

Weierstraß-Institut für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint

ISSN 0946 – 8633

Discrete random walk on large spherical grids generated by spherical means for PDEs

K.K. Sabelfeld^{1,2}, I.A. Shalimova², and A.I. Levykin²

submitted: October 4th 2001

¹ Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstraße 39, D – 10117 Berlin, Germany
E-mail: sabelfel@wias-berlin.de

² Institute of Computational Mathematics
and Mathematical Geophysics, Russian Acad. Sci.
Lavrentieva str., 6, 630090 Novosibirsk, Russia

No. 970

Berlin 2004



1991 *Mathematics Subject Classification.* 65C05, 76F99.

This work is partly supported by the RFBR Grant HIII - 03-01-00914, the program of Leading Scientific Schooles under Grant N 1271.2003.1, and the NATO Linkage Grant N 978912.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: <http://www.wias-berlin.de/>

Abstract

A new general stochastic-deterministic approach for a numerical solution of boundary value problems of potential and elasticity theories is suggested. It is based on the use of the Poisson-like integral formulae for overlapping spheres. An equivalent system of integral equations is derived and then approximated by a system of linear algebraic equations. We develop two classes of special Monte Carlo iterative methods for solving these systems of equations which are a kind of stochastic versions of the Chebyshev iteration method and successive overrelaxation method (SOR). In the case of classical potential theory this approach accelerates the convergence of the well known Random Walk on Spheres method (RWS). What is however much more important, this approach suggests a first construction of a fast convergent finite-variance Monte Carlo method for the system of Lamé equations.

1 Introduction

There are three main classes of stochastic numerical methods for solving PDEs:

(1) *Methods based on probabilistic representations* of solutions in the form of expectations over diffusion stochastic processes; central problem is here the construction of effective numerical solutions of relevant stochastic differential equations governing the above diffusion processes [3].

(2) *Random Walk on Spheres methods* based on the Monte Carlo calculation of the iterations of Green functions for standard domains like a sphere, ellipsoid, cylinder, etc., e.g., see [1], [10], [8]. This approach is very close to the first one; it was developed for equations with constant coefficients, because for such equations it is possible to find the Green function explicitly. There are two different justification methods in this approach: the first one is based on the interpretation of the random walk on spheres process as a martingale [1] on a Markov chain embedded into the Wiener process; the second method exploits the representation of the solution in the form of a Neumann series and the standard Markov Chain Monte Carlo technique for evaluation of this series, e.g., see [10].

(3) *Random Walk on Boundary methods* based on the boundary integral equations of the potential theory. These methods, suggested in 1982 by K. Sabelfeld in [9], were generalized in [4], [14] to all classical, both interior and exterior boundary value problems of the electrostatic, heat and elastic potential theory. Note that in this case, there are no difficulties with the boundary conditions and exterior problems. Important is also that the dimension of the problem is actually reduced by one dimension since the phase space of the integral equations is the boundary of the domain. As to the disadvantages of this class of methods, - often, the variance analysis is very complicated.

The advantage of the methods (1) is that they are well theoretically developed, and can be applied to quite general scalar second order PDEs with variable coefficients. However

the list of drawbacks is very serious: (1) the probabilistic representations are possible only for scalar elliptic and parabolic equations; so a system of elliptic equations like the Lamé equation is out of the question; (2) there are a lot of difficulties in relation with the boundaries in the case of non-Dirichlet boundary conditions. Even for the Dirichlet conditions, considerable difficulties arise when approximating the random process near the boundary: one should take care that in each step, the process is inside the domain. This implies a rapid diminishing of the integration step when approaching the boundary, which in turn rapidly increases the computational cost; (3) exterior boundary value problems are hard or better to say impossible to solve by a numerical simulation of diffusion processes in unbounded regions.

In this paper, we deal with a new class of Markov chain simulation technique which we present for simplicity for the Laplace equation, but which is applicable also for systems of elliptic equations. In a sense, the method uses the advantages of the approaches 1 and 2: here we use the Green functions for standard subdomains like a sphere, and simultaneously reformulate the original differential boundary value problem in the form of a system of integral equations of Fredholm type with a well defined deterministic phase space, in contrast to the standard walk on spheres method where the spheres are random, and we cannot there use the advantage of the Fredholm theory, e.g., see the basic approach described in our book [12]. In [13], we have extended this approach by using the Poisson integral formula for overlapping spheres, and considered the relevant system of integral equations. The kernel of the Poisson integral formula was the generating transition probability density function of the Markov chain. The iterative procedure was actually a randomized method of simple iterations. Generally, this iterative procedure diverges in the case of Lamé equation. Therefore, we turn to a different iteration method. To this end, we switch to a discrete approximation of our system of integral equations. Surprisingly, this not only has complicated the method, but in contrary, we have obtained a convenient fast convergent method with a finite variance.

2 2D Dirichlet problem and the Poisson kernel

2.1 Integral formulation

For simplicity, we will explain here the main idea of the method for the two-dimensional Dirichlet problem for the Laplace equation. It turns out that even in this simple case where the conventional random walk on spheres works as well, the new method converges much faster.

Let us consider the boundary value problem

$$\Delta u(x) = 0, x \in D, \quad u(y) = \varphi, \quad y \in \Gamma = \partial D, \quad (2.1)$$

where the domain D consists of two overlapping discs $K(x_0^{(1)}, R_1)$ and $K(x_0^{(2)}, R_2)$:

$$D = K(x_0^{(1)}, R_1) \cup K(x_0^{(2)}, R_2); \quad K(x_0^{(1)}, R_1) \cap K(x_0^{(2)}, R_2) \neq \emptyset, \quad (2.2)$$

We denote by γ_1 the part of the circle $S(x_0^{(1)}, R_1)$ which belongs to the second disc while

Γ_1 is the part of the circle $S(x_0^{(1)}, R_1)$ not belonging to the second disc; analogously γ_2 and Γ_2 are defined. So the boundary of the domain D consists of Γ_1 and Γ_2 .

The regular solution to the harmonic equation satisfies the spherical mean value relation in each of the two discs:

$$u(x) = \frac{R^2 - r^2}{2\pi R} \int_{S(x_0, R)} \frac{u(y) dS_y}{|x - y|^2}. \quad (2.3)$$

Here $R = R_1$ in the first, and $R = R_2$ in the second disc, while $r = r_i = |x - x_0^{(i)}|$ is the distance from x to the circle's center, $i = 1, 2$.

