JOURNAL OF COMPUTATIONAL PHYSICS 113, 35–51 (1994)

A Fast Direct Algorithm for the Solution of the Laplace Equation on Regions with Fractal Boundaries

P. JONES*

Department of Mathematics, Yale University, New Haven, Connecticut 06520

J. MA† AND V. ROKHLIN†

Department of Computer Science, Yale University, New Haven, Connecticut 06520

Received August 5, 1993

An algorithm is presented for the rapid direct solution of the Laplace equation on regions with fractal boundaries. In a typical application, the numerical simulation has to be on a very large scale involving at least tens of thousands of equations with as many unknowns, in order to obtain any meaningful results. Attempts to use conventional techniques have encountered insurmountable difficulties, due to excessive CPU time requirements of the computations involved. Indeed, conventional direct algorithms for the solution of linear systems require order $O(N^3)$ operations for the solution of an $N \times N$-problem, while classical iterative methods require order $O(N^2)$ operations, with the constant strongly dependent on the problem in question. In either case, the computational expense is prohibitive for large-scale problems. The direct algorithm of the present paper requires $O(N)$ operations with a constant dependent only on the geometry of the boundary, making it considerably more practical for large-scale problems encountered in the computation of harmonic measure of fractals, complex iteration theory, potential theory, and growth phenomena such as crystallization, electrodeposition, viscous fingering, and diffusion-limited aggregation.

1. INTRODUCTION

During the last decade, the numerical solution of the Laplace equation on regions with fractal boundaries has been becoming increasingly popular both in mathematics and physics. In mathematics, examples include harmonic measure of fractals, complex iteration theory, and potential theory. In physics, examples include growth phenomena such as crystallization, electrodeposition, viscous fingering, and diffusion-limited aggregation, where the harmonic measure governs the growth of the fractal surfaces [25]. Thus, much recent work has been focused on the study of the metric properties of harmonic measure on fractals [1, 7, 14, 17]. Several attempts have been made during the last several years to solve such problems numerically (see [1]).

Two approaches to the study of the metric properties of the harmonic measure on fractals have received most of the attention in recent years.

1. Viewing the harmonic measure as the relative hitting probability at points on the surface and using the Monte Carlo method to conduct computer simulations on parallel machines such as the connection machine (see [1]).

2. Formulating the problem as an integral equation of the first kind and solving the resulting equation numerically.

While the first approach has produced some significant results (see [1]), the computation becomes prohibitively expensive when high accuracy is desired, due to the slow convergence of the Monte Carlo method (as is well known, the error of a Monte Carlo simulation decays like $1/\sqrt{N}$, where N is the number of trials).

On the other hand, the second approach has also encountered insurmountable difficulties, due to excessive CPU time requirements of the computations involved. Indeed, in order to obtain mathematically meaningful results, systems of linear equations have to be solved, involving at least tens of thousands of equations with as many unknowns. Conventional direct methods require order $O(N^3)$ operations for the solution of an $N \times N$ linear system, while classical iterative methods require order $O(N^2)$ operations, with the constant strongly dependent on the problem in question. In either case, the computational expense is prohibitive for large-scale problems.

We present a direct algorithm for the rapid solution of the Laplace equation on regions with fractal boundaries. The algorithm requires $O(N)$ operations with a constant dependent only on the geometrical properties of the fractal boundaries, where N is the number of elements in the discretization of the fractal. The scheme is sufficiently fast that

* Research supported in part by NSF under Grant DMS 9213595.
† Research supported in part by DARPA under Contract F49620/91/C/0084, and in part by ONR under Grant N00014-89-J-1527.

0021-9991/94 $6.00

problems involving $N \sim 10^6$ can be solved, even without the use of supercomputers, and admits far-reaching generalizations.

The plan of the paper is as follows. In Section 2, we begin with the definition of the problems to be addressed; in Section 3, we summarize certain mathematical and numerical facts to be used in this paper; in Section 4, we develop the mathematical apparatus used in the construction of the fast algorithm by borrowing terminology from the standard scattering theory for the Helmholtz equation; in Section 5, we present the description of the fast algorithm, and in Section 6, we illustrate the performance of the algorithm by numerical examples. Finally, in Section 7, we discuss our conclusions.

2. STATEMENT OF THE PROBLEMS

In this section, we define the problems to be addressed, namely, the boundary value problems for the Laplace equation on regions with fractals of Cantor type as the boundaries.

A fractal of Cantor type is a classical example of fractals (see, for example, [2, 10, 18]), which can be generated recursively by dividing a square in the plane into four squares (boxes) with a ratio of sides as a parameter.

Given a real number a $(0 < a < \frac{1}{2})$, we define a sequence of sets (see Fig. 1),

$$C_0^a = \{\text{the unit square}\}, \tag{1}$$

$$C_1^a = \{4 \text{ corner boxes with } a \text{ as their sizes}\}, \tag{2}$$

$$\cdots$$

$$C_l^a = \{4^l \text{ corner boxes with } a^l \text{ as their sizes}\}, \tag{3}$$

$$\cdots$$

where l is an integer. The sequence $\{C_0^a, C_1^a, C_2^a, \ldots\}$ decreases monotonically:

$$C_0^a \supseteq C_1^a \supseteq C_2^a \supseteq \cdots.$$

The Cantor set C^a associated with the given ratio a is defined as the intersection of the sequence (see Fig. 2):

$$C^a = \bigcap_{l \in \{0,1,2,\ldots,\infty\}} C_l^a. \tag{4}$$

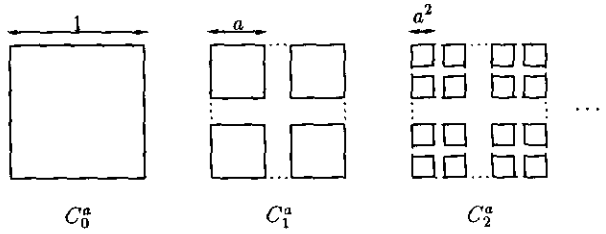

FIG. 1. Sequence $\{C_0^a, C_1^a, C_2^a, \ldots\}$.

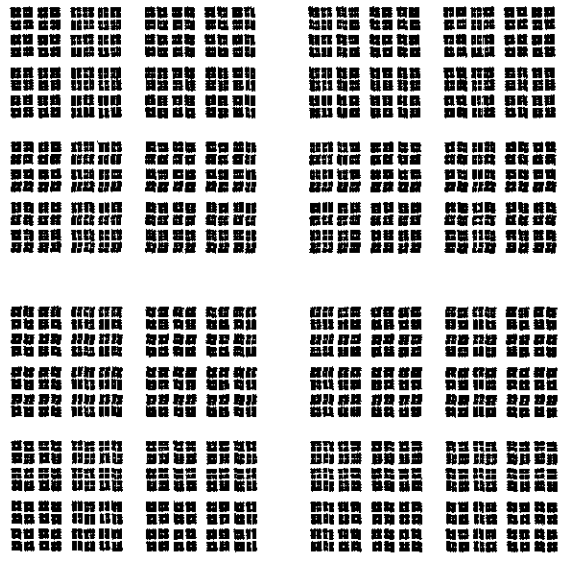

FIG. 2. Cantor set C^a—a fractal.

For a given integer $l \geq 1$, we define the lth approximation to the Cantor set C^a as a set

$$A_l = \{\text{the centers of all boxes in } C_l^a\}. \tag{5}$$

We will refer to the 4^l boxes of C_l^a generated during the lth step of the above process as level l boxes. Thus, there is one box on level 0, and it coincides with the unit square. The level $l+1$ is obtained from the level l by subdividing each box on the level l into four corner boxes (see Fig. 1).

We will also impose a tree structure on the hierarchical structure of C^a, so that if *ibox* is a fixed box at level l, the four boxes at level $l+1$ obtained by subdivision of *ibox* are considered its children, while the four child boxes are considered brothers.

Given a Cantor set C^a, we will consider the exterior Dirichlet problem for the Laplace equation defined by the formulae

$$\begin{aligned} &\Delta u = 0 \qquad \text{for} \quad x \in R^2 \setminus C^a, \\ &u|_{C^a} = f. \end{aligned} \tag{6}$$

To ensure the uniqueness of the solution of problem (6), the far-field condition

$$\lim_{|x| \to \infty} |u(x)| < \infty \tag{7}$$

is normally imposed.

The proof of the following theorem can be found, for example, in [24]; and Remark 2.1 is a well-known fact (see [9]), for example).

THEOREM 2.1. *The boundary value problem* (6) *is a well-posed problem.*

Remark 2.1. Suppose that $\lim_{l\to\infty} f_l = f$ on C^a and u_l is the solution of Dirichlet problem with boundary condition defined on a set of boxes C_l^a (see (3)):

$$\begin{aligned} &\Delta u = 0 \qquad \text{for} \quad x \in R^2 \backslash C_l^a, \\ &u|_{C_l^a} = f_l. \end{aligned} \tag{8}$$

and satisfies the far-field condition (7). Then u_l converges to the solution u of the boundary value problem (6).

Remark 2.2. As is well known (see [9], for example), the boundary value problems (8) can be formulated as an integral equation of the first kind by representing the solution as the logarithmic potential of the charge distribution on the boundary C_l^a. Denoting by σ the charge distribution over the boundary C_l^a, we obtain the integral equation

$$\int_{C_l^a} \ln|x-t|\, d\sigma(t) = f_l(x) - f_l(\infty) \tag{9}$$

with $x \in C_l^a$, where the integration is in the sense of Borel measure. Detailed discussion of the uniqueness of the solution of (9) can be found, for example, in [24].

