6.27 From 6.89, we have:

$$p(\mathbf{t}_N|\boldsymbol{\theta}) = \int p(\mathbf{t}_N|\mathbf{a}_N)\, p(\mathbf{a}_N|\boldsymbol{\theta})\, d\mathbf{a}_N$$

From 6.79, we have:

$$p(\mathbf{t}_N|\mathbf{a}_N) = \prod_{n=1}^{N} \sigma(a_n)^{t_n}\, (1-\sigma(a_n))^{1-t_n}$$

and we know that:

$$p(\mathbf{a}_N|\boldsymbol{\theta}) = \mathcal{N}(\mathbf{a}_N|\mathbf{0}, \mathbf{C}_N) = \frac{1}{(2\pi)^{N/2}|\mathbf{C}_N|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{a}_N^T \mathbf{C}_N^{-1} \mathbf{a}_N\right\}$$

$$\Longrightarrow p(\mathbf{t}_N|\boldsymbol{\theta}) = \int \left(\prod_{n=1}^{N} \sigma(a_n)^{t_n}\, (1-\sigma(a_n))^{1-t_n}\right) \left(\frac{1}{(2\pi)^{N/2}|\mathbf{C}_N|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{a}_N^T \mathbf{C}_N^{-1} \mathbf{a}_N\right\}\right) d\mathbf{a}_N$$

Using 4.135, we have the Laplace approximation of the posterior:

$$p(\mathbf{t}_N|\boldsymbol{\theta}) \approx (p(\mathbf{t}_N|\mathbf{a}_N^\star)\, p(\mathbf{a}_N^\star|\boldsymbol{\theta}))\, \frac{(2\pi)^{N/2}}{|\mathbf{A}|^{1/2}}$$

where from 4.132 :

$$\mathbf{A} = -\nabla\nabla \ln\left(p(\mathbf{t}_N|\mathbf{a}_N)\, p(\mathbf{a}_N|\boldsymbol{\theta})\right)|_{\mathbf{a}_N = \mathbf{a}_N^\star}$$

We can get this result directly from 6.82,

$$\mathbf{A} = \left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)$$

$$\Longrightarrow p(\mathbf{t}_N|\boldsymbol{\theta}) \approx (p(\mathbf{t}_N|\mathbf{a}_N^\star)\, p(\mathbf{a}_N^\star|\boldsymbol{\theta}))\, \frac{(2\pi)^{N/2}}{|\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)|^{1/2}}$$

$$\implies \ln p(\mathbf{t}_N|\boldsymbol{\theta}) \approx \ln p(\mathbf{t}_N|\mathbf{a}_N^\star) + \ln p(\mathbf{a}_N^\star|\boldsymbol{\theta}) + \frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)|$$

$$= \Psi(\boldsymbol{a}_N^\star) + \frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)|$$

which is the same as the result in 6.90.

Taking partial derivative of this log-likelihood w.r.t θ_j,

$$\frac{\partial \ln p(\mathbf{t}_N|\boldsymbol{\theta})}{\partial \theta_j} = \frac{\partial}{\partial \theta_j}\left(\ln p(\mathbf{t}_N|\mathbf{a}_N^\star) + \ln p(\mathbf{a}_N^\star|\boldsymbol{\theta}) + \frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)|\right)$$

Let's consider ther derivative of $(\ln p(\mathbf{t}_N|\mathbf{a}_N^\star) + \ln p(\mathbf{a}_N^\star|\boldsymbol{\theta}))$.

