PHYS 512 Problem Set 4

Teophile Lemay, 281081252

1

1.1 a)

Newton's method is a way to perform a non-linear least squares fit to some data by using the gradient of the model in parameter space to update the fit. Fitting the Lorentzian

$$d(t) = \frac{a}{1 + \frac{(t-t_0)^2}{w^2}}$$

requires the following derivatives:

$$\frac{\partial d}{\partial a} = \frac{1}{1 + \frac{(t-t_0)^2}{w^2}}$$

$$\frac{\partial d}{\partial t_0} = \frac{\frac{2a(t-t_0)}{w^2}}{\left(1 + \frac{(t-t_0)^2}{w^2}\right)^2}$$

$$\frac{\partial d}{\partial w} = \frac{\frac{2a(t-t_0)^2}{w^3}}{\left(1 + \frac{(t-t_0)^2}{w^2}\right)^2}$$

which make up the gradient of d in parameter space:

$$\boldsymbol{\nabla} d = \begin{pmatrix} \frac{\partial d}{\partial a} \\ \frac{\partial d}{\partial t_0} \\ \frac{\partial d}{\partial w} \end{pmatrix}$$

For any parameter guess $\mathbf{m}$ evaluating the gradient over an interval and comparing to the data, Newton's method gives a parameter update according to

$$\boldsymbol{\nabla} d(\mathbf{m})^T N^{-1} \boldsymbol{\nabla} d(\mathbf{m}) \delta \mathbf{m} = \boldsymbol{\nabla} d(\mathbf{m})^T N^{-1} (\text{data} - d(\mathbf{m}))$$

$$\delta \mathbf{m} = (\boldsymbol{\nabla} d(\mathbf{m})^T N^{-1} \boldsymbol{\nabla} d(\mathbf{m}))^{-1} \boldsymbol{\nabla} d(\mathbf{m})^T N^{-1} (\text{data} - d(\mathbf{m}))$$

$$\delta \mathbf{m} = (\boldsymbol{\nabla} d(\mathbf{m})^T \boldsymbol{\nabla} d(\mathbf{m}))^{-1} \boldsymbol{\nabla} d(\mathbf{m})^T (\text{data} - d(\mathbf{m}))$$

Where N is a noise matrix which can be omitted by setting it to identity. Iterating Newton's method until the parameter update $\delta \mathbf{m}$ becomes smaller than some threshold gives the best fit parameters.

Figure 1 shows the result of a Newton's method fit of the data in "sidebands.npz" to the Lorentzian using the analytical gradient derived above. The iterations were stopped once the magnitude of the parameter update δm reached the threshold of 10^{-10} which took 13 steps.
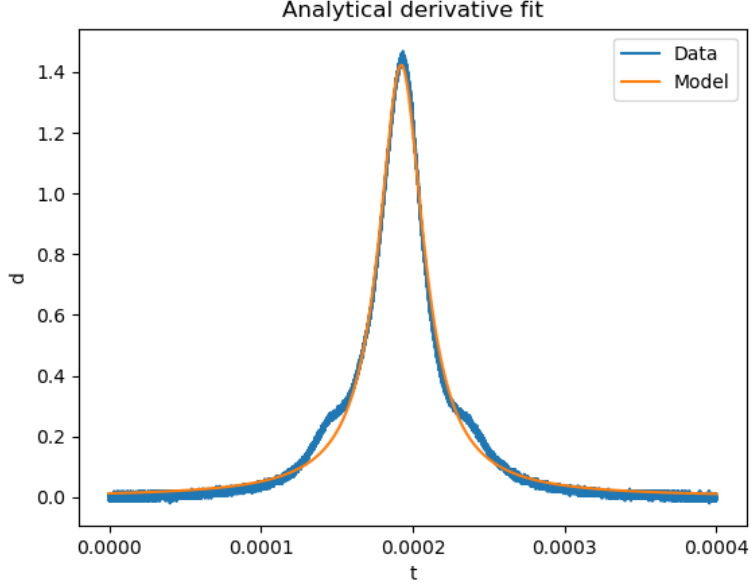
1.2 b)

For non-linear least squares fitting, we know

$$\boldsymbol{\nabla} A^T N^{-1} \boldsymbol{\nabla} A \delta m = \boldsymbol{\nabla} A^T N^{-1}(d - A(m))$$

where A is the model, N is a diagonal matrix such that $N_{i,i} = \sigma_i^2$ (assuming random Gaussian noise in data), d is the measurement data, and m are model parameters. Setting $m \to m_t$ the "true" parameters such that $A(m_t) = d_t$ where d_t is the noiseless data, the equation above can be rearranged to

$$\delta m = (\boldsymbol{\nabla} A^T N^{-1} \boldsymbol{\nabla} A)^{-1} \boldsymbol{\nabla} A^T N^{-1}(d - d_t) \ .$$

Obviously, $d - d_t = n$ is the noise in the data so

$$\delta m = (\boldsymbol{\nabla} A^T N^{-1} \boldsymbol{\nabla} A)^{-1} \boldsymbol{\nabla} A^T N^{-1} n$$

Following the same steps as for linear least squares fitting, the error in parameters $\left\langle (\delta m)^2 \right\rangle$ is

$$\left\langle (\delta m)^2 \right\rangle = (\boldsymbol{\nabla} A^T N^{-1} \boldsymbol{\nabla} A)^{-1} \ .$$

Since the model does not fit the data very well at all points, (e.g. figure 1 at $t \approx 0.00015$, and $t \approx 0.00025$), I chose to approximate the noise-free values of the data as the result of passing the data through a Gaussian filter. This assumes that the noise only makes up the high frequency oscillations in the data, and the filtered data passes approximately in the middle of the noise throughout the measured interval (Figure 2).