3.21 Eigenvalue expansion of a real, symmetric matrix $\mathbf{A}$ is given by:

$$\mathbf{A} = \sum_{i=1}^{D} \lambda_i \mathbf{u}_i \mathbf{u}_i^T$$

where D are its dimensions, λ refers to the eigenvalue and $\mathbf{u}_i$ the eigenvector.

Proving 3.117:

$$\text{L.H.S} = \frac{d}{d\alpha} ln|\mathbf{A}| = \frac{d}{d\alpha} ln\left(\prod_{i=1}^{D} \lambda_i\right) \qquad \text{from C.47}$$

$$= \frac{d}{d\alpha} \sum_{i=1}^{D} ln(\lambda_i)$$

$$= \sum_{i=1}^{D} \frac{1}{\lambda_i} \frac{d\lambda_i}{d\alpha}$$

$$\text{R.H.S} = Tr\left(\mathbf{A}^{-1} \frac{d}{d\alpha} \mathbf{A}\right)$$

$$= Tr\left(\left(\sum_{i=1}^{D} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T\right)\left(\frac{d}{d\alpha} \sum_{j=1}^{D} \lambda_j \mathbf{u}_j \mathbf{u}_j^T\right)\right)$$

$$= Tr\left(\sum_{i=1}^{D} \sum_{j=1}^{D} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T \frac{d}{d\alpha} \lambda_j \mathbf{u}_j \mathbf{u}_j^T\right)$$

$$= Tr\left(\sum_{i=1}^{D} \sum_{j=1}^{D} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T \left(\mathbf{u}_j \mathbf{u}_j^T \frac{d\lambda_j}{d\alpha} + \lambda_j \mathbf{u}_j \frac{d\mathbf{u}_j^T}{d\alpha} + \lambda_j \frac{d\mathbf{u}_j}{d\alpha} \mathbf{u}_j^T\right)\right)$$

$$= Tr\left(\left(\sum_{i=1}^{D} \sum_{j=1}^{D} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T \mathbf{u}_j \mathbf{u}_j^T \frac{d\lambda_j}{d\alpha}\right) + \left(\sum_{i=1}^{D} \sum_{j=1}^{D} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T \lambda_j \mathbf{u}_j \frac{d\mathbf{u}_j^T}{d\alpha}\right) + \left(\sum_{i=1}^{D} \sum_{j=1}^{D} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T \lambda_j \frac{d\mathbf{u}_j}{d\alpha} \mathbf{u}_j^T\right)\right)$$

$$= Tr\left(\sum_{i=1}^{D} \sum_{j=1}^{D} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T \mathbf{u}_j \mathbf{u}_j^T \frac{d\lambda_j}{d\alpha}\right) + Tr\left(\sum_{i=1}^{D} \sum_{j=1}^{D} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T \lambda_j \mathbf{u}_j \frac{d\mathbf{u}_j^T}{d\alpha}\right) + Tr\left(\sum_{i=1}^{D} \sum_{j=1}^{D} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T \lambda_j \frac{du_j}{d\alpha} \mathbf{u}_j^T\right)$$

Evaluating these 3 terms separately,

$$Tr\left(\sum_{i=1}^{D}\sum_{j=1}^{D}\frac{1}{\lambda_i}\mathbf{u}_i\mathbf{u}_i^T\mathbf{u}_j\mathbf{u}_j^T\frac{d\lambda_j}{d\alpha}\right) = Tr\left(\sum_{i=1}^{D}\sum_{j=1}^{D}\frac{1}{\lambda_i}\mathbf{u}_i(\mathbf{u}_i^T\mathbf{u}_j)\mathbf{u}_j^T\frac{d\lambda_j}{d\alpha}\right)$$

$$= Tr\left(\sum_{i=1}^{D}\frac{1}{\lambda_i}\mathbf{u}_i\mathbf{u}_i^T\frac{d\lambda_i}{d\alpha}\right)$$

$$= \left(\sum_{i=1}^{D}\frac{1}{\lambda_i}Tr\left(\mathbf{u}_i\mathbf{u}_i^T\right)\frac{d\lambda_i}{d\alpha}\right)$$

$$= \left(\sum_{i=1}^{D}\frac{1}{\lambda_i}(1)\frac{d\lambda_i}{d\alpha}\right)$$

