Numerical implementation of the Maxwell Green function

Sean Molesky
*Polytechnique Montréal**

This document summarizes the operating principles of Gila—a julia package implementing the three dimensional electromagnetic Green function.

Introduction Convergence in norm for volume integral methods is only guaranteed when the individual elements of the basis implementation respects all the continuity conditions of the quantities being approximated, as imposed by the kernel [1]. For electric, displacement, or magnetic fields, this means that a great deal of care must be taken in the selecting the partition functions and mesh of the domain [2]. A principal advantages of the fluctuating volume current method is that the integral kernel makes no demands on continuity, allowing the use of any set of discontinuous basis functions. In turn, this leads to better numerical performance for inhomogeneous and anisotropic materials [3].

I. CONCEPTUAL FOUNDATION

Notation—The formulation of macroscopic electromagnetics employed by Gila supposes that the response properties of each body in question are defined by a linear relative permittivity response $\mathbf{X}_\epsilon$ and a relative permeability response $\mathbf{X}_\mu$. As per standard convention, ω is used for the radial frequency $\mathsf{k}_\circ$ is used for the free space wave vector $2\pi/\lambda$, The letter $z = \sqrt{\mu_\circ/\epsilon_\circ}$ is used for the impedance of free space. In deference to flexibility, the electromagnetic field and polarization current densities are treated as the six component vectors $\mathbf{f} = \{\mathbf{e}, \mathbf{h}\}$ and $\mathbf{p} = \{\mathbf{j}, \mathbf{m}\}$. Throughout, capital bold letters are used to denote linear operators, with symbols reused to denote closely related meanings. For examples, $\mathbf{Z}$ is used to denote the constant linear operator defined by scalar multiplication by z. Both scalar and matrix multiplication, as context dictates, are implied by juxtaposition.

A. volume integral relations

Taking i, s and t subscripts to denote the incident, scattered, and total quantities—principally bare currents, those induced by scattering, and the combination of these two components respectively—

$$\frac{\mathsf{i}}{\mathsf{k}_\circ}\begin{bmatrix}\mathbf{j}_\mathsf{s}\\ -\mathbf{m}_\mathsf{s}\end{bmatrix} = \begin{bmatrix} z^{-1}\mathbf{X}_{\mathsf{j}e} & z\mathbf{X}_{\mathsf{j}\mathsf{h}} \\ z^{-1}\mathbf{X}_{\mathsf{m}e} & z\mathbf{X}_{\mathsf{m}\mathsf{h}}\end{bmatrix}\begin{bmatrix}\mathbf{e}_\mathsf{t}\\ \mathbf{h}_\mathsf{t}\end{bmatrix}$$
$$\frac{\mathsf{i}}{\mathsf{k}_\circ}\mathbf{p}_\mathsf{s} = \mathbf{X}\mathbf{f}_\mathsf{t}$$

defines the bound (induced or scattered) polarization current densities produced by the total electromagnetic field. Like the scattered electromagnetic field, the scattered current thus represents a self consistent solution of the electrodynamics in the presence of a linear scatterer. That such a transformation is possible follows as a consequence of the general invertability of an operator with a positive definite anti-symmetric component. Directly, given some computational domain Ω—assumed to include the a full set of boundary conditions—and let

$$\mathbf{M}_\circ = -\begin{bmatrix}\mathbf{Z}^{-1} & -\mathsf{i}\boldsymbol{\nabla}_\mathsf{c}^{\mathsf{k}_\circ}\\ \mathsf{i}\boldsymbol{\nabla}_\mathsf{c}^{\mathsf{k}_\circ} & \mathbf{Z}\end{bmatrix} \tag{1}$$

be the vacuum Maxwell operator, with $\boldsymbol{\nabla}_\mathsf{c}^{\mathsf{k}_\circ}$ denoting the vector curl divided by $\mathsf{k}_\circ$. Under these definitions the differential formulation of the Maxwell equations then becomes

$$(\mathbf{M}_\circ - \mathbf{X})\,\mathbf{f}_\mathsf{t} = \frac{\mathsf{i}}{\mathsf{k}_\circ}\mathbf{p}_\mathsf{i}. \tag{2}$$

Take $\mathbf{G}_\circ$ to be the vacuum Green function, i.e. the inverse of $\mathbf{M}_\circ$. Decomposing $\mathbf{f}_\mathsf{t}$ as $\mathbf{f}_\mathsf{s} + \mathbf{f}_\mathsf{i}$, by using the fact that $\mathbf{M}_\circ\mathbf{f}_\mathsf{i} = \mathsf{i}\mathbf{p}_\mathsf{i}/\mathsf{k}_\circ$, Eq. (2) shows that

$$\mathbf{f}_\mathsf{s} = \mathbf{G}_\circ\mathbf{X}\mathbf{f}_\mathsf{t}. \tag{3}$$

