The fast Gauss transform with all the proofs.

VIKAS C. RAYKAR
University of Maryland, CollegePark
vikas@umiacs.umd.edu

In this report we discuss some technical aspects of the fast Gauss transform (FGT). The FGT is a special case of the single level FMM for the Gaussian potential. We give detailed derivation of the correct error bounds which is rather terse in the original paper. We also briefly discuss about the data structures to implement the space subdivision in higher dimensional spaces.

[**FGT**]: April 8, 2006

Contents

1. FAST GAUSS TRANSFORM

Here we give a detailed analysis of the fast Gauss transform (FGT) based on the techniques and the notation used in the FMM primer [Raykar 2005]. The FGT is a special case of the single level FMM for the Gaussian potential. For each *target* *point* $\{y_j \in \mathrm{R}^d\}_{j=1,\ldots,M}$ the discrete Gauss transform is defined as,

$$\Phi(y_j) = \sum_{i=1}^{N} q_i e^{-\|y_j - x_i\|^2/h^2}. \tag{1}$$

where $\{q_i \in \mathrm{R}^+\}_{i=1,\ldots,N}$ are the *source weights*, $\{x_i \in \mathrm{R}^d\}_{i=1,\ldots,N}$ are the *source points*, i.e., the center of the Gaussians, and $h \in \mathrm{R}^+$ is the *source scale* or *bandwidth*. $\Phi(y_j)$ is the total contribution at y_j of N Gaussians centered at x_i each with bandwidth h. Each Gaussian is weighted by the term q_i. The computational complexity to evaluate the discrete Gauss transform at M target points is $O(MN)$.

The Fast Gauss Transform [Greengard and Strain 1991] is an approximation algorithm that reduces the computational complexity to $O(M+N)$. The constant depends on the desired precision. Given any $\epsilon > 0$, it computes an approximation $\hat{\Phi}(y_j)$ to $\Phi(y_j)$ such that the maximum absolute error relative to the total weight $Q = \sum_{i=1}^{N} q_i$,

$$\max_{y_j} \left[\frac{|\hat{\Phi}(y_j) - \Phi(y_j)|}{Q} \right] \leq \epsilon. \tag{2}$$

The algorithm is based on the single level FMM described above. The Gaussian is factorized via Hermite and Taylor series. The error bound derived in the original paper was shown to be incorrect and a new bound was derived in [Baxter and Roussos 2002].

The space is subdivided into a number of boxes and the sources and targets are assigned to different boxes. For all sources belonging to a particular box, the S-expansion (Hermite series) coefficients are consolidated at the center of each source box. For each target box the S-expansion (Hermite series) coefficients at each source box are translated via the S|R translation operator as R-expansion (Taylor series) coefficients at the center of the target box. Since the Gaussian decays very rapidly only a few neighboring source boxes will have influence on the target box.

2. HERMITE POLYNOMIALS

The Hermite polynomial $H_n(y)$ is defined by the Rodrigues formula

$$H_n(y) = (-1)^n e^{y^2} \frac{d^n}{dy^n} \left(e^{-y^2} \right), \;\; y \in \mathbf{R}. \tag{3}$$

Following are the first few Hermite polynomials.

$$\begin{aligned} H_0(y) &= 1 \\ H_1(y) &= 2y \\ H_2(y) &= 4y^2 - 2 \\ H_3(y) &= 8y^3 - 12y \end{aligned} \tag{4}$$

The generating function for the Hermite polynomials is

$$e^{2yx - x^2} = \sum_{n=0}^{\infty} \frac{x^n}{n!} H_n(y). \tag{5}$$

Multiplying both sides of the preceding equation by e^{-y^2} yields

$$e^{-(y-x)^2} = \sum_{n=0}^{\infty} \frac{x^n}{n!} h_n(y), \tag{6}$$

where the Hermite functions $h_n(y)$ are defined as,

$$h_n(y) = e^{-y^2} H_n(y). \tag{7}$$

We need a shifted and the scaled version of this formula.

$$\begin{aligned} e^{-(y-x)^2/h^2} &= e^{-[(y-x_0)-(x-x_0)]^2/h^2} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{x-x_0}{h}\right)^n h_n\left(\frac{y-x_0}{h}\right). \end{aligned} \tag{8}$$

This formula tells us how to evaluate the Gaussian field at the target y due to a source at x, as a Hermite expansion centered at a nearby source x_0. This series converges rapidly and for a given precision, only a certain number of p terms need be retained.

Interchanging x and y we have the following expansion.

$$e^{-(y-x)^2/h^2} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{y-y_0}{h}\right)^n h_n\left(\frac{x-y_0}{h}\right). \tag{9}$$

This formula tells us how to evaluate the Gaussian field at the target y due to a source at x, as a Taylor expansion centered at a nearby target y_0.

The following recurrence relation will be useful in computation of the Hermite functions.

$$h_{n+1}(y) = 2y h_n(y) - 2n h_{n-1}(y), \;\; y \in \mathbf{R}. \tag{10}$$

The Cramer's inequality [Hille 1926] for Hermite polynomials will be useful in deriving the error bounds.

$$|H_n(y)| \leq K 2^{n/2} \sqrt{n!} e^{y^2/2}, \tag{11}$$

where $K < 1.09$. Based on the Cramer's inequality we have the following useful bound for Hermite functions.

$$\frac{1}{n!} |h_n(y)| \leq K 2^{n/2} \frac{1}{\sqrt{n!}} e^{-y^2/2}. \tag{12}$$

