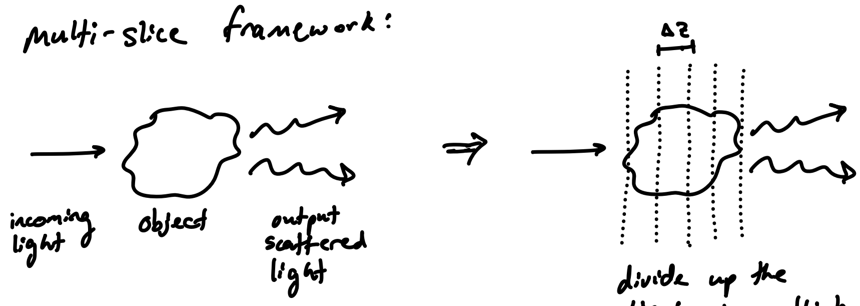
This is a derivation for the multi-slice mathematical framework for refractive-index reconstruction from electric-field measurements.
This follows the methodology introduced in Kamilov's paper:
"Optical Tomographic Image Reconstruction based on Beam propagation and Sparse regularization"

We make some slight tweaks, but the main incentive for this document is for beginner students to see a more thorough step-by-step of the derivation.

— Shwetadwip Chowdhury

multi-slice framework:



divide up the object into multiple slices and lets beam propagate through. Each layer of the object acts as a transmission mask.

Definitions:

$\vec{x}$: 3D $\Delta n(\vec{r})$ distribution
We want to solve for this
$\vec{x} \in \mathbb{R}^n$

$\vec{y}$: 2D complex-valued electric field measurements.
$\vec{y} \in \mathbb{C}^m$, where $m < n$

S: nonlinear and nonconvex forward model that predicts 2D efield measurement from known $\vec{x}$

$$\vec{y} = S(\vec{x})$$

Remember that the $S\{\dots\}$ operator is $\mathbb{R}^n \rightarrow \mathbb{C}^m$

propagation kernel by some distance Δz

$$S_k(\vec{x}) = \mathrm{diag}\big(P_k(\vec{x})\big) \cdot H \cdot S_{k-1}(\vec{x})$$

Efield at k-th layer, as predicted by $S\{\dots\}$

transmission mask at k-th layer of object

Efield at k-1 layer, as predicted by $S\{\dots\}$

$$P_k(\vec{x}) = \exp(j k_0 \Delta z \cdot \vec{x}_k)$$

corresponds to object's RI values at the object's k-th layer

Say object has a total of K layers
Then to solve for $\vec{x}$, write the following optimization

$$\hat{x} = \underset{\vec{x}}{\mathrm{argmin}} \left\{ D(\vec{x}) + \tau R(\vec{x}) \right\}$$

Data fidelity term, which we try to minimize

Regularizer

$$D(\vec{x}) = \frac{1}{L}\sum_{\ell=1}^{L} D_\ell(\vec{x})$$

$$= \frac{1}{2L}\sum_{\ell=1}^{L} \| \vec{y}_K^{\,\ell} - S_K^{\ell}(\vec{x}) \|_{\ell^2}^2$$

here, $S_K(\vec{x})$ is the forward model's prediction of the electric field at last object layer.

y_K is the raw measurement of the electric field at the last object layer.

ℓ is the acquisition number. This is an important parameter since you will have to take various measurements to solve this high-dimensional problem.

Definition:

$\ell 2$ norm

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \|\vec{x}\|_{\ell 2}^2 = x_1^2 + x_2^2 + \dots x_n^2$$

$$= \vec{x}^H \cdot \vec{x}$$

$$= \langle \vec{x}, \vec{x} \rangle$$

From ℓ^2 norm definition, we see that

$$D(\vec{x}) = \frac{1}{L}\sum_{\ell=1}^{L} D_\ell(\vec{x}) = \frac{1}{2L}\sum_{\ell=1}^{L} \| \vec{y}_K^{\,\ell} - S_K^{\ell}(\vec{x}) \|_{\ell^2}^2$$

is minimized whenever each component $D_\ell(\vec{x})$ is minimized. So lets find the minimum of just each individual $D_\ell(\vec{x})$, for which we just need to find gradient of:

$$D_\ell(\vec{x}) = \frac{1}{2}\|\vec{y}_K - S_K(\vec{x})\|_{\ell 2}^2$$

for notational sake, I will refer to $D_\ell(\vec{x})$ as $D(\vec{x})$, and generally ignore the "ℓ" parameter.