For a given integer $l \geq 1$, suppose that $N = 4^l$, and $A_l = \{z_i \mid i = 1, 2, \ldots, N\}$ is the lth approximation to the Cantor set C^a. Then, by the definition of Borel integration, the integral equation (9) is discretized as a linear system of equations

$$A\sigma = b \qquad \text{with} \quad A \in R^{N\times N} \text{ and } b \in R^N, \tag{10}$$

where

$$A_{ij} = \log|z_i - z_j| \qquad \text{for} \quad i \neq j, \tag{11}$$

$$A_{ii} = \iint_\Omega \ln r\, dx\, dy \Big/ \iint_\Omega dx\, dy, \tag{12}$$

with $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$, and a square $\Omega \in C_l^a$ centered at (x_0, y_0). The following theorem can be easily verified.

THEOREM 2.2. *Defined by* (10), (11), *and* (12), *the matrix A is nonsingular.*

In the rest of this paper, we focus on the numerical solution of linear systems of the type defined by (10), (11), and (12).

3. MATHEMATICAL AND NUMERICAL PRELIMINARIES

In this section, we summarize certain well known mathematical and numerical facts to be used in the rest of this paper. They can be found, for example, in [5, 8, 11, 24].

3.1. Boundary Value Problems for the Laplace Equation

Suppose that $\Gamma \subset R^2$ is a Jordan curve, parameterized by its length γ: $[0, L] \to R^2$, and Ω is the region bounded by Γ, so that $\partial\bar{\Omega} = \Gamma$. Suppose further that N: $[0, L] \to R^2$ is the interior normal to Γ. For an integrable function f: $[0, L] \to R^1$, we will be considering the following problems:

(A) Interior Dirichlet problem,

$$\begin{aligned} &\Delta\Phi(x) = 0 \qquad &&\text{for} \quad x \in \Omega \\ &\Phi(x) = f(\gamma^{-1}(x)) \qquad &&\text{for} \quad x \in \Gamma; \end{aligned} \tag{13}$$

(B) Exterior Neumann problem,

$$\begin{aligned} &\Delta\Psi(x) = 0 \qquad &&\text{for} \quad x \in R^2 \backslash \Omega \\ &\frac{\partial}{\partial N}\Psi(x) = f(\gamma^{-1}(x)) \qquad &&\text{for} \quad x \in \Gamma \end{aligned} \tag{14}$$

with Ψ satisfying the far-field condition

$$0 < \lim_{|x|\to\infty} \frac{|u(x)|}{\log|x|} < \infty, \qquad \lim_{|x|\to\infty} |u(x) - c\log|x|\,| = 0, \tag{15}$$

where $c \neq 0$ is a constant.

As is well known, each of the above two problems has a unique solution for any continuous right-hand side f and piecewise smooth boundary Γ (see, for example, [8]).

3.2. Single and Double Layer Potentials

Suppose that a point charge of unit intensity is located at the point $x_0 \in R^2$. Then, for any $x \in R^2$ with $x \neq x_0$, the potential due to this charge is described by the expression

$$\phi_{x_0}(x) = -\ln(\|x - x_0\|). \tag{16}$$

The potential of a dipole of unit intensity located at x_0 and oriented at the direction $h \in R^2$ ($\|h\| = 1$) is described by the formula

$$\phi_{x_0,h}(x) = \frac{h(x-x_0)}{\|x-x_0\|^2}. \tag{17}$$

It is well known that the potential ϕ_{x_0} due to a point charge at $x_0 \in R^2$ (defined by formula (16)) is harmonic in

any region excluding the source point x_0. Moreover, for any harmonic function u: $R^2 \to R^1$, there exists an analytic function w: $\mathbb{C} \to \mathbb{C}$ such that $u(x, y) = \mathrm{Re}(w(x, y))$. In the rest of this paper, we will make no distinction between points in R^2 and points in $\mathbb{C}$. In complex terms, the potentials ϕ_{x_0} and $\phi_{x_0,h}$ defined by the expressions (16) and (17), respectively, assume the forms

$$x_{z_0}(z) = \mathrm{Re}(-\ln(z - z_0))$$

and

$$\phi_{z_0,h}(z) = \mathrm{Re}\left(\frac{h}{z - z_0}\right),$$

where $z = x + iy$ and $z_0 = x_0 + iy_0$. Following the standard practice, we will refer to the analytic function $\ln(z - z_0)$ as the potential at the point $z \in \mathbb{C}$ due to a charge located at the point z_0.

For an integrable function μ: $[0, L] \to R^1$, the potential of a single layer with density μ is given by

$$\Psi(x) = \int_0^L \phi_{\gamma(t)}(x)\, \mu(t)\, dt, \tag{18}$$

and the potential of a double layer with the dipole density μ is given by

$$\Phi(x) = \int_0^L \phi_{\gamma(t),N(t)}(x)\, \mu(t)\, dt. \tag{19}$$

3.3. Integral Equations of the Classical Potential Theory

In the classical potential theory, the interior Dirichlet problem (13) is solved by representing Φ as the potential of a double layer, and the exterior Neumann problem (14) is solved by representing Ψ as the potential of a single layer. The analysis of the single layer and double layer potentials in the vicinity of the boundaries results in two integral equations of the second kind.

(A1) Interior Dirichlet problem,

$$\pi\mu(x) + \int_0^L \frac{\partial}{\partial N(t)} \log \|\gamma(x) - \gamma(t)\|\, \mu(t)\, dt = f(x). \tag{20}$$

(B1) Exterior Neumann problem,

$$-\pi\mu(x) + \frac{\partial}{\partial N(x)} \int_0^L \log \|\gamma(x) - \gamma(t)\|\, \mu(t)\, dt = f(x). \tag{21}$$

3.4. Galerkin Method for the Solution of Integral Equations

As is well known, the classical Galerkin method can be used for the numerical solution of integral equations of the form

$$\mu(x) + \int_a^b K(x, t)\, \mu(t)\, dt = f(x). \tag{22}$$

Given an orthonormal basis $\{P_n(x)\}$ in $L^2[a, b]$, the function μ_n: $[0, L] \to R^1$ defined by

$$\mu_n(x) = \sum_{j=1}^{n} c_j P_j(x) \tag{23}$$

and satisfying

$$(r, \mu_n) = 0 \tag{24}$$

will be used to approximate the solution of the integral equation (22), where the error function $r(x)$ in formula (24) is defined via

$$r(x) = \mu_n(x) + \int_0^L K(x, t)\, \mu_n(t)\, dt - f(x). \tag{25}$$

The above procedure results in a linear system $Bc = b$ defined by

$$B_{ij} = \int_0^L \int_0^L K(x, t)\, P_i(x)\, P_j(t)\, dx\, dt \quad \text{for} \quad i \neq j, \quad 1 \leqslant i, j \leqslant n, \tag{26}$$

and for $1 \leqslant i \leqslant n$, $B_{ii} = 1$,

$$b_i = \int_0^L f(x)\, P_i(x)\, dx. \tag{27}$$

Lemma 3.1 below introduces the well-known Legendre polynomials, and Lemmas 3.2 and 3.3 present the integrals (26) in Cartesian and polar coordinate respectively for integral equations (20) and (21).

LEMMA 3.1. *The Legendre polynomials $P_n(x)$ defined by*

$$\begin{aligned} P_0(x) &= \frac{1}{\sqrt{2}}, \\ P_n(x) &= \sqrt{\frac{2n+1}{2}}\, \frac{1}{2^n n!}\, \frac{d^n}{dx^n}(x^2 - 1)^n, \qquad n = 1, 2, \ldots, \end{aligned} \tag{28}$$

are orthonormal polynomials on interval $[-1, 1]$.

LEMMA 3.2. *Suppose that Ω is a square with vertices* $(1, 1)$, $(1, -1)$, $(-1, -1)$, $(-1, 1)$, *the boundary of Ω is denoted by Γ, and $\{P_1(x), P_2(x), \ldots, P_n(x), \ldots\}$ are the Legendre polynomials (see Lemma 3.1). Suppose further that μ is the solution of Eq.* (20) *or* (21), *and for each of the two horizontal sides of Ω,*

$$\mu_n(x) = \sum_{i=0}^{n} \xi_i P_i(x), \tag{29}$$

for each of the two vertical sides of Ω,

$$\mu_n(y) = \sum_{i=0}^{n} \eta_i P_i(y). \tag{30}$$

Then the integrals in (26) *are of the form*

$$I_{ij} = \int_{-1}^{1}\int_{-1}^{1} \frac{(y+1)}{(x+1)^2+(y+1)^2} P_i(x)\, P_j(y)\, dx\, dy, \tag{31}$$

for any two adjacent sides of the square Ω, *and*

$$J_{ij} = \int_{-1}^{1}\int_{-1}^{1} \frac{2}{(x-y)^2+4} P_i(x)\, P_j(y)\, dx\, dy, \tag{32}$$

for non-adjacent sides of the square Ω, *where* $1 \leqslant i, j \leqslant n$.

The following lemma is obtained immediately by considering the integrals I_{ij} in (31) in polar coordinates.

LEMMA 3.3. *Suppose that for* $1 \leqslant i, j \leqslant n$, *integrals* I_{ij} *are defined by formula* (31). *Then*

$$I_{ij} = \left\{\int_0^{\pi/4}\int_0^{2/\cos\theta} + \int_{\pi/4}^{\pi/2}\int_0^{2/\sin\theta}\right\} P_i(r\cos\theta - 1) \times P_j(r\sin\theta - 1)\sin\theta\, dr\, d\theta. \tag{33}$$

Remark 3.1. While the Integrals I_{ij} in the formula (31) are singular functions in the Cartesian coordinates, they are smooth (C^∞) functions in the polar coordinates. Thus, the integrals (33) can be efficiently evaluated via Gaussian quadratures in polar coordinates.

3.5. Ranks of Interactions

In this section, we consider electrostatic interactions and establish Theorem 3.1, which is the principal analytical tool of this paper.