[Baxter and Roussos 2002] use a slightly tighter version of this

$$\frac{1}{n!} |h_n(y)| \leq 2^{n/2} \frac{1}{\sqrt{n!}} e^{-y^2/2}. \tag{13}$$

3. MULTI-INDEX NOTATION

In order to generalize the above relations to higher dimensions we use the multi-index notation. A multi-index $\alpha = (\alpha_1, \alpha_2 \ldots, \alpha_d)$ is a d-tuple of nonnegative integers. The length of the multi-index α is defined as $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_d$. If p is an integer, we say $\alpha \geq p$ of $\alpha_i \geq p$ for $1 \leq i \leq d$. The factorial of α is defined as $\alpha! = \alpha_1!\alpha_2! \ldots \alpha_d!$. For any multi-index $\alpha \in \mathbf{N}^d$ and $x = (x_1, x_2, \ldots, x_d) \in \mathbf{R}^d$ the d-variate monomial x^α is defined as $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_d^{\alpha_d}$. x^α is of degree n if

$|\alpha| = n$. Let $x, y \in \mathbf{R}^d$ and $v = x \cdot y = x_1 y_1 + \ldots + x_d y_d$. Then using the multi-index notation v^n can be written as,

$$v^n = \sum_{|\alpha|=n} \frac{n!}{\alpha!} x^\alpha y^\alpha. \tag{14}$$

The α^{th} derivative with respect to x is defined as,

$$\frac{d^\alpha}{dx^\alpha} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}}. \tag{15}$$

The multidimensional Hermite function is defined as

$$h_\alpha(y) = e^{-\|y\|^2} H_\alpha(y) = h_{\alpha_1}(y_1) h_{\alpha_2}(y_2) \ldots h_{\alpha_d}(y_d). \tag{16}$$

The Hermite expansion of the multivariate Gaussian can be written as,

$$e^{-(y-x)^2/h^2} = \sum_{\alpha \geq 0} \frac{1}{\alpha!} \left(\frac{x - x_0}{h} \right)^\alpha h_\alpha \left(\frac{y - x_0}{h} \right). \tag{17}$$

The Taylor expansion of the multivariate Gaussian can be written as,

$$e^{-(y-x)^2/h^2} = \sum_{\beta \geq 0} \frac{1}{\beta!} h_\beta \left(\frac{x - y_0}{h} \right) \left(\frac{y - y_0}{h} \right)^\beta. \tag{18}$$

4. TAYLOR EXPANSION OF THE HERMITE FUNCTION

For the translation operators we will also need the Taylor expansion of the Hermite function $h_\alpha(y)$ about an arbitrary point $y_0 \in \mathbf{R}^d$.

$$h_\alpha(y) = \sum_{\beta \geq 0} \frac{(y - y_0)^\beta}{\beta!} D^\beta h_\alpha(y_0). \tag{19}$$

$$h_\alpha(y) = (-1)^\alpha D^\alpha e^{-\|y\|^2}. \tag{20}$$

Hence,

$$D^\beta h_\alpha(y) = (-1)^\beta h_{\alpha+\beta}(y). \tag{21}$$

$$h_\alpha(y) = \sum_{\beta \geq 0} \frac{(y - y_0)^\beta}{\beta!} (-1)^\beta h_{\alpha+\beta}(y). \tag{22}$$

5. FACTORIZATION [S- AND R-EXPANSION]

The potential or field at y due to a source at x_i is given by

$$\Phi(y, x_i) = e^{-\|y - x_i\|^2/h^2}. \tag{23}$$

Note that $\Phi(y, x_i)$ is a regular potential. All the sources and targets are assumed to lie in the unit box $B_0 = [0, 1]^d$. B_0 is subdivided into smaller boxes with sides of length $\sqrt{2}rh$, parallel to the axes, with a fixed $r \leq 1/\sqrt{2}$. The spatial domain enclosed by the l^{th} box is denoted as $I_1(l)$ and the center of the box is denoted by x_c^l. $I_1^c(l)$ is the complement of $I_1(l)$. Since the potential is regular we need not worry about singular points.

For the l^{th} box $\forall x_i \in I_1(l)$ and $\forall y \in I_1^c(l)$ we need an S-expansion (far-field) of the form

$$\Phi(y, x_i) = \sum_{\alpha \geq 0} a_\alpha(x_i, x_c^l) S_\alpha(y - x_c^l). \tag{24}$$

The FGT uses the Hermite expansion of the Gaussian as the S-expansion.

$$\Phi(y, x_i) = \sum_{\alpha \geq 0} \frac{1}{\alpha!} \left(\frac{x_i - x_c^l}{h} \right)^\alpha h_\alpha \left(\frac{y - x_c^l}{h} \right). \tag{25}$$

$$a_\alpha(x_i, x_c^l) = \frac{1}{\alpha!} \left(\frac{x_i - x_c^l}{h} \right)^\alpha, \quad S_\alpha(y - x_c^l) = h_\alpha \left(\frac{y - x_c^l}{h} \right). \tag{26}$$

For the l^{th} box $\forall x_i \in I_1^c(l)$ and $\forall y \in I_1(l)$ we need an R-expansion (local) of the form

$$\Phi(y, x_i) = \sum_{\beta \geq 0} b_\beta(x_i, x_c^l) R_\beta(y - x_c^l). \tag{27}$$

The FGT uses the Taylor expansion of the Gaussian (which is obtained by interchanging y and x_i in the Hermite expansion) as the R-expansion.

$$\Phi(y, x_i) = \sum_{\beta \geq 0} \frac{1}{\beta!} h_\beta \left(\frac{x_i - x_c^l}{h} \right) \left(\frac{y - x_c^l}{h} \right)^\beta. \tag{28}$$

$$b_\beta(x_i, x_c^l) = \frac{1}{\beta!} h_\beta \left(\frac{x_i - x_c^l}{h} \right), \quad R_\beta(y - x_c^l) = \left(\frac{y - x_c^l}{h} \right)^\beta. \tag{29}$$