Therefore,

$$\begin{aligned}[\mathbf{Id} - \mathbf{G}_\circ\mathbf{X}]\,\mathbf{f}_\mathsf{t} &= \mathbf{f}_\mathsf{i}\\ [\mathbf{Id} - \mathbf{X}\mathbf{G}_\circ]\,\mathbf{p}_\mathsf{s} &= -\mathsf{i}\mathsf{k}_\circ\mathbf{X}\mathbf{f}_\circ + \mathbf{X}\mathbf{G}_\circ\mathbf{p}_\mathsf{i}\\ [\mathbf{Id} - \mathbf{X}\mathbf{G}_\circ]\,\mathbf{p}_\mathsf{t} &= \mathbf{p}_\mathsf{i} + \mathbf{p}_\circ,\end{aligned} \tag{4}$$

where, for later convenience, we have used the identities $\mathbf{f}_\mathsf{i} = \mathbf{f}_\circ + (\mathsf{i}/\mathsf{k}_\circ)\,\mathbf{G}_\circ\mathbf{p}_\mathsf{i}$ with $\mathbf{f}_\circ$ representing a possible incoming radiative field—free solution entering through the boundary of the computational domain—and $\mathbf{p}_\circ = -\mathsf{i}\mathsf{k}_\circ\mathbf{X}\mathbf{f}_\circ$. Recall that a linear operator $\mathbf{A}$, when acting on a Hilbert space, can be decomposed into symmetric and anti-symmetric components as $\mathbf{A} = \mathbf{A}^\mathsf{h} + \mathsf{i}\mathbf{A}^\mathsf{s}$, where

$$\mathbf{A}^\mathsf{s} = \frac{\mathbf{A} + \mathbf{A}^\dagger}{2}, \qquad \mathbf{A}^\mathsf{a} = \frac{\mathbf{A} - \mathbf{A}^\dagger}{2\mathsf{i}}, \tag{5}$$

and a $\dagger$ superscript denotes the operator adjoint. Accordingly, if $\mathbf{A}^\mathsf{a}$ is either positive or negative definite ($\mathbf{A}^\mathsf{a} \succ 0$ or $\mathbf{A}^\mathsf{a} \prec 0$), then the kernel of $\mathbf{A}$ is empty, and $\mathbf{A}$ is invertible. Within the volume occupied by scattering material the anti-symmetric component of $\mathbf{X}$, and by extension the anti-symmetric component of $\mathbf{T} = \mathbf{X}^{-1} - \mathbf{G}_\circ$, is positive definite due to its physical connection with the positive

* sean.molesky@polymtl.ca

dissipation of power [4]. Based on these observations, it follows that the $(\mathbf{Id} - \mathbf{X}\mathbf{G}_\circ)$ appearing on the left-hand side of (4) is also invertible, leading to the definition

$$\mathbf{W} = [\mathbf{Id} - \mathbf{X}\mathbf{G}_\circ]^{-1}. \tag{6}$$

One of the central use cases intended for Gila is to implement iterative solution methods for evaluating Eq. (6) for specific input vectors.

B. analytic form

The implementation of the vacuum Green function supplied by Gila follows largely from its traditionally defined analytic form. Shifting to Fourier space, this result will be derived by "inverting" Eq. (2) at a complex frequency $\omega + \mathrm{i}\delta$, represented as $\zeta\omega = \omega + \mathrm{i}\delta$. Under this transformation, retaining all other conventions,

$$\breve{\mathbf{M}}_\circ = -\begin{bmatrix} \zeta\mathbf{Z}^{-1} & \times_{\mathbf{k}} \\ -\times_{\mathbf{k}} & \zeta\mathbf{Z} \end{bmatrix}, \tag{7}$$

where $\times_{\mathbf{k}}$ is the "cross-product" operator defined locally at any $\mathbf{k}$ vector index in terms of the cartesian components of $\mathbf{k}$—assumed to be normalized by $\mathsf{k}_\circ$ in all following expressions—as

$$\times_{\mathbf{k}} = \begin{bmatrix} 0 & -\mathsf{k}_z & \mathsf{k}_y \\ \mathsf{k}_z & 0 & -\mathsf{k}_x \\ -\mathsf{k}_y & \mathsf{k}_x & 0 \end{bmatrix}. \tag{8}$$

Employing block matrix inversion,

$$\breve{\mathbf{G}}_\circ = \begin{bmatrix} \zeta\mathbf{Z} & -\times_{\mathbf{k}} \\ \times_{\mathbf{k}} & \zeta\mathbf{Z}^{-1} \end{bmatrix} \begin{bmatrix} \breve{\mathbf{G}}^\circ_\circ & \mathbf{0} \\ \mathbf{0} & \breve{\mathbf{G}}^\circ_\circ \end{bmatrix}, \tag{9}$$

where $\breve{\mathbf{G}}^\circ_\circ = \left[\left(\mathsf{k}^2 - \zeta^2\right)\mathbf{Id} - \mathbf{k}\otimes\mathbf{k}\right]^{-1}$ and $\mathsf{k} = \|\mathbf{k}\|_{\mathrm{E}}$ (the norm of $\mathbf{k}$) [5]. Hence, noting that through the Fourier transform one may equally switch from $\times_{\mathbf{k}}$ back to $-\mathrm{i}\boldsymbol{\nabla}^{\mathsf{k}_\circ}_{\mathbf{c}}$, calculating the vacuum Green function in real space amounts to calculating the inverse Fourier transform of $\breve{\mathbf{G}}^\circ_\circ = \frac{1}{\mathsf{k}^2-\zeta^2}\left[\mathbf{Id} - \zeta^{-2}\mathbf{k}\otimes\mathbf{k}\right]$. To carry out this task, we will make use of the following lemma.