For two sets of points $\{x_1, x_2, ..., x_m\}$ and $\{y_1, y_2, y_n\}$, we define the interaction matrix of the two sets as a $m \times n$ matrix given by

$$(\phi_{y_j}(x_i)) = \begin{pmatrix} \phi_{y_1}(x_1) & \phi_{y_2}(x_1) & \cdots & \phi_{y_n}(x_1) \\ \phi_{y_1}(x_2) & \phi_{y_2}(x_2) & \cdots & \phi_{y_n}(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_{y_1}(x_m) & \phi_{y_2}(x_m) & \cdots & \phi_{y_n}(x_m) \end{pmatrix} \tag{34}$$

$$= \begin{pmatrix} \ln(x_1-y_1) & \ln(x_1-y_2) & \cdots & \ln(x_1-y_n) \\ \ln(x_2-y_1) & \ln(x_2-y_2) & \cdots & \ln(x_2-y_n) \\ \vdots & \vdots & & \vdots \\ \ln(x_m-y_1) & \ln(x_m-y_2) & \cdots & \ln(x_m-y_n) \end{pmatrix}. \tag{35}$$

Lemma 3.4 below can be easily proved by expanding $\ln(1-\omega)$ into Taylor series with respect to ω; Lemma 3.5 is the reformulation of Lemma 3.4 and will be used in the proof of Theorem 3.1.

LEMMA 3.4. *Let a unit point charge be located at* z_0. *Then for any* z *such that* $|z| > |z_0|$,

$$\phi_{z_0}(z) = \ln(z - z_0) = a_0 \ln(z) + \sum_{k=1}^{\infty} \frac{a_k}{z^k}, \tag{36}$$

where

$$a_0 = 1, \qquad a_k = (-)\, z_0^k/k. \tag{37}$$

Furthermore, for any $p \geqslant 1$,

$$\left|\phi_{z_0}(z) - a_0 \ln(z) - \sum_{k=1}^{p} \frac{a_k}{z^k}\right| \leqslant \left(\frac{1}{c-1}\right)\left(\frac{1}{c}\right)^p, \tag{38}$$

where

$$c = |z/z_0|. \tag{39}$$

LEMMA 3.5. *Suppose that the expansion* (36) *is truncated after* p *terms* ($p \geqslant 1$), *and the error of the truncated expansion by* $\varepsilon_{z_0}^p(z)$,

$$\varepsilon_{z_0}^p(z) = \phi_{z_0}(z) - a_0 \ln(z) - \sum_{k=1}^{p} (a_k/z^k).$$

Then for any $p \geqslant 1$,

$$\phi_{z_0}(z) = \varepsilon_{z_0}^p(z) + u_p v_p^{\mathrm{T}} \tag{40}$$

with the vectors u_p *and* v_p *defined by*

$$u_p = \left(\ln z, \frac{1}{z}, \frac{1}{z^2}, ..., \frac{1}{z^p}\right) \tag{41}$$

$$v_p = \left(1, \frac{-z_0}{1}, \frac{-z_0^2}{2}, ..., \frac{-z_0^p}{p}\right). \tag{42}$$

Furthermore,

$$|\varepsilon_{z_0}^p(z)| \leqslant \left(\frac{1}{c-1}\right)\left(\frac{1}{c}\right)^p \tag{43}$$

with c *defined by* (39).

Clearly, the truncation error $\varepsilon_{z_0}^p(z)$ in (43) decays exponentially as the function of p. Thus few terms in (36) are needed to achieve any given accuracy. The following theorem follows immediately from formulae (40) and (43).

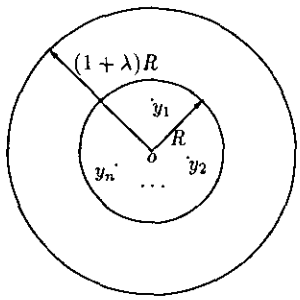

FIG. 3. Points $\{x_i\}$ outside the circle with radius $(1+\lambda)R$.

THEOREM 3.1. *Let n unit charges be located within the circle $|y|<R$ at points $\{y_1, ..., y_n\}$, $\lambda>0$ be some real number and $\{x_1, x_2, ..., x_m\}$ be another set of points such that $|x_i|>(1+\lambda)R$ for all $1\leqslant i\leqslant m$ (see Fig. 3). Then the interaction matrix of the two sets $\{x_i\}$ and $\{y_j\}$ has the decomposition*

$$(\phi_{y_j}(x_i)) = \begin{pmatrix} \ln x_1 & \frac{1}{x_1} & \frac{1}{x_1^2} & \cdots & \frac{1}{x_1^p} \\ \ln x_2 & \frac{1}{x_2} & \frac{1}{x_2^2} & \cdots & \frac{1}{x_2^p} \\ \vdots & \vdots & \vdots & & \vdots \\ \ln x_m & \frac{1}{x_m} & \frac{1}{x_m^2} & \cdots & \frac{1}{x_m^p} \end{pmatrix}$$

$$\times \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \frac{-y_1}{1} & \frac{-y_2}{1} & \cdots & \frac{-y_m}{1} \\ \frac{-y_1^2}{2} & \frac{-y_2^2}{2} & \cdots & \frac{-y_m^2}{2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{-y_1^p}{p} & \frac{-y_2^p}{p} & \cdots & \frac{-y_m^p}{p} \end{pmatrix} + E^p, \qquad (44)$$

where the truncation error $E^p=(\varepsilon_{y_j}^p(x_i))$ is bounded by the expression

$$|\varepsilon_{y_j}^p(x_i)| \leqslant \left(\frac{1}{\lambda}\right)\left(\frac{1}{1+\lambda}\right)^p \qquad (45)$$

for $1\leqslant i\leqslant m$ and $1\leqslant j\leqslant n$.

Inequality (45) states that every element of the matrix of truncation error E^p decays exponentially as the function of p. Thus for any given accuracy, the interaction matrix of the two sets $\{x_i\}$ and $\{y_j\}$ can be decomposed into the product of two matrices of low rank.

For two sets of points, $x_1, x_2, ..., x_m\in\mathbb{C}$ and $y_1, y_2, ..., y_n\in\mathbb{C}$, we say that the two sets are *well-separated* (see Fig. 4) if there exist points $x_0, y_0\in\mathbb{C}$, and $R>0$ such that

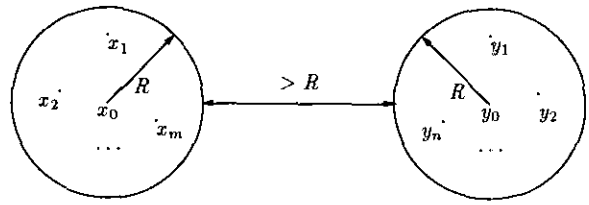

FIG. 4. Well-separated sets in the plane.

$$|x_i - x_0| < R \quad \text{for} \quad 1\leqslant i\leqslant m,$$
$$|y_j - y_0| < R \quad \text{for} \quad 1\leqslant j\leqslant n,$$
$$|x_0 - y_0| > 3R.$$

Theorem 3.1 implies that with any prescribed precision, the interaction matrix of two well-separated sets can be decomposed into a product of two matrices of low rank, and the rank is bounded by a constant depending only on the separation of the two sets (see formula (45)).

3.6. Interactions in Cantor Sets

Suppose that D_1 and D_2 are two subsets of a Cantor set C^a and that A_l is the lth approximation of C^a (see (5)), then we will refer to the interaction between $D_1\cap A_l$ and $D_2\cap A_l$ as the interacion between D_1 and D_2.

Theorem 3.2 below states that for any given ratio a, the interactions in Cantor set C^a are of low rank; and the ranks are bounded by a constant dependent only on the ratio a in the generation of the Cantor set. Lemma 3.6 is obvious and will be used in the proof of Theorem 3.2.

LEMMA 3.6. *Suppose that D_1, D_2 are two subsets of a Cantor set C^a with ratio a, and each of them are divided into four pieces of equal size (child boxes) (see Figure 5). Suppose further that A is a child box of D_1, and B is a child box of D_2. Then the rank of the interaction matrix between subsets D_1 and D_2 is at most four times as large as that between boxes A and B.*

THEOREM 3.2. *For a real a $(0<a<\frac{1}{2})$, and an integer $l\geqslant 1$, suppose that C^a is the Cantor set associated with the ratio a, and C_l^a is the set of boxes generated at the lth level*

$$C_l^a = \{D_1, D_2, ..., D_{4^l}\}. \qquad (46)$$

Then for any given precision, the rank of the interaction matrix between any two boxes D_i and D_j $(i\neq j)$ is bounded by a constant dependent only on the ratio a for generating the Cantor set C^a. The constant does not depend on the sizes of boxes and the numbers of points inside the boxes. In other words, the matrix of interactions between any two boxes at

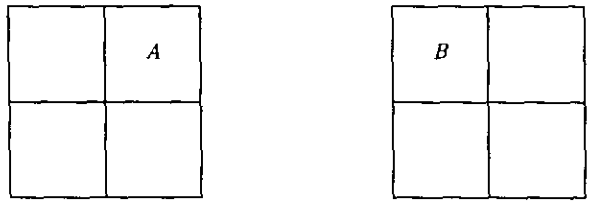

FIG. 5. Two subsets of C^a.

any level of Cantor set C^a is of fixed rank, to any prescribed precision.

Proof. The theorem follows immediately from Theorem 3.1 if any two sets D_i and D_j $(i \neq j)$ are well separated.

Suppose that D_i and D_j are not well separated from each other. Then we divide each of them into four pieces (child box) of equal size (see Fig. 5). Suppose that A is a child box of D_i, and B is a child box of D_j (see Fig. 5). Then the theorem follows from Lemma 3.6 and Theorem 3.1, if A and B are well separated from each other.