It is not difficult to find out that the function

$$p(y; x) = \frac{R_1^2 - |x - x_0^{(1)}|^2}{2\pi R_1} \cdot \frac{1}{|x - y|^2} \quad (2.4)$$

is a probability density function of the variable $y \in S(x_0^{(1)}, R_1)$, for all $x \in K(x_0^{(1)}, R_1)$. This immediately follows from the representation of the solution $u = 1$ to the Dirichlet problem for the Laplace equation $\Delta u(x) = 0$, $u(y) = 1$ through the Poisson integral. From the probabilistic representation of the Dirichlet boundary value problem considered the density $p(y; x)$ coincides with the pdf of the first passage on $S(x_0^{(1)}, R_1)$ of a Wiener process starting at $x \in K(x_0^{(1)}, R_1)$.

It is possible to find explicitly the distribution function $P(x \rightarrow y \in \gamma)$ - the probability for a particle starting at $x \in K(x_0, R)$, with $r = |x - x_0|$, to reach an arc $\gamma \in S(x_0, R)$ defined by the limit angles α_1 and α_2 , say, $\alpha_1 < \alpha_2$, since,

$$\begin{aligned} P(x \rightarrow y \in \gamma) &= \frac{R^2 - r^2}{2\pi R} \int_{\gamma} \frac{dS_y}{|x - y|^2} = \frac{1}{\pi} \operatorname{arctg} \left\{ \frac{R + r}{R - r} \operatorname{tg} \frac{\alpha}{2} \right\} \Bigg|_{\alpha_1}^{\alpha_2} \\ &= \frac{1}{\pi} \operatorname{arctg} \left\{ \frac{R + r}{R - r} \operatorname{tg} \frac{\alpha_2}{2} \right\} - \frac{1}{\pi} \operatorname{arctg} \left\{ \frac{R + r}{R - r} \operatorname{tg} \frac{\alpha_1}{2} \right\}. \end{aligned} \quad (2.5)$$

Let us now write down the Poisson formulae for both discs in the form

$$u(x) = \frac{R_1^2 - r_1^2}{2\pi R_1} \int_{S(x_0^{(1)}, R_1)} \frac{u(y)}{|x - y|^2} dS_y, \quad u(y) = \frac{R_2^2 - r_2^2}{2\pi R_2} \int_{S(x_0^{(2)}, R_2)} \frac{u(x')}{|y - x'|^2} dS_{x'}. \quad (2.6)$$

Let us introduce the notation: $v_1(x) = u(x)$ for $x \in \gamma_2$, and $v_2(x) = u(x)$ for $x \in \gamma_1$. Then, (2.6) reads

$$v_1(x) = \int_{\gamma_1} p(y; x) v_2(y) dS_y + f_1(x), \quad v_2(y) = \int_{\gamma_2} p(x'; y) v_1(x') dS_{x'} + f_2(y), \quad (2.7)$$

where

$$f_1(x) = \int_{\Gamma_1} p(y, x) \varphi(y) dS_y, \quad f_2(y) = \int_{\Gamma_2} p(x', y) \varphi(x') dS_{x'} . \quad (2.8)$$

It is convenient to rewrite the system (2.7) in the matrix-integral form:

$$v = \mathbf{G}v + F \quad (2.9)$$

where $v = (v_1, v_2)^T$, $F = (f_1, f_2)^T$, and \mathbf{G} is the matrix-integral operator which acts on v as follows

$$\mathbf{G}v(x, y) = \begin{pmatrix} 0 & \int_{\gamma_1} p(y, x) v_2(y) dS_y \\ \int_{\gamma_2} p(x', y) v_1(x') dS_{x'} & 0 \end{pmatrix} .$$

The system of integral equations (2.9) with the integral operator \mathbf{G} has nice properties. First of all, the L_1 -norm of \mathbf{G} is less than 1, for any configuration of the two overlapping discs, since $\int_{S(x, R)} p(y; x) dS_y = 1$. Hence $(E - \mathbf{G})^{-1}$ exists and is represented as a convergent Neumann series. This also follows from the next assertion which presents a nice property of the Poisson kernel.

Lemma 1 *For any $x \in \gamma_2$ and any $y \in \gamma_1$*

$$\int_{\gamma_1} p(y; x) dS_y = \int_{\gamma_2} p(y'; y) dS_{y'} = 1 - \frac{\theta_1^*}{\pi} - \frac{\theta_2^*}{\pi} ,$$

where the angles θ_1^* and θ_2^* are defined as follows: $2\theta_1^*$ is the angle of view of the arc γ_1 from the centre of the first circle, and $2\theta_2^*$ is the angle of view of the arc γ_2 from the centre of the second circle.

Proof. For any $x \in \gamma_2$ we have obviously:

$$p(y; x) = \frac{\cos(\psi)}{\pi|x-y|} - \frac{1}{2\pi R_1} \quad (2.10)$$

which follows from the cosine relation: $|R_1^2 - r_1|^2 + |x - y|^2 = 2R_1|x - y| \cos(\psi)$ where ψ is the angle between $x - y$ and $x - x_0^{(1)}$.

Using the relation (2.10) we can write:

$$\int_{\gamma_1} p(y; x) dS_y = \frac{1}{\pi} \left\{ \int_{\gamma_1} \frac{\cos(\psi)}{|x-y|} dS_\psi \right\} - \frac{1}{2\pi R_1} \int_{\gamma_1} dS .$$

In the right-hand side, the first integral in the braces is the double layer potential integral which is equal (e.g., see [10]) to the angle of view of the arc γ_1 from the point x , i.e., to $(2\pi - 2\theta_2^*)/2$. The second integral is simply θ_1^*/π . This proves the Lemma 1, since exactly the same result is obviously obtained for the second circle, when $y \in \gamma_1$.

In constructions of iterative numerical procedures, we will need the information about the principal eigen-value of the integral operator. In the next theorem we find this eigen-value explicitly.