Jordan's lemma—Take $\mathsf{f} : \mathbb{C} \to \mathbb{C}$ to be an analytic function on the upper half-plane

$$\mathrm{U} = \{z \in \mathbb{C} \mid \Im\, z \geqslant 0\},$$

and let

$$\mathrm{C}_{\mathsf{R}} = \left\{z \in \mathbb{C} \mid z = \mathsf{R}\mathrm{e}^{\mathrm{i}\theta} \wedge \theta \in [0,\pi]\right\},$$

with $\mathsf{R} > 0$ to denote a semi-circle of radius R confined to the upper half plane. If for each $\mathsf{R} > 0$ there is a positive constant M_{R} such that $z \in \mathrm{C}_{\mathsf{R}} \Rightarrow \|\mathsf{f}(z)\| \leqslant \mathrm{M}_{\mathsf{R}}$ and $\mathrm{M}_{\mathsf{R}} \to 0$ as $\mathsf{R} \to \infty$, then

$$\lim_{\mathsf{R}\to\infty} \int_{\mathrm{C}_{\mathsf{R}}} \mathrm{d}z\ \mathsf{f}(z)\, \mathrm{e}^{\mathrm{i}\alpha z}$$

whenever $\alpha > 0$.

Proof. To begin, notice that $\sin\theta$ is a concave function on $\theta \in [0,\pi]$, and that $\sin\frac{\pi}{2} = 1$. As, such

$$(\forall \alpha \in [0,1])\ \ \sin\left(\alpha\frac{\pi}{2}\right) \geqslant \alpha \Rightarrow$$
$$\left(\forall \theta \in \left[0,\frac{\pi}{2}\right]\right)\ \ \sin(\theta) \geqslant \frac{2}{\pi}\theta.$$

Thus, recalling that exp is a monotonic function on $\mathbb{R}$, if $\mathsf{R} > 0$, this inequality implies that $\exp(-\mathsf{R}\sin\theta) \leqslant \exp(-2\mathsf{R}\theta/\pi)$ so that

$$\int_0^{\pi/2} \mathrm{d}\theta\ \mathrm{e}^{-\mathsf{R}\sin\theta} \leqslant \int_0^{\pi/2} \mathrm{d}\theta\ \mathrm{e}^{-2\mathsf{R}\theta/\pi} = \frac{\pi}{2\mathsf{R}}\left(1 - \mathrm{e}^{-\mathsf{R}}\right) \leqslant \frac{\pi}{2\mathsf{R}}.$$

Transforming to polar coordinates

$$\int_{\mathrm{C}_{\mathsf{R}}} \mathrm{d}z\ \mathsf{f}(z)\, \mathrm{e}^{\mathrm{i}\alpha z} =$$
$$\mathsf{R}\int_0^{\pi} \mathrm{d}\theta\ \mathrm{i}\mathrm{e}^{\mathrm{i}\theta}\mathsf{f}\left(\mathsf{R}\mathrm{e}^{\mathrm{i}\theta}\right)\exp\left[\mathrm{i}\alpha\mathsf{R}\left(\cos\theta + \mathrm{i}\sin\theta\right)\right].$$

Therefore, because $\|\mathrm{i}\mathrm{e}^{\mathrm{i}\theta}\mathsf{f}\left(\mathsf{R}\mathrm{e}^{\mathrm{i}\theta}\right)\exp\left(\mathrm{i}\alpha\mathsf{R}\cos\theta\right)\| \leqslant \mathrm{M}_{\mathsf{R}}$,

$$\int_{\mathrm{C}_{\mathsf{R}}} \mathrm{d}z\ \mathsf{f}(z)\, \mathrm{e}^{\mathrm{i}\alpha z} \leqslant \frac{\pi}{\alpha}\mathrm{M}_{\mathsf{R}}.$$

Having established this result, the limit that $\mathsf{R} \to \infty$ can be taken, proving the lemma. □

Let $\mathbf{s} = \frac{\mathbf{r}-\tilde{\mathbf{r}}}{\lambda}$, with the associated scalar magnitude denotes as s, be the wavelength scaled separation between a spatial points $\mathbf{r}$ and $\tilde{\mathbf{r}}$—leaving an overall factor of $(2\pi)^2$ compared to the standard definition of the inverse Fourier transform. Supposing that ζ is given a $\propto \mathsf{k}$ functional dependence in the limit of large k [6] by taking partial derivatives on $\mathbf{s}$ the outer product appearing in the definition of $\breve{\mathbf{G}}^\circ_\circ$ can be factored as

$$\mathbf{G}^\circ_\circ = \mathbf{Id}^{\uparrow}\mathbf{g}_\circ + \boldsymbol{\nabla}^{\mathsf{k}_\circ}_{\mathbf{d}}\mathbf{g}_\circ\boldsymbol{\nabla}^{\mathsf{k}_\circ\dagger}_{\mathbf{d}} \tag{10}$$

