Generalized fast multipole method

2010 IOP Conf. Ser.: Mater. Sci. Eng. 10 012230

(http://iopscience.iop.org/1757-899X/10/1/012230)

IOP Conf. Series: Materials Science and Engineering **10** (2010) 012230 doi:10.1088/1757-899X/10/1/012230

Generalized Fast Multipole Method

Pierre-David Létourneau, Cristopher Cecka, and Eric Darve
496 Lomita Mall, Stanford CA 94305, USA

E-mail: darve@stanford.edu

Abstract. The fast multipole method (FMM) is a technique allowing the fast calculation of long-range interactions between N points in $O(N)$ or $O(N \ln N)$ steps with some prescribed error tolerance. The FMM has found many applications in the field of integral equations and boundary element methods, in particular by accelerating the solution of dense linear systems arising from such formulations. Original FMMs required analytical expansions of the kernel, for example using spherical harmonics or Taylor expansions. In recent years, the range of applicability and the ease of use of FMMs has been extended by the introduction of black box [1] or kernel independent techniques [2]. In these approaches, the user only provides a subroutine to numerically calculate the interaction kernel. This allows changing the definition of the kernel with minimal change to the computer program. In this talk we will present a novel kernel independent FMM, which leads to diagonal multipole-to-local operators. This results in a significant reduction in the computational cost [1], in particular when high accuracy is needed. The approach is based on Cauchy's integral formula and the Laplace transform. We will present a numerical analysis of the convergence, methods to choose the parameters in the FMM given some tolerance, and the steps required to build a multilevel scheme from the single level formulation. Numerical results are given for benchmark calculations to demonstrate the accuracy as a function of the number of multipole coefficients, and the computational cost of the different steps in the method.

1. Introduction

The fast multipole method (FMM) is a general class of methods to reduce the cost of computing:

$$\phi_i = \sum_{j=1}^{N} K(r_i, r_j)\, \sigma_j, \quad 1 \le i \le N \tag{1}$$

when N is large. The basic approximation strategy in the FMM is a low-rank approximation of the kernel of the type:

$$K(r, r^*) = \sum_{m=1}^{p} \sum_{q=1}^{p} u_m(r)\, T_{mq}\, v_q(r^*) + \varepsilon$$

With a low-rank approximation of this type one can construct an $O(N)$ or $O(N \ln N)$ method to calculate the sum in Eq. (1).

Fast multipole methods have been derived for many different types of kernels, including the electrostatic kernel $1/r$ and Helmholtz kernel e^{ikr}/r. Efforts have been made to extend the method to more general kernels [2]. The authors for example have developed a fast technique applicable to any non-oscillatory kernels [1]. In this paper, we consider the creation of an FMM with two goals in mind:

(i) The method should be applicable to a wide class of kernels.

(ii) The multipole-to-local operator T_{mq} should be diagonal.

Even though requirement (i) is satisfied by the method in [1], requirement (ii) is not. The second requirement allows in principle reducing the computational cost of the method by reducing the number of operations involved in the multiplication of v_q by T_{mq}. This has the potential to significantly reduce the cost of these general FMM schemes to improve their applicability and efficiency.

Some of the applications we have in mind include the use of radial basis functions such as r, r^n (n odd), $r^n \log r$ (n even), $\exp(-cr^2)$, $\sqrt{r^2+c^2}$, $1/\sqrt{r^2+c^2}$, $1/(r^2+c^2)$, ..., for interpolation schemes. These are popular schemes for mesh deformation [3] and graphics applications [4]. Interpolation using these functions requires generalized FMMs capable of handling a wide class of functions.

The method we are proposing is based on Cauchy's integral formula for analytic functions. This is the class of kernels that is covered by this new scheme. In this paper, we are only presenting an outline of the method. Additional details will be presented in future publications. This new technique offers a general framework to construct FMMs. Depending on the specific nature of the kernel, various optimizations and modifications can be made to this general scheme to improve its computational efficiency and accuracy.

2. Cauchy's integral formula and low-rank approximations

For simplicity, let us assume that $K(r, r^*)$ is of the type $K(r-r^*)$ (translation independent). In addition we will start with the 1D case. The 2D and 3D extensions will be discussed later on.

