The log-likelihood ℓ can be written as a function of θ, and also implicitly as a function of β via $\theta = g(\beta' x)$. Specifically,

$$\ell(\theta) = \sum_i (y_i \theta_i - \gamma(\theta_i))/\phi + \tau(y_i, \phi).$$

Viewing this as a function of the random effects, we have

$$\eta = x\beta + z\gamma,$$

where γ is the vector of random effects for a particular group, and $\theta = g(\eta)$.

We can differentiate the log-likelihood to get:

$$\partial\ell/\partial\gamma = (\partial\eta/\partial\gamma)^T \cdot (\partial\theta/\partial\eta)^T \cdot \partial\ell/\partial\theta$$

where $\partial\eta/\partial\gamma = z$ is $n \times q$ (with q being the number of realized random effects in a group) $\partial\theta/\partial\eta$ is $n \times n$, and $\partial\ell/\partial\theta$ is $n \times 1$.

Furthermore, $\partial\ell/\partial\theta = (y - \gamma'(\theta))/\phi$ is the Pearson residual. To determine $\partial\theta/\partial\eta$, let $f(\mu) \equiv g^{-1}(\gamma'^{-1}(\mu))$ be the link function. Then

$$f^{-1}(\eta) = \gamma'(g(\eta)),$$

and differentiating this yields

$$f^{-1\prime}(\eta) = \gamma''(g(\eta)) \cdot g'(\eta) = v(\eta) \cdot g'(\eta).$$

Therefore we have

$$g'(\eta) = f^{-1\prime}(\eta)/v(\eta),$$

so

$$\partial\ell/\partial\gamma^T = z^T [f^{-1\prime}(\eta)/v(\eta)] \cdot (y - \mu)/\phi,$$

where $\eta = \beta' x$. Also note that $f^{-1\prime}(\eta)/v(\eta)$ is the $n \times n$ diagonal matrix consisting of $f^{-1\prime}(\eta_i)/v(\eta_i)$ for $i = 1, \ldots, n$.

If the link is canonical, then $\partial\theta/\partial\eta^T = I$ and the term can be dropped.

The second derivative is the sum of two terms

$$(\partial\eta/\partial\gamma)^T \cdot (\partial\theta/\partial\eta)^T \cdot \partial^2\ell/\partial\theta\theta^T \cdot \partial\theta/\partial\eta \cdot \partial\eta/\partial\gamma = z^T \cdot (\partial\theta/\partial\eta)^T \cdot \partial^2\ell/\partial\theta\theta^T \cdot \partial\theta/\partial\eta \cdot z,$$

and

$$\sum_i \partial\ell/\partial\theta_i \cdot (\partial\eta/\partial\gamma)^T \cdot \partial^2\theta_i/\partial\eta\eta^T \cdot \partial\eta/\partial\gamma = \sum_i \partial\ell/\partial\theta_i \cdot \partial^2\theta_i/\partial\eta_i^2 \cdot z_i^T z_i,$$

where z_i is the i^{th} row of z. This in turn is equal to

$$z^T \cdot \operatorname{diag}\left(\partial\ell/\partial\theta_i \cdot \partial^2\theta_i/\partial\eta_i^2\right) \cdot z.$$

The second derivative involves

$$\partial^2\theta_i/\partial\eta_i^2 = \frac{v(\eta_i)f^{-1\prime\prime}(\eta_i) - f^{-1\prime}(\eta_i)v'(\eta_i)}{v(\eta_i)^2}.$$

Note that all other elements of the matrix $\partial^2\theta_i/\partial\eta\eta^T$ are zero.

Note that above the variance function v is differentiated with respect to η, but we usually write v as a function of μ. If we change notation so that v' is now the derivative of v with respect to μ, we can rewrite this as

$$\partial^2\theta_i/\partial\eta_i\eta_i^T = \frac{v(\eta_i)f^{-1\prime\prime}(\eta_i) - (f^{-1\prime}(\eta_i))^2 v'(f^{-1}(\eta_i))}{v(\eta_i)^2}.$$

To evaluate these expressions we need $f^{-1\prime}$, $f^{-1\prime\prime}$, and the variance function v, which are all already available. We also need to add v', which is not currently available.

To summarize, the Hessian has the form $z'Az$, where A is a diagonal matrix with elements

$$[f^{-1\prime}(\eta_i)]^2/v(\eta_i) + \partial\ell/\partial\theta_i \cdot \frac{v(\eta_i)f^{-1\prime\prime}(\eta_i) - (f^{-1\prime}(\eta_i))^2 v'(f^{-1}(\eta_i))}{v(\eta_i)^2}.$$