Since $\mathbf{a}_N^\star$ depends on θ_j, we apply the chain rule :

$$= \frac{\partial \ln p(\mathbf{t}_N \mid \mathbf{a}_N^\star)}{\partial \mathbf{a}_N^\star}\cdot\frac{\partial \mathbf{a}_N^\star}{\partial \theta_j} + \frac{\partial \ln p(\mathbf{t}_N \mid \mathbf{a}_N^\star)}{\partial \theta_j} + \frac{\partial \ln p(\mathbf{a}_N^\star \mid \boldsymbol{\theta})}{\partial \mathbf{a}_N^\star}\cdot\frac{\partial \mathbf{a}_N^\star}{\partial \theta_j} + \frac{\partial \ln p(\mathbf{a}_N^\star \mid \boldsymbol{\theta})}{\partial \theta_j}$$

$\dfrac{\partial \ln p(\mathbf{t}_N \mid \mathbf{a}_N^\star)}{\partial \theta_j} = 0$ since $\ln p(\mathbf{t}_N|\mathbf{a}_N^\star)$ has no direct dependence on $\boldsymbol{\theta}$.

$$= \left(\frac{\partial \ln p(\mathbf{t}_N \mid \mathbf{a}_N^\star)}{\partial \mathbf{a}_N^\star} + \frac{\partial \ln p(\mathbf{a}_N^\star \mid \boldsymbol{\theta})}{\partial \mathbf{a}_N^\star}\right)\cdot\frac{\partial \mathbf{a}_N^\star}{\partial \theta_j} + \frac{\partial \ln p(\mathbf{a}_N^\star \mid \boldsymbol{\theta})}{\partial \theta_j}$$

By definition of $\mathbf{a}_N^\star$, it is the mode of the posterior:

$$\mathbf{a}_N^\star = \arg\max_{\mathbf{a}_N}\left(\ln p(\mathbf{t}_N \mid \mathbf{a}_N) + \ln p(\mathbf{a}_N \mid \boldsymbol{\theta})\right)$$

So the gradient of the log-posterior with respect to $\mathbf{a}_N$ vanishes at the mode as the gradient is 0:

$$\frac{\partial}{\partial \mathbf{a}_N}\left(\ln p(\mathbf{t}_N \mid \mathbf{a}_N) + \ln p(\mathbf{a}_N \mid \boldsymbol{\theta})\right)\Big|_{\mathbf{a}_N=\mathbf{a}_N^\star} = 0$$

Hence, the derivative of the log-posterior simplifies to:

$$\frac{\partial}{\partial \theta_j}\left(\ln p(\mathbf{t}_N \mid \mathbf{a}_N^\star) + \ln p(\mathbf{a}_N^\star \mid \boldsymbol{\theta})\right) = \frac{\partial \ln p(\mathbf{a}_N^\star \mid \boldsymbol{\theta})}{\partial \theta_j}$$

Finding the derivative of $\ln p(\mathbf{a}_N^\star|\boldsymbol{\theta})$, and only focusing on direct dependence on $\boldsymbol{\theta}$, and ignoring any derivatives dependent on $\boldsymbol{\theta}$ through $\mathbf{a}_N^\star$:

$$\frac{\partial \ln p(\mathbf{a}_N^\star|\boldsymbol{\theta})}{\partial \theta_j} = \frac{\partial}{\partial \theta_j} \ln\left(\frac{1}{(2\pi)^{N/2}|\mathbf{C}_N|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{a}_N^{\star T}\mathbf{C}_N^{-1}\mathbf{a}_N^\star\right\}\right)$$

$$= \frac{\partial}{\partial \theta_j}\left(-\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{C}_N| - \frac{1}{2}\mathbf{a}_N^{\star T}\mathbf{C}_N^{-1}\mathbf{a}_N^\star\right)$$

$$= 0 - \frac{1}{2}\cdot\frac{\partial \ln|\mathbf{C}_N|}{\partial \theta_j} - \frac{1}{2}\cdot\frac{\partial\left(\mathbf{a}_N^{\star T}\mathbf{C}_N^{-1}\mathbf{a}_N^\star\right)}{\partial \theta_j}$$

$$= -\frac{1}{2}\cdot Tr\left(\mathbf{C}_N^{-1}\left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\right) - \frac{1}{2}\cdot\left(\mathbf{a}_N^{\star T}\left(\frac{\partial \mathbf{C}_N^{-1}}{\partial \theta_j}\right)\mathbf{a}_N^\star\right)$$

$$= -\frac{1}{2}\cdot Tr\left(\mathbf{C}_N^{-1}\left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\right) - \frac{1}{2}\cdot\left(-\mathbf{a}_N^{\star T}\mathbf{C}_N^{-1}\left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\mathbf{C}_N^{-1}\mathbf{a}_N^\star\right)$$

$$= \frac{1}{2}\cdot\left(\mathbf{a}_N^{\star T}\mathbf{C}_N^{-1}\left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\mathbf{C}_N^{-1}\mathbf{a}_N^\star\right) - \frac{1}{2}\cdot Tr\left(\mathbf{C}_N^{-1}\left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\right)$$