Derivation of gradient of Data-fidelity term:

$$D(\vec{x}) = \frac{1}{2} \left\| \vec{y}_k - S_k(\vec{x}) \right\|_{\ell_2}^2$$

$$= \frac{1}{2}\left[\left(\vec{y}_K - S_K(\vec{x})\right)^H \cdot \left(\vec{y}_K - S_K(\vec{x})\right) \right]$$

$$= \frac{1}{2}\left[\left(\vec{y}_K^H - S_K(\vec{x})^H\right) \cdot \left(\vec{y}_K - S_K(\vec{x})\right) \right]$$

$$= \frac{1}{2}\left[\vec{y}_K^H \cdot \vec{y}_K - \vec{y}_K^H S_K(\vec{x}) - S_K(\vec{x})^H \cdot \vec{y}_K + S_K(\vec{x})^H \cdot S_K(\vec{x}) \right] \quad \ldots (1)$$

$$= \frac{1}{2}\langle \vec{y}_K, \vec{y}_K \rangle - \mathrm{Re}\left\{ S_K(\vec{x}), \vec{y}_K^H \right\} + \frac{1}{2}\langle S_K(\vec{x}), S_K(\vec{x}) \rangle$$

Define:

$$\frac{\partial}{\partial x_j} S_k(\vec{x}) = \begin{bmatrix} \frac{\partial}{\partial x_j}\left[S_k(\vec{x})\right]_1 \\ \vdots \\ \frac{\partial}{\partial x_j}\left[S_k(\vec{x})\right]_m \end{bmatrix} \qquad \nabla D(\vec{x}) = \left[\frac{\partial}{\partial x_1} D(\vec{x}) \quad \cdots \quad \frac{\partial}{\partial x_n} D(\vec{x}) \right]$$

substitute into each column

Hessian Matrix

$$\frac{\partial}{\partial \vec{x}} S_k(\vec{x}) = \left[\frac{\partial}{\partial x_1} S_K(\vec{x}) \quad \cdots \quad \frac{\partial}{\partial x_n} S_K(\vec{x}) \right] = \begin{bmatrix} \frac{\partial}{\partial x_1}\left[S_k(\vec{x})\right]_1 & \cdots & \frac{\partial}{\partial x_n}\left[S_k(\vec{x})\right]_1 \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1}\left[S_k(\vec{x})\right]_m & \cdots & \frac{\partial}{\partial x_n}\left[S_k(\vec{x})\right]_m \end{bmatrix}_{m \times n}$$

Starting from ①

$$D(\vec{x}) = \frac{1}{2}\cdot\left[\vec{y}_k^H\cdot\vec{y}_k - \vec{y}_k^H S_k(\vec{x}) - S_k(\vec{x})^H\cdot\vec{y}_k + S_k(\vec{x})^H\cdot S_k(\vec{x})\right]$$

$$= \frac{1}{2}\cdot\left[\vec{y}_k^H\cdot\vec{y}_k - 2\,\mathrm{Re}\{\vec{y}_k^H S_k(\vec{x})\} + S_k(\vec{x})^H\cdot S_k(\vec{x})\right]$$

$$= \frac{1}{2}\vec{y}_k^H\cdot\vec{y}_k - \mathrm{Re}\{\vec{y}_k^H\cdot S_k(\vec{x})\} + \frac{1}{2} S_k(\vec{x})^H\cdot S_k(\vec{x})$$

$$\frac{\partial}{\partial x_j}D(\vec{x}) = \underbrace{\frac{1}{2}\frac{\partial}{\partial x_j}\vec{y}_k^H\cdot\vec{y}_k}_{=0,\text{ since } \vec{y}_k \text{ is not dependent on } \vec{x}} - \frac{\partial}{\partial x_j}\cdot\mathrm{Re}\{\vec{y}_k^H\cdot S_k(\vec{x})\} + \frac{1}{2}\frac{\partial}{\partial x_j}\left[S_k(\vec{x})^H\cdot S_k(\vec{x})\right]$$

$$= -\mathrm{Re}\left\{\frac{\partial}{\partial x_j}\vec{y}_k^H\cdot S_k(\vec{x})\right\} + \frac{1}{2}\underbrace{\frac{\partial}{\partial x_j}\left[S_k(\vec{x})^H\cdot S_k(\vec{x})\right]}_{\text{how do we solve this?}}$$

lets do derivation for single function value

Real valued

$$s(x) = a(x) + j\,b(x)$$

$$s^H(x) = s^*(x) = a(x) - j\,b(x)$$

Therefore:

$$s^H(x)\cdot s(x) = a^2(x) + b^2(x)$$

$$\frac{\partial}{\partial x}\left[s^H(x)\cdot s(x)\right] = \frac{\partial}{\partial x}\left[a^2(x) + b^2(x)\right]$$

$$= 2a(x)\frac{\partial}{\partial x}a(x) + 2b(x)\frac{\partial}{\partial x}b(x) \quad \ldots ②$$