Theorem 1. *The integral operator \mathbf{G} is a Fredholm operator with the kernels $p(y; x)$, $p(x'; y)$, continuous on $x \in \gamma_2$, $y \in \gamma_1$, with integrable singularities at the points of intersection of γ_1 and γ_2 of the type $p(y; x) \asymp \frac{\sin(\theta_1^* + \theta_2^*)}{\pi |x - y|}$ as $x \rightarrow y$. The eigen-values of \mathbf{G} , λ_i , are all real, and $\lambda_i = \pm \sigma_i \rho(\mathbf{G})$ where $\rho(\mathbf{G})$ is the spectral radius of \mathbf{G} ($\sigma_i \leq 1$ are positive constants) given explicitly by*

$$\rho(\mathbf{G}) = 1 - \frac{\theta_1^*}{\pi} - \frac{\theta_2^*}{\pi}.$$

Integral equation (2.9) has a unique solution which solves the Dirichlet problem (2.1).

Proof. First let us show that the singularities have the form $p(y; x) \asymp \frac{\sin(\theta_1^* + \theta_2^*)}{\pi |x - y|}$ as $x \rightarrow y$. Simple geometrical considerations show that $R^2 - r^2 = |x - y| \cdot b$, where $b = |x - y^*|$, and y^* is the second point of intersection of the line $x - y$ with the circle $S(x_0, R)$. Thus $p(y; x) \leq \frac{b}{2\pi R|x - y|}$. Now, as $x \in \gamma_2 \rightarrow y \in \gamma_1$, we have in the limit $b = 2R \sin(\theta_1^* + \theta_2^*)$.

Let us now consider the eigen-value problem. Note that the integral operator \mathbf{G} is not symmetric, but we can show that it can be symmetrized. Indeed, introducing the new functions $w_1(x) = v_1(x) \times \sqrt{R_1/[R_1^2 - r_1^2]}$, $w_2(y) = v_2(y) \times \sqrt{R_2/[R_2^2 - r_2^2]}$ we come to the eigenvalue problem for the symmetric integral equation

$$\lambda w = \hat{\mathbf{G}}w$$

where

$$\hat{\mathbf{G}} = \begin{pmatrix} 0 & \int_{\gamma_1} \frac{g(x, y)}{|x - y|^2} w_2(y) dS_y \\ \int_{\gamma_2} \frac{g(x, y)}{|y - x|^2} w_1(x) dS_x & 0 \end{pmatrix}.$$

Here $g(x, y)$ is a symmetric function: $g(x, y) = \frac{1}{2\pi} \sqrt{(R_1^2 - r_1^2)(R_2^2 - r_2^2)}/[R_1 R_2]$.

Thus the eigenvalues λ are real, moreover, they are concentrated in the interval $(-\rho, \rho)$ symmetrically relative to the origin, where $\rho = \rho(\mathbf{G}) < 1$ is the spectral radius. Indeed, if λ is an eigen-value with the corresponding eigen-function (ψ_1, ψ_2) , then $-\lambda$ is also an eigen-value with the corresponding eigen-function $(\psi_1, -\psi_2)$.

Let us now evaluate the spectral radius of our system of integral equations. Taking the eigen-function as a constant $(1, 1)^T$, we see that the corresponding eigenvalue is given by

$$\lambda_0 = \int_{\gamma_i} p(y; x) dS_y = 1 - \frac{\theta_1^*}{\pi} - \frac{\theta_2^*}{\pi}, \quad i = 1, 2 \quad (2.11)$$

which does not obviously depend on x .

It is not difficult to show that $\rho(\mathbf{G}) = \lambda_0$. Indeed, let λ be an arbitrary eigen-value, and $(\psi_1, \psi_2)^T$ - the corresponding eigen-function. For any $x \in \gamma_2$ we can write $|\lambda| |\psi_1(x)| \leq |\psi_2(y^*)| \lambda_0$, where y^* is a point where $|\psi_2|$ reaches its maximum. For any $y \in \gamma_1$ we have

analogously: $|\lambda||\psi_2(y)| \leq |\psi_1(x^*)|\lambda_0$, where x^* is the point of maximum of $|\psi_1|$. From these two inequalities we get the desired result: $|\lambda|^2 \leq \lambda_0^2$. Thus $\rho(\mathbf{G}) = \lambda_0$.

Finally, the equivalence of the integral equation and the Dirichlet problem is obvious: the solution of the Dirichlet problem satisfies the integral equation whose solution is unique.

2.2 Approximating system of linear algebraic equations

Suppose we want to approximate the system of integral equations (2.9) by a system of algebraic equations. To this end, we take uniform meshes x_1, \dots, x_{m_1+1} on the arc γ_1 and y_1, \dots, y_{m_2+1} on γ_2 generating by the uniform polar angles distributions (the end points are included); these meshes subdivide γ_1 and γ_2 in the set of arcs $\gamma_1^{(i)}$, $i = 1, \dots, m_1$ and $\gamma_2^{(i)}$, $i = 1, \dots, m_2$, respectively. Since the Poisson kernel $p(y, x)$ has a weak singularity, it is convenient to take the approximation in the form:

$$\int_{\gamma_1} p(y, y_k)v_2(y)dS_y = \sum_{i=1}^{m_1} p_i^{(1)}(x_i, y_k)v_2(x_i), \quad k = 1, \dots, m_2,$$

and analogously,

$$\int_{\gamma_2} p(x', x_k)v_1(x')dS_{x'} = \sum_{i=1}^{m_2} p_i^{(2)}(y_i, x_k)v_1(y_i), \quad k = 1, \dots, m_1,$$

where

$$p_i^{(1)}(x_i, y_k) = \int_{\gamma_1^{(i)}} p(y, y_k)dS_y, \quad p_i^{(2)}(y_i, x_k) = \int_{\gamma_2^{(i)}} p(x', x_k)dS_{x'}. \quad (2.12)$$

These coefficients can be evaluated explicitly, using the formula (2.5). The same approximation is used to calculate the right hand sides f_1 and f_2 in all grid points. Thus we come to a discrete approximation of (2.9) in the form of the following system of linear algebraic equations:

$$v^{(k)} = \sum_{i=1}^{m_1+m_2} a_{ki} v^{(i)} + F^{(k)}, \quad k = 1, 2, \dots, m_1 + m_2 \quad (2.13)$$

where v^k , $k = 1, \dots, m_1$ are the approximations to the function v_1 , and v^k , $k = m_1 + 1, \dots, m_1 + m_2$ are the approximations to the function v_2 . The same for the functions $F^{(k)}$.