We now therefore consider a kernel $K(x)$, $x \in \mathbb{R}$, and assume that, in some region R around the point x, it is an analytic (holomorphic) function, that is the function $K(z)$, $z \in \mathbb{C}$, is complex differentiable at every point in R. Then, by Cauchy's integral formula:

$$K(x) = \frac{1}{2\pi i} \oint_{\Gamma} \frac{K(z)}{z-x}\, dz$$

The curve Γ is closed and contains x.

Assume that $\mathrm{Re}(z-x) > 0$ (Re is the real part of a complex number) then:

$$\frac{1}{z-x} = \int_0^\infty e^{-s(z-x)}\, ds$$

The reason why we want to use this formula is that after approximating these integrals using a numerical quadrature, we will obtain a low rank approximation of the kernel, and later on a fast method.

Since Γ encloses x it is not always the case that $\mathrm{Re}(z-x) > 0$. However, by applying rotations in the complex plane we get:

$$\frac{1}{z-x} = -\int_0^\infty e^{s(z-x)}\, ds \qquad \text{if } \mathrm{Re}(z-x) < 0$$

$$\frac{1}{z-x} = -i\int_0^\infty e^{is(z-x)}\, ds \qquad \text{if } \mathrm{Im}(z-x) > 0$$

$$\frac{1}{z-x} = i\int_0^\infty e^{-is(z-x)}\, ds \qquad \text{if } \mathrm{Im}(z-x) < 0$$

Let us decompose Γ into four curves such that on Γ_1 $\mathrm{Re}(z-x)>0$, on Γ_2 $\mathrm{Im}(z-x)>0$, on Γ_3 $\mathrm{Re}(z-x)<0$, and on Γ_4 $\mathrm{Im}(z-x)<0$. Then:

$$\begin{aligned} K(x) = \quad & \frac{1}{2\pi i}\int_{\Gamma_1} K(z)\int_0^\infty e^{-s(z-x)}\,ds dz - \frac{1}{2\pi}\int_{\Gamma_2} K(z)\int_0^\infty e^{is(z-x)}\,ds dz \\ & - \frac{1}{2\pi i}\int_{\Gamma_3} K(z)\int_0^\infty e^{s(z-x)}\,ds dz + \frac{1}{2\pi}\int_{\Gamma_4} K(z)\int_0^\infty e^{-is(z-x)}\,ds dz \end{aligned} \tag{2}$$

This is illustrated in Figure 1.

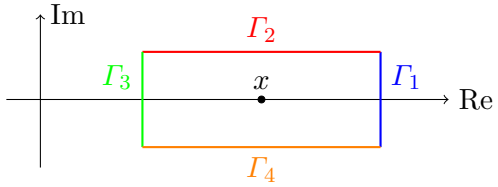

Figure 1. Schematic of the four contour curves around x.

How can we use this formula to construct a fast $O(N)$ method? We need to approximate $K(x-y)$ using a low rank decomposition. Let us consider the first integral along Γ_1. The variable x is now replaced by $x-y$. We assume we have obtained a quadrature along s, with weights w_q and points s_q, that approximates the integral along s. Then the contribution from Γ_1 to $K(x-y)$ can be approximated as:

$$\sum_q \left[\frac{w_q}{2\pi i}\int_{\Gamma_1} K(z)e^{-s_q z}\,dz\right] e^{s_q x}\, e^{-s_q y}$$

Denote $u_q(x) = e^{s_q x}$, $v_q(y) = e^{-s_q y}$, $T_{qq} = (w_q/2\pi i)\int_{\Gamma_1} K(z)\,e^{-s_q z}\,dz$, and we see that this approximation is a low-rank approximation of the type (1). The multipole-to-local operator T_{qq} is diagonal. We immediately point out that this is not sufficient to construct an FMM scheme since we have not provided a method to gather multipole expansions from leaf nodes and propagate them up the tree, and a method to scatter local expansions from the root of the tree down to the leaves. However this formula provides the starting point for our analysis.