6. FAST GAUSS TRANSFORM

Let the target point y belong to the n^{th} box, i.e., $y \in I_1(n)$. We need to evaluate the total field at y due sources in all boxes. Since the Gaussian decays very rapidly only a few boxes close to the target box can contribute more than $Q\epsilon$ to the field at y. If we include only $(2n+1)^d$ nearest boxes, then the error due to ignoring all other boxes is bounded by $Qe^{-2r^2n^2}$. This implies $n \geq \sqrt{\frac{\ln 1/\epsilon}{2r^2}}$. Suppose we want compute the field at all $y_j \in I_1(n)$ due to all sources $x_i \in I_1(l)$. There are four different ways of doing this. Let there be N_B sources in $I_1(l)$ and M_C targets in $I_1(n)$. Given the sources in one box and the targets in a neighboring box, the computation is performed using one of the following four methods depending on the number of sources and targets in these boxes: Direct evaluation is used if the number of sources and targets are small. If the sources are clustered in a box then they can be transformed into Hermite expansion about the center of the box. This expansion is directly evaluated at each target in the target box if the number of the targets is small. If the targets are clustered then the sources or their expansion are converted to a local Taylor series which is then evaluated at each target in the box. In practice a cutoff $N_F = O(p^{d-1})$ and $M_L = O(p^{d-1})$ is introduced.

—If $N_B < N_F$ then the source box $I_1(l)$ sends out N_B Gaussians.

—If $N_B \geq N_F$ then the source box $I_1(l)$ sends out a Hermite expansion.

—If $M_C < M_L$ then the target box $I_1(n)$ evaluates all the fields sent to it immediately.

—If $M_C \geq M_L$ then the target box $I_1(n)$ transforms all fields sent to it into Taylor series, accumulates the coefficients, and then evaluates the Taylor series.

6.1 Direct Evaluation

For the l^{th} box let $\Phi^l(y)$ be the potential at $y \in I_1^c(l)$ due to all sources $x_i \in I_1(l)$. For all $y \in I_1(n)$

$$\Phi^l(y) = \sum_{x_i \in I_1(l)} q_i \Phi(y, x_i) = \sum_{x_i \in I_1(l)} q_i e^{-\|y-x_i\|^2/h^2}. \tag{30}$$

The computational cost is $O(N_B M_C)$.

6.2 S-expansion (Hermite) at the center of each source box

We form S-expansion (Hermite) of the potential about the source box center x_c^l. For each box we consolidate the S-expansion coefficients due to all sources in that box into a single rapidly converging Hermite expansion about the center of the box. For all $y \in I_1(n)$

$$\begin{aligned}
\Phi^l(y) &= \sum_{x_i \in I_1(l)} q_i \Phi(y, x_i) \\
&= \sum_{x_i \in I_1(l)} q_i \left[\sum_{\alpha \geq 0} a_\alpha(x_i, x_c^l) S_\alpha(y - x_c^l) \right] \\
&= \sum_{\alpha \geq 0} \left[\sum_{x_i \in I_1(l)} q_i a_\alpha(x_i, x_c^l) \right] S_\alpha(y - x_c^l) \\
&= \sum_{\alpha \geq 0} A_\alpha^l S_\alpha(y - x_c^l) = \sum_{\alpha \geq 0} A_\alpha^l h_\alpha \left(\frac{y - x_c^l}{h} \right)
\end{aligned} \tag{31}$$

where

$$A_\alpha^l = \sum_{x_i \in I_1(l)} q_i a_\alpha(x_i, x_c^l) = \frac{1}{\alpha!} \sum_{x_i \in I_1(l)} q_i \left(\frac{x_i - x_c^l}{h} \right)^\alpha. \tag{32}$$

We retain only the first p^d coefficients. The computational cost is $O(p^d N_B) + O(p^d M_C)$. The error due to truncation of the Hermite series is bounded by

$$E_H(p) = \left| \sum_{\alpha \geq p} A_\alpha^l h_\alpha \left(\frac{y - x_c^n}{h} \right) \right| \leq \frac{Q_l}{(1-r)^d} \sum_{k=0}^{d-1} \binom{d}{k} (1 - r^p)^k \left(\frac{r^p}{\sqrt{p!}} \right)^{d-k} \tag{33}$$

where $Q_l = \sum_{x_i \in I_1(l)} q_i$ and $r < 1$. See Appendix 1 for the derivation.
The total potential at y due to all sources is

$$\widehat{\Phi}(y) = \sum_{\forall l} \widehat{\Phi}^l(y) = \sum_{\forall l} \sum_{\alpha < p} A_\alpha^l h_\alpha \left(\frac{y - x_c^l}{h} \right). \tag{34}$$

Since there are at most $(2n+1)^d$ source boxes within range the total computational complexity will be $O(p^dN) + O((2n+1)^d p^d M)$. The error

$$\frac{|\widehat{\Phi}(y) - \Phi(y)|}{Q} \leq \frac{1}{(1-r)^d} \sum_{k=0}^{d-1} \binom{d}{k} (1-r^p)^k \left(\frac{r^p}{\sqrt{p!}}\right)^{d-k} \tag{35}$$

where $r < 1$.