Note that written in a 2×2 -block matrix form, say, $v = Av + f$, the matrix A has zero diagonal blocks, while the block A_{12} relates the values of the function v_1 with the values of function v_2 , and converse relation yields the matrix A_{21} .

We use also different approximations, in particular, based on linear interpolations of v and F , and a refinement 12-point Gauss approximation formula in the end points of the arc.

2.3 A chain of n overlapping discs

Generalization to the case of n circles is straightforward, and the relevant system of integral equations is written in a block matrix form, whose general structure is shown below in (2.14) for the case when the domain is a chain of overlapping discs (ChOD) where each disc has intersections with only two immediate neighbour discs (which do not overlap). Thus introducing $k = 2(n - 1)$ functions v_i and writing the Poisson formulae in each disc we come to a system of k integral equations $v = \mathbf{G} v + f$ where the kernel of the matrix integral operator \mathbf{G} is a $k \times k$ -matrix G which has the following structure: in the first row, only the kernel G_{12} is not zero, the second and the third rows have the following non-zero kernels: G_{21}, G_{24} and G_{31}, G_{34} . The same for the rows 4 and 5: the non-zero entries are G_{43}, G_{46} and G_{53}, G_{56} , etc., so that the j -th' row non-zero entries (j is even) are $G_{j,j-1}$ and $G_{j,j+2}$, while the $j + 1$ -th' row non-zero entries are $G_{j+1,j-1}$ and $G_{j+1,j+2}$. The last row has only one non-zero entry: $G_{k,k-1}$. For illustration, in (2.14) we present the kernel matrix G for 5 circles, so that $k = 8$.

$$G = \begin{pmatrix} 0 & G_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ G_{21} & 0 & 0 & G_{24} & 0 & 0 & 0 & 0 \\ G_{31} & 0 & 0 & G_{34} & 0 & 0 & 0 & 0 \\ 0 & 0 & G_{43} & 0 & 0 & G_{46} & 0 & 0 \\ 0 & 0 & G_{53} & 0 & 0 & G_{56} & 0 & 0 \\ 0 & 0 & 0 & 0 & G_{65} & 0 & 0 & G_{68} \\ 0 & 0 & 0 & 0 & G_{75} & 0 & 0 & G_{78} \\ 0 & 0 & 0 & 0 & 0 & 0 & G_{87} & 0 \end{pmatrix}. \quad (2.14)$$

Note that the kernel matrix G is cyclic, of index 2, and the general structure of the eigen-values can be extracted from the following representation of the characteristic determinant:

$$\det(\lambda E - G) = \lambda^{k-2r} \prod_{i=1}^r (\lambda^2 - \sigma_i \rho^2(G))$$

where $r \geq 1$ is an integer number, and $\sigma_1 = 1$, $|\sigma_r| \leq |\sigma_{r-1}| \leq \dots \leq \sigma_1$. So this type of matrix always has real eigenvalues $\pm \rho$.

Let us turn to the approximation of the above system of integral equations by a system of linear algebraic equations. In this case, instead of $2n - 2$ integral equations for the functions v_1, \dots, v_{2n-2} , we will have an approximating system of linear algebraic equations ($u = Au + f$) with a 5-diagonal block matrix (see (2.14)) with zero diagonal blocks, and in each block-row there are only two non-zero blocks which relate the values of the function u on the arc inside one disk with the values defined on two arcs belonging to two neighbor discs. The first and last rows have only one non-zero block because they have only one neighbor.

To ensure that the system of linear algebraic equations is a good approximation to the exact system of integral equations it is enough to prove that $(E - A)^{-1}$ exists. This is ensured by the fact that our matrices are all substochastic, and their spectral radii are all less than 1, see the error estimations given by (2.15).

So let us use the numbering of the arcs quite simple: the first arc γ_1 is the arc inside the

first disc $l = 1$ which overlaps only with the second disc $l = 2$, where two further arcs lie: γ_2 and γ_3 , etc. This numbering results in the structure of the matrix shown in (2.14). The relevant numbering of the functions is used: the functions $v_j(x)$, $j = 2l - 2$, $j = 2l - 1$ are defined on the arcs γ_j which belong to the disc l where $j = 2l - 2$ and $j = 2l - 1$, $l = 2, \dots, n - 1$. Analogously for the first and last discs.

Let us define the difference between $v_j(x_i)$, $j = 2l - 2$, the exact solution of the system of integral equations taken in l -th disc, at a point $x_i \in \gamma_j$, and the approximation $u_i^{(j)}$ taken as the i -th component of the solution of our linear equation (i th row in the j -th block of the matrix A): $\epsilon_i^{(j)} = u_i^{(j)} - v_j(x_i)$. Hence the error vector ϵ has in j -th block the components $\epsilon_i^{(j)}$.

Let us also define the errors δ_j and δ_j^f , the errors of approximation of the Poisson integrals in l -th disc ($j = 2l - 2$) over γ_j and Γ_j , respectively. Thus we can write for the i -th row in the l th disc ($j > 1$):

$$\begin{aligned}
u_i^{(j)} - v_j(x_i) &= \sum_{k=1}^{2n-2} a_{ik}^{j-1} u_k^{(j-1)} - \int_{\gamma_{j-1}} p(y; x_i) v_{j-1}(y) dS(y) \\
&+ \sum_{k=1}^{2n-2} a_{ik}^{j+2} u_k^{(j+2)} - \int_{\gamma_{j+2}} p(y; x_i) v_{j+2}(y) dS(y) + \delta_j^f(x_i) \\
&= \sum_{k=1}^{2n-2} a_{ik}^{j-1} (u_k^{(j-1)} - v_{(j-1)}(x_k)) + \sum_{k=1}^{2n-2} a_{ik}^{j+2} (u_k^{(j+2)} - v_{(j+2)}(x_k)) \\
&+ \left\{ \sum_{k=1}^{2n-2} a_{ik}^{j-1} (v_{(j-1)}(x_k)) - \int_{\gamma_{j-1}} p(y; x_i) v_{j-1}(y) dS(y) \right\} \\
&+ \left\{ \sum_{k=1}^{2n-2} a_{ik}^{j+2} (v_{(j+2)}(x_k)) - \int_{\gamma_{j+2}} p(y; x_i) v_{j+2}(y) dS(y) \right\} + \delta_j^f(x_i).
\end{aligned} \tag{2.15}$$

The difference for the case $j = 2l - 1$ is treated analogously, we need only to replace $j - 1$ with $j - 2$, and $j + 2$ with $j + 1$.

