6.21 We already know that:

$$p(t_{N+1}|\mathbf{t}, \mathbf{x}_{N+1}) = \mathcal{N}(\mathbf{k}^T\mathbf{C}_N^{-1}\mathbf{t},\, c - \mathbf{k}^T\mathbf{C}_N^{-1}\mathbf{k})$$

Considering the mean:

$$\mathbf{k}^T\mathbf{C}_N^{-1}\mathbf{t}$$

$$= \mathbf{k}(\mathbf{x})^T(\beta^{-1}\mathbf{I}_N + \mathbf{K})^{-1}\mathbf{t}$$

Since vector $\mathbf{k}$ has n elements $k(\mathbf{x}_n, \mathbf{x})$, $\mathbf{k} = \alpha^{-1}\mathbf{\Phi}\boldsymbol{\phi}(\mathbf{x})$, giving us:

$$= (\alpha^{-1}\mathbf{\Phi}\boldsymbol{\phi}(\mathbf{x}))^T(\beta^{-1}\mathbf{I}_N + \alpha^{-1}\mathbf{\Phi}\mathbf{\Phi}^T)^{-1}\mathbf{t}$$

$$= (\alpha^{-1}\mathbf{\Phi}\boldsymbol{\phi}(\mathbf{x}))^T(\beta)(\mathbf{I}_N + (\beta/\alpha)\mathbf{\Phi}\mathbf{\Phi}^T)^{-1}\mathbf{t}$$

$$= \beta\alpha^{-1}\boldsymbol{\phi}(\mathbf{x})^T\left(\mathbf{\Phi}^T(\mathbf{I}_N + (\beta/\alpha)\mathbf{\Phi}\mathbf{\Phi}^T)^{-1}\right)\mathbf{t}$$

Applying C.6, we get:

$$= \beta\alpha^{-1}\boldsymbol{\phi}(\mathbf{x})^T\left((\mathbf{I}_M + (\beta/\alpha)\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T\right)\mathbf{t}$$

$$= \beta\alpha^{-1}\boldsymbol{\phi}(\mathbf{x})^T\alpha(\alpha\mathbf{I}_M + \beta\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T\mathbf{t}$$

$$= \beta\boldsymbol{\phi}(\mathbf{x})^T\mathbf{S}_N^{-1}\mathbf{\Phi}^T\mathbf{t}$$

$$= \boldsymbol{\phi}(\mathbf{x})^T(\beta\mathbf{S}_N^{-1}\mathbf{\Phi}^T\mathbf{t})$$

$$= \boldsymbol{\phi}(\mathbf{x})^T\mathbf{m}_N$$

Considering the variance:

$$c - \mathbf{k}^T\mathbf{C}_N^{-1}\mathbf{k}$$

$$= \left(k(\mathbf{x}, \mathbf{x}) + \beta^{-1}\right) - (\alpha^{-1}\mathbf{\Phi}\boldsymbol{\phi}(\mathbf{x}))^T\left(\beta^{-1}\mathbf{I}_N + \alpha^{-1}\mathbf{\Phi}\mathbf{\Phi}^T\right)^{-1}(\alpha^{-1}\mathbf{\Phi}\boldsymbol{\phi}(\mathbf{x}))$$

$$= \left(\alpha^{-1}\boldsymbol{\phi}(\mathbf{x})^T\boldsymbol{\phi}(\mathbf{x}) + \beta^{-1}\right) - (\alpha^{-1}\boldsymbol{\Phi}\boldsymbol{\phi}(\mathbf{x}))^T \left(\beta^{-1}\mathbf{I}_N + \alpha^{-1}\boldsymbol{\Phi}\boldsymbol{\Phi}^T\right)^{-1} (\alpha^{-1}\boldsymbol{\Phi}\boldsymbol{\phi}(\mathbf{x}))$$

$$= \beta^{-1} + \boldsymbol{\phi}(\mathbf{x})^T\alpha^{-1}\boldsymbol{\phi}(\mathbf{x}) - \boldsymbol{\phi}(\mathbf{x})^T \left(\alpha^{-1}\boldsymbol{\Phi}^T \left(\beta^{-1}\mathbf{I}_N + \alpha^{-1}\boldsymbol{\Phi}\boldsymbol{\Phi}^T\right)^{-1} \alpha^{-1}\boldsymbol{\Phi}\right) \boldsymbol{\phi}(\mathbf{x})$$

$$= \beta^{-1} + \boldsymbol{\phi}(\mathbf{x})^T \left(\alpha^{-1}\mathbf{I}_M - \alpha^{-1}\mathbf{I}_M\boldsymbol{\Phi}^T \left(\beta^{-1}\mathbf{I}_N + \boldsymbol{\Phi}\alpha^{-1}\mathbf{I}_M\boldsymbol{\Phi}^T\right)^{-1} \boldsymbol{\Phi}\alpha^{-1}\mathbf{I}_M\right) \boldsymbol{\phi}(\mathbf{x})$$

If we let

$$\mathbf{A} = \alpha\mathbf{I}_M$$
$$\mathbf{B} = \boldsymbol{\Phi}^T$$
$$\mathbf{C} = \boldsymbol{\Phi}$$
$$\mathbf{D} = \beta^{-1}\mathbf{I}_N$$

Then, using C.7,

$$\alpha^{-1}\mathbf{I}_M - \alpha^{-1}\mathbf{I}_M\boldsymbol{\Phi}^T \left(\beta^{-1}\mathbf{I}_N + \boldsymbol{\Phi}\alpha^{-1}\mathbf{I}_M\boldsymbol{\Phi}^T\right)^{-1} \boldsymbol{\Phi}\alpha^{-1}\mathbf{I}_M = \left(\alpha\mathbf{I}_M + \boldsymbol{\Phi}^T \left(\beta^{-1}\mathbf{I}_N\right)^{-1} \boldsymbol{\Phi}\right)^{-1}$$

$$= \left(\alpha\mathbf{I}_M + \beta\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1}$$

$$= \mathbf{S}_N$$

Therefore the variance becomes:

$$= \beta^{-1} + \boldsymbol{\phi}(\mathbf{x})^T\mathbf{S}_N\boldsymbol{\phi}(\mathbf{x})$$

The results are identical to the ones in equation 3.58.