Now consider:

$$2\, s^H(x)\cdot\frac{\partial}{\partial x}s(x)$$

$$= 2\left[a(x) - j\,b(x)\right]\cdot\frac{\partial}{\partial x}\left[a(x) + j\,b(x)\right]$$

$$= 2\left[a(x) - j\,b(x)\right]\cdot\left[\frac{\partial}{\partial x}a(x) + j\frac{\partial}{\partial x}b(x)\right]$$

$$= 2\left[a(x)\frac{\partial}{\partial x}a(x) + j\,a(x)\frac{\partial}{\partial x}b(x) - j\,b(x)\frac{\partial}{\partial x}a(x) + b(x)\frac{\partial}{\partial x}b(x)\right]$$

$$= \underbrace{2a(x)\frac{\partial}{\partial x}a(x) + 2b(x)\frac{\partial}{\partial x}b(x)}_{\text{Real component}} + 2j\underbrace{\left[a(x)\frac{\partial}{\partial x}b(x) - b(x)\frac{\partial}{\partial x}a(x)\right]}_{\text{Imaginary component}} \quad ③$$

If we compare ② and ③, we see that:

$$\frac{\partial}{\partial x}\left[s^H(x)\cdot s(x)\right] = \mathrm{Re}\left\{2\, s^H(x)\cdot \frac{\partial}{\partial x} s(x)\right\} \quad \cdots ④$$

We use this property below

$$\frac{\partial}{\partial x_j} D(\vec{x}) = -\mathrm{Re}\left\{\frac{\partial}{\partial x_j}\, \vec{y}_K^{\,H}\cdot S_K(\vec{x})\right\} + \frac{1}{2}\frac{\partial}{\partial x_j}\left[S_K(\vec{x})^H\cdot S_K(\vec{x})\right]$$

$$= -\mathrm{Re}\left\{\frac{\partial}{\partial x_j}\, \underline{\vec{y}_K^{\,H}}\cdot S_K(\vec{x})\right\} + \mathrm{Re}\left\{S_K^H(\vec{x})\cdot\frac{\partial}{\partial x_j} S_K(\vec{x})\right\}$$

constant with respect to x_j

$$= -\mathrm{Re}\left\{\vec{y}_K^{\,H}\cdot\frac{\partial}{\partial x_j} S_K(\vec{x})\right\} + \mathrm{Re}\left\{S_K^H(\vec{x})\cdot\frac{\partial}{\partial x_j} S_K(\vec{x})\right\}$$

$$= \mathrm{Re}\left\{\vec{S}_K^H(\vec{x})\cdot\frac{\partial}{\partial x_j} S_K(\vec{x}) - \vec{y}_K^{\,H}\cdot\frac{\partial}{\partial x_j} S_K(\vec{x})\right\}$$

$$\frac{\partial}{\partial x_j} D(\vec{x}) = \mathrm{Re}\left\{\left(S_K^H(\vec{x}) - \vec{y}_K^{\,H}\right)\cdot\frac{\partial}{\partial x_j} S_K(\vec{x})\right\}$$

$$\nabla D(\vec{x}) = \frac{\partial}{\partial \vec{x}} D(\vec{x})$$

$$= \left[\frac{\partial}{\partial x_1} D(\vec{x}) \quad \frac{\partial}{\partial x_2} D(\vec{x}) \quad \cdots \quad \frac{\partial}{\partial x_n} D(\vec{x})\right]$$

$$= \mathrm{Re}\left\{\left(S_K(\vec{x}) - y_K\right)^H\cdot\frac{\partial}{\partial \vec{x}} S_K(\vec{x})\right\}$$

$$\left[\nabla D(\vec{x})\right]^H = \mathrm{Re}\left\{\left[\frac{\partial}{\partial \vec{x}} S_K(\vec{x})\right]^H\cdot\left(S_K(\vec{x}) - \vec{y}_K\right)\right\}$$

This will be used for gradient descent protocol