where

$$\mathbf{g}_\circ(\mathbf{s}) = \frac{1}{2\pi}\int_0^\infty \mathrm{d}\mathsf{k}\ \mathsf{k}^2 \int_0^\pi \mathrm{d}\theta\ \mathrm{i}\sin\theta\ \frac{\mathrm{e}^{2\pi\mathrm{i}\mathsf{k}\mathsf{s}\cos\theta}}{(\mathsf{k}^2 - \zeta^2)} \tag{11}$$
$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} \mathrm{d}\mathsf{k}\ \frac{\mathsf{k}\mathrm{e}^{2\pi\mathrm{i}\mathsf{k}\mathsf{s}}}{\mathsf{s}\,(\mathsf{k}^2 - \zeta^2)} = \mathrm{i}\ \frac{\mathrm{e}^{2\pi\mathrm{i}\zeta\mathsf{s}}}{4\pi\mathsf{s}},$$

$\nabla^{\mathsf{k}_\circ}_{\mathbf{d}}$ is the divergence operator, and $\mathbf{Id}^{\uparrow}$ is the operator transforming a scalar field into a three dimensional vector field, for some index $\mathbf{x}$, by the rule

$$\mathbf{Id}^{\uparrow}\mathsf{f}(\mathbf{x}) \mapsto \begin{bmatrix} \mathsf{f}(\mathbf{x}) & 0 & 0 \\ 0 & \mathsf{f}(\mathbf{x}) & 0 \\ 0 & 0 & \mathsf{f}(\mathbf{x}) \end{bmatrix}. \tag{12}$$

Because we have defined $\breve{\mathbf{G}}_{\circ}^{\circ} = \left[\left(k^2 - \zeta^2\right)\mathbf{Id} - \mathbf{k}\otimes\mathbf{k}\right]^{-1}$, by making the appropriate associations for representations of operators in Fourier space,

$$\left[\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ} - \zeta^2\mathbf{Id}\right]\breve{\mathbf{G}}_{\circ}^{\circ} = \mathbf{Id}. \tag{13}$$

Hence,

$$\begin{aligned}\zeta^2\breve{\mathbf{G}}_{\circ}^{\circ} &= \boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\left(\mathbf{Id}^{\uparrow}\mathbf{g}_{\circ} + \boldsymbol{\nabla}_{\mathbf{d}}^{k_\circ}\mathbf{g}_{\circ}\boldsymbol{\nabla}_{\mathbf{d}}^{k_\circ\dagger}\right) - \mathbf{Id} \\ \breve{\mathbf{G}}_{\circ}^{\circ} &= \frac{1}{\zeta^2}\left(\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\mathbf{Id}^{\uparrow}\mathbf{g}_{\circ} - \mathbf{Id}\right),\end{aligned} \tag{14}$$

where in going from the second to the third line we have used the fact that $\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\boldsymbol{\nabla}_{\mathbf{d}}^{k_\circ} = \boldsymbol{\nabla}_{\mathbf{d}}^{k_\circ}\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ} = \mathbf{0}$. For implementation purposes, it may also be helpful to note that for a complex frequency Eq. (4) becomes

$$\begin{aligned}&\left[(\mathbf{Id} + z\mathbf{X}) - z\mathbf{X}\begin{bmatrix}\left(\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\right)^2 & i\frac{z}{\zeta^2}\left(\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\right)^3 \\ -i\frac{z}{\zeta^2}\left(\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\right)^3 & \left(\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\right)^2\end{bmatrix}\mathbf{Id}^{\uparrow}\mathbf{g}_{\circ}\right]\mathbf{p}_t \\ &= \mathbf{p}_i + \mathbf{p}_{\circ}.\end{aligned} \tag{15}$$

II. INTEGRAL FORMS

Appealing to a reformulation in terms of differential forms, or the careful use of vector calculus identities, the $\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}$ operator appearing above can be used to transform the volume integrals implicit in Eq. (14) into surface integrals. Applying this transformation

$$\begin{aligned}&\mathbf{q}_{m,j}^{b\dagger}\left[\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\boldsymbol{\nabla}_{\mathbf{c}}^{k_\circ}\mathbf{Id}^{\uparrow}\mathbf{g}_{\circ}\right]\mathbf{q}_{l,i}^{a} = \sum_{h,k}\left(\hat{\mathbf{n}}_h\times\hat{\mathbf{j}}\right)\cdot\left(\hat{\mathbf{n}}_k\times\hat{\mathbf{i}}\right) \\ &\oiint_h\oiint_k \mathbf{g}_{\circ}\left(\mathbf{r}_m^b - \mathbf{r}_n^a + \left(\mathbf{v}_m^b - \mathbf{v}_l^a\right)\right).\end{aligned} \tag{16}$$

where a and b are used as volume labels, m and l are used as cube labels, j and i are used as direction labels, and $\hat{\mathbf{n}}_h$ and $\hat{\mathbf{n}}_k$ are the faces of the target cube and source cube respectively. Using the cube face storage convention describe in sec. V, using the face pair numbering $6*(\text{source face index}-1)+\text{target face index}$, the face contributions for a given source and target cube pair are