Thus written in the matrix form these relations are

$$\epsilon = A\epsilon + \delta + \delta^f. \tag{2.16}$$

Let $\Delta\varphi = \max_i(\varphi_{i+1} - \varphi_i)$ be the maximum difference taken over the all angular meshes. For simplicity we take a simple estimations $\|\delta\| < C_1\Delta\varphi$ and $\|\delta^f\| < C_2\Delta\varphi$. Therefore, we have

$$\|\epsilon\| \leq \|(E - A)^{-1}\| (C_1 + C_2)\Delta\varphi. \tag{2.17}$$

3 Stochastic iteration methods

In this section we present various Monte Carlo iterative procedures for solving linear systems of equations, generally being integral equations, with specific details for system of

linear algebraic equations. First we present a general iteration method with random parameters which in the deterministic limit tends to the iterative procedure with Chebyshev parameters. In the next subsection we describe a randomized version of the successive overrelaxation method (SOR). Both classes of methods will be used then to solve our system of linear equations on circles.

3.1 A stochastic iterative procedure with optimal random parameters

Assume we have to solve a linear, generally, integral equation of the second kind:

$$u(x) = \mu \int_X k(x, y)u(y)dy + f(x) \quad (3.1)$$

or in the operator form: $u = \mu Ku + f$; X is the Euclidean space.

Standard Monte Carlo Markov chain algorithms (MCMC) for solving this kind of equations usually require that $\rho(|\mu K|) < 1$, where the integral operator $|K|$ is defined by its kernel $|k(x, y)|$, ρ is the spectral radius.

The first extension beyond the conventional Neumann series approach was suggested by K. Sabelfeld in [9] and then developed in the book [10]. This approach is based on a spectral transformation of the parameter μ . This generates different iteration procedures which are convergent even if $\rho(|\mu||K|) > 1$. However the main question - when the variance of the relevant Monte Carlo estimator is finite - was resolved under quite restrictive assumptions.

Here we deal with the following iterative procedure for the equation (3.1), with $\mu = 1$, starting with $u_0 = 0, u_1 = \beta_0 f$:

$$u_{j+1} = \alpha_j u_j + \beta_j (f + K u_j), \quad j = 1, 2, \dots \quad (3.2)$$

where α_j, β_j are some positive constants which we choose so that $\alpha_j + \beta_j = 1$.

Simple analysis shows that if we assume that the eigen-functions ϕ_l (defined through $\phi_k = \lambda_k K \phi_k$) form a complete system in the space $L_2(X)$ of square-integrable functions on a space X , then the following estimation of the error can be made.

Let the initial error be $\epsilon_0 = \sum_{i=1}^{\infty} c_i \phi_i$, then

$$\epsilon_n = \sum_{i=1}^{\infty} c_i \left[\prod_{j=1}^n (\alpha_j + \beta_j \lambda_i^{-1}) \right] \phi_i = \sum_{i=1}^{\infty} c_i \left[\prod_{j=1}^n \left(1 - \beta_j \frac{\lambda_i - 1}{\lambda_i} \right) \right] \phi_i. \quad (3.3)$$

Hence if for all λ_i there exists a set of numbers β such that

$$\left| 1 - \beta \frac{\lambda_k - 1}{\lambda_k} \right| = q_k < 1 - \delta, \quad \delta > 0$$

then for all β_j belonging to this set the method converges.

It is possible to construct different Monte Carlo estimators following this iterative procedure. We prefer to construct biased estimators: first, we fix n , the number of iterations

we will perform, and choose the numbers $\beta_0, \beta_1, \beta_2, \dots, \beta_{n-1}$ in the interval $[0, 1]$. How to make such a choice optimal, we will discuss later.

Then we proceed as follows. First, we need to have a reversed indexation, so let $\beta'_k = \beta_{n-k}$, $k = 1, \dots, n$.

Start a Markov chain from the point where the solution should be found, say, x_0 , and take the current state as $X = x_0$. The current value of the iteration index is $j = 1$. Take the initial value of the weight as $Q = 1$. The initial value of the random estimator is $\xi = f(x_0)\beta'_1$.

1. Sample uniformly in $(0, 1)$ a random number $rand$ and check if $rand > \beta'_j$. If so, then calculate the random estimator

$$\xi = \xi + Qf(X)\beta'_{j+1}$$

and go to the next iteration which means that we put $j = j + 1$ and go to 1 if $j < n$.

2. Otherwise we simulate the transition from the current state X to the next state Y according to the transition density $p(x, y)$. Then recalculate the weight $Q = Qk(X, Y)/p(X, Y)$, and the random estimator is scored as $\xi = \xi + Q\beta'_{j+1}f(Y)$. The current state is now $i = k$, $X = Y$; we turn to the next iteration again by putting $j = j + 1$ and goto 1 if $j < n$.

After n steps we finish the evaluation of our random estimator $\xi(x_0)$. It is not difficult to show that the constructed random estimator ξ is unbiased: $u_n(x_0) = E\xi$.

Indeed, for $i = 1, 2$ this is obvious since $u_0 = 0$, $u_1 = \beta_0 f$. The next step is also the next step in the Markov chain method since

$$u_{j+2} = (\alpha_{j+1}E + \beta_{j+1}K)(\alpha_j E + \beta_j K)u_j + (\alpha_{j+1}E + \beta_{j+1}K)\beta_j f + \beta_{j+1}f,$$

and so on, we have after n -steps:

$$\begin{aligned} u_{j+n} &= \left\{ \prod_{i=1}^n (\alpha_{j+n-i}E + \beta_{j+n-i}K) \right\} u_j \\ &+ \sum_{k=1}^{n-1} \left[\left\{ \prod_{i=1}^{n-k} (\alpha_{j+n-i}E + \beta_{j+n-i}K) \right\} \beta_{j+k-1}f \right] + \beta_{j+n-1}f. \end{aligned} \quad (3.4)$$

3.1.1 Optimal random parameters β_k

The parameters α_i and β_i can be chosen according to the Chebyshev iteration method which is based on polynomials which are uniformly close to zero. However in our method, it is quite natural to choose these parameters randomly, according to a minimization of the probabilistic error (see [5]). Remarkably, the Chebyshev choice of parameters follows from this probabilistic approach.