Suppose that sets A and B are not well separated. Then we keep dividing them into pieces until the smallest pieces are well separated. Due to Lemma 3.6 and Theorem 3.1, the rank of the interaction matrix between any two boxes at any level of the Cantor set is bounded by

$$p \cdot 4^k \qquad \text{with} \quad k = \left| \frac{\ln(1-2a)/(\sqrt{2}-1)}{\ln a} \right|, \tag{47}$$

where p is the rank of the interaction matrix between two well-separated pieces. ∎

Remark 3.2. Clearly, the estimate (47) is an extremely pessimistic one. In the following section, we obtain much sharper numerical estimates (see Section 4.3).

4. SCATTERING THEORY FOR THE LAPLACE EQUATION

In this section, we borrow terminology from the standard scattering theory for the Helmholtz equation and refer to the result as scattering theory for the Laplace equation. In the following, we first introduce the concept of scattering matrix, and then we present a merging scheme for generating scattering matrices recursively.

Throughout this section, Γ will denote a Jordan curve, parameterized by its length γ: $[0, L] \to R^2$. The region bounded by Γ will be denoted by Ω, and D will denote a compact subset of Ω (Fig. 6). In addition, $G: \Omega \times \Omega \to R^1$ will denote the Green's function for domain Ω, and N: $[0, L] \to R^2$ will denote the interior normal to Γ. For a compact set $E \subset R^2$, $\mathcal{M}(E)$ will denote the set of all non-negative Borel measures on E.

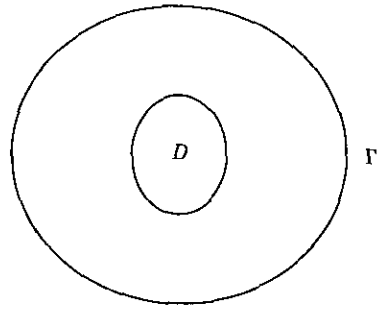

FIG. 6. Compact set D in domain Ω with boundary Γ.

4.1. Scattering Matrices

Any function $\Phi: \bar{\Omega} \to R$ harmonic inside Ω and continuous on $\bar{\Omega}$ will be referred to as *incoming potential.* As is well known, for any continuous function $\varphi: \Gamma \to R$, there exists a unique function $\Phi: \bar{\Omega} \to R$ harmonic on Ω and continuous on $\bar{\Omega}$ such that $\Phi|_\Gamma = \varphi$. Therefore, we will abose the notation by referring to the function $\varphi: \Gamma \to R$ as the incoming potential.

Suppose that $q \in L^2(\Omega)$ and σ is a Borel measure over D. Given a function $K \in L^2(R^2 \times R^2)$, the function

$$\Psi(x) = \int_D K(x, t)\, q(t)\, d\sigma(t) \qquad \text{for} \quad x \in R^2 \backslash D \tag{48}$$

will be referred to as *outgoing potential.* Similarly, we will call its restriction $\psi = \Psi|_\Gamma$ onto Γ an outgoing potential. Outside the domain $\bar{\Omega}$, function Ψ will also be referred to as *scattered potential.*

Remark 4.1. Particularly, we are interested in the case when $K(x, t) = \ln \|x - t\|$, and $q(t)$ is the characteristic function of D. Then the outgoing potential

$$\Psi(x) = \int_D \ln \|x - t\|\, d\sigma(t) \tag{49}$$

is a function harmonic in $R^2 \backslash D$ and satisfies the far field condition (15).

We define three operators

$$L: L^2(\Gamma) \to L^2(D),$$
$$P: \mathcal{M}(D) \to L^2(R^2 \backslash D),$$
$$S: \mathcal{M}(D) \to L^2(\Gamma),$$

via the formulae

$$L(\varphi)(x) = \frac{1}{2\pi} \int_0^L \varphi(\gamma(t)) \cdot \frac{\partial}{\partial N_t} G(x, \gamma(t)) \cdot dt, \tag{50}$$

$$P(\sigma)(x) = \int_D K(x, t)\, q(t)\, d\sigma(t)$$

$$\text{for} \quad x \in R^2 \backslash D, \tag{51}$$

$$S(\sigma)(x) = \int_D K(x, t)\, q(t)\, d\sigma(t) \qquad \text{for} \quad x \in \Gamma. \tag{52}$$

We will be considering equations of the form

$$P(\sigma) = f, \tag{53}$$

with $f \in L^2(D)$. A special case of Eq. (53) is the integral equation (9) defined in Section 2, with D a Cantor set, $K(x, t) = \ln \|x - t\|$, and $q(t)$ the characteristic function of D.

DEFINITION 4.1. The operator $\alpha: L^2(\Gamma) \to L^2(\Gamma)$ defined by the expression

$$\alpha = S \cdot P^{-1} \cdot L \tag{54}$$

will be referred to as the *scattering matrix* associated with the triple (D, Γ, K).

Remark 4.2. Given an incoming potential φ on the boundary Γ, the operation of α on φ can be viewed as consisting of three steps:

1. The operator L constructs a function $f = L\varphi: \bar{\Omega} \to R$ harmonic over the compact set D and such that $(L\varphi)|_\Gamma = \varphi$.

2. The operator P^{-1} constructs the solution $\sigma = P^{-1}f$ of Eq. (53) from the harmonic function $f = L\varphi$. The charge distribution $\sigma = P^{-1}L\varphi$ will be referred to as the *induced charge distribution.*

3. The operator S defined by (52) constructs an outgoing potential $\psi = S\sigma$ on Γ from the induced charge distribution σ. The outhgoing potential ψ will be referred to as the *induced outgoing potential.*

Thus, the scattering matrix α converts an incoming potential φ into the induced outgoing potential ψ,

$$\psi = \alpha\varphi. \tag{55}$$

4.2. Recursive Generation of Scattering Matrices

When a compact subset D of domain Ω is a union of mutually disjoint compact sets $\{D_i\}$, the scattering matrix of D can be obtained by merging the scattering matrices of $\{D_i\}$. We begin with introducing the requisite notation and then present a merging scheme for the recursive generation of scattering matrices.

4.2.1. *Notation*

Suppose that $\Delta = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_m\} \subset \Omega$ is a set of closed Jordan curves. Each $\Gamma_i \in \Delta$ is parameterized by its length

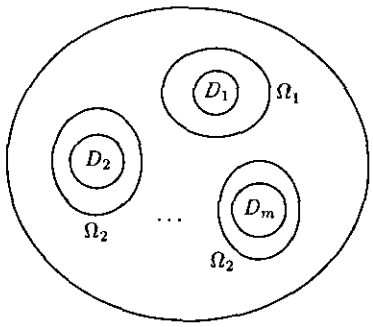

FIG. 7. $D_i \subset \Omega_i$ for $1 \leqslant i \leqslant m$.

$\gamma_i: [0, L_i] \to R^2$, $\Omega_i \subset \Omega$ is the region bounded by Γ_i, and D_i is a compact subset of Ω_i (see Fig. 7). Suppose further that for $1 \leqslant i \leqslant m$, D_i is a compact subset of Ω_i, function $G_i: \Omega_i \times \Omega_i \to R^1$ is the Green's function for domain Ω_i, and function $N_i: [0, L] \to R^2$ is the interior normal to Γ_i.

Assuming that the domains $\{\Omega_i\}$ are mutually disjoint, we will consider a special compact subset of domain Ω defined by

$$D = \bigcup_{i=1}^{m} D_i.$$

In addition to operators L, P, and S defined by (50), (51), and (52) in the preceding section for (D, Γ), we will require the operators for (D_i, Γ_i) with $1 \leqslant i \leqslant m$,

$$L_{ii}: L^2(\Gamma) \to L^2(\Omega_i),$$
$$P_i: \mathcal{M}(D_i) \to L^2(R^2 \backslash D_i),$$
$$S_{ii}: \mathcal{M}(D_i) \to L^2(\Gamma_i),$$

defined by

$$L_{ii}(\varphi)(x) = \frac{1}{2\pi} \int_0^{L_i} \varphi(\gamma_i(t)) \cdot \frac{\partial}{\partial N_i} G_i(x, \gamma_i(t)) \cdot dt, \tag{56}$$

$$P_i(\sigma)(x) = \int_{D_i} K(x, t)\, q(t)\, d\sigma(t) \qquad \text{for} \quad x \in R^2 \backslash D, \tag{57}$$

$$S_{ii}(\sigma)(x) = \int_{D_i} K(x, t)\, q(t)\, d\sigma(t) \qquad \text{for} \quad x \in \Gamma_i. \tag{58}$$

We will consider equations of the form

$$P_i(\sigma) = f_i \tag{59}$$

with $f_i \in L^2(D_i)$, $1 \leqslant i \leqslant m$.

DEFINITION 4.2. A function $\varphi_i \in L^2(\Gamma_i)$ will be called the *total incoming potential* if for any $x \in D_i$,

$$P_i(\sigma|_{D_i})(x) = L_{ii}(\varphi_i)(x), \tag{60}$$

where operators P_i and L_{ii} are defined by (57) and (56), respectively, and $\sigma|_{D_i}$ is the restriction of the charge distribution σ (defined by (53)) to the compact subset $D_i \subset D$.

We are now applying the definition of scattering matrix to each subset D_i for $1 \leqslant i \leqslant m$ (see Definition 4.1). Suppose that for any i $(1 \leqslant i \leqslant m)$, function φ_i is the total incoming potential on Γ_i, function ψ_i is the outgoing potential induced by φ_i, and operator α_i is the scattering matrix for the domain D_i. Then we have

$$\alpha_i = S_{ii} P_i^{-1} L_{ii}, \tag{61}$$

$$\psi_i = \alpha_i \cdot \varphi_i \tag{62}$$

(see (54) and (55)).