$$\begin{aligned}\hat{\mathbf{i}}\hat{\mathbf{i}} &= \{15,-16,-21,22,29,-30,-35,36\} \\ \hat{\mathbf{j}}\hat{\mathbf{i}} &= \{-13,14,19,-20\} \\ \hat{\mathbf{k}}\hat{\mathbf{i}} &= \{-25,26,31,-32\} \\ \hat{\mathbf{i}}\hat{\mathbf{j}} &= \{-3,4,9,-10\} \\ \hat{\mathbf{j}}\hat{\mathbf{j}} &= \{1,-2,-7,8,29,-30,-35,36\} \\ \hat{\mathbf{k}}\hat{\mathbf{j}} &= \{-27,28,33,-34\} \\ \hat{\mathbf{i}}\hat{\mathbf{k}} &= \{-5,6,11,-12\} \\ \hat{\mathbf{j}}\hat{\mathbf{k}} &= \{-17,18,23,-24\} \\ \hat{\mathbf{k}}\hat{\mathbf{k}} &= \{1,-2,-7,8,15,-16,-21,22\}.\end{aligned} \tag{17}$$

To reproduce the physics of non-local response, we implement the model proposed in [7]. To that end, we assign a charge to each cubic cell, through the domain function c_n^a. The interaction between charges is then

$$c_m^b\mathbf{C}c_l^a, \tag{18}$$

where $\mathbf{C}$ is integration Kernel that corresponds to the Coulomb potential. Using that

$$\begin{aligned}\frac{1}{x} &= k_\circ\boldsymbol{\nabla}_{\mathbf{d}}^{k_\circ}\frac{\mathbf{x}}{2x} \\ \hat{\mathbf{n}}_h\cdot k_\circ\frac{\mathbf{x}}{2x} &= k_\circ^2\boldsymbol{\nabla}_{\mathbf{d}}^{k_\circ}\left[\hat{\mathbf{n}}_h\cdot\frac{\mathbf{x}\otimes\mathbf{x}}{6x}\right],\end{aligned} \tag{19}$$

successive applications of the divergence theorem allow one to express the volume integrals as surface integrals. Defining

$$\mathbf{D}(\mathbf{x}) = k_\circ^2\frac{\mathbf{x}\otimes\mathbf{x}}{6x}, \tag{20}$$

we have

$$\begin{aligned}c_m^b\mathbf{C}c_l^a &= -c_m^b a_m^b\sum_{h,k} \\ &\oiint_h\oiint_k \hat{\mathbf{n}}_h\cdot\mathbf{D}\left(\mathbf{r}_m^b - \mathbf{r}_l^a + \left(\mathbf{v}_m^b - \mathbf{v}_l^a\right)\right)\cdot\hat{\mathbf{n}}_k\end{aligned} \tag{21}$$

Like before, a and b denote the cuboid where the charge is located, n and l are cube/voxel labels, $\mathbf{r}_m^b$ and $\mathbf{r}_l^a$ are the positions of the centers of the cubes, $\mathbf{v}_m^b$ and $\mathbf{v}_l^a$ are the positions to the integration surfaces with respect to the center of each cube, and $\hat{\mathbf{n}}_h$ and $\hat{\mathbf{n}}_l$ are the normal vectors to the integration surfaces. I'm not sure that the notation is how you best like it. Also, this paragraph is redundant, but I'm just making sure that I understand everything.

Additionally, the interaction between cubic charges and currents is given by

$$\begin{aligned}\mathbf{e}_m^b\mathbf{q}_l^a &= -k_\circ\boldsymbol{\nabla}^{k_\circ}V_m^b\mathbf{q}_l^a \\ &= -k_\circ\sum_h\left(\mathbf{q}_l^a\cdot\hat{\mathbf{n}}_h\right)\oiint_h V_m^b\left(\mathbf{r}_m^b - \mathbf{r}_l^a - \mathbf{v}_l^a\right),\end{aligned} \tag{22}$$

where $V_m^b(\mathbf{r})$ is the electrostatic potential at $\mathbf{r}$ due to the homogeneous charge distribution in the parallelepiped labeled by b and m. Its analytical expression can be found in [8]. The formula therein expresses the coordinates of $\mathbf{r}$ from the $(0,0,0)$ vertex. In terms of the coordinates with respect to the center of the cubic cell b, m, we have

$$\begin{aligned}V_m^b(\mathbf{r}) = \sum_{i,j,k=0}^{1}P_{\{x_i,y_j,z_k\}}^c\Bigg[&x_i y_j\,\mathrm{arctanh}\left(\frac{z_k}{r_{ijk}}\right) \\ &-\frac{x_i^2}{2}\arctan\left(\frac{y_j z_k}{x_i r_{ijk}}\right)\Bigg],\end{aligned} \tag{23}$$

where $P_{\{t_1,t_2,t_3\}}^c$ permutes cyclically the variables t_i in the expression upon which it acts, in square brackets. In Eq. (23), $\mathbf{r}_i = \mathbf{r} + (-1)^{(i+1)}\mathbf{l}$, where the spatial components of $\mathbf{l}$ are the lengths of the parallelepiped in each direction, and $r_{ijk} = \sqrt{x_i^2 + y_j^2 + z_k^2}$.