6.3 R-expansion (Taylor) at the center of each target box

We form R-expansion (Taylor) of the potential about the target box center x_c^n. For each source box $I_1(l)$ we consolidate the R-expansion coefficients due to all sources in that box into a single Taylor expansion about the center of the target box $I_1(n)$. For all $y \in I_1(n)$

$$\begin{aligned}
\Phi^l(y) &= \sum_{x_i \in I_1(l)} q_i \Phi(y, x_i) = \sum_{x_i \in I_1(l)} q_i \left[\sum_{\beta \geq 0} b_\beta(x_i, x_c^n) R_\beta(y - x_c^n) \right] \\
&= \sum_{\beta \geq 0} \left[\sum_{x_i \in I_1(l)} q_i b_\beta(x_i, x_c^n) \right] R_\beta(y - x_c^n) \\
&= \sum_{\beta \geq 0} B_\beta^{ln} R_\beta(y - x_c^n) = \sum_{\beta \geq 0} B_\beta^{ln} \left(\frac{y - x_c^n}{h}\right)^\beta
\end{aligned} \tag{36}$$

where,

$$B_\beta^{ln} = \sum_{x_i \in I_1(l)} q_i b_\beta(x_i, x_c^n) = \frac{1}{\beta!} \sum_{x_i \in I_1(l)} q_i h_\beta \left(\frac{x_i - x_c^n}{h}\right). \tag{37}$$

We retain only the first p^d coefficients. The computational cost is $O(p^d N_B) + O(p^d M_C)$. The error due to truncation of the Taylor series is bounded by

$$E_T(p) = \left| \sum_{\beta \geq p} B_\beta^{ln} \left(\frac{y - x_c^n}{h}\right)^\beta \right| \leq \frac{Q_l}{(1-r)^d} \sum_{k=0}^{d-1} \binom{d}{k} (1-r^p)^k \left(\frac{r^p}{\sqrt{p!}}\right)^{d-k} \tag{38}$$

where $Q_l = \sum_{x_i \in I_1(l)} q_i$ and $r < 1$. See Appendix 2 for the derivation.
The total potential at $y \in I_1(n)$ due to all sources is

$$\widehat{\Phi}(y) = \sum_{\forall l} \widehat{\Phi}^l(y) = \sum_{\forall l} \sum_{\beta < p} B_\beta^{ln} \left(\frac{y - x_c^n}{h}\right)^\beta = \sum_{\beta < p} \left[\sum_{\forall l} B_\beta^{ln} \right] \left(\frac{y - x_c^n}{h}\right)^\beta. \tag{39}$$

Since there are at most $(2n+1)^d$ source boxes within range the total computational complexity will be $O((2n+1)^d p^d N) + O(p^d M)$. The error

$$\frac{|\widehat{\Phi}(y) - \Phi(y)|}{Q} \leq \frac{1}{(1-r)^d} \sum_{k=0}^{d-1} \binom{d}{k} (1-r^p)^k \left(\frac{r^p}{\sqrt{p!}}\right)^{d-k} \tag{40}$$

where $r < 1$.

6.4 S|R translation

We accumulate all sources in a box via a truncated Hermite expansion and then translate the S-expansion coefficients into R-expansion Taylor coefficients at the center of the target box containing the evaluation point. Let $y \in I_1(n)$, i.e., the target point belongs to the n^{th} box with center x_c^n. Now we want to write the potential at y due to all sources in the l^{th} box $\Phi^l(y)$ (expanded using an S expansion around x_c^l) as an R expansion around x_c^n.

$$\Phi^l(y) = \sum_{\beta \geq 0} C_\beta^{ln} R_\beta(y - x_c^n) = \sum_{\beta \geq 0} C_\beta^{ln} \left(\frac{y - x_c^n}{h} \right)^\beta \tag{41}$$

where $C_\beta^{ln} = \sum_{\alpha \geq 0} (S|R)_{\beta\alpha}(x_c^n - x_c^l) A_\alpha^l$ and A_α^l are the S-expansion coefficients around x_c^l. The S|R-translation operator can be derived using the Taylor expansion of the Hermite function [See Section 4].

$$h_\alpha \left(\frac{y - x_c^l}{h} \right) = \sum_{\beta \geq 0} \frac{(-1)^{|\beta|}}{\beta!} \left(\frac{y - x_c^n}{h} \right)^\beta h_{\alpha+\beta} \left(\frac{x_c^n - x_c^l}{h} \right). \tag{42}$$

The S-expansion is given by

$$\begin{aligned}
\Phi^l(y) &= \sum_{\alpha \geq 0} A_\alpha^l h_\alpha \left(\frac{y - x_c^l}{h} \right) = \sum_{\alpha \geq 0} A_\alpha^l \left[\sum_{\beta \geq 0} \frac{(-1)^{|\beta|}}{\beta!} \left(\frac{y - x_c^n}{h} \right)^\beta h_{\alpha+\beta} \left(\frac{x_c^n - x_c^l}{h} \right) \right] \\
&= \sum_{\beta \geq 0} \left[\frac{(-1)^{|\beta|}}{\beta!} \sum_{\alpha \geq 0} A_\alpha^l h_{\alpha+\beta} \left(\frac{x_c^n - x_c^l}{h} \right) \right] \left(\frac{y - x_c^n}{h} \right)^\beta \\
&= \sum_{\beta \geq 0} \left[\frac{(-1)^{|\beta|}}{\beta!} \sum_{\alpha \geq 0} A_\alpha^l (-1)^{|\alpha|+|\beta|} h_{\alpha+\beta} \left(\frac{x_c^l - x_c^n}{h} \right) \right] \left(\frac{y - x_c^n}{h} \right)^\beta \\
&= \sum_{\beta \geq 0} \left[\frac{1}{\beta!} \sum_{\alpha \geq 0} A_\alpha^l (-1)^{|\alpha|} h_{\alpha+\beta} \left(\frac{x_c^l - x_c^n}{h} \right) \right] \left(\frac{y - x_c^n}{h} \right)^\beta \\
&= \sum_{\beta \geq 0} C_\beta^{ln} \left(\frac{y - x_c^n}{h} \right)^\beta
\end{aligned} \tag{43}$$

where,

$$C_\beta^{ln} = \frac{1}{\beta!} \sum_{\alpha \geq 0} A_\alpha^l (-1)^{|\alpha|} h_{\alpha+\beta} \left(\frac{x_c^l - x_c^n}{h} \right). \tag{44}$$