We will also require operators for $1 \leqslant i, j \leqslant m$,

$$L_i: L^2(\Gamma) \to L^2(\Gamma_i),$$
$$S_i: L^2(\Gamma_i) \to L^2(\Gamma),$$
$$L_{ji}: L^2(\Gamma_i) \to L^2(\Gamma_j), \qquad i \neq j,$$

defined by

$$L_i(\varphi)(x) = \frac{1}{2\pi} \int_0^L \varphi(\gamma(t)) \cdot \frac{\partial}{\partial N_t} G(x, \gamma(t)) \cdot dt \quad \text{for} \quad x \in \Gamma_i, \tag{63}$$

$$S_i \psi_i = \int_{D_i} K(x, t)\, q(t)\, d\sigma(t) \qquad \text{for} \quad x \in \Gamma, \tag{64}$$

$$L_{ji} \psi_i = \int_{D_i} K(x, t)\, q(t)\, d\sigma(t) \qquad \text{for} \quad x \in \Gamma_j. \tag{65}$$

In other words, the operator S_i converts the outgoing potential ψ_i on Γ_i into the scattered potential on Γ, and the operator L_{ji} converts the ougoing potential ψ_i on Γ_i into the scattered potential on Γ_j.

DEFINITION 4.3. Suppose that the operator

$$U: \begin{pmatrix} L^2(\Gamma_1) \\ L^2(\Gamma_2) \\ \vdots \\ L^2(\Gamma_m) \end{pmatrix} \to \begin{pmatrix} L^2(\Gamma_1) \\ L^2(\Gamma_2) \\ \vdots \\ L^2(\Gamma_m) \end{pmatrix}$$

is defined by

$$U = \begin{pmatrix} I & -L_{12}\alpha_2 & \cdots & -L_{1m}\alpha_m \\ -L_{21}\alpha_1 & I & \cdots & -L_{2m}\alpha_m \\ \vdots & \vdots & \cdots & \vdots \\ -L_{m1}\alpha_1 & -L_{m2}\alpha_2 & \cdots & I \end{pmatrix}. \tag{66}$$

Then the operator

$$S_p: L^2(\Gamma) \to \begin{pmatrix} L^2(\Gamma_1) \\ L^2(\Gamma_2) \\ \vdots \\ L^2(\Gamma_m) \end{pmatrix}$$

defined by the formula

$$S_p = U^{-1} \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{pmatrix} \tag{67}$$

will be referred to as the *splitting matrix*, provided the operator U in (66) is nonsingular.

4.2.2. *A Merging Scheme for Scattering Matrices*

Theorem 4.1 is the principal tool of this paper; it describes a scheme to obtain the scattering matrix of domain D $(= \bigcup_{i=1}^m D_i)$ from the scattering matrices of subdomains $D_1, D_2, \ldots, D_m$. Lemma 4.1 is used in the proof of Theorem 4.1.

LEMMA 4.1. *Suppose that φ is an incoming potential on the boundary Γ and φ_i is the total incoming potential on Γ_i defined by formula* (60). *If the operator U defined by* (66) *is nonsingular, then we have*

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_m \end{pmatrix} = S_p \cdot \varphi. \tag{68}$$

In other words, the splitting matrix S_p converts the incoming potential φ on Γ into the total incoming potentials $\{\varphi_1, \varphi_2, \ldots, \varphi_m\}$ on boundaries $\{\Gamma_1, \Gamma_2, \ldots, \Gamma_m\}$.

Proof. For any i $(1 \leqslant i \leqslant m)$, the total incoming potential φ_i on the boundary Γ_i equals the sum of the potential from Γ and the scattered potentials from $\{\Gamma_j\}$ with $j \neq i$ and $1 \leqslant j \leqslant m$. That is,

$$\varphi_i = L_i \varphi + \sum_{j \neq i} L_{ij} \psi_j. \tag{69}$$

Combining (62) with (69), we have

$$\varphi_i = L_i \varphi + \sum_{j \neq i} L_{ij} \alpha_j \varphi_j. \tag{70}$$

Viewing the above equations as a $m \times m$ linear system, we obtain (68). ∎

THEOREM 4.1. (Recursive generation of scattering matrices). *Given scattering matrices* $\{\alpha_i\}$ *for subdomains* $\{D_i\}$, *the scattering matrix* α *of domain* $D = \bigcup_{i=1}^{m} D_i$ *is given by the formula*

$$\alpha = (S_1\alpha_1 \quad S_2\alpha_2 \quad \cdots \quad S_m\alpha_m)\, S_p, \tag{71}$$

where operators $\{S_i\}$ *are defined by* (64), *and the splitting matrix* S_p *is defined by* (67).

Proof. Suppose that φ is an incoming potential on Γ, and σ is the charge distribution induced by φ. Then due to (48), the induced outgoing potential ψ assumes the form

$$\begin{aligned}\psi(x) &= \int_D K(x, t)\cdot q(t)\cdot d\sigma(t) \\ &= \sum_{i=1}^{m} \int_{D_i} K(x, t)\cdot q(t)\cdot d\sigma(t)\end{aligned} \tag{72}$$

for any $x \in \Gamma$.

Combining (64) with (72), we obtain

$$\psi(x) = \sum (S_i \cdot \psi_i)(x). \tag{73}$$

Substituting ψ_i in (73) by $\alpha_i\varphi_i$ (see (62)), we have

$$\begin{aligned}\psi &= \sum S_i\cdot\psi_i = \sum S_i\cdot\alpha_i\cdot\varphi_i \\ &= (S_1\alpha_1 \quad S_2\alpha_2 \quad \cdots \quad S_m\alpha_m)\begin{pmatrix}\varphi_1\\ \varphi_2\\ \vdots\\ \varphi_m\end{pmatrix}.\end{aligned} \tag{74}$$

The conclusion of the theorem follows from the combination of (74), (68), and (54). ∎

4.3. Scattering Matrices in Cantor Sets

Suppose that D is a compact subset of Cantor set C^a, and the set D is enclosed in a domain Ω with its boundary denoted by Γ (see Fig. 8). The boundary Γ will be referred to as a frame boundary.

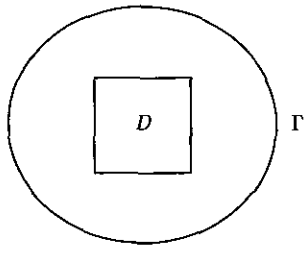

FIG. 8. A subset $D \subset C^a$ enclosed by Γ.

To specifically deal with the problem defined in Section 2 (see (10), (11), (12)), we consider scattering matrices in Cantor sets, with $K(x, t) = \ln \|x - t\|$ and $q(t)$ as the characteristic function of D. Then an incoming potential Φ is harmonic in Ω, and an outgoing potential Ψ is harmonic in $R^2\backslash D$ and satisfies the far-field condition (15) (see Remark 4.1).

Furthermore, the operators $P: L^2(D) \to L^2(R^2\backslash D)$ defined by formula (51) and $S: L^2(D) \to L^2(\Gamma)$ defined by formula (52) assume the form

$$P(\sigma)(x) = \int_D \ln \|x - t\|\, d\sigma(t), \tag{75}$$

$$S(\sigma)(x) = \int_D \ln \|x - t\|\, d\sigma(t). \tag{76}$$

In this section, we construct an analytical apparatus for the efficient representation of scattering matrices of subsets of Cantor sets. In Section 4.3.1 we represent incoming and outgoing potentials in terms of single and double layer distributions; in Section 4.3.2 we describe a merging scheme for the recursive generation of scattering matrices; and in Section 4.3.3 we discuss the discretization of scattering matrices and other related operators.

4.3.1. *Representation of Potentials*

Theorem 4.2 below described the representation of incoming and outgoing potentials in terms of single and double layer distributions. Theorem 4.3 follows immediately from Theorem 4.2. Lemmas 4.2 and 4.3 are obvious (see Section 3.3) and are used in the formulation of Theorem 4.2.

LEMMA 4.2. *Suppose that* φ *is an incoming potential on the boundary* Γ *and* ψ *is an outgoing potential on* Γ. *Then the incoming and outgoing potentials* Φ *and* Ψ *are respectively the solutions of the following two boundary value problems:*

(AA) *Interior Dirichlet problem* (*incoming potential*),

$$\begin{aligned}\Delta\Phi(x) &= 0 && \text{for} \quad x \in \Omega \\ \Phi(x) &= \varphi(\gamma^{-1}(x)) && \text{for} \quad x \in \Gamma\end{aligned} \tag{77}$$

(BB) *Exterior Neumann problem* (*outgoing potential*),

$$\begin{aligned}\Delta\Psi(x) &= 0 && \text{for} \quad x \in R^2\backslash\Omega \\ \frac{\partial}{\partial N}\Psi(x) &= \psi(\gamma^{-1}(x)) && \text{for} \quad x \in \Gamma\end{aligned} \tag{78}$$

with Ψ *satisfying the far-field condition* (15).