III. USER INTERFACE

Gila works exclusively with rectangular cuboids, referred alternatively as volumes or domains, composed of small cubic cells.

IV. COMPUTATIONAL MODEL

! need to investigate SVD tradeoff + CUDA multiplication speed up.

V. CONVENTIONS

· Cube faces.

The relative separation between the centers of a pair of cubes is specified by grid coordinates. In addition to this information, to compute the surface integrals described above, "local" relative separation information between pairs of faces of the pair of cubes is also need. This information is generated by the cubeFaces function in the GilaCirc module. The storage convention reference a cube in the upper-forward-right quadrant of a set of Cartesian vectors, using `U` and `L` labels to denote "upper" and "lower" coordinate values relative to a given plane and associated normal direction.

$$\texttt{cat}(\texttt{yzL}, \texttt{yzU}, \texttt{xzL}, \texttt{xzU}, \texttt{xyL}, \texttt{xyU}, \texttt{dims} = 3)$$

Point specifications for each face follow a counterclockwise convention when viewed along the positive normal axis of the corresponding plane.

· Green function interaction elements.

The storage format of the Green function interaction elements for every pair of cubes is $[[\texttt{ii}, \texttt{ji}, \texttt{ki}]^{\mathsf{T}}; [\texttt{ij}, \texttt{jj}, \texttt{kj}]^{\mathsf{T}}]; [\texttt{ik}, \texttt{jk}, \texttt{kk}]^{\mathsf{T}}]$ during the writing process to facilitate numbering, debugging, and possible future extensions. However, by reciprocity, the inclusion of both transpose element pairs is redundant, even for asymmetrically sized cubes. Hence, once the writing process in terminate, the storage format is condensed to $[[\texttt{ii}, \texttt{jj}, \texttt{kk}]^{\mathsf{T}}; [\texttt{ij}, \texttt{ik}, \texttt{jk}]^{\mathsf{T}}]$;

· Scaling conventions.

Vectors in Gila are taken to be coefficients defined relative to a unit strength basis of cuboids; rows are taken to be coefficients defined relative to the associated linear functionals mapping a unique cuboid and direction combination to unity. Hence, when used in integral expressions, the use of vectors leads to length scaling proportionalities, while the use of linear functionals causes inverse length scaling proportionalities.

Based on this convention, the rectangular surface integrals occurring in Eq. (16) scale inversely to the characteristic length opposite to the face being integrated over in target cuboid, and proportionally to the characteristic lengths of the face being integrated over in the source cuboid. However, because the Green function kernel expression of Eq. (16) themselves scale proportionally to the characteristic length opposite to the face being integrated over in both the source and target, the overall scaling of Eq. (16) is inversely proportional to the volume of the source cuboid in the absence of singular contributions.

There is also a potential for confusion when observing that the weak singular integrals calculated via DIRECTFN are scaled by inverse volume of the cuboid. As opposed to the Gaussian quadrature subroutine used elsewhere, in order to provide greater flexibility, DIRECTFN works with the relative unit length scales, as opposed to normalizing all integrals to a fixed length. This choice causes the results of DIRECTFN to scale proportionally to the product of the source and target face areas. Multiplication by the inverse characteristic volume of the cuboid is thus a result of the Gila linear functional convention, resulting in the same scaling characteristics described above.

VI. GPU COMMANDS

· Get GPU operating information.

terminal command: nvidia-smi

1. Toeplitz and Circulant Forms

In the main text, we state that the translation invariance of the Green function results in the $\mathcal{P}$ matrix having a symmetric block Toeplitz form. Later, we then expanded this matrix to a block circulant form to further aid calculations. In this appendix, we provide additional details explaining these two statements. We begin with the symmetric block Toeplitz form. In our construction of the problem, there are five levels of indexing:

1. The two body indices (self and external interactions).

2. The $\{\mathsf{x}, \mathsf{y}, z\}$ the voxel number indices of the two basis functions with the cuboid girds (one index pair for each Cartesian coordinate).

3. The $\left\{\hat{\mathsf{i}}, \hat{\mathsf{j}}\right\}$ pair of Cartesian direction indices.

The qualification of symmetric block Toeplitz means that at each level of indexing the matrix subblocks have a Toeplitz form. For example, starting at the highest

level of indexing for a particular self body interaction, the matrix blocks are indexed by the $x_n x_m$ voxel values. These matrix blocks have a Toeplitz form

$$\mathcal{P} = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{11} & x_{12} & x_{13} \\ x_{31} & x_{21} & x_{11} & x_{12} \\ x_{41} & x_{31} & x_{21} & x_{11} \end{pmatrix}. \tag{24}$$

(in this example we consider a $4 \times 4 \times 4$ cubiod grid.) If we choose one of these matrices, for instance x_{14}, then the y index subblocks are again Toeplitz

$$x_{14} = \begin{pmatrix} y_{11}^{14} & y_{12}^{14} & y_{13}^{14} & y_{14}^{14} \\ y_{21}^{14} & y_{11}^{14} & y_{12}^{14} & y_{13}^{14} \\ y_{31}^{14} & y_{21}^{14} & y_{11}^{14} & y_{12}^{14} \\ y_{41}^{14} & y_{31}^{14} & y_{21}^{14} & y_{11}^{14} \end{pmatrix}.$$