3. Connection with Fourier and Laplace transforms

The formula (2) can be viewed as an extension of the Fourier and Laplace transforms. Specifically, when the kernel satisfies some additional conditions at infinity, Eq. (2) can be related to Fourier and Laplace transforms. To understand the connection, let us assume that the paths Γ_i are straight segments:

$$\begin{aligned} K(x) = \quad & \frac{1}{2\pi}\int_{-a}^{a} K(c+it)\int_0^\infty e^{-ist}e^{-s(c-x)}\,ds dt + \frac{1}{2\pi}\int_b^c K(t+ia)\int_0^\infty e^{ist}e^{-isx-sa}\,ds dt \\ & + \frac{1}{2\pi}\int_{-a}^{a} K(b+it)\int_0^\infty e^{ist}e^{s(b-x)}\,ds dt + \frac{1}{2\pi}\int_b^c K(t-ia)\int_0^\infty e^{-ist}e^{isx-sa}\,ds dt \end{aligned} \tag{3}$$

See Fig. 2 for the notations.

The Fourier or Laplace transforms can be recognized provided the kernel K satisfies some additional assumptions. Assume that $\lim_{c\to\pm\infty} K(c+it)/c = 0$, then the first and third integral in (3) become vanishingly small, e.g.:

$$\lim_{c\to\infty} K(c+it)\int_0^\infty e^{-ist}e^{-s(c-x)}\,ds = 0$$

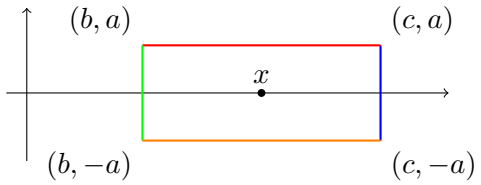

Figure 2. Notations for straight segment contours.

If $K(t+ia)$ is in $\mathrm{L}^1(\mathbb{R})$ with respect to t, then we can choose $b=-\infty$ and $c=\infty$:

$$K(x) = \frac{1}{2\pi}\Big(\int_{-\infty}^{\infty} K(t+ia)\int_0^{\infty} e^{ist}e^{-isx-sa}\,ds\,dt + \int_{-\infty}^{\infty} K(t-ia)\int_0^{\infty} e^{-ist}e^{isx-sa}\,ds\,dt\Big)$$
$$= \frac{1}{2\pi}\Big(\int_0^{\infty} e^{-isx-sa}\int_{-\infty}^{\infty} K(t+ia)e^{ist}\,dt\,ds + \int_0^{\infty} e^{isx-sa}\int_{-\infty}^{\infty} K(t-ia)e^{-ist}\,dt\,ds\Big)$$

After taking the limit $a\to 0$ and doing a change of variable $s\to 2\pi s$, we get:

$$K(x) = \int_{-\infty}^{0} e^{2\pi isx}\int_{-\infty}^{\infty} K(t)e^{-2\pi ist}\,dt\,ds + \int_0^{\infty} e^{2\pi isx}\int_{-\infty}^{\infty} K(t)e^{-2\pi ist}\,dt\,ds$$
$$= \int_{-\infty}^{\infty} e^{2\pi isx}\int_{-\infty}^{\infty} K(t)e^{-2\pi ist}\,dt\,ds$$

This is the formula for the Fourier transform and its inverse transform.

Similarly let us assume that $K(z)\to 0$ when $|z|\to\infty$ and $\mathrm{Re}(z)\geq b$. In that case we can choose $a\to\infty$ and $c\in\Theta(a)$ (bounded above and below by a, i.e., c and a go to infinity at the same rate). Then:

$$\lim_{\substack{a\to\infty\\ c\to\infty}} \frac{1}{2\pi}\int_{-a}^{a} K(c+it)\int_0^{\infty} e^{-ist}e^{-s(c-x)}\,ds\,dt = 0$$

and similarly for the second and fourth integral in Eq. (3). We are only left with the third integral:

$$K(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} K(b+it)\int_0^{\infty} e^{ist}e^{s(b-x)}\,ds\,dt$$

If we change the order of integration, assume that $K(b+it)$ is in $\mathrm{L}^1(\mathbb{R})$ with respect to t, and do the change of variable $z=b+it$, then:

$$K(x) = \frac{1}{2\pi i}\int_0^{\infty} e^{-sx}\int_{b-i\infty}^{b+i\infty} K(z)e^{sz}\,dz\,ds$$

We recognize the Laplace transform and its inverse.