To analyze the error, it is convenient to work with the operator $B = E - K$. We introduce the corresponding polynomial by

$$P_n(t) = \prod_{i=1}^n (1 - \beta_i t) \quad (3.5)$$

which relates, in view of (3.4), the errors through

$$\epsilon_{j+n} = P_n(B)\epsilon_j . \quad (3.6)$$

It is the general idea, in the iterative methods, to make the polynomial $P_n(t)$ as closer to zero as possible, and in the deterministic approach the problem was solved by Markov and Chebyshev. In our case, our parameters β_k and $\alpha_k = 1 - \beta_k$ are random, and it is natural to measure the error in the probabilistic sense, as well presented in [5]. Let us follow this approach.

So let us assume that we have chosen n random numbers β_1, \dots, β_n which are equally and independently distributed on some interval. Then the polynomial P_n is a random variable, and we can write:

$$\ln |P_n| = \sum_{k=1}^n \ln p_k \quad (3.7)$$

where $p_k = |1 - \beta_k t|$. Note that the random numbers $\ln p_k$ are equally and independently distributed, so we can apply the central limit theorem. This implies that the distribution of $\ln p_k$ tends to the Gaussian distribution, as n increases:

$$K(x) = \frac{1}{\sqrt{2\pi nD}} \exp \left\{ -\frac{(x - na)^2}{2nD} \right\}$$

where $a = \langle \ln |1 - \beta_k t| \rangle$ is the expectation, and D - the variance of $\ln p_k$. Standard considerations yield

$$P(|P_n| > \epsilon) \approx 1 - \Phi \left(\frac{\ln(\epsilon) - na}{\sqrt{2nD}} \right)$$

where Φ is the function $\Phi = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$. From this follows that to ensure that the probability of deviation of $P_n(t)$ tends to zero we have to require that the expectation a is negative.

Let $\varphi(x)$ be the distribution density of β_k which is defined on the interval $[M^{-1}, m^{-1}]$ where m and M are the lower and upper boundaries of the spectrum of the operator $E - K$. Thus we have

$$a = \int_{1/M}^{1/m} \ln |tx - 1| \varphi(x) dx, \quad D = \int_{1/M}^{1/m} (\ln |tx - 1| - a)^2 \varphi(x) dx . \quad (3.8)$$

It is natural to assume that the expectation a does not depend on t , which implies that

$$\frac{da}{dt} = \int_{1/M}^{1/m} \frac{x \varphi(x)}{tx - 1} dx = 0 . \quad (3.9)$$

A density function on $[M^{-1}, m^{-1}]$ which solves (3.9) has the form:

$$\varphi(x) = \frac{1}{\pi x \sqrt{(1 - mx)(Mx - 1)}}. \quad (3.10)$$

This gives

$$a = -\ln \frac{\sqrt{M} + \sqrt{m}}{\sqrt{M} - \sqrt{m}}, \quad (3.11)$$

and

$$D < \pi^2 + 8 \ln 2 \sqrt{\frac{m}{M}} + O\left(\left(\frac{m}{M}\right)^{3/2}\right).$$

From this, the following estimation of the number of iterations n required to reach the error ϵ can be derived:

$$n > \frac{\ln(\epsilon)}{a}$$

where the expectation a is given by (3.11).

Sampling from the density φ is simple: by the inversion method we find first the simulation formula for the random number β_k^{-1} , which finally yields

$$\beta_k = \frac{2}{(M - m) \cos(\pi \text{rand}_k) + M + m} \quad (3.12)$$

where rand_k are random numbers uniformly and independently distributed on $(0, 1)$. A variance reduction can be achieved by the following modification: the interval is uniformly divided into n subintervals, and then, change in the simulation formula (3.12) β_k with $(j - \xi_j)/n$ where j are integer numbers which cyclically vary with period n as $j = 1, \dots, n$ and ξ_j are random numbers uniformly distributed on $(0, 1)$. Remarkably, if ξ_j are changed with their expectations 0.5, we come to the method with optimal Chebyshev parameters, see for details [5].

3.2 SOR method

Again, let us explain the main idea in the simple case of two overlapping discs and the governing system of integral equations (2.9). The matrix integral operator \mathbf{G} can be represented as $\mathbf{G} = \mathcal{L} + \mathcal{U}$ where \mathcal{L} and \mathcal{U} are the lower and upper - triangular operators, respectively:

$$\mathcal{L}v = \begin{pmatrix} 0 & 0 \\ \int_{\gamma_2} p(x', y) v_1(x') dS_{x'} & 0 \end{pmatrix}, \quad \mathcal{U}v = \begin{pmatrix} 0 & \int p(y, x) v_2(y) dS_y \\ \gamma_1 & 0 \end{pmatrix}.$$

Introducing a scalar parameter ω we rewrite our equation $v = \mathbf{G}v + F$ in the form:

$$v = (E - \omega \mathcal{L})^{-1} [(1 - \omega)E + \omega \mathcal{U}]v + \omega (E - \omega \mathcal{L})^{-1} F. \quad (3.13)$$

This is a general form of the SOR method (e.g., see [2]). In the case we deal with we note that $(E - \omega\mathcal{L})^{-1} = E + \omega\mathcal{L}$, therefore, our equation has the following simple form:

$$v = \mathbf{T}v + d$$

where

$$\mathbf{T} = (E + \omega\mathcal{L})[(1 - \omega)E + \omega\mathcal{U}], \quad d = \omega(E + \omega\mathcal{L})F.$$

If the integral operator \mathbf{T} were a contracting operator, we could apply a standard collision estimator. This can be applied directly to the integral form, or to the approximating system of linear algebraic equations. Here it is convenient again to use a Markov chain of length n , to evaluate the n -th approximation. Note that the structure of the operator \mathbf{T} has a nice probabilistic sense: in each step, we either stay in the current state with probability $1 - \omega$, or, otherwise, make a transition according the operators \mathcal{U} and \mathcal{L} , successively. Note that in the case of matrix operators, there are well known interrelations between the spectra of \mathbf{G} and \mathbf{T} , e.g., see [2], [6] which can be used to analyse the convergence and variance of stochastic methods.