$$= \left(\sum_{i=1}^{D}\frac{1}{\lambda_i}\frac{d\lambda_i}{d\alpha}\right)$$

Next,

$$Tr\left(\sum_{i=1}^{D}\sum_{j=1}^{D}\frac{1}{\lambda_i}\mathbf{u}_i\mathbf{u}_i^T\lambda_j\mathbf{u}_j\frac{d\mathbf{u}_j^T}{d\alpha}\right) = Tr\left(\sum_{i=1}^{D}\sum_{j=1}^{D}\frac{\lambda_j}{\lambda_i}\mathbf{u}_i(\mathbf{u}_i^T\mathbf{u}_j)\frac{d\mathbf{u}_j^T}{d\alpha}\right)$$

$$= Tr\left(\sum_{i=1}^{D}\mathbf{u}_i\frac{d\mathbf{u}_i^T}{d\alpha}\right)$$

Finally,

$$Tr\left(\sum_{i=1}^{D}\sum_{j=1}^{D}\frac{1}{\lambda_i}\mathbf{u}_i\mathbf{u}_i^T\lambda_j\frac{d\mathbf{u}_j}{d\alpha}\mathbf{u}_j^T\right) = Tr\left(\sum_{i=1}^{D}\sum_{j=1}^{D}\frac{\lambda_j}{\lambda_i}\mathbf{u}_i\mathbf{u}_i^T\frac{d\mathbf{u}_j}{d\alpha}\mathbf{u}_j^T\right)$$

$$= Tr\left(\sum_{i=1}^{D}\sum_{j=1}^{D}\frac{\lambda_j}{\lambda_i}\frac{d\mathbf{u}_j}{d\alpha}\mathbf{u}_j^T\mathbf{u}_i\mathbf{u}_i^T\right) \qquad \text{since } \mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$$

$$= Tr\left(\sum_{i=1}^{D}\sum_{j=1}^{D}\frac{\lambda_j}{\lambda_i}\frac{d\mathbf{u}_j}{d\alpha}(\mathbf{u}_j^T\mathbf{u}_i)\mathbf{u}_i^T\right)$$

$$= Tr\left(\sum_{i=1}^{D} \frac{d\mathbf{u}_i}{d\alpha}\mathbf{u}_i^T\right)$$

Summing all 3, we get:

$$\left(\sum_{i=1}^{D} \frac{1}{\lambda_i}\frac{d\lambda_i}{d\alpha}\right) + Tr\left(\sum_{i=1}^{D} \mathbf{u}_i\frac{d\mathbf{u}_i^T}{d\alpha}\right) + Tr\left(\sum_{i=1}^{D} \frac{d\mathbf{u}_i}{d\alpha}\mathbf{u}_i^T\right)$$

$$= \left(\sum_{i=1}^{D} \frac{1}{\lambda_i}\frac{d\lambda_i}{d\alpha}\right) + Tr\left(\sum_{i=1}^{D} \mathbf{u}_i\left(\frac{d\mathbf{u}_i^T}{d\alpha}\right) + \frac{d\mathbf{u}_i}{d\alpha}\mathbf{u}_i^T\right)$$

$$= \left(\sum_{i=1}^{D} \frac{1}{\lambda_i}\frac{d\lambda_i}{d\alpha}\right) + Tr\left(\sum_{i=1}^{D} \left(\frac{d\mathbf{u}_i\mathbf{u}_i^T}{d\alpha}\right)\right)$$

$$= \left(\sum_{i=1}^{D} \frac{1}{\lambda_i}\frac{d\lambda_i}{d\alpha}\right) + \left(\sum_{i=1}^{D} \left(\frac{d\left(Tr(\mathbf{u}_i\mathbf{u}_i^T)\right)}{d\alpha}\right)\right) \qquad \text{since } d(Tr(\mathbf{X})) = Tr(d\mathbf{X})$$

$$= \left(\sum_{i=1}^{D} \frac{1}{\lambda_i}\frac{d\lambda_i}{d\alpha}\right) + \left(\sum_{i=1}^{D} \left(\frac{d\,(1)}{d\alpha}\right)\right)$$

$$= \left(\sum_{i=1}^{D} \frac{1}{\lambda_i}\frac{d\lambda_i}{d\alpha}\right) + 0$$

$$= \left(\sum_{i=1}^{D} \frac{1}{\lambda_i}\frac{d\lambda_i}{d\alpha}\right)$$

$$= \text{L.H.S}$$

Now we use this result to derive 3.92, starting from 3.86.