There are two sources of truncation. First we truncate the Hermite series coefficients A_α^l and then we truncate the coefficients C_β^{ln}. The error due to truncation of both the Taylor series and the Hermite series is bounded by

$$E_{TH}(p) \leq E_T(p) + \frac{Q_l}{(1 - \sqrt{2}r)^{2d}} \left[\sum_{k=0}^{d-1} \binom{d}{k} \left(1 - (\sqrt{2}r)^p \right)^k \left(\frac{(\sqrt{2}r)^p}{\sqrt{p!}} \right)^{d-k} \right]^2 \tag{45}$$

where $Q_l = \sum_{x_i \in I_1(l)} q_i$ and $r < 1/\sqrt{2}$. See Appendix 3 for the derivation. The total potential at $y \in I_1(n)$ due to all sources is

$$\widehat{\Phi}(y) = \sum_{\forall l} \widehat{\Phi}^l(y) = \sum_{\forall l} \sum_{\beta < p} \widehat{C}_\beta^{ln} \left(\frac{y - x_c^n}{h} \right)^\beta = \sum_{\beta < p} \left[\sum_{\forall l} \widehat{C}_\beta^{ln} \right] \left(\frac{y - x_c^n}{h} \right)^\beta . \quad (46)$$

where,

$$\widehat{C}_\beta^{ln} = \frac{1}{\beta!} \sum_{\alpha < p} A_\alpha^l (-1)^{|\alpha|} h_{\alpha+\beta} \left(\frac{x_c^l - x_c^n}{h} \right). \quad (47)$$

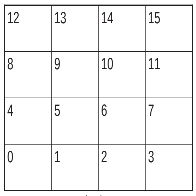
The total computational complexity will be $O((2n+1)^d d p^{d+1} N_{box}) + O(p^d M) + O(p^d N)$.

7. ALGORITHM

A formal description of the algorithm and detailed analysis of the computational cost can be seen in the original paper [Greengard and Strain 1991].

8. DATA STRUCTURES

The first step of the algorithm is the spatial subdivision of the unit hypercube into N_{side}^d boxes of side $\sqrt{2}rh$ were $r < 1/2$. The boxes are numbered from 1 to N_{side}^d in the order of increasing dimensionality d. Fig. 1 shows an example of spatial ordering in two dimensions.



(a)

Fig. 1. Example of spatial ordering in two dimensions.

We need the following operations to implement the FGT.

(1) Given $x = (x_1, x_2, \ldots, x_d) \in \mathbf{R}^d$ find the corresponding box number to which x belongs.

$$BoxNumber = \left[\sum_{j=1}^{d} \lfloor x_j N_{side} \rfloor N_{side}^{j-1} \right]. \quad (48)$$

(2) Given the box number find the center of the box. First write $BoxNumber$ in base N_{side}.

$$(BoxNumber)_{(10)} = (a_d a_{d-1} \ldots a_2 a_1)_{(N_{side})}. \quad (49)$$

The j^{th} coordinate of the box center is given by

$$BoxCenter_j = (\sqrt{2}rh)(a_j + \frac{1}{2}). \quad (50)$$

(3) List the $(2n+1)^d$ neighbors of a given box. First write $BoxNumber - 1$ in base N_{side}.

$$(BoxNumber)_{(10)} = (a_d a_{d-1} \ldots a_2 a_1)_{(N_{side})}. \tag{51}$$

Increment or decrement each a_j as follows

$$b_{ji} = a_j + i \text{ for } i = 0, \pm1, \pm2, \ldots, \pm n. \tag{52}$$

The base N_{side} representation of the $(2n+1)^d$ neighbors is as follows

$$(b_{di_d} \ldots b_{2i_2} b_{1i_1})_{(N_{side})}. \tag{53}$$

where $(i_d, \ldots, i_2, i_1) \in \{0, \pm1, \pm2, \ldots, \pm n\}^d$. The box number is given by converting it into base 10 representation.

$$(b_{di_d} \ldots b_{2i_2} b_{1i_1})_{(N_{side})} = (BoxNumber)_{(10)} \tag{54}$$

REFERENCES

Baxter, B. J. C. and Roussos, G. 2002. A new error estimate of the fast gauss transform. *SIAM Journal of Scientific and Statistical Computing 24,* 1, 257–259. 2, 3, 13

Greengard, L. and Strain, J. 1991. The fast Gauss transform. *SIAM Journal of Scientific and Statistical Computing 12,* 1, 79–94. 2, 9, 13, 14

Hille, E. 1926. A class of reciprocal functions. *The Annals of Mathematics, 2nd Ser. 27,* 4 (Jun.), 427–464. 3

Raykar, V. C. 2005. A short primer on the fast multipole method. *Available online on the author's website*. 1