Finding the derivative of $-\frac{1}{2}\ln|\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)|$:

$$-\frac{\partial}{\partial \theta_j}\left(\frac{1}{2}\ln|\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)|\right) = -\frac{1}{2}\frac{\partial}{\partial \theta_j}\ln|\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)|$$

$$= -\frac{1}{2}Tr\left(\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)^{-1}\frac{\partial\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)}{\partial \theta_j}\right)$$

$$= -\frac{1}{2} Tr\left(\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)^{-1} \left(\frac{\partial \mathbf{W}_N}{\partial \theta_j} + \frac{\partial \mathbf{C}_N^{-1}}{\partial \theta_j} \right) \right)$$

$$= -\frac{1}{2} Tr\left(\left(\mathbf{W}_N + \mathbf{C}_N^{-1}\right)^{-1} \left(\left(\sum_{n=1}^{N} \frac{\partial \mathbf{W}_N}{\partial a_n^\star} \cdot \frac{\partial a_n^\star}{\partial \theta_j} \right) - \mathbf{C}_N^{-1} \frac{\partial \mathbf{C}_N}{\partial \theta_j} \mathbf{C}_N^{-1} \right) \right)$$

$$= -\frac{1}{2} Tr\left(\left(\mathbf{C}_N^{-1} \left(\mathbf{C}_N \mathbf{W}_N + \mathbf{I}_N\right)\right)^{-1} \left(\left(\sum_{n=1}^{N} \frac{\partial \mathbf{W}_N}{\partial a_n^\star} \cdot \frac{\partial a_n^\star}{\partial \theta_j} \right) - \mathbf{C}_N^{-1} \frac{\partial \mathbf{C}_N}{\partial \theta_j} \mathbf{C}_N^{-1} \right) \right)$$

$$= -\frac{1}{2} Tr\left(\left(\mathbf{I}_N + \mathbf{C}_N \mathbf{W}_N\right)^{-1} \mathbf{C}_N \left(\left(\sum_{n=1}^{N} \frac{\partial \mathbf{W}_N}{\partial a_n^\star} \cdot \frac{\partial a_n^\star}{\partial \theta_j} \right) - \mathbf{C}_N^{-1} \frac{\partial \mathbf{C}_N}{\partial \theta_j} \mathbf{C}_N^{-1} \right) \right)$$

$$\begin{aligned} &= -\frac{1}{2} Tr\left(\left(\mathbf{I}_N + \mathbf{C}_N \mathbf{W}_N\right)^{-1} \mathbf{C}_N \left(\sum_{n=1}^{N} \frac{\partial \mathbf{W}_N}{\partial a_n^\star} \cdot \frac{\partial a_n^\star}{\partial \theta_j} \right) \right) \\ &\quad + \frac{1}{2} Tr\left(\left(\mathbf{I}_N + \mathbf{C}_N \mathbf{W}_N\right)^{-1} \mathbf{C}_N \, \mathbf{C}_N^{-1} \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j} \right) \mathbf{C}_N^{-1} \right) \end{aligned}$$

$$\begin{aligned} &= -\frac{1}{2} Tr\left(\left(\mathbf{I}_N + \mathbf{C}_N \mathbf{W}_N\right)^{-1} \mathbf{C}_N \left(\sum_{n=1}^{N} \frac{\partial \mathbf{W}_N}{\partial a_n^\star} \cdot \frac{\partial a_n^\star}{\partial \theta_j} \right) \right) \\ &\quad + \frac{1}{2} Tr\left(\left(\mathbf{I}_N + \mathbf{C}_N \mathbf{W}_N\right)^{-1} \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j} \right) \mathbf{C}_N^{-1} \right) \end{aligned}$$

$\displaystyle\sum_{n=1}^{N} \frac{\partial \mathbf{W}_N}{\partial a_n^\star} \cdot \frac{\partial a_n^\star}{\partial \theta_j}$ is a sum of diagonal matrices, each of which is multiplied by a scalar.