Going deeper, the z index subblock of a given y block also show this same structure

$$y_{12}^{14} = \begin{pmatrix} z_{11}^{12;14} & z_{12}^{12;14} & z_{13}^{12;14} & z_{14}^{12;14} \\ z_{21}^{12;14} & z_{11}^{12;14} & z_{12}^{12;14} & z_{13}^{12;14} \\ z_{31}^{12;14} & z_{21}^{12;14} & z_{11}^{12;14} & z_{12}^{12;14} \\ z_{41}^{12;14} & z_{31}^{12;14} & z_{21}^{12;14} & z_{11}^{12;14} \end{pmatrix},$$

as do the 3×3 matrices associated with each z block. At each level, the Toeplitz blocks can be embedded in circulant blocks by treating either first or last element as a reflective boundary. The meaning of this statement again becomes clear in example. Taking $\mathcal{P}$ as in (24), at the outermost level the embedded matrix becomes

$$\mathcal{P}_{\mathcal{C}} = \left(\begin{array}{cccc|cccc} x_{11} & x_{12} & x_{13} & x_{14} & x_{aa} & x_{41} & x_{31} & x_{21} \\ x_{21} & x_{11} & x_{12} & x_{13} & x_{14} & x_{aa} & x_{41} & x_{31} \\ x_{31} & x_{21} & x_{11} & x_{12} & x_{13} & x_{14} & x_{aa} & x_{41} \\ x_{41} & x_{31} & x_{21} & x_{11} & x_{12} & x_{13} & x_{14} & x_{aa} \\ \hline x_{aa} & x_{41} & x_{31} & x_{21} & x_{11} & x_{12} & x_{13} & x_{13} \\ x_{14} & x_{aa} & x_{41} & x_{31} & x_{21} & x_{11} & x_{12} & x_{13} \\ x_{13} & x_{14} & x_{aa} & x_{41} & x_{31} & x_{21} & x_{11} & x_{12} \\ x_{12} & x_{13} & x_{14} & x_{aa} & x_{41} & x_{31} & x_{21} & x_{11} \end{array}\right), \tag{25}$$

where the x_{ij} entries that appear are also embedded in circulant subblocks (expanded forms of the identically named x_{ij} as used previously). Unlike the block Toeplitz form, which is not Toeplitz matrix, the block circulant form is circulant. The usefulness of this procedure stems from the special properties of circulant matrices. Specifically, the eigenvectors are the Fourier vectors and the eigenvalues are

$$\lambda_j = c_0 + c_{n-1}\omega_j + c_{n-2}\omega_j^2 + \ldots + c_1\omega_j^{n-1}, \tag{26}$$

where $\omega_j = e^{2\pi i j/n}$, and c_i are the entires along a single row of the circulant matrix. This lets us the expand the individual tensor components of the matrix, $(N_x^s + N_x^t) \times \left(N_y^s + N_y^t\right) \times (N_z^s + N_z^t)$ with s standing for the source object and t the target, as

$$\mathcal{P}_{C}^{\hat{i},\hat{j}} = F^* \left(F\left(c\right)\right) F, \tag{27}$$

where F is the unnormalized Fourier transform, $e^{-2\pi ij/n}$ convention, and c is the first column. Using the fast Fourier transform algorithm, these operation are much more computationally efficient than standard matrix multiplication, taking the number of actions that needs to be performed from $\left(N_x N_y N_z\right)^2$ to approximately $N_x N_y N_z \lg\left(N_x N_y N_z\right)$. The basic premise of the fast Fourier transform is discussed in the next appendix.

2. Removal of embedding overhead

Given the circulant embedding described above, the operation that must generally be carried out to compute the action of $\mathbf{G}_\circ$ on a source vector $\mathbf{s}$ is

$$\mathbf{P}\mathbf{F}_c^\dagger \tilde{\mathbf{G}}_\circ \mathbf{F}_c \mathbf{E}\mathbf{s}, \tag{28}$$

where $\mathbf{F}_c$ is the Fourier transform operator in the "circulant space", $\tilde{\mathbf{G}}_\circ$ is the Fourier transform of the first column of $\mathbf{G}_\circ$, $\mathbf{E}$ is the embedding operation, and $\mathbf{P}$ is the projection operation. The computational overhead that would seem to be implied by the embedding and projection operations can be removed by treating these operations in the Fourier basis.

Suppose that the dimension of $\mathbf{s}$ is n, and let $\mathbf{c}_k$ denote the Fourier transform coefficients of the step function on the first n dimensions of the circulant space. Take $\mathbf{s}_c$ to be the vector obtained by repeating $\mathbf{s}$. In matrix form, the Eq.(28) is equivalent to

3. Fast Fourier Transform

The N roots of unity of the exponential function furnish all irreducible representations of the cyclic group C_N. For N discrete spacings, they also provide the possible values of the Fourier kernel. This connection gives the matrix form of the Fourier transformation special properties that can be used to dramatically improve the efficiency of computations.