4. Construction of fast methods

The problems of moving information up and down the tree in the fast multipole method and finding a suitable quadrature (s_k, w_k) are related. There are probably many ways to approach this question. We chose to use a spectral decomposition. The function e^{sx} for example is not suitable, without making some modifications. Assume for example that we have a cluster C of particles x_j with intensities σ_j. Then the multipole coefficients are of the form:

$$M_q(C) = \sum_{j,\, x_j\in C} e^{-s_q(C)x_j}\,\sigma_j$$

where $s_q(C)$ denotes a quadrature adapted to cluster C. Let us assume that D is the parent cluster of C in the traditional FMM tree decomposition. Then we need to be able to calculate the contribution to $M_q(D)$ of particles in C that is $\sum_{j,\, x_j \in C} e^{-s_q(D) x_j}\, \sigma_j$. However the quadrature points $s_q(D)$ are different from $s_q(C)$. A procedure to gather multipole expansions up the tree is therefore needed. Similarly we need a procedure to scatter down the tree local expansions.

These procedures can be developed if we use Gaussian functions instead of exponentials. Let us perform the change of variable $s \to s^2$ in Γ_1 for example

$$\int_0^\infty \Big[\frac{s}{2\pi i} \int_{\Gamma_1} K(z) e^{-s^2 z}\, dz\Big]\, e^{s^2(x-y)}\, ds$$

The functions $e^{s^2 x}$ are still not suitable. Depending on the sign of x they may be bounded or not. In addition this formulation is not translation invariant, which should be expected. We therefore introduce a few additional parameters to reformulate this integral appropriately. Let us denote c_x the center of cluster C and c_y the center of the cluster E containing y. Then the integral becomes:

$$\int_0^\infty \Big[\frac{s}{2\pi i} \int_{\Gamma_1} K(z) e^{-s^2(z+c_y-c_x)}\, dz\Big]\, e^{s^2(x-c_x)} e^{s^2(c_y-y)}\, ds$$

We need to enforce that the Gaussian functions decay. To reduce the number of notations, we now assume that the clusters containing x and y have the same size. We add one more parameter l given by:

$$l = \min_{z \in \Gamma_1} \frac{\mathrm{Re}(z) + c_y - c_x}{2} \tag{4}$$

With this parameter:

$$\int_0^\infty \Big[\frac{s}{2\pi i} \int_{\Gamma_1} K(z) e^{-s^2(z+c_y-c_x-2l)}\, dz\Big]\, e^{-s^2(l-x+c_x)} e^{-s^2(l-c_y+y)}\, ds$$

This expansion must apply to any $x \in C$ and $y \in E$. Let us denote R the radius of each cluster. Since we must have by definition $\mathrm{Re}(z) > x - y$ for all x and y, we have $\mathrm{Re}(z) > c_x + R - (c_y - R) = 2R + c_x - c_y$. Consequently, from Eq. (4) $l > R$. This implies that $l > x - c_x$ and $l > c_y - y$. The two Gaussian functions $e^{-s^2(l-x+c_x)}$ and $e^{-s^2(l-c_y+y)}$ therefore decay as $s \to \infty$.

5. Error analysis

We shortly discuss the error analysis and the construction of the scheme. Let us consider as an example the kernel $\sqrt{x^2 + c^2}$ which is representative of a commonly found radial basis function. Let us assume that $c \in \mathbb{R}$. This kernel has two branch cuts starting at $\pm\, i\, c$ on the imaginary axis. The contours of integration Γ_i must therefore avoid those branch cuts. This kernel does not have a Fourier or Laplace transform as described previously. Therefore we must resort to a contour curve formed by the four pieces Γ_1 through Γ_4.

