Recall that at each step in the Fisher iterations we are solving

$$X^t W X \beta = X^t W Z,$$

where $W = diag[W(\mu_1), \ldots, W(\mu_n)]$. We denote $J = X^t W X$. This corresponds to estimating the parameter β of the following model:

$$Z_i = x_i^t \beta + \epsilon_i, \quad \epsilon_i \sim N(0, 1/W(\mu_i)),$$

and ϵ_i independent. If we multiply this equation by $W^{1/2} = diag[\sqrt{W(\mu_1)}, \ldots, \sqrt{W(\mu_n)}]$, we get a regular fixed variance linear model:

$$\tilde{Z}_i = \tilde{x}_i^t \beta + \tilde{\epsilon}_i,$$

where $\tilde{Z}_i = \sqrt{W(\mu_i)} Y_i$, $\tilde{x}_i = \sqrt{W(\mu_i)} x_i$, $\tilde{\epsilon}_i \sim N(0, 1)$. The solution to this last model is the same as for the weighted model:

$$\beta = (\tilde{X}^t \tilde{X})^{-1} \tilde{X}^t \tilde{Z} = J^{-1} X^t W Z$$

and the corresponding hat matrix

this is correct

$$\tilde{H} = \tilde{X}(\tilde{X}^t \tilde{X})^{-1} \tilde{X}^t = W^{1/2} X J^{-1} X^t W^{1/2}.$$

This is the hat matrix we use to identify leverages for a glm.