Let us discuss the case of $n > 2$ overlapping discs, each disc having no more than 2 neighbors (the ChOD-domain). The remarkable property $(E - \omega\mathcal{L}^{-1} = E + \omega\mathcal{L}$ does not hold for our system of equations with the chosen indexation of the arcs generating the matrix G (for illustration, see, e.g., the matrix G in (2.14) in the case of 5 discs). However it can be shown that this indexation can always be chosen so that the property $(E - \omega\mathcal{L})^{-1} = E + \omega\mathcal{L}$ holds true.

We define now a family of domains for which this theory works. First, let us define a closed subset of discs as follows: the first disc in this subset overlaps with the second disc, the second with the third, etc, and the last disc overlaps again with the first disc; note that in this subset, each disc has only two overlapping discs.

Let us define an S_2 -disc domain as follows: each disc may overlap with arbitrary number of discs, however each intersection is a result of overlapping only of two discs. Further, S_2 -disc domain may include a subset of a closed set of discs, but the number of discs in such a subset should be even.

Thus we will deal in the following theorem with the family of S_2 -disc domains, which is quite general.

Theorem 2. *Assume that D is an S_2 -disc domain. Then the indexation of arcs can be chosen so that the matrix G is cyclic, of index 2, and the property $(E - \omega\mathcal{L})^{-1} = E + \omega\mathcal{L}$ holds true.*

4 Discrete random walks

In this section we present stochastic algorithms which are applied to the discrete approximation of the relevant integral equations. Hence the stochastic algorithms are based on discrete versions of the iteration methods described (I - a nonstationary iteration method (3.2), and II - SOR method). These algorithms can be considered as a Random Walk approach for solving the relevant system of linear algebraic equations on the basis of relevant iteration method which is different from the conventional Monte Carlo method based on

the convergent Neumann series. Note that for a chosen iteration method, different randomization schemes can be constructed.

4.1 Discrete Random Walk for the iteration method (3.2)

Let us consider a system of linear algebraic equations (SLAE) which approximates the relevant system of integral equations for our domain. The SLAE can be written in the form (2.13), or in the form related to the appropriate indexation.

So we have to construct a Monte Carlo estimator for a system of m linear algebraic equations

$$x_i = \sum_{j=1}^m a_{ij}x_j + b_i, \quad i = 1, \dots, m$$

or in the matrix form,

$$x = Ax + b. \quad (4.1)$$

We assume that Max , min , the maximal and minimal eigen-values of the matrix A are known or at least estimated.

As discussed in section 3.1, we will construct unbiased random estimators ξ_n for u_n , the n -th approximation to the solution x , and more precisely, to its l -th component u_n^l .

First of all, we have to choose a nonnegative transition density matrix $p_{ij} = p(i \rightarrow j)$; $i, j = 1, \dots, m$; $\sum_{j=1}^m p_{ij} = 1$ for all i , and consistent with the matrix A , i.e., $p_{ij} \neq 0$ if $a_{ij} \neq 0$. It is convenient to take

$$p_{ij} = \frac{|a_{ij}|}{\sum_{j=1}^m |a_{ij}|}.$$

This ensures that the random walk will be concentrated only on non-zero elements which is important since we deal with sparse block matrices. We will not have absorptions in our random walk.

The first variant of the algorithm can be presented as follows:

1. Choose n random parameters according to the formula:

$$\beta_i = \frac{2}{Max + min + (Max - min) \cos(\pi \text{rand}(i))}, \quad i = 1, \dots, n$$

where $\text{rand}(i)$, $i = 1, \dots, n$ are independent samples generated by a rand-generator. The initial score is set to zero: $S = 0$. Calculate the initial value of the estimator as $v = b_l \beta_1$.

2. The initial weight $Q = 1$, and the initial number of iteration $j = 1$; fix the initial state as $i = l$.

3. Take a sample $\gamma_j = \text{rand}(j)$;

if $\gamma_j > \beta_j$, then calculate $v = v + Q b_i \beta_{j+1}$, and make the next iteration, i.e., $j = j + 1$ and go to 3. if j , the number of iterations is less than n ; otherwise make a score: $S = S + v$, and start the new statistics from 2.

4. Otherwise, if $\gamma_j < \beta_j$, we simulate the transition from the old state i to the new state k according to the density $p_{ik} = p(i \rightarrow k)$.

Recalculate the weight and the random estimator: $Q = Q a_{ik}/p_{ik}$, $v = v + Q \beta_{j+1} b_k$; then, renew the state as $i = k$, and go to the next iteration, i.e., $j = j + 1$, and go to p.3, if j , the number of iterations is less than n ; otherwise make a score: $S = S + v$, and start the new statistics from p.1.

Averaging the estimator over statistics of size N gives the result: $u_n^l \approx S/N$.

4.2 Discrete Random Walk method based on SOR

Here we present two variants of the random walk algorithm. In the first version, we assume that D is a S_2 -disc domain, so that $(E - \omega L)^{-1} = E + \omega L$, and hence our system (4.1) can be rewritten in the form

$$x = Tx + f, \quad (4.2)$$

where $T = (E + \omega L)((1 - \omega)E + \omega U)$, and $f = (E + \omega L)b$.

The first algorithm for calculation of n -th approximation is based on a direct randomized calculation of the finite number of iterations of the matrix T , i.e., by evaluation of the Neumann series $f + Tf + T^2f + \dots + T^n f$. As in the previous section, we do not introduce absorption in our Markov chain. So to calculate the component x_l of the solution to (4.2), we suggest the following algorithm

1. Fix n , the number of iterations to be made, and choose the parameter ω , say, equal to 1 as in Seidel's method. The initial score is set to zero: $S = 0$. Calculate the matrix T , and the vector f .
2. Set the initial weight $Q = 1$, the number of iteration $j = 1$, and the current state of the Markov chain $i = l$. The initial value of the estimator is set as $v = f_l$,
3. Simulate the transition from the state i to the new state k according to the density $p_{ik} = p(i \rightarrow k)$ which is chosen, e.g., as in the method of the previous section:

$$p_{ij} = \frac{|t_{ij}|}{\sum_{j=1}^m |t_{ij}|}.$$

Recalculate the weight and the random estimator:

$$Q = Q t_{ik}/p_{ik}, \quad v = v + Q f_k;$$

then, renew the state as $i = k$, and go to the next iteration, i.e., $j = j + 1$, and go to p.3, if j , the number of iterations is less than n ; otherwise make a score: $S = S + v$, and start the new statistics from p.2.