The n^{th} element of $\dfrac{\partial \mathbf{W}_N}{\partial a_n^\star}$ is given by:

$$\frac{\partial \sigma(a_n^\star)(1-\sigma(a_n^\star))}{\partial a_n^\star} = \sigma(a_n^\star) \frac{\partial (1-\sigma(a_n^\star))}{\partial a_n^\star} + (1-\sigma(a_n^\star)) \frac{\partial \sigma(a_n^\star)}{\partial a_n^\star}$$

$$= \sigma(a_n^\star)(-\sigma(a_n^\star)(1-\sigma(a_n^\star))) + (1-\sigma(a_n^\star))\sigma(a_n^\star)(1-\sigma(a_n^\star))$$

$$= \sigma(a_n^\star)(1-\sigma(a_n^\star))(1-2\sigma(a_n^\star))$$

Substituting this back, and expanding the trace of the first term into a sum of diagonal values, we get:

$$= -\frac{1}{2}\sum_{n=1}^{N}\Big[(\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N)^{-1}\mathbf{C}_N\Big]_{nn}\ \sigma(a_n^\star)(1-\sigma(a_n^\star))(1-2\sigma(a_n^\star))\left(\frac{\partial a_n^\star}{\partial\theta_j}\right)$$
$$+\frac{1}{2}Tr\left((\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N)^{-1}\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\mathbf{C}_N^{-1}\right)\Bigg)$$

The terms arising from the explicit dependence on $\boldsymbol{\theta}$ are :

$$\frac{1}{2}\cdot\left(\mathbf{a}_N^{\star T}\mathbf{C}_N^{-1}\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\mathbf{C}_N^{-1}\mathbf{a}_N^\star\right) - \frac{1}{2}\cdot Tr\left(\mathbf{C}_N^{-1}\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\right)$$
$$+\frac{1}{2}Tr\left((\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N)^{-1}\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\mathbf{C}_N^{-1}\right)$$

Combining the trace terms, we get :

$$\frac{1}{2}Tr\left((\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N)^{-1}\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\mathbf{C}_N^{-1}\right) - \frac{1}{2}Tr\left(\mathbf{C}_N^{-1}\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\right)$$

Using cyclic property of Trace, this becomes:

$$= \frac{1}{2}Tr\left((\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N)^{-1}\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\mathbf{C}_N^{-1}\right) - \frac{1}{2}Tr\left(\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\mathbf{C}_N^{-1}\right)$$
$$= \frac{1}{2}Tr\left((\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N)^{-1}\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\mathbf{C}_N^{-1} - \left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\mathbf{C}_N^{-1}\right)$$
$$= \frac{1}{2}Tr\left(\left((\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N)^{-1} - \mathbf{I}_N\right)\left(\frac{\partial\mathbf{C}_N}{\partial\theta_j}\right)\mathbf{C}_N^{-1}\right)$$

Using Woodbury Identity,

$$\left(\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N\right)^{-1} - \mathbf{I}_N$$

$$= -\mathbf{C}_N \left(\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N\right)^{-1} \mathbf{W}_N$$

Substituting back the result, we get:

$$-\frac{1}{2}Tr\left(\mathbf{C}_N \left(\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N\right)^{-1} \mathbf{W}_N \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\mathbf{C}_N^{-1}\right)$$

Using cyclic property of Trace, this becomes:

$$\begin{aligned} &= -\frac{1}{2}Tr\left(\mathbf{C}_N^{-1}\mathbf{C}_N \left(\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N\right)^{-1} \mathbf{W}_N \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\right) \\ &= -\frac{1}{2}Tr\left(\left(\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N\right)^{-1} \mathbf{W}_N \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\right) \end{aligned}$$

Finally, the terms arising from the explicit dependence on $\boldsymbol{\theta}$ are given by:

$$\frac{1}{2}\left(\mathbf{a}_N^{\star T}\mathbf{C}_N^{-1}\left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\mathbf{C}_N^{-1}\mathbf{a}_N^{\star}\right) - \frac{1}{2}Tr\left(\left(\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N\right)^{-1} \mathbf{W}_N \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)\right)$$

which is the same as the result in 6.91.

