e-015	2.03e-013	5.13e-014
1e-05	64	1.11e-016	7.11e-013	2.22e-013
1e-05	128	2.22e-016	7.70e-013	2.36e-013
1e-05	256	2.22e-016	8.02e-013	2.44e-013
1e-05	512	2.22e-016	8.18e-013	2.48e-013
1e-07	64	3.33e-016	8.88e-016	4.32e-016
1e-07	128	4.44e-016	1.33e-015	5.59e-016
1e-07	256	4.44e-016	2.55e-015	7.62e-016
1e-07	512	4.44e-016	3.00e-015	1.05e-015

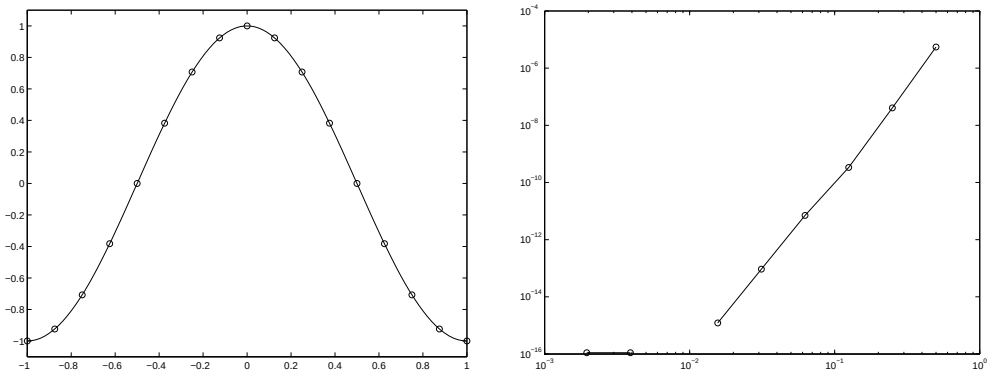

Figure 2: Left: Optimal state solution for $\nu = 10^{-3}$, $N = 8$, with the Nyström formula. Circles are the interpolation points. Right: convergence history of $\|y_N - y_{exact}\|_\infty$ for $\nu = 2^{-n}, N = 16$ for $n = 1 \cdots 10$.

Test case 2

In this test series, we investigate the convergence performance of the CGS and MG iteration. In this case we consider an application corresponding to a kernel that represents the covariance function of an Ornstein-Uhlenbeck stochastic process at the equilibrium, that arises in statistical communication theory [10, 14]. We have $K(x,t) = -e^{-\alpha|x-t|}/2$, where $\alpha > 0$ represents the characteristic correlation time of the process. Here, $y(x)$ represents a signal and $u(x)$ a control for the signal. In this case, the norm of the kernel $K(x,t)$ as defined in (2.12) is approximately equal to 0.142, so that we can state existence and uniqueness of solution for a given control.

In addition, we take $g(x) = 2/\pi, u(x) = \sin(\pi x)$ and a target function which is discon-

Table 4: Case 2. Results with the CGS scheme; $\alpha = 1$.

ν	N	$\|y-z\|$	$\|\|res\|\|_2$	ρ	N_{iter}
1e-03	512	6.66e-04	9.03e-015	7.02e-04	6
1e-03	256	6.66e-04	8.89e-015	7.01e-04	6
1e-03	128	6.66e-04	8.60e-015	6.99e-04	6
1e-04	512	6.67e-05	1.32e-013	4.10e-05	5
1e-04	256	6.67e-05	1.30e-013	4.09e-05	5
1e-04	128	6.66e-05	1.26e-013	4.06e-05	5
1e-05	512	6.67e-06	8.33e-014	4.10e-06	4
1e-05	256	6.67e-06	8.27e-014	4.09e-06	4
1e-05	128	6.66e-06	8.12e-014	4.06e-06	4
1e-06	512	6.67e-07	8.56e-016	4.21e-07	4
1e-06	256	6.67e-07	8.32e-016	4.12e-07	4
1e-06	128	6.66e-07	8.08e-016	4.04e-07	4

Table 5: Case 2. Results with the multigrid scheme; $\alpha = 1$ and $m_1 = 1, m_2 = 1$ pre- and post-smoothing sweeps.

ν	N	$\|y-z\|$	$\|\|res\|\|_2$	ρ	N_{cycle}
1e-03	512	6.66e-04	3.74e-016	1.07e-05	3
1e-03	256	6.66e-04	3.65e-016	1.04e-05	3
1e-03	128	6.66e-04	3.48e-016	1.00e-05	3
1e-04	512	6.67e-05	3.52e-013	1.04e-06	2
1e-04	256	6.67e-05	3.51e-013	1.03e-06	2
1e-04	128	6.66e-05	3.48e-013	1.03e-06	2
1e-05	512	6.67e-06	3.52e-015	1.03e-07	2
1e-05	256	6.67e-06	3.52e-015	1.03e-07	2
1e-05	128	6.66e-06	3.48e-015	1.02e-07	2
1e-06	512	6.67e-07	1.62e-016	4.76e-08	2
1e-06	256	6.67e-07	1.27e-016	3.75e-08	2
1e-06	128	6.66e-07	1.08e-016	3.18e-08	2

tinuous as typical in the modeling of signals. We choose

$$z(x) = \lfloor 5[x(x-1)/\pi + (1+1/\pi^2)\sin(\pi x)]\rfloor/5,$$

where $\lfloor\ \rfloor$ is the floor function.