The error analysis relies on a Fourier series approximation of the Gaussian functions $g(s) = e^{-s^2(l-x+c_x)}$ and $e^{-s^2(l-c_y+y)}$. For brevity, we only outline the main points of the analysis. A detailed and complete analysis was done, confirming the schematic analysis below. The error analysis requires an integration from $-\infty$ to ∞ for s, otherwise $g(s)$ is effectively "discontinuous" and its Fourier spectrum decays slowly. We therefore modify the multipole-to-local operator and consider:

$$\int_{-\infty}^\infty \Big[\frac{s\, u(s)}{2\pi i} \int_{\Gamma_1} K(z) e^{-s^2(z+c_y-c_x-2l)}\, dz\Big]\, e^{-s^2(l-x+c_x)} e^{-s^2(l-c_y+y)}\, ds$$

where $u(s)$ is the Heaviside function. The Fourier spectrum of $g(s)$ decays rapidly. For example let us consider some tolerance ε and a bound L such that, when $|s| > L$, the function $g(s)$ is smaller than ε. On $[-L, L]$, we can expand $g(s)$ using a Fourier series. The Fourier series itself decays like a Gaussian function. There is an integer K such that all the Fourier coefficients of $g(s)$ beyond K are smaller than ε: $|\hat{g}_k| < \varepsilon$ for $|k| > K$. Let us now denote:

$$T(s) = \frac{s\ u(s)}{2\pi i} \int_{\Gamma_1} K(z) e^{-s^2(z+c_y-c_x-2l)}\ dz$$

and $\hat{T}_k$ its kth Fourier coefficient. As a consequence of Parseval's theorem, the function $T^{\rm lb}(s) = \sum_{k=-2K}^{2K} \hat{T}_k\ e^{2\pi i k s/2L}$ is such that:

$$\int_{-\infty}^{\infty} T(s)\ e^{-s^2(l-x+c_x)} e^{-s^2(l-c_y+y)}\ ds = \int_{-\infty}^{\infty} T^{\rm lb}(s)\ e^{-s^2(l-x+c_x)} e^{-s^2(l-c_y+y)}\ ds + O(\varepsilon)$$

The key property of this change from T to $T^{\rm lb}$ is that the integral

$$\int_{-\infty}^{\infty} T^{\rm lb}(s)\ e^{-s^2(l-x+c_x)} e^{-s^2(l-c_y+y)}\ ds$$

can be computed accurately, with error ε, using a trapezoidal rule with $4K+1$ points only. In addition, since the function $g(s)$ is even, we really only need $2K+1$ points. This number is essential. This is the number of quadrature points s_q mentioned in Section 4. It drives the computational cost of the FMM.

Let us now try to understand the parameters that determine K. The function $g(s)$ is given again by $g(s) = e^{-s^2(l-x+c_x)}$. The term $-x + c_x$ goes from $-R$ to R. Therefore the coefficient of the Gaussian is bounded by:

$$s^2(l-R) \leq s^2(l-x+c_x) \leq s^2(l+R)$$

When $s^2(l-x+c_x) \sim s^2(l-R)$, this corresponds to a slow decay and L needs to be large. When $s^2(l-x+c_x) \sim s^2(l+R)$, the Fourier spectrum decays slowly leading to a large bandwidth. As a result, the number of quadrature points can be shown to be of order:

$$\ln(\max_{\Gamma_1} |K(z)|/\varepsilon)\ \sqrt{\frac{l+R}{l-R}}$$

Therefore we should try to make $l - R$ as large as possible. This means the path Γ_1 should be moved to the right as much as possible. The path Γ_3 is similar. The Gaussian function is $e^{-s^2(l_3-(c_x-x))}$ with

$$l_3 = \min_z \frac{c_x - c_y - \mathrm{Re}(z)}{2}$$

The number of quadrature points is of order:

$$\ln(\max_{\Gamma_3} |K(z)|/\varepsilon)\ \sqrt{\frac{l_3+R}{l_3-R}}$$

so that the path Γ_3 should be moved to the left. Finally the path Γ_2 has $e^{-s^2(l_2-i(c_x-x))}$ with

$$l_2 = \min_z \frac{\mathrm{Im}(z)}{2}$$

The number of quadrature points is of order:

$$\ln(\max_{\Gamma_2}|K(z)|/\varepsilon)\,\sqrt{1+\frac{R^2}{l_2^2}}$$

so that the path Γ_2 should be moved up (large l_2). For kernel Ks that are real valued the path Γ_4 is the complex conjugate of Γ_2 so the same analysis applies.