Averaging the estimator over statistics of size N gives the result.

Another version of this algorithm is the following. Let $\mathcal{L}_{i,j}$ and $\mathcal{U}_{i,j}$ be the entries of the triangular matrices $E + \omega L$ and $(1 - \omega)E + \omega U$, respectively. Two new transition densities are defined by

$$p_{\mathcal{L}}(i \rightarrow i') = \frac{|\mathcal{L}_{i,j}|}{\sum_{j=1}^m |\mathcal{L}_{i,j}|}, \quad p_{\mathcal{U}}(i' \rightarrow k) = \frac{|\mathcal{U}_{i,j}|}{\sum_{j=1}^m |\mathcal{U}_{i,j}|}. \quad (4.3)$$

According to the representation $T = (E + \omega L)((1 - \omega)E + \omega U)$, make the transition from the state i to state k in two steps: first, sample the transition from i to a state i' according to the matrix $E + \omega L$ (i.e., the transition $i \rightarrow i'$ is sampled from the pdf $p_{\mathcal{L}}(i \rightarrow i')$ defined in (4.3)), and then make the transition $i' \rightarrow k$ according to the matrix $((1 - \omega)E + \omega U)$ (i.e., the transition $i' \rightarrow k$ is sampled from the pdf $p_{\mathcal{U}}(i' \rightarrow k)$ defined in (4.3)). In each step the weight is recalculated, so that in the first step $Q = Q_{\mathcal{L}_{i,i'}/p_{\mathcal{L}}(i \rightarrow i')}$ and then, $Q = Q_{\mathcal{U}_{i',k}/p_{\mathcal{U}}(i' \rightarrow k)}$, with the final random estimator $v = v + Q f_k$.

It is worth mentioning that in the simulation of these discrete distributions, it is quite useful to apply the highly economical algorithm described in [7], because in the preprocessing stage, we fix the number of the grid points and then the cost of the algorithm will not depend of the matrices size - it needs only one sample made by the random generator and a couple of "if" operators. This makes possible to work with very large matrices L and U .

5 Conclusion and extensions

We have presented in this paper only the main idea, and the details are given for simplicity only for the case eminently loved in the probability theory and stochastic numerics - the Dirichlet problem for the Laplace equation. Our main motivation however were systems of elliptic equations and equations of higher order where both the classical probabilistic approach and Markov Chain Monte Carlo methods fail.

The method presented is quite general, and we mention here two examples of extensions we have made. The first example is the Dirichlet problem for the system of Lamé equations:

$$\mu \Delta \mathbf{u}(x) + (\lambda + \mu) \text{grad div } \mathbf{u}(x) = 0, \quad x \in D, \quad \mathbf{u}(y) = \mathbf{g}, \quad y \in \Gamma$$

where λ and μ are the Lamé constants of elasticity.

All the considerations can be extended starting from the generalized Poisson formula:

$$\mathbf{u}(x) = \int_{S(x_0, R)} p(y; x) B \mathbf{u}(y) dS(y) \quad (5.1)$$

where the matrix B is given in a simple and explicit form, see [12].

Second example is the biharmonic problem:

$$\Delta^2 u(x) = 0, \quad x \in D, \quad u \Big|_{\Gamma} = g_0, \quad \frac{\partial u}{\partial n} \Big|_{\Gamma} = g_1.$$

Here n is the exterior normal vector to the boundary Γ .

The generalized Poisson formula in this case has the same form as (5.1) where the vector \mathbf{v} includes the function u and its derivatives.

These two examples and more will be presented in a forthcoming paper [11].

References

- [1] Ermakov, S.M., Nekrutkin, W.W. and Sipin, A.S. *Random Processes for Classical Equations of Mathematical Physics*. vol. **34** of Math. Appl. (Soviet Series). Kluwer, Dodrecht, 1989
- [2] Hadjimos A., Plemmons R. Optimal p-cyclic SOR. *Numer. Math.*, **67**, 475-490 (1994)
- [3] P.E. Kloeden and E. Platen. Numerical solution of stochastic differential equations. Springer. Berlin - Heidelberg, 1992.
- [4] O.A.Kurbanmuradov, N.A.Simonov and K.K.Sabelfeld. The walk on boundary algorithms. Nauka, Novosibirsk, 1989, 162 p. (in Russian)
- [5] G.I. Marchuk and V.I. Lebedev. Numerical methods in the neutron transport theory. Atomizdat, Moscow, 1971 (in Russian).
- [6] Niethanner W. The successive overrelaxation method (SOR) and Markov chains. *Annals of Operations Research*, **103**, 351-358 (2001).
- [7] Ripley B. Stochastic simulation. John Wiley & Sons, NY, 1987
- [8] Royster, W.C. A Poisson integral formula for the ellipse and some applications. *Proc. AMS.*, 15(1964), 661-670.
- [9] Sabelfeld, K.K. Vector Monte Carlo algorithms for solving systems of elliptic equations of the second order and the Lamé equation. *Dokl.Akad.Nauk SSSR*, **262**, 1982, No.5, 1076–1080 (in Russian).
- [10] Sabelfeld, K.K. *Monte Carlo Methods in Boundary Value Problems*. Springer-Verlag, Berlin – Heidelberg – New York, 1991.
- [11] K. Sabelfeld, A. Levykin, I. Shalimova. Spherical integral equations and discrete stochastic iterative methods. *Monte Carlo Methods and Applications*, **11**, N1, 2005.
- [12] Sabelfeld, K.K. and Shalimova, I.A. *Spherical Means for PDEs*. VSP, The Netherlands, Utrecht, 1997
- [13] Sabelfeld, K.K. and Shalimova I.A. Random walk methods for static elasticity problems. *Monte Carlo Methods and Applications*, **8**, 2002, N2.
- [14] Sabelfeld, K.K. and Simonov, N.A. *Random Walks on Boundary for Solving PDEs*. VSP, The Netherlands, Utrecht, 1994