9. APPENDICES

9.1 Appendix 1 [Hermite series truncation error]

The error $E_H(p)$ due to truncating the series after p^d terms is

$$
\begin{aligned}
E_H(p) &= \left|\sum_{\alpha\geq p} A_\alpha^l h_\alpha\left(\frac{y-x_c^l}{h}\right)\right| \leq \sum_{\alpha\geq p}\left|A_\alpha^l\right|\left|h_\alpha\left(\frac{y-x_c^l}{h}\right)\right| \\
&= \sum_{\alpha\geq p}\left|A_\alpha^l\right|\prod_{j=1}^{d}\left|h_{\alpha_j}\left(\frac{(y)_j-(x_c^l)_j}{h}\right)\right| \quad [(y)_j \text{ is the } j^{th} \text{ coordinate of } y.] \\
&= \sum_{\alpha\geq p}\left|\frac{1}{\alpha!}\sum_{x_i\in I_1(l)} q_i\left(\frac{x_i-x_c^l}{h}\right)^\alpha\right|\prod_{j=1}^{d}\left|h_{\alpha_j}\left(\frac{(y)_j-(x_c^l)_j}{h}\right)\right| \\
&\leq \sum_{\alpha\geq p}\left|\sum_{x_i\in I_1(l)} q_i\right|\left(\frac{r}{\sqrt{2}}\right)^\alpha\prod_{j=1}^{d}\frac{1}{\alpha_j!}\left|h_{\alpha_j}\left(\frac{(y)_j-(x_c^l)_j}{h}\right)\right| \quad [\text{Since the box side is } \sqrt{2}rh] \\
&\leq \sum_{\alpha\geq p}\left[\sum_{x_i\in I_1(l)}|q_i|\right]\prod_{j=1}^{d} r^{\alpha_j}2^{-\alpha_j/2}\frac{1}{\alpha_j!}\left|h_{\alpha_j}\left(\frac{(y)_j-(x_c^l)_j}{h}\right)\right| \\
&= Q_l\sum_{\alpha\geq p}\prod_{j=1}^{d} r^{\alpha_j}2^{-\alpha_j/2}\frac{1}{\alpha_j!}\left|h_{\alpha_j}\left(\frac{(y)_j-(x_c^l)_j}{h}\right)\right| \quad [\text{Define } Q_l=\sum_{x_i\in I_1(l)}|q_i|] \\
&\leq Q_l\sum_{\alpha\geq p}\prod_{j=1}^{d}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}2^{-\alpha_j/2}2^{-\alpha_j/2}e^{-((y)_j-(x_c^l)_j)^2/2h^2} \quad [\text{Cramer's inequality}] \\
&\leq Q_l\sum_{\alpha\geq p}\prod_{j=1}^{d}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j} = Q_l\left[\sum_{\alpha\geq 0}\prod_{j=1}^{d}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}-\sum_{\alpha< p}\prod_{j=1}^{d}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}\right] \\
&= Q_l\left[\prod_{j=1}^{d}\left\{\sum_{\alpha_j\geq 0}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}\right\}-\prod_{j=1}^{d}\left\{\sum_{\alpha_j< p}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}\right\}\right] \\
&= Q_l\left[\left\{\sum_{\alpha_j< p}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}+\sum_{\alpha_j\geq p}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}\right\}^d-\left\{\sum_{\alpha_j< p}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}\right\}^d\right] \\
&= Q_l\sum_{k=0}^{d-1}\binom{d}{k}\left(\sum_{\alpha_j< p}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}\right)^k\left(\sum_{\alpha_j\geq p}\frac{1}{\sqrt{\alpha_j!}}r^{\alpha_j}\right)^{d-k} \quad [\text{Binomial theorem}] \\
&\leq Q_l\sum_{k=0}^{d-1}\binom{d}{k}\left(\sum_{\alpha_j< p} r^{\alpha_j}\right)^k\left(\frac{r^p}{\sqrt{p!}}\sum_{\alpha_j\geq 0} r^{\alpha_j}\right)^{d-k} \\
&\leq Q_l\sum_{k=0}^{d-1}\binom{d}{k}\left(\frac{1-r^p}{1-r}\right)^k\left(\frac{r^p}{\sqrt{p!}}\frac{1}{1-r}\right)^{d-k} \quad [\text{Geometric series } r<1.] \\
&= \frac{Q_l}{(1-r)^d}\sum_{k=0}^{d-1}\binom{d}{k}(1-r^p)^k\left(\frac{r^p}{\sqrt{p!}}\right)^{d-k}
\end{aligned}
\tag{55}
$$

The error bound in the original FGT paper [Greengard and Strain 1991] was shown to be incorrect and a new bound was proposed by [Baxter and Roussos 2002]. This is the new error estimate as in [Baxter and Roussos 2002] but derived in a slightly different way.

9.2 Appendix 2 [Taylor series truncation error]

The error $E_T(p)$ due to truncating the series after p^d terms is (derivation very similar to $E_H(p)$)

$$
\begin{aligned}
E_T(p) &= \left| \sum_{\beta \geq p} B_\beta^{ln} \left(\frac{y - x_c^n}{h} \right)^\beta \right| \\
&\leq \sum_{\beta \geq p} |B_\beta^{ln}| \left| \left(\frac{y - x_c^n}{h} \right)^\beta \right| \\
&= \sum_{\beta \geq p} \frac{1}{\beta!} \left| \sum_{x_i \in I_1(l)} q_i h_\beta \left(\frac{x_i - x_c^n}{h} \right) \right| \left| \left(\frac{y - x_c^n}{h} \right)^\beta \right| \quad \text{[Eq. 37.]} \\
&\leq \sum_{\beta \geq p} \left[\sum_{x_i \in I_1(l)} |q_i| \left\{ \frac{1}{\beta!} \left| h_\beta \left(\frac{x_i - x_c^n}{h} \right) \right| \right\} \right] \left| \left(\frac{y - x_c^n}{h} \right)^\beta \right| \\
&\leq \sum_{\beta \geq p} \left[\sum_{x_i \in I_1(l)} |q_i| \left\{ \frac{1}{\beta!} \left| h_\beta \left(\frac{x_i - x_c^n}{h} \right) \right| \right\} \right] r^\beta 2^{-\beta/2} \\
&= \sum_{\beta \geq p} \left[\sum_{x_i \in I_1(l)} |q_i| \left\{ \prod_{j=1}^{d} \frac{1}{\beta_j!} \left| h_{\beta_j} \left(\frac{(x_i)_j - (x_c^n)_j}{h} \right) \right| \right\} \right] r^\beta 2^{-\beta/2} \\
&\leq \sum_{\beta \geq p} \left[\sum_{x_i \in I_1(l)} |q_i| \left\{ \prod_{j=1}^{d} \frac{1}{\sqrt{\beta_j!}} 2^{\beta_j/2} \right\} \right] r^\beta 2^{-\beta/2} \\
&= Q_l \sum_{\beta \geq p} \prod_{j=1}^{d} \frac{1}{\sqrt{\beta_j!}} r^{\beta_j} \\
&\leq \frac{Q_l}{(1-r)^d} \sum_{k=0}^{d-1} \binom{d}{k} (1 - r^p)^k \left(\frac{r^p}{\sqrt{p!}} \right)^{d-k} \qquad (56)
\end{aligned}
$$