As a conclusion all paths must be essentially moved away from the point $c_x - c_y$ on the real axis. A constraint is that the poles $\pm ic$ cannot be enclosed by the path. This is therefore an optimization problem, which can be solved numerically. Note that the kernel plays role, for example through the term $\max|K(z)|$. It is not essential to solve this problem with great accuracy since the cost and accuracy of the scheme only weakly depend on picking the optimal path. Any "reasonable" choice away from the branch and from $c_x - c_y$ is typically sufficient.

The choice of optimal contour does depend on the kernel. For example if we choose $1/\sqrt{x^2+c^2}$, the kernel also has branch cuts starting at $\pm ic$. In addition the kernel is unbounded as $z \to \pm ic$. However since $1/\sqrt{z^2+c^2}$ decays as $|z| \to \infty$, the contours Γ_1, Γ_2 and Γ_4 can be moved to infinity as explain in Section 3 for the Laplace transform. The optimal choice in this case is Γ_3 extending from $b - i\infty$ to $b + i\infty$, $b > 0$. The number of quadrature points is then of order:

$$\ln(\max_{\Gamma_3}|K(z)|/\varepsilon)\,\sqrt{\frac{c_x - c_y + 2R}{c_x - c_y - 2R}}$$

This implies that $c_x - c_y > 2R$. This condition is equivalent to saying that the two clusters need to be well separated (R is their radius).

6. Interpolation and anterpolation

Going back to the question of Section 4, we now propose an approach. Consider the multipole coefficients of the form:

$$M_q(C) = \sum_{j,\, x_j \in C} e^{-s_q(C)^2(l_C - x_j + c_C)}\,\sigma_j$$

We want to approximate the multipole coefficients for the parent cluster D:

$$M_q(D) = \sum_{j,\, x_j \in C} e^{-s_q(D)^2(l_D - x_j + c_D)}\,\sigma_j$$

In general since the radius of D is twice that of C, the Gaussian function decays much faster, and therefore the quadrature points $s_q(D)$ tend to cluster near 0 when compared to the quadrature points $s_q(C)$.

Let us assume that the quadrature points $s_q(C)$ have a spacing $\Delta s(C)$ while those for D have a spacing $\Delta s(D)$. As an example we can assume that $\Delta s(D) = \Delta s(C)/2$, but the method can be extended to more general relations. In that case, we need to interpolate the function

$$M_C(s_q) = \sum_{j,\, x_j \in C} e^{-s_q(C)^2(l_C - x_j + c_C)}\,\sigma_j$$

and calculate its values at points with a spacing of $\Delta s(D)$ instead of $\Delta s(C)$. This can be done by performing a fast Fourier transform of $M_C(s_q)$, padding with zeros, and performing an inverse transform. The coefficients $M_q(D)$ are obtained by keeping those samples in the desired interval: $[-L(D), L(D)]$ (using the notations of Section 5). The process is completed by a multiplication

with $e^{-s_q(D)^2(l_D+c_D-l_C-c_C)}$. The coefficient $l_D + c_D - l_C - c_C$ can be shown to be positive if $l_D - R_D > l_C - R_C$, which is the typical situation as clusters become larger. Indeed:

$$l_D + c_D - l_C - c_C > l_D - l_C - R_C = l_D - R_D - l_C + R_C > 0$$

since $R_D = 2R_C$.

This procedure can be used to gather multipole coefficients going up the tree in the fast multipole method. When going down, we need to scatter local coefficients from the root of the tree down to the leaves. The procedure is similar, even though the justification is more involved. In that case, we have functions with high frequency components in s corresponding to the fast decay away from 0, for large clusters. However the final integration is against $e^{-s^2(l-c_y+y)}$ whose Fourier spectrum decays rapidly. Therefore the high frequencies in the local multipole expansions can be removed when moving down the tree. This is done in a manner similar to the steps during the gathering phase (upward) of the FMM. First the local coefficients are multiplied by $e^{-s_q(D)^2(l_D-c_D-l_C+c_C)}$, if again D is the parent cluster of C. With the analysis above, we have that $l_D - c_D - l_C + c_C > 0$. Then the local expansion is padded with zeros to change its interval of definition from $[-L(D), L(D)]$ to $[-L(C), L(C)]$. Finally since we only need the low-frequency components, we Fourier transform the coefficients, remove the high frequencies that do not contribute, and inverse Fourier transform the coefficients.