LEMMA 4.3. *Suppose that* φ *is an incoming potential on the boundary* Γ *and* ψ *is an outgoing potential on* Γ. *Suppose*

further that a dipole distribution μ_d and a charge distribution μ_s are respectively the solutions of the two integral equations,

$$\pi\mu_d(x)+\int_0^L \frac{\partial}{\partial N(t)}\log\|\gamma(x)-\gamma(t)\|\,\mu_d(t)\,dt=\varphi(x) \quad (79)$$

$$-\pi\mu_s(x)+\frac{\partial}{\partial N(x)}\int_0^L \log\|\gamma(x)-\gamma(t)\|\,\mu_s(t)\,dt=\psi(x). \quad (80)$$

Then the incoming and outgoing potentials can be represented by

$$\Phi(x)=\int_0^L \frac{\partial}{\partial N(t)}\log\|\gamma(x)-\gamma(t)\|\,\mu_d(t)\,dt, \quad (81)$$

$$\Psi(x)=\int_0^L \log\|\gamma(x)-\gamma(t)\|\,\mu_s(t)\,dt. \quad (82)$$

Theorem 4.2 below can be viewed as a reformulation of Lemma 4.3 in the operator notation. First, we define four operators

$$Q_d: L^2(\Gamma)\to L^2(\Omega),$$
$$Q_s: L^2(\Gamma)\to L^2(R^2\backslash\Omega),$$
$$P_d: L^2(\Gamma)\to L^2(\Gamma),$$
$$P_s: L^2(\Gamma)\to L^2(\Gamma),$$

via

$$Q_d(\mu_d)(x)=\int_0^L \frac{\partial}{\partial N(t)}\log\|\gamma(x)-\gamma(t)\|\,\mu_d(t)\,dt, \quad (83)$$

$$Q_s(\mu_s)(x)=\int_0^L \log\|\gamma(x)-\gamma(t)\|\,\mu_s(t)\,dt, \quad (84)$$

$$P_d(\mu_d)(x)=\pi\mu_d(x)+\int_0^L \frac{\partial}{\partial N(t)}\log\|\gamma(x)-\gamma(t)\|\times\mu_d(t)\,dt, \quad (85)$$

$$P_s(\mu_s)(x)=-\pi\mu_s(x)+\frac{\partial}{\partial N(x)}\int_0^L \log\|\gamma(x)-\gamma(t)\|\times\mu_s(t)\,dt. \quad (86)$$

THEOREM 4.2. *Suppose that φ is an incoming potential on the boundary Γ and ψ is an outgoing potential on Γ. Then the incoming and outgoing potentials can be represented via*

$$\Phi=Q_dP_d^{-1}\varphi, \quad (87)$$

$$\Psi=Q_sP_s^{-1}\psi, \quad (88)$$

where operators Q_s, Q_d, P_s, and P_d are defined by (84), (83), (86), *and* (85), *respectively.*

THEOREM 4.3. *Suppose that $L: L^2(\Gamma)\to L^2(D)$ is the operator defined by* (50). *Then*

$$L\varphi=(Q_dP_d^{-1}\varphi)|_D, \quad (89)$$

with the operators P_d and Q_d defined by (85) *and* (83), *respectively.*

4.3.2. *Recursive Generation of Scattering Matrices*

Suppose that $\{D_1, D_2, D_3, D_4\}\subset C^a$ are four subsets (boxes) resulting from the subdivision of a bigger subset (box), and the set D is the union the four subsets $\{\{D_i\}$,

$$D=\bigcup_{i=1}^{4} D_i. \quad (90)$$

Suppose further that D is enclosed in a square Ω with its boundary denoted by Γ. The square Ω will be referred to as frame domain (box) while Γ will be referred to as the frame boundary.

Suppose that for any integer i $(1\leqslant i\leqslant m)$, the set D_i is enclosed in a square Ω_i with its boundary denoted by Γ_i see Fig. 9). Within the tree structure of the Cantor set C^a (see Section 2), we will refer to the frame boxes of neighbor boxes as *neighbor frame boxes* and the frame box of a parent box as *parent frame box.*

In this section, we obtain the scattering matrix for D from scattering matrices for domains $\{D_1, D_2, D_3, D_4\}$. First, we need the representations of operators

$$L_i: L^2(\Gamma)\to L^2(\Gamma_i),$$
$$S_i: L^2(\Gamma_i)\to L^2(\Gamma),$$
$$L_{ji}: L^2(\Gamma_i)\to L^2(\Gamma_j),$$

defined by formulae (63), (64), and (65), respectively.

Theorems 4.4 and 4.5 below follow from Theorem 4.2 immediately, Lemma 4.4 is an immediate consequence of Lemma 4.1, and Theorem 4.6 is a consequence of Theorem 4.1 and Lemma 4.4. Theorems 4.4 and 4.5 describe

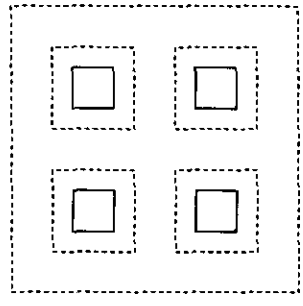

FIG. 9. Four subsets of C^a and their frame boxes.

representations of the operators L_i, S_i, and L_{ij} defined by formulae (63), (64), and (65). Theorem 4.6 describes a merging scheme for the recursive generation of scattering matrices for subsets of Cantor sets.

THEOREM 4.4. *Suppose that* $1 \leqslant i \leqslant m$, *and* $L_i: L^2(\Gamma) \to L^2(\Gamma_i)$ *is the operator defined by* (63). *Then*

$$L_i \varphi = (Q_d P_d^{-1} \varphi)|_{\Gamma_i}, \tag{91}$$

with operators Q_d *and* P_d *defined by* (83) *and* (85), *respectively.*

We will define the operators $Q_{s,i}$ and $P_{s,i}$ and observe that they are similar to the operators Q_s in (84) and P_s in (86) with region Ω replaced with region Ω_i,

$$Q_{s,i}: L^2(\Gamma_i) \to L^2(R^2 \backslash \Omega_i),$$
$$P_{s,i}: L^2(\Gamma_i) \to L^2(\Gamma_i),$$

by

$$Q_{s,i}(\sigma_s)(x) = \int_0^{L_i} \log \|\gamma_i(x) - \gamma_i(t)\| \, \sigma_s(t) \, dt, \tag{92}$$

$$P_{s,i}(\sigma_s)(x) = -\pi \sigma_s(x) + \frac{\partial}{\partial N_i(x)} \int_0^{L_i} \log \|\gamma_i(x) - \gamma_i(t)\| \times \sigma_s(t) \, dt. \tag{93}$$

THEOREM 4.5. *Suppose that for* $1 \leqslant i, j \leqslant 4$, *the operator*

$$S_i: L^2(\Gamma_i) \to L^2(\Gamma)$$

is defined by (64) *and the operator*

$$L_{ji}: L^2(\Gamma_i) \to L^2(\Gamma_j)) \qquad \text{for} \quad i \neq j$$

is defined by (65). *Then*

$$S_i \psi = (Q_{s,i} P_{s,i}^{-1} \psi)|_{\Gamma}, \tag{94}$$

$$L_{ji} \psi = (Q_{s,i} P_{s,i}^{-1} \psi)|_{\Gamma_j}, \tag{95}$$

with operators $Q_{s,i}$ *and* $P_{s,i}$ *defined by* (92) *and* (93), *respectively.*

Remark 4.3. Suppose that the operator

$$U: \begin{pmatrix} L^2(\Gamma_1) \\ L^2(\Gamma_2) \\ L^2(\Gamma_3) \\ L^2(\Gamma_4) \end{pmatrix} \to \begin{pmatrix} L^2(\Gamma_1) \\ L^2(\Gamma_2) \\ L^2(\Gamma_3) \\ L^2(\Gamma_4) \end{pmatrix}$$

is defined by

$$U = \begin{pmatrix} I & -L_{12}\alpha_2 & -L_{13}\alpha_3 & -L_{14}\alpha_4 \\ -L_{21}\alpha_1 & I & -L_{23}\alpha_3 & -L_{24}\alpha_4 \\ -L_{31}\alpha_1 & -L_{32}\alpha_2 & I & -L_{24}\alpha_4 \\ -L_{41}\alpha_1 & -L_{24}\alpha_2 & -L_{43}\alpha_3 & I \end{pmatrix}. \tag{96}$$

Then the splitting matrix (see (67))

$$S_p: L^2(\Gamma) \to \begin{pmatrix} L^2(\Gamma_1) \\ L^2(\Gamma_2) \\ L^2(\Gamma_3) \\ L^2(\Gamma_4) \end{pmatrix}$$

is given by

$$S_p = U^{-1} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \end{pmatrix}, \tag{97}$$

with L_i defined by (91) for $1 \leqslant i \leqslant 4$ and L_{ij} defined by (95) for $1 \leqslant i, j \leqslant 4$.

LEMMA 4.4. *Suppose that* φ *is an incoming potential on the boundary* Γ *and that for* $1 \leqslant i \leqslant 4$, φ_i *is the total incoming potential on* Γ_i, *as defined by* (60). *Then*

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \end{pmatrix} = S_p \cdot \varphi, \tag{98}$$

with the splitting matrix S_p *defined by* (97). *In other words, the splitting matrix converts the incoming potential* φ *on* Γ *into the total incoming potentials* $\{\varphi_i\}$ *on boundaries* $\{\Gamma_i\}$.

THEOREM 4.6. *Suppose that for all* $1 \leqslant i \leqslant 4$ *the scattering matrix for the compact set* D_i *is denoted by* α_i (*see* (90)). *Then the scattering matrix* α *for the set* $D = \bigcup_{i=1}^4 D_i$ *is given by the merging formula*

$$\alpha = (S_1\alpha_1 \quad S_2\alpha_2 \quad S_3\alpha_3 \quad S_4\alpha_4) \, S_p, \tag{99}$$

with operator S_i *defined by formula* (94) *for* $1 \leqslant i \leqslant 4$ *and the splitting matrix* S_p *defined by formula* (97).

Remark 4.4. Due to the self-similarity in Cantor sets (see Section 2), we only need to compute one scattering matrix per level, even though there are 4^l subsets at the lth level.