In Table 4, results obtained with the CGS iteration are reported. We see that tracking improves as ν decreases. We obtain robust tracking despite z is discontinuous. Next, we consider the same setting and apply the multigrid scheme given by Algorithm 3.2. Results with the multigrid scheme are reported in Table 5. We obtain that both the CGS and the MG scheme converge efficiently to the solution and the observed convergence rate is weakly dependent on the mesh size. As ν becomes smaller, convergence rates improve thus showing robustness. We see that multigrid convergence rates are two order of magnitude better than the CGS convergence rates. Moreover, the observed rates are in good agreement with the corresponding estimates by local Fourier analysis. In Fig. 3, we depict the computed optimal state and control solutions.

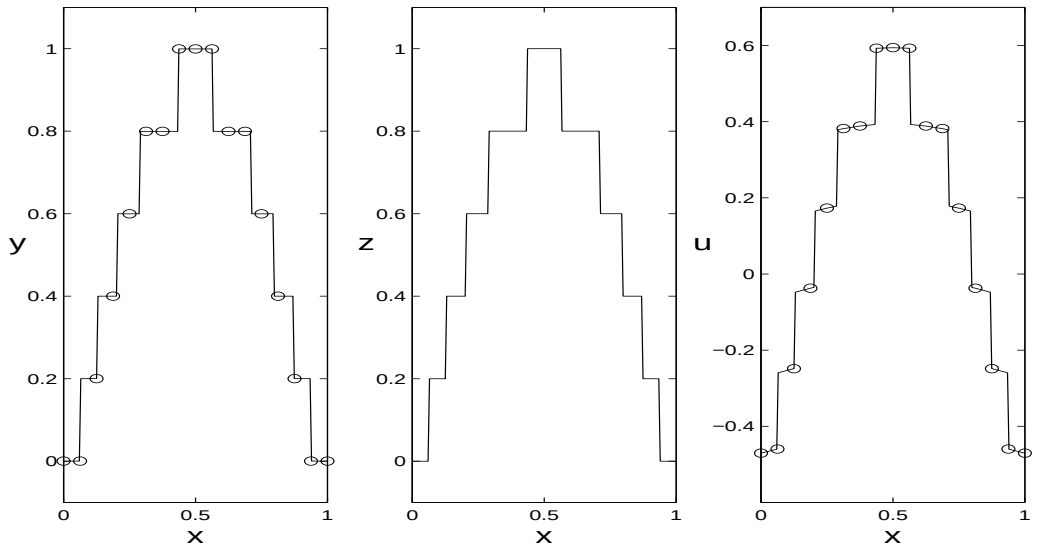

Figure 3: Case 2. Optimal solution y_N (left), target z, and control u_N (right) for $\alpha = 1$, $\nu = 10^{-3}$, $N = 8$, with the Nyström formula. Circles are the interpolation points.

6. Conclusion

The formulation of optimal control problems governed by Fredholm integral equations of second kind was presented, proving existence and uniqueness of optimal solutions. To solve the corresponding optimality systems, an efficient computational framework was presented based on a collective Gauss-Seidel scheme and a multigrid scheme. Optimal computational performance and robustness of these iterative schemes was proved by local Fourier analysis and demonstrated by results of numerical experiments.

References

[1] K.E. Atkinson, *The Numerical Solution of Integral Equations of the Second Kind*, Cambridge Monographs on Applied and Computational Mathematics, 2009.

[2] A. Borzì and K. Kunisch, *The numerical solution of the steady state solid fuel ignition model and its optimal control*, SIAM J. Sci. Comp., 22 (2000), pp. 263–284.

[3] A. Borzì and V. Schulz, *Multigrid methods for PDE optimization*, SIAM Review, 51 (2009), pp. 361–395..

[4] J.H. Bramble, *Multigrid Methods*, Pitman Research Notes in Mathematics Series, Essex, 1993.

[5] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Mathematics of Computation, 31 (1977), pp. 333–390.

[6] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer, 1998.

[7] W. Hackbusch, *Integral Equations, Theory and numerical treatment* , Birkhäuser, 1995.

[8] W. Hackbusch, *Multi-grid Methods and Applications*, Springer-Verlag, New York, 1985.

[9] K. Ito and K. Kunisch, *Lagrange Multiplier Approach to Variational Problems and Applications*, SIAM, 2008.

[10] T. KAILAT, *Some Integral Equations with "Nonrational" Kernels*, IEEE Trans. Inf. Th. Vol. IT-12, 4 (1966), pp. 442–447.
[11] M.L. KRASNOV, A.I. KISELEV AND G.I. MAKARENKO, *Integral Equations*, Mir, Moscow, 1983.
[12] J.L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.
[13] P. NEITTAANMÄKI AND D. TIBA, *Optimal Control of Nonlinear Parabolic Systems*, Marcel Dekker, New York, 1994.
[14] S. PRIMAK, V. IA KONTOROVICH, V. KONTOROVITCH, AND V. LYANDRES, *Stochastic Methods and Their Applications to Communications : Stochastic Differential Equations Approach*, Wiley, 2004.
[15] U. TROTTENBERG, C. OOSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press, London, 2001.
[16] J. VAN BLADEL, *Electromagnetic Fields*, IEEE Press, 1985.