9.3 Appendix 3 [S|R truncation error]

Using the expression for A^l_α, C^{ln}_β can be simplified as follows.

$$
\begin{aligned}
C^{ln}_\beta &= \frac{1}{\beta!}\sum_{\alpha\geq 0} A^l_\alpha (-1)^{|\alpha|} h_{\alpha+\beta}\left(\frac{x^l_c - x^n_c}{h}\right) \\
&= \frac{1}{\beta!}\sum_{\alpha\geq 0}\left[\frac{1}{\alpha!}\sum_{x_i\in I_1(l)} q_i \left(\frac{x_i - x^l_c}{h}\right)^\alpha\right](-1)^{|\alpha|} h_{\alpha+\beta}\left(\frac{x^l_c - x^n_c}{h}\right) \\
&= \frac{1}{\beta!}\sum_{x_i\in I_1(l)} q_i \left[\sum_{\alpha\geq 0}\frac{1}{\alpha!}\left(\frac{x_i - x^l_c}{h}\right)^\alpha (-1)^{|\alpha|} h_{\alpha+\beta}\left(\frac{x^l_c - x^n_c}{h}\right)\right] \\
&= \frac{1}{\beta!}\sum_{x_i\in I_1(l)} q_i h_\beta\left(\frac{x_i - x^n_c}{h}\right). \qquad (57)
\end{aligned}
$$

Note that this is exactly the same as B^{ln}_β the R-expansion coefficients, which should be the case. There is a error in the original paper [Greengard and Strain 1991] where there is an extra $(-1)^{|\beta|}$ term.
The error $E_{TH}(p)$ due to truncating both the series after p^d terms is

$$
\begin{aligned}
E_{TH}(p) &= \left|\sum_{\beta\geq p}\widehat{C}^{ln}_\beta\left(\frac{y - x^n_c}{h}\right)^\beta\right| \\
&= \left|\sum_{\beta\geq p}\left(C^{ln}_\beta + (\widehat{C}^{ln}_\beta - C^{ln}_\beta)\right)\left(\frac{y - x^n_c}{h}\right)^\beta\right| \\
&= \left|\sum_{\beta\geq p} C^{ln}_\beta\left(\frac{y - x^n_c}{h}\right)^\beta + \sum_{\beta\geq p}(\widehat{C}^{ln}_\beta - C^{ln}_\beta)\left(\frac{y - x^n_c}{h}\right)^\beta\right| \\
&\leq \left|\sum_{\beta\geq p} C^{ln}_\beta\left(\frac{y - x^n_c}{h}\right)^\beta\right| + \left|\sum_{\beta\geq p}(\widehat{C}^{ln}_\beta - C^{ln}_\beta)\left(\frac{y - x^n_c}{h}\right)^\beta\right| \quad \text{[Triangle inequality.]} \\
&= E_T(p) + \left|\sum_{\beta\geq p}(\widehat{C}^{ln}_\beta - C^{ln}_\beta)\left(\frac{y - x^n_c}{h}\right)^\beta\right| \qquad (58)
\end{aligned}
$$

$\widehat{C}^{ln}_{\beta}$ can be simplified as follows

$$
\begin{aligned}
\widehat{C}^{ln}_{\beta} &= \frac{1}{\beta!}\sum_{\alpha<p} A^{l}_{\alpha}(-1)^{|\alpha|}h_{\alpha+\beta}\left(\frac{x^l_c - x^n_c}{h}\right) \\
&= \frac{1}{\beta!}\sum_{\alpha<p}\left[\frac{1}{\alpha!}\sum_{x_i\in I_1(l)} q_i\left(\frac{x_i - x^l_c}{h}\right)^{\alpha}\right](-1)^{|\alpha|}h_{\alpha+\beta}\left(\frac{x^l_c - x^n_c}{h}\right) \\
&= \frac{1}{\beta!}\sum_{x_i\in I_1(l)} q_i\left[\sum_{\alpha<p}\frac{1}{\alpha!}\left(\frac{x_i - x^l_c}{h}\right)^{\alpha}(-1)^{|\alpha|}h_{\alpha+\beta}\left(\frac{x^l_c - x^n_c}{h}\right)\right] \\
&= \frac{1}{\beta!}\sum_{x_i\in I_1(l)} q_i\left[\left(\sum_{\alpha\geq 0} - \sum_{\alpha\geq p}\right)\frac{1}{\alpha!}\left(\frac{x_i - x^l_c}{h}\right)^{\alpha}(-1)^{|\alpha|}h_{\alpha+\beta}\left(\frac{x^l_c - x^n_c}{h}\right)\right] \\
&= C^{ln}_{\beta} - \frac{1}{\beta!}\sum_{x_i\in I_1(l)} q_i\left[\sum_{\alpha\geq p}\frac{1}{\alpha!}\left(\frac{x_i - x^l_c}{h}\right)^{\alpha}(-1)^{|\alpha|}h_{\alpha+\beta}\left(\frac{x^l_c - x^n_c}{h}\right)\right] \\
&= C^{ln}_{\beta} + (\widehat{C}^{ln}_{\beta} - C^{ln}_{\beta})
\end{aligned}
\tag{59}
$$