The only term arising from the dependence of $\mathbf{a}_N^{\star}$ on $\boldsymbol{\theta}$ is:

$$-\frac{1}{2}\sum_{n=1}^{N}\left[\left(\mathbf{I}_N + \mathbf{C}_N\mathbf{W}_N\right)^{-1}\mathbf{C}_N\right]_{nn} \sigma(a_n^{\star})(1-\sigma(a_n^{\star}))(1-2\sigma(a_n^{\star}))\left(\frac{\partial a_n^{\star}}{\partial \theta_j}\right)$$

which is the same as the result in 6.92.

From 6.84, we have:

$$\mathbf{a}_N^\star = \mathbf{C}_N(\mathbf{t}_N - \boldsymbol{\sigma}_N)$$

$$\implies \frac{\partial \mathbf{a}_n^\star}{\partial \theta_j} = \frac{\partial}{\partial \theta_j}(\mathbf{C}_N(\mathbf{t}_N - \boldsymbol{\sigma}_N))$$

$$= \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)(\mathbf{t}_N - \boldsymbol{\sigma}_N) + \mathbf{C}_N \left(\frac{\partial(\mathbf{t}_N - \boldsymbol{\sigma}_N)}{\partial \theta_j}\right)$$

$$= \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)(\mathbf{t}_N - \boldsymbol{\sigma}_N) - \mathbf{C}_N \left(\frac{\partial \boldsymbol{\sigma}_N}{\partial \theta_j}\right)$$

$$= \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)(\mathbf{t}_N - \boldsymbol{\sigma}_N) - \mathbf{C}_N \left(\sum_{n=1}^{N} \frac{\partial \boldsymbol{\sigma}_N}{\partial a_n^\star} \cdot \frac{\partial a_n^\star}{\partial \theta_j}\right)$$

$\displaystyle\sum_{n=1}^{N} \frac{\partial \boldsymbol{\sigma}_N}{\partial a_n^\star} \cdot \frac{\partial a_n^\star}{\partial \theta_j}$ is a vector with the n^{th} given by : $\sigma(a_n^\star)(1 - \sigma(a_n^\star))\left(\frac{\partial a_n^\star}{\partial \theta_j}\right)$

$$\implies \sum_{n=1}^{N} \frac{\partial \boldsymbol{\sigma}_N}{\partial a_n^\star} \cdot \frac{\partial a_n^\star}{\partial \theta_j} = \mathbf{W}_N \left(\frac{\partial \mathbf{a}_N^\star}{\partial \theta_j}\right)$$

Substituting this back, we get:

$$\frac{\partial \mathbf{a}_N^\star}{\partial \theta_j} = \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)(\mathbf{t}_N - \boldsymbol{\sigma}_N) - \mathbf{C}_N \mathbf{W}_N \left(\frac{\partial \mathbf{a}_N^\star}{\partial \theta_j}\right)$$

$$\implies (\mathbf{I}_N + \mathbf{C}_N \mathbf{W}_N) \frac{\partial \mathbf{a}_N^\star}{\partial \theta_j} = \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)(\mathbf{t}_N - \boldsymbol{\sigma}_N)$$

$$\implies \frac{\partial \mathbf{a}_N^\star}{\partial \theta_j} = (\mathbf{I}_N + \mathbf{C}_N \mathbf{W}_N)^{-1} \left(\frac{\partial \mathbf{C}_N}{\partial \theta_j}\right)(\mathbf{t}_N - \boldsymbol{\sigma}_N)$$

which is the same as the result in 6.94.