There are two ways to think about the the matrix form of the Fourier transform for N cubes. First, it is the character table of C_N: the element at index $\{l, m\}$ is $e^{2\pi i(lm)/N}$, where l and m range from 0 to $N-1$. As such, the columns and rows of the Fourier transform matrix (DFT) are orthogonal under the conjugate inner product, and by including a prefactor of $1/\sqrt{N}$ become orthonormal. It is also the projection onto the different irreducible representations. As rotations are abelian, any irreducible representation is a one dimensional subspace, and so the total space of functions over N discrete points can be decomposed into these vectors. The matrix of the Fourier transform is the collection of inner products with these vectors.

The central idea of the fast Fourier transform (FFT) is that if $\mathsf{N} = 2^{\mathfrak{n}}$ then there is a set of group morphism $\mathfrak{m}_{\mathfrak{i}}$ equating representations on $\mathsf{N}_{\mathfrak{i}}$ with representations on $\mathsf{N}_{\mathfrak{i}+1} \oplus \mathsf{N}_{\mathfrak{i}+1}$ where $\mathsf{N}_{\mathfrak{i}+1} = \mathsf{N}_{\mathfrak{i}}/2$. The existence of these transformations can be understood in two steps. Begin by picturing the N sampling points of a vector as the N roots of unity in the complex plane, and subdivide the collection of points into two sets by selections alternating. Take the inner product with the same irreducible vector on each of the these subspaces, say $e^{2\pi i m/(N/2)}$. By then multiplying the result from one of the two subspaces with any rotation of the group $e^{\pi i m/(N/2)}$ smaller than a $e^{i\pi}$, effectively the inner product with $e^{2\pi i(m)/N}$. Using this basis and one shifted by π, which must be distinct, all vectors of the larger group are accounted for. In matrix form

$$\mathsf{F}_{\mathfrak{i}+1} = \begin{pmatrix} \mathsf{I}_{N/2} & \mathsf{D} \\ \mathsf{I}_{N/2} & -\mathsf{D} \end{pmatrix} \begin{pmatrix} \mathsf{F}_{\mathfrak{i}} & 0 \\ 0 & \mathsf{F}_{\mathfrak{i}} \end{pmatrix} \sigma,$$

where σ is the matrix of odd-even permutations, and D a diagonal matrix with entries $e^{2\pi i n/N}$ for $\mathfrak{n} \leqslant (\mathsf{N}-1)/2$.

The complete algorithm uses factorization recursively. With careful consideration, the form of the self referential σ matrix can be determined without the need for direct computation, meaning that prior to multiplying with the D matrices the algorithm is linear. For a vector of size $2^{\mathfrak{l}}$, $\mathfrak{l}$ steps are then required to form the proper Fourier transformed vector, each requiring N multiplications. As a result, the number of steps needed by the algorithm is

$$\mathsf{FFT} \propto \mathsf{N} \ln_2(\mathsf{N}),$$

in the limit of large N. Note however that this relies on specific vector sizes. For most FFT libraries similar algorithms are implemented if $\mathfrak{n}$ is any combination of powers of the primes $\{2, 3, 5, 7\}$ with increasing performance for smaller repeated primes (ideally $2^{\mathfrak{n}}$).

4. §Particulars

• In order to simplify the Green function action, Gla enforces that the number of cells in each Cartesian dimension of a volume be divisible by two. If the user enters an odd cell number, one additional cell is added, and the cell scale is redefined.

• I have flipped the information storage convention for the reduced green function when the number of target cells is larger than the number of source cells. After the first element, the entries correspond to the target volume loop, stored in standard ascending, rather than reversed order.

[1] M. Van Beurden and S. Van Eijndhoven, Well-posedness of domain integral equations for a dielectric object in homogeneous background, Journal of Engineering Mathematics **62**, 289 (2008).

[2] M. Van Beurden and S. Van Eijndhoven, Gaps in present discretization schemes for domain integral equations, in *Electromagnetics in Advanced Applications, 2007. ICEAA 2007. International Conference on* (IEEE, 2007) pp. 673–675.

[3] J. Markkanen, P. Yla-Oijala, and A. Sihvola, Discretization of volume integral equation formulations for extremely anisotropic materials, IEEE Transactions on Antennas and Propagation **60**, 5195 (2012).

[4] L. D. Landau and E. M. Lifshitz, *Statistical physics: volume 5*, Vol. 5 (Elsevier, 2013).

[5] Equivalent forms of the block matrix inverse show that the order of $\times_{\mathbf{k}}$ and $\breve{\mathbf{G}}^{\circ}_{\circ}$ appearing the off-diagonal blocks may be switched. This an interesting, necessary, observation about how the Green function separates longitudinal and transverse evolution.

[6] This assumption of regularization is motivated by the fact that on arbitrarily small length scales our present mathematical model is very unlikely to match physical reality.

[7] T. Giovannini, L. Bonatti, P. Lafiosca, L. Nicoli, M. Castagnola, P. G. Illobre, S. Corni, and C. Cappelli, Do we really need quantum mechanics to describe plasmonic properties of metal nanostructures?, ACS Photonics **9**, 3025 (2022).

[8] J. Waldvogel, The newtonian potential of a homogeneous cube, Zeitschrift für angewandte Mathematik und Physik ZAMP **27**, 867 (1976).