Equation 3.86 gives us:

$$\ln p(\mathbf{t}|\alpha,\beta) = \frac{M}{2}\ln\alpha + \frac{N}{2}\ln\beta - E(\mathbf{m}_N) - \frac{1}{2}\ln|\mathbf{A}| - \frac{N}{2}\ln(2\pi)$$

$$= \frac{M}{2}\ln\alpha + \frac{N}{2}\ln\beta - \left(\frac{\beta}{2}||\mathbf{t} - \mathbf{\Phi}\mathbf{m}_N||^2 + \frac{\alpha}{2}\mathbf{m}_N^T\mathbf{m}_N\right) - \frac{1}{2}\ln|\mathbf{A}| - \frac{N}{2}\ln(2\pi)$$

Taking derivative w.r.t α,

$$\frac{d\ln p(\mathbf{t}|\alpha,\beta)}{d\alpha} = \frac{M}{2\alpha} + 0 - \frac{d}{d\alpha}\left(\frac{\beta}{2}||\mathbf{t}-\mathbf{\Phi m}_N||^2 + \frac{\alpha}{2}\mathbf{m}_N^T\mathbf{m}_N\right) - \frac{1}{2}Tr\left(\mathbf{A}^{-1}\frac{d}{d\alpha}\mathbf{A}\right) - 0$$

$$= \frac{M}{2\alpha} - \frac{d}{d\alpha}\left(\frac{\beta}{2}||\mathbf{t}-\mathbf{\Phi m}_N||^2 + \frac{\alpha}{2}\mathbf{m}_N^T\mathbf{m}_N\right) - \frac{1}{2}Tr\left(\mathbf{A}^{-1}\frac{d}{d\alpha}\mathbf{A}\right)$$

Solving for the second term:

$$\frac{d}{d\alpha}\left(\frac{\beta}{2}||\mathbf{t}-\mathbf{\Phi m}_N||^2 + \frac{\alpha}{2}\mathbf{m}_N^T\mathbf{m}_N\right)$$

$$= \frac{d}{d\alpha}\left(\frac{\beta}{2}||\mathbf{t}-\mathbf{\Phi m}_N||^2\right) + \frac{d}{d\alpha}\left(\frac{\alpha\mathbf{m}_N^T\mathbf{m}_N}{2}\right)$$

$$= \frac{d}{d\mathbf{m}_N}\left(\frac{\beta}{2}||\mathbf{t}-\mathbf{\Phi m}_N||^2\right)\frac{d\mathbf{m}_N}{d\alpha} + \frac{\alpha}{2}\left(\frac{d(\mathbf{m}_N^T\mathbf{m}_N)}{d\alpha}\right) + \left(\frac{\mathbf{m}_N^T\mathbf{m}_N}{2}\right)\frac{d\alpha}{d\alpha}$$

$$= \frac{d}{d\mathbf{m}_N}\left(\frac{\beta}{2}||\mathbf{t}-\mathbf{\Phi m}_N||^2\right)\frac{d\mathbf{m}_N}{d\alpha} + \frac{\alpha}{2}\left(\frac{d(\mathbf{m}_N^T\mathbf{m}_N)}{d\mathbf{m}_N}\right)\frac{d\mathbf{m}_N}{d\alpha} + \frac{\mathbf{m}_N^T\mathbf{m}_N}{2}$$

$$= \left(\beta\mathbf{\Phi}^T(\mathbf{\Phi m}_N - \mathbf{t})\right)^T\frac{d\mathbf{m}_N}{d\alpha} + \frac{\alpha}{2}(2\mathbf{m}_N)\frac{d\mathbf{m}_N}{d\alpha} + \frac{\mathbf{m}_N^T\mathbf{m}_N}{2}$$

$$= \left(\beta\mathbf{\Phi}^T(\mathbf{\Phi m}_N - \mathbf{t}) + \alpha\mathbf{m}_N\right)^T\frac{d\mathbf{m}_N}{d\alpha} + \frac{\mathbf{m}_N^T\mathbf{m}_N}{2}$$