4.3.3. *Discretization of Scattering Matrices*

From the preceding sections, it is clear that the construction of scattering matrices for subsets of Cantor sets, either

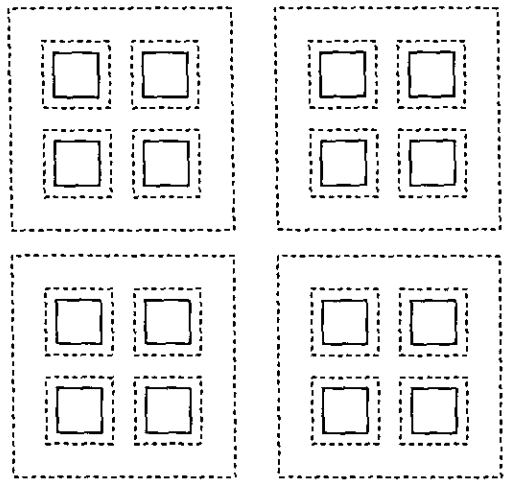

FIG. 10. Frame boxes for Cantor set C^a.

directly or recursively, depends on the choice of frame boundaries (see (54) and (99)). For Cantor set C^a, we recursively generate frame boxes for subsets of a Cantor set C^a such that frame boxes at the same level are mutually disjoint and the distance between a box and its frame box equals the distance between two neighbor frame boxes (see Fig. 10).

Based on the above choice of frame boxes, Tables I and II list the number of Legendre nodes needed on frame boundaries for the representation of incoming and outgoing potentials to single and double precision, respectively.

Remark 4.5. *Two observations can be easily made from Tables I and II:*

1. *The number of Legendre nodes needed to obtain double precision is only twice the number of nodes needed for single precision.*

2. *The number of nodes needed increases rapidly with the increase in ratio a.*

With the choice of frame boundaries and Legendre nodes on the boundaries, we are ready to discretize operators P in (75), S in (76), P_d in (85), $P_{s,i}$ in (93), Q_d in (83), $Q_{s,i}$ in (92), L in (89), L_i in (91), S_i in (94), L_{ij} in (95), and S_p in (97) with $1 \leqslant i, j \leqslant 4$.

There are three types of discretization:

1. Operators P and S in (75) and (76) are discretized as follows. Suppose that $\{z_1, z_2, ..., z_m\}$ is the approximation

TABLE I

Number of Legendre Nodes for Single Precision
$\varepsilon = 10^{-7}$ (Absolute Error)

Ratio a	0.1	0.2	0.3	0.35	0.4	0.45
Number of nodes per side	12	18	30	46	66	100

TABLE II

Number of Legendre Nodes for Double Precision
$\varepsilon = 10^{-14}$ (Absolute Error)

Ratio a	0.1	0.2	0.3	0.35	0.4	0.45
Number of nodes per side	30	44	60	80	120	200

to the compact subset $D \subset \Omega$, and $\{x_1, x_2, ..., x_p\}$ are the Legendre nodes on the boundary Γ of domain Ω. The discretization of P defined by (75) is the matrix $\tilde{P}$ defined by

$$(\tilde{P})_{ij} = \ln \|z_i - z_j\| \quad \text{for} \quad i \neq j, \quad 1 \leqslant i, j \leqslant m, \tag{100}$$

$$(\tilde{P})_{ii} = A_{ii}, \tag{101}$$

where A_{ii} is defined by (12). Similarly, discretizing the operator S defined by (76), we otain the matrix $\tilde{S}$ defined by

$$(\tilde{S})_{ij} = \ln \|x_i - z_j\| \quad \text{for} \quad 1 \leqslant i \leqslant p, \quad 1 \leqslant j \leqslant m. \tag{102}$$

2. Operators P_d in (85) and $P_{s,i}$ in (93) are discretized by the Galerkin method described in Section 3.4, and operators Q_d in (83) and $Q_{s,i}$ in (92) are discretized by the Gaussian quadrature rule based on Legendre nodes.

3. Operators L, L_i, S_i, L_{ij}, and S_p with $1 \leqslant i, j \leqslant 4$, defined in formulae (89), (91), (94), (95), and (97), are combinations of operators P_d in (85), Q_d in (83), $Q_{s,i}$ in (92), and $P_{s,i}$ in (93). Therefore, the discretization of the operators L, L_i, S_i, L_{ij}, and S_p need not be defined separately. Therefore, a scattering matrix can be computed either directly via formula (54) or recursively via formula (99).

Theorem 4.7 is the immediate consequence of Theorem 2.2.

THEOREM 4.7. *With the discretization described above, the matrix U defined by* (96) *is non-singular. Therefore, the splitting matrix in* (97) *is uniquely defined.*

5. THE FAST DIRECT ALGORITHM

In this section, we describe a direct algorithm for the rapid solution of the Laplace equation on regions with fractal boundaries. The algorithm exploits the fact that for any given ratio a, interactions at any level in the Cantor set C^a are of low rank (the ranks depend only on the constant ratio a for generating the Cantor set and do not depend on the sizes of boxes and number of points inside). The low rank of interactions is reflected in the coefficient matrix in Eq. (10) as the low rank of its off-diagonal submatrices (see Fig. 11). Thus, we can recursively compress these matrices of low rank without actually generating them.

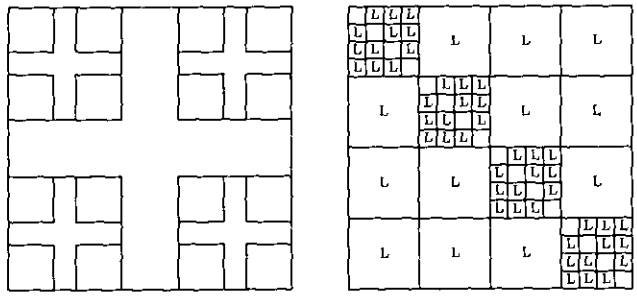

FIG. 11. Approximation of Cantor set and coefficient matrix.

To be more specific, let us consider four subsets (boxes) in a Cantor set C^a, depicted in Fig. 12. They are boxes of size d, and the distance between any two of them is $(1-2a)d$. The interactions between them are of low rank (see Section 3.6) and can be represented via scattering matrices (see Section 4.3).

Starting with the hierarchical structure of a Cantor set C^a (see Section 2), we proceed by introducing a set of frame boxes arranged in a tree structure (see Section 4.3). For a given precision ε, we determine the number of Legendre nodes needed on frame boundaries for the representation of potentials (see Tables I and II).

To describe the algorithm, we need the following notation:

L, number of levels in the approximation of the Cantor set C^a (see Section 2).

N, $N=4^L$, the size of the approximation of the Cantor set C^a (see Section 2).

p, number of Legendre nodes on each side of frame boundaries for the representation of potentials to a given precision ε (see Section 4.3.3).

L_1, number of a level on which a scattering martrix is computed directly.

m, $=4^{L-L_1}$, size of linear systems to be solved directly.

A_f, restriction of matrix A defined by (11) onto a subset *ibox* at level L_1. In other words, $(A_f)_{ij}=\ln\|z_i-z_j\|$ and $(A_f)_{ii}=A_{ii}$, where points $\{z_1, ..., z_m\}\subset ibox$, and A_{ii} is defined by (12).

The fast direct algorithm is a two-pass procedure. In the first (bottom-up) pass, we compute the scattering matrix for level L_1 directly via formula (54) and scattering matrices for all coarser levels (level number $<L_1$) by using the merging scheme described in Theorem 4.6. In the second (top-down) pass, we generate total incoming potentials on frame boundaries up to level L_1 by using Lemma 4.4. Finally, we solve 4^{L_1} small-scale linear systems of size $m\times m$ directly at level L_1. The following is a formal description of the algorithm.

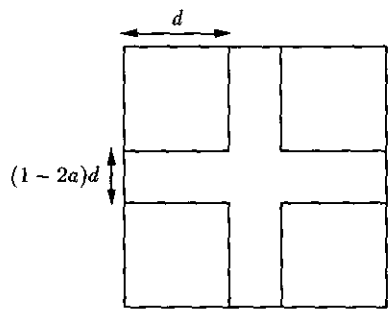

FIG. 12. Four subsets of Cantor set C^a.

THE ALGORITHM.

Initialization

Comment (Computations in the initialization are done once for all.)

Step 1

Comment (Given a real number a $(0<a<\frac{1}{2})$ and an integer L, construct the approximation of Cantor set C^a and its frame boxes)

do $lev=0, 1, 2, ..., L$
  do $ibox=1, 2, ..., 4^{lev}$
    Divide each box into four corner boxes according to the constant a (see Section 2).
    Construct the frame box for *ibox* (see Section 4.3).
  endo
endo
do $lev=L$
  do $ibox=1, 2, ..., 4^L$
    Compute the center of *ibox*.
  enddo
enddo

Step 2

Comment (For a given precision ε, precompute operators P_d^{-1}, P_s^{-1}, and A_f^{-1})

do
  Determine number of Legendre nodes p (see Tables I and II)
  Generate the inverse of operators P_d and P_s defined by (85) and (86), via the classical Galerkin method (see Section 3.4).
  Compute the inverse of $m\times m$ matrix A_f directly.
enddo

Upward Pass

Comment (Compute scattering matrices and splitting matrices)

Step 3

do $lev=L_1$
  Compute the discretized operators L and S defined by (89) and (76).
  Compute the scattering matrix directly via formula (54): $\alpha=SA_f^{-1}L$.
endo

Step 4

do $lev = L_1 - 1, L_1 - 2, ..., 1, 0$
  do $i = 1, ..., 4$
    Compute operators L_i and S_i (defined by formulae (91) and (94)) via the Gaussian quadrature rule based on Legendre nodes.
    do $j = 1, ..., 4$
      Compute L_{ij} defined by (95) via the Gaussian quadrature rule.
    endo
  endo
  Compute the splitting matrix S_p by using formula (97).
  Compute the scattering matrix α via the merging scheme in Theorem 4.6.
  endo

Downward Pass

Comment (Splitting matrices are now available. Compute total incoming potentials on all frame boundaries up to level L_1)

Step 5

  do $lev = 1, 2, ..., L_1$
    do $ibox = 1, 2, ..., 4^{lev}$
      Compute total incoming potential φ_i by formula (98)
  enddo
endo

Step 6

do $ibox = 1, 2, ..., 4^{L_1}$
  Solve $m \times m$ linear system directly by computing $\sigma_i = A_f^{-1} L \varphi_i$, where operator L has been computed at Step 3.
endo

Remark 5.1. Suppose that ibox is a fixed box at level l. Then in the splitting process, the total incoming potential on the frame boundary of *ibox* can be computed independently from those on the other frame boundaries of boxes at the same level. Thus, we can obtain a part of the solution independently from the rest of the solution if only a part of the solution is desired.