7. Extension to multiple dimensions

The extension to multiple dimensions is beyond the scope of this paper. Let us briefly mention that the Cauchy integral formula can be extended to multiple variables

$$K(x,y,z) = \frac{1}{(2\pi i)^3} \oint_{\Gamma_x} \oint_{\Gamma_y} \oint_{\Gamma_z} \frac{K(c_1,c_2,c_3)}{(c_1-x)(c_2-y)(c_3-z)}\, dc_1\, dc_2\, dc_3$$

The contours must be such that poles or branch cuts of K are avoided. General strategies have been found to construct such contours in the case for example of poles or branch cuts in the $\mathrm{Re}(c) < 0$ half-plane. The procedures are essentially the same as the algorithms and analyses presented above except that they need to be applied to each dimension in turn, x, y, and z. As an example, for $K(x,y,z) = 1/\sqrt{x^2+y^2+z^2+c^2}$, $c \in \mathbb{R}$, we showed that, when applying the FMM in the $+x$ direction, the variable x should use an infinite vertical line for Γ_3 (Laplace transform case), while the y and z variables should use infinite horizontal lines for Γ_2 and Γ_4 (Fourier transform case). The five other directions are treated in a similar way, by permuting the variables accordingly ($-x$, $+y$, $-y$, $+z$, $-z$). This is analogous to the formulation for $1/r$ of Greengard et al. [5], which uses a combination of functions of the type e^{sx} and e^{isx} and requires six different expansions for $+x$, $-x$, $+y$, $-y$, $+z$, and $-z$. See also Darve et al. [6, 7] for a similar type of approach in the context of the Helmholtz kernel. These formulations can be recovered from our approach. This derivation is however beyond the scope of this manuscript.

8. Preliminary Numerical Results

In this section, we present some preliminary numerical results. These numerical results are preliminary in the sense that we have not implemented all the possible optimizations and therefore the accuracy and efficiency can be further improved. However these results provide a proof-of-concept of the overall approach. We considered two kernels: the multiquadric $\sqrt{1+x^2}$ and inverse multi-quadric $1/\sqrt{1+x^2}$ in one dimension. These kernels have branch cuts starting at $\pm i$. There are many different options to construct a fast scheme and we will present only some of them. These options primarily depend on the choice of contour for the integration (see Figure 3):

(i) Contour Γ_a: when the function is smooth at the origin, we do not always need to build an FMM oct-tree. It is possible, in many cases, to use a single FMM expansion for the entire computational domain. This reduces the number of multipole-to-local transfers from 189 (in 3D) to only 1. This is therefore a much faster technique when it is applicable. It fails in two cases. When the kernel is singular at $x = 0$, we must use an oct-tree. In addition, from the error analysis for contour Γ_a, we see that when the clusters is large compared to 1 (for this choice of kernels) the number of quadrature points in GFMM is large. In that case, we need to switch to contour Γ_b.

(ii) Contour Γ_b: this contour requires the construction of an oct-tree. This is essentially because it does not enclose the origin and therefore clusters must be well-separated. The advantage of using Γ_b is that the number of multipole coefficients is largely independent of the size of the cluster, and consequently problems of very large size can be treated. The drawback is that the number of multipole-to-local transfers is larger.

(iii) Contour Γ_c: this is the Fourier transform case. This contour is essentially a variant of Γ_a. It is applicable when the kernel is integrable. The advantage is that the number of multipole coefficients is in general smaller than with Γ_a since in effect we have removed the two vertical segments in Γ_a. The fact that the horizontal lines extend to infinity only marginally increases the number of coefficients.

(iv) Contour Γ_d: this is the Laplace transform case. This contour is a variant of Γ_b and requires an oct-tree. Again compared to Γ_b the number of coefficients is reduced since we have only one segment in the contour instead of four as in Γ_b.

We present numerical results for $\sqrt{1+x^2}$ and $1/\sqrt{1+x^2}$ using contours Γ_a and Γ_b. These results are still suboptimal. The technique to optimize the contour is based on a simulated annealing. We believe that better ways can be used that will reduce the number of quadrature points further. In the left panel of Figure 4, we present results for Γ_a. **The variable l denotes the size of the cluster.** GFMM was applied to four adjacent clusters with the same size. Particles were randomly distributed in the interval. The error was computed using the relative L^2 norm. When the cluster is large, the contour becomes very elongated. In that case, the number of quadrature points must increase. The right panel of Figure 4 shows results using contour Γ_b. The latter does not suffer from a breakdown at large sizes. However as was pointed out earlier, it requires building an oct-tree with well-separated clusters. This is not required for Γ_a.