$$= \left(\beta\mathbf{\Phi}^T\mathbf{\Phi m}_N - \beta\mathbf{\Phi}^T\mathbf{t} + \alpha\mathbf{m}_N\right)^T\frac{d\mathbf{m}_N}{d\alpha} + \frac{\mathbf{m}_N^T\mathbf{m}_N}{2}$$

$$= \left((\beta\mathbf{\Phi}^T\mathbf{\Phi} + \alpha\mathbf{I})\mathbf{m}_N - \beta\mathbf{\Phi}^T\mathbf{t}\right)^T\frac{d\mathbf{m}_N}{d\alpha} + \frac{\mathbf{m}_N^T\mathbf{m}_N}{2}$$

$$= \left(\mathbf{Am}_N - \beta\mathbf{\Phi}^T\mathbf{t}\right)^T\frac{d\mathbf{m}_N}{d\alpha} + \frac{\mathbf{m}_N^T\mathbf{m}_N}{2}$$

Since $\mathbf{m}_N = \beta\mathbf{A}^{-1}\mathbf{\Phi}^T\mathbf{t}$, $\quad\beta\mathbf{\Phi}^T\mathbf{t} = \mathbf{Am}_N$, and the above term becomes

$$= (\mathbf{Am}_N - \mathbf{Am}_N)^T\frac{d\mathbf{m}_N}{d\alpha} + \frac{\mathbf{m}_N^T\mathbf{m}_N}{2}$$

$$= \frac{\mathbf{m}_N^T \mathbf{m}_N}{2}$$

Solving for the third term:

$$\frac{1}{2} Tr\left(\mathbf{A}^{-1} \frac{d}{d\alpha} \mathbf{A}\right)$$

$$= \frac{1}{2} Tr\left(\mathbf{A}^{-1} \frac{d}{d\alpha} (\alpha \mathbf{I} + \beta \mathbf{\Phi}^T \mathbf{\Phi})\right)$$

$$= \frac{1}{2} Tr\left(\mathbf{A}^{-1} \mathbf{I}\right)$$

$$= \frac{1}{2} Tr\left(\mathbf{A}^{-1}\right)$$

$$= \frac{1}{2}\left(\sum_{i=1}^{M} \frac{1}{\alpha + \lambda_i}\right)$$

Adding the three terms, we get:

$$\frac{d \ln p(\mathbf{t}|\alpha, \beta)}{d\alpha} = \frac{M}{2\alpha} - \frac{\mathbf{m}_N^T \mathbf{m}_N}{2} - \frac{1}{2}\left(\sum_{i=1}^{M} \frac{1}{\alpha + \lambda_i}\right)$$

Setting the derivative to 0, we get stationary points of the evidence function w.r.t α:

$$0 = \frac{M}{2\alpha} - \frac{\mathbf{m}_N^T \mathbf{m}_N}{2} - \frac{1}{2}\left(\sum_{i=1}^{M} \frac{1}{\alpha + \lambda_i}\right)$$

$$\implies 0 = \frac{M}{\alpha} - \mathbf{m}_N^T \mathbf{m}_N - \sum_{i=1}^{M} \frac{1}{(\alpha + \lambda_i)}$$

$$\implies 0 = M - \alpha \mathbf{m}_N^T \mathbf{m}_N - \sum_{i=1}^{M} \frac{\alpha}{(\alpha + \lambda_i)}$$

$$\implies \alpha \mathbf{m}_N^T \mathbf{m}_N = M - \sum_{i=1}^{M} \frac{\alpha}{(\alpha + \lambda_i)}$$

$$\implies \alpha \mathbf{m}_N^T \mathbf{m}_N = \sum_{i=1}^{M} 1 - \sum_{i=1}^{M} \frac{\alpha}{(\alpha + \lambda_i)}$$

$$\implies \alpha \mathbf{m}_N^T \mathbf{m}_N = \sum_{i=1}^{M} \frac{(\alpha + \lambda_i) - \alpha}{(\alpha + \lambda_i)}$$

$$\implies \alpha \mathbf{m}_N^T \mathbf{m}_N = \sum_{i=1}^{M} \frac{\lambda_i}{(\alpha + \lambda_i)}$$

$$\implies \alpha \mathbf{m}_N^T \mathbf{m}_N = \gamma$$

$$\implies \alpha = \frac{\gamma}{\mathbf{m}_N^T \mathbf{m}_N}.$$