A brief analysis of the algorithmic complexity is given below.

Step Number	Operation count	Explanation
Step 1	$O(N)$	$4N$ boxes (squares) are involved. Each box is determined by its center and size.
Step 2	$O(m^3 + p^3)$	Operators P_d and P_s are of size $4p \times 4p$, and the operator A_f is of size $m \times m$.
Step 3	$O(mp + m^2 p)$	Operator L is of size $m \times 4p$. Operator S is of size $4p \times m$. Operator A_f^{-1} is of size $m \times m$.
Step 4	$O(p^3 \log N)$	Operator L_i, S_i, and L_{ij} is of size $4p \times 4p$. S_p is of size $16p \times 4p$. Operator α is of size $4p \times 4p$. There are $\log N$ levels.
Step 5	$O(p^3 N)$	The computation of the total incoming potential φ_i on each frame boundary requires p^3 operations. There are 4^{L_1+1} ($< N$) frame boundaries involved.
Step 6	$O(m^2 pN)$	Operator A_f^{-1} is of size $m \times m$. Operator L is of size $m \times 4p$. Potential φ_i is a vector of size $4p$. Computations of $A_f^{-1} L \varphi_i$ are done 4^{L_1} ($< N$) times.

The time complexity of the algorithm is, therefore,

$$(\beta_1 p^3 + \beta_2 m^2 p) \cdot N + \beta_3 p^3 \cdot \log N, \tag{103}$$

where the constant m is normally chosen to be $m \sim 256$, the constant p depends on the geometry of a given fractal boundary, and the choice of frame boundaries (see Tables I and II) and the constants β_1, β_2, and β_3 depend on the computer system, implementation, language, etc.

The remark below can be obtained easily from formulae (88), (86), (84), and (55).

Remark 5.2. Given scattering matrices and total incoming potentials, the evaluation of the potential

$$\Psi(x) = \int_{C^a} \ln \|x - t\| \, d\sigma(t) \tag{104}$$

at any point $x \in R^2 \backslash C^a$ requires at most $O(\log N)$ operations, where σ is the charge distribution over C^a.

6. NUMERICAL RESULTS

A computer program has been implemented utilizing the algorithm of this paper, and it is capable of computing either whole or part of the solution and of evaluating the potential at any point.

For this paper, we considered fractal boundaries of Cantor type with ratios 0.1, 0.3, and 0.45. In the first two experiments ($a = 0.1, 0.3$), the potentials on frame boundaries are represented to double precision, while in the third experiment ($a = 0.45$) they are represented to 10 digits. The size of linear systems inverted directly at the final stage has been chosen to be $m = 256$. All calculations have been conducted on a Sparc II workstation.

TABLE III

Comparison of Timings

N	Levels	T_{alg} (min)	$T_{CG\&FMM}$ (h)	T_{GE} (est.)
4,096	6	6	0.4	19.1 h
16,384	7	9	3.3	51 days
65,536	8	11	26.4 (est.)	9 years
262,144	9	14	211.2 (est.)	572 years
1,048,576	10	19	1689.6 (est.)	366,283 years

Note. $a = 0.1$, $p = 30$, $m = 256$, $\varepsilon = 10^{-14}$

TABLE V

Comparison of Timings

N	Levels	T_{alg} (min)	$T_{CG\&FMM}$ (h)	T_{GE} (est.)
4,096	6	122	1.6	19.1 h
16,384	7	181	13.0	51 days
65,536	8	240	103.7 (est.)	9 years
262,144	9	297	829.3 (est.)	572 years
1,048,576	10	360	6634.2 (est.)	366,283 years

Note. $a = 0.45$, $p = 80$, $m = 256$, $\varepsilon = 10^{-10}$

The results are summarized in Tables III, IV, and V. The first column is the size of the approximation to a Cantor set. The second column is the number of levels in the generation of Cantor set. The third column is the actual CPU time of the fast direct algorithm of the preceding section. The fourth column is either the CPU time or estimated CPU time of the combined algorithm, the conjugate gradient (CG) algorithm combined with the fast multipole method (FMM) (see [21]). The last column is the estimated timing for the Gaussian elimination (it is given here only for comparison purposes).

The following observations can be made from Tables III, IV, and V:

1. Although our fast algorithm asymptotically requires $O(N)$ operations, the actual running time of the algorithm as observed from the numerical experiments seems to behave like $\log N$, due to the fact that the constant β_3 in (103) is rather large compared to the constants β_1 and β_2. The constant β_3 in (103) can be substantially reduced either by a better choice of frame boundaries, or by improving the representation of potentials.

2. For any $N \geq 4096$, our algorithm is faster than the combined algorithm (the CG method combined with the FMM, see [21]).

3. The performance of our algorithm deteriorates with the increase in ratio a as is expected (see Tables I and II).

TABLE IV

Comparison of Timings

N	Levels	T_{alg} (min)	$T_{CG\&FMM}$ (h)	T_{GE} (est.)
4,096	6	45	1.2	19.1 h
16,384	7	67	8.3	51 days
65,536	8	88	66.7 (est.)	9 years
262,144	9	110	533.8 (est.)	572 years
1,048,576	10	134	4270.3 (est.)	366,283 years

Note. $a = 0.3$, $p = 60$, $m = 256$, $\varepsilon = 10^{-14}$

The following is a recapitulation of the other notation to be used in the illustration of our numerical experiments:

a —ratio for generating a Cantor set.

p —number of Legendre nodes on each side of frame boundaries.

m —size of linear systems inverted directly.

ε —precision of incoming and outgoing potentials.

N —size of the linear system (10) to be solved.

7. CONCLUSIONS

An $O(N)$ direct algorithm has been constructed for the rapid solution of the Laplace equation on regions with fractal boundaries, where N is the number of elements in the discretization of the fractal. In the algorithm, operators of low rank are recursively compressed and the inverse is constructed in a compressed form so that it can be applied to a vector rapidly. The algorithm is capable of generating only a part of the solution if desired. Evaluation of the potential at any point requires $O(\log N)$ operations. Numerical examples presented in Section 6 indicate that even very large-scale problems result in acceptable CPU time requirements. In the presented paper, a two-dimensional version of the algorithm is described. Generalizing this result to three dimensions is fairly straightforward and will be reported at a later date.

REFERENCES

1. E. Aurell, M. Benedicks, P. Jones, and S. Grossman, *Harmonic Measure on a Fractal*, Annual Review of Swedish National Research Council, May 1992 (unpublished).
2. M. Barnsley, *Fractals Everywhere* (Academic Press, New York/London, 1988).
3. G. Beylkin, R. Coifman, and V. Rokhlin, *Commun. Pure Appl. Math.* **14**, 141 (1991).
4. P. Billingsley, *Ergodic Theory and Information* (Wiley, New York/London/Sydney, 1965).
5. G. Birkhoff and R. Lynch, *Numerical Solution of Elliptic Problems*, SIAM Studies in Applied Mathematics (SIAM, Philadelphia, 1984).

6. J. Carrier, L. Greengard, and V. Rokhlin, *SIAM J. Sci. Stat. Comput.* **9**, No. (4) (1988).
7. L. Carleson, *Ann. Acad. Sci. Fenn.* **10**, 113 (1985).
8. R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol. II* (Wiley Interscience, New York, 1953).
9. J. L. Doob, *Classical Potential Theory and Its Probabilistic Counterparts* (Springer-Verlag, New York, 1983).
10. G. A. Edgar, *Measure, Topology, and Fractal Geometry* (Springer-Verlag, New York, 1990).
11. C. A. J. Fletcher, *Computational Galerkin Methods* (Springer-Verlag, New York, 1984).
12. L. Greengard and V. Rokhlin, *J. Comput. Phys.* **73**, 325 (1987).
13. L. Greengard and V. Rokhlin, *Commun. Pure Appl. Math.* **14**, 419 (1991).
14. P. Jones and T. Wolff, *Acta Math.* **161** (1988).
15. P. Lax and R. Phillips, *Scattering Theory* (Academic Press, New York, 1967).
16. J. Ma, Ph.D. thesis, Yale University, November 1992.
17. N. G. Makarov, in *Proceedings, Int. Cong. Math., Berkeley, CA, 1986.*
18. B. Mandelbrot, *Fractals: Form, Chance and Dimension* (Freeman, San Francisco, 1977).
19. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
20. I. G. Petrovsky, *Lectures on Partial Differential Equations* (Dover, New York, 1991).
21. V. Rokhlin, *J. Comput. Phys. 60* 187 (1985).
22. J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis* (Springer-Verlag, New York, 1980).
23. H. Takayasu, *Fractals in the Physical Sciences* (Manchester Univ. Press, Manchester/New York, 1990).
24. M. Tsuji, *Potential Theory in Modern Function Theory* (Stevens, New York, 1959).
25. T. Vicsek, *Fractal Growth Phenomena* (World Scientific, Singapore, 1989).
26. J. Wermer, *Potential Theory*, Lecture Notes in Mathematics, Vol. 408 (Springer-Verlag, New York, 1981).