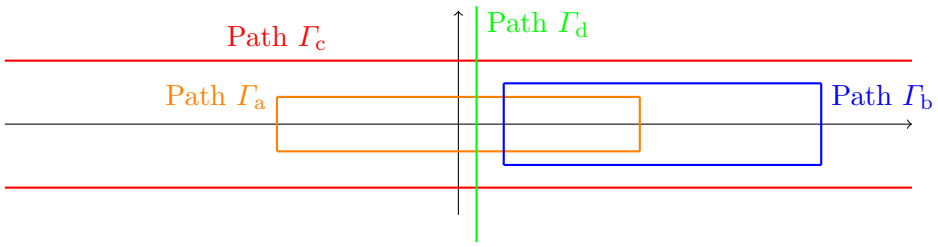

Figure 3. Different integration contours.

9. Conclusion

We have presented a framework to develop generalized fast multipole methods (GFMM), in particular methods that are applicable to radial basis functions for interpolation schemes. This new approach has the potential to greatly speed up traditional FMMs since the multipole-to-local operators are diagonal. We have developed a numerical analysis of the error and schemes

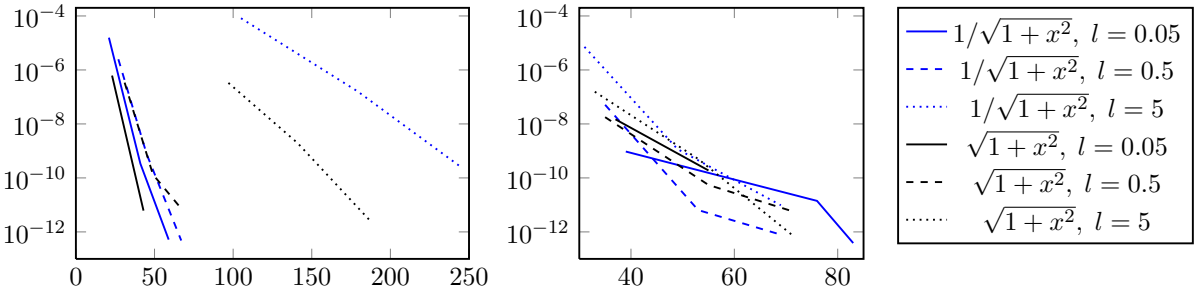

Figure 4. Numerical results with $1/\sqrt{x^2+1}$ and $\sqrt{x^2+1}$. The number of quadrature points is on the x axis, while the relative error with the L^2 norm is shown on the y axis. Left panel: contour Γ_{a}. Right panel: contour Γ_{b}.

to select the parameters in the method. The contour of integration in the complex plane needs to be optimized depending on the kernel and the size of the clusters. Numerical algorithms can be used for this optimization. This algorithm shares similarities with previously published methods [5, 6, 7]. The type of kernels that can be treated by the new approach has been extended significantly compared to [5, 6, 7]. It is probably the case that most kernels found in practical applications can be treated by this method. An interesting aspect of this method is the fact that a diagonal multipole-to-local operator allows treating oscillatory kernels as well, for example of the type of $\exp(ikr)/r$. This is not the case for many black box or kernel independent FMMs.

References

[1] Fong W and Darve E 2009 *Journal of Computational Physics* **228** 8712–8725
[2] Ying L, Biros G and Zorin D 2004 *Journal of Computational Physics* **196** 591–626
[3] De Boer A, Van der Schoot M and Bijl H 2007 *Computers and Structures* **85** 784–795
[4] Carr J, Beatson R, Cherrie J, Mitchell T, Fright W, McCallum B and Evans T 2001 *Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (ACM) p 76
[5] Greengard L and Rokhlin V 2008 *Acta numerica* **6** 229–269
[6] Darve E and Havé P 2004 *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* **362** 603–628
[7] Darve E and Havé P 2004 *Journal of Computational Physics* **197** 341–363