$$
\begin{aligned}
&\left|\sum_{\beta\geq p}(\widehat{C}_\beta^{ln}-C_\beta^{ln})\left(\frac{y-x_c^n}{h}\right)^\beta\right| \leq \sum_{\beta\geq p}\left|(\widehat{C}_\beta^{ln}-C_\beta^{ln})\right|\left|\left(\frac{y-x_c^n}{h}\right)^\beta\right| \\
&= \sum_{\beta\geq p}\frac{1}{\beta!}\left|\sum_{x_i\in I_1(l)} q_i\left[\sum_{\alpha\geq p}\frac{1}{\alpha!}\left(\frac{x_i-x_c^l}{h}\right)^\alpha(-1)^{|\alpha|}h_{\alpha+\beta}\left(\frac{x_c^l-x_c^n}{h}\right)\right]\right|\left|\left(\frac{y-x_c^n}{h}\right)^\beta\right| \\
&\leq \sum_{\beta\geq p}\frac{1}{\beta!}\left[\sum_{x_i\in I_1(l)}|q_i|\left|\sum_{\alpha\geq p}\frac{1}{\alpha!}\left(\frac{x_i-x_c^l}{h}\right)^\alpha(-1)^{|\alpha|}h_{\alpha+\beta}\left(\frac{x_c^l-x_c^n}{h}\right)\right|\right]\left|\left(\frac{y-x_c^n}{h}\right)^\beta\right| \\
&\leq \sum_{\beta\geq p}\frac{1}{\beta!}\left[\sum_{x_i\in I_1(l)}|q_i|\sum_{\alpha\geq p}\frac{1}{\alpha!}\left|\left(\frac{x_i-x_c^l}{h}\right)^\alpha\right|\left|h_{\alpha+\beta}\left(\frac{x_c^l-x_c^n}{h}\right)\right|\right]\left|\left(\frac{y-x_c^n}{h}\right)^\beta\right| \\
&\leq \sum_{\beta\geq p}\frac{1}{\beta!}\left[\sum_{x_i\in I_1(l)}|q_i|\sum_{\alpha\geq p}\frac{1}{\alpha!}r^\alpha 2^{-\alpha/2}\left|h_{\alpha+\beta}\left(\frac{x_c^l-x_c^n}{h}\right)\right|\right]r^\beta 2^{-\beta/2} \\
&\leq \sum_{\beta\geq p}\frac{1}{\beta!}\left[\sum_{x_i\in I_1(l)}|q_i|\sum_{\alpha\geq p}\frac{1}{\alpha!}r^\alpha 2^{-\alpha/2}\left\{\prod_{j=1}^d\left|h_{\alpha_j+\beta_j}\left(\frac{(x_c^l)_j-(x_c^n)_j}{h}\right)\right|\right\}\right]r^\beta 2^{-\beta/2} \\
&\leq \sum_{\beta\geq p}\frac{1}{\beta!}\left[\sum_{x_i\in I_1(l)}|q_i|\sum_{\alpha\geq p}\frac{1}{\alpha!}r^\alpha 2^{-\alpha/2}\left\{\prod_{j=1}^d\left|\sqrt{(\alpha_j+\beta_j)!}2^{\alpha_j/2}2^{\beta_j/2}e^{-((x_c^l)_j-(x_c^n)_j)^2/2h^2}\right|\right\}\right]r^\beta 2^{-\beta/2} \\
&\leq Q_l\sum_{\alpha\geq p}\sum_{\beta\geq p}\frac{1}{\alpha!}\frac{1}{\beta!}\left[r^\alpha\left\{\prod_{j=1}^d\sqrt{(\alpha_j+\beta_j)!}\right\}\right]r^\beta \\
&\leq Q_l\sum_{\alpha\geq p}\sum_{\beta\geq p}\prod_{j=1}^d r^{\alpha_j+\beta_j}\frac{\sqrt{(\alpha_j+\beta_j)!}}{\alpha_j!\beta_j!} \\
&= Q_l\sum_{\alpha\geq p}\sum_{\beta\geq p}\prod_{j=1}^d\sqrt{\frac{1}{\alpha_j!\beta_j!}}r^{\alpha_j+\beta_j}\sqrt{\frac{(\alpha_j+\beta_j)!}{\alpha_j!\beta_j!}} \\
&\leq Q_l\sum_{\alpha\geq p}\sum_{\beta\geq p}\prod_{j=1}^d\sqrt{\frac{1}{\alpha_j!\beta_j!}}r^{\alpha_j+\beta_j}\sqrt{2}^{\alpha_j+\beta_j} \leq Q_l\sum_{\alpha\geq p}\sum_{\beta\geq p}\prod_{j=1}^d\sqrt{\frac{1}{\alpha_j!\beta_j!}}\left(\sqrt{2}r\right)^{\alpha_j+\beta_j} \\
&\leq Q_l\sum_{\alpha\geq p}\prod_{j=1}^d\sqrt{\frac{1}{\alpha_j!}}\left(\sqrt{2}r\right)^{\alpha_j}\left[\sum_{\beta\geq p}\prod_{j=1}^d\sqrt{\frac{1}{\beta_j!}}\left(\sqrt{2}r\right)^{\beta_j}\right] \\
&\leq Q_l\sum_{\alpha\geq p}\prod_{j=1}^d\sqrt{\frac{1}{\alpha_j!}}\left(\sqrt{2}r\right)^{\alpha_j}\left[\frac{1}{(1-\sqrt{2}r)^d}\sum_{k=0}^{d-1}\binom{d}{k}\left(1-(\sqrt{2}r)^p\right)^k\left(\frac{(\sqrt{2}r)^p}{\sqrt{p!}}\right)^{d-k}\right] \\
&\leq \frac{Q_l}{(1-\sqrt{2}r)^{2d}}\left[\sum_{k=0}^{d-1}\binom{d}{k}\left(1-(\sqrt{2}r)^p\right)^k\left(\frac{(\sqrt{2}r)^p}{\sqrt{p!}}\right)^{d-k}\right]^2 \quad [r<1/\sqrt{2}] \qquad (60)
\end{aligned}
$$