We now focus on the following time-dependent PDE:

$$\mathcal{L}u(x,t) := \partial_t u(x,t) - \nabla\cdot(a(x)\nabla u(x,t)) = f(x,t), \quad x\in\Omega,\ t\in[0,T] \tag{1}$$
$$u(x,t) = 0, \qquad x\in\partial\Omega,\ t\in[0,T] \tag{2}$$
$$u(x,0) = u_0(x), \qquad x\in\Omega \tag{3}$$

where the function $f\sim\mathcal{N}(\bar{f},K)$. The solution $u$ is thus also random.

We will set up a prior on the solution $u$ to the above problem. To do so we first let $v_h\in S_h$ be some approximation of the initial condition $u_0(x)$ in the FEM space $S_h$. To be more specific we will assume that $v_h(x)=\Phi(x)^*\boldsymbol{\gamma} := \sum_{i=1}^{J}\phi_i(x)\gamma_i$. Note that $\Phi(x) := (\phi_1(x),\dots,\phi_J(x))^T$. We take the prior on $u$ to be:

$$u\sim\mathcal{N}(m_0,V_0) \tag{4}$$

where $m_0(x,t) := v_h(x) = \Phi(x)^*\boldsymbol{\gamma}$ ($m_0$ is constant in time). The prior covariance operator $V_0$ is defined as:

$$(V_0 g)(x,t) = \int_\Omega\int_0^T k^{(0)}_{x,t,y,s}\, g(y,s)\mathrm{d}s\mathrm{d}y \tag{5}$$

where $k^{(0)}_{x,t,y,s}$ is defined as follows:

$$k^{(0)}_{x,t,y,s} := \sum_{i=1}^{J}\lambda_j\phi_j(x)\phi_j(y)k^{(0)}(t,s) \tag{6}$$

We now introduce a uniform time grid:

$$t_n = n\delta,\ \ n=0,1,\dots,N$$

where $\delta$ is the spacing between consecutive times and $N=\frac{T}{\delta}$ (*assume that $N$ is an integer*). The time kernel $k^{(0)}(t,s)$ will be taken to be:

$$k^{(0)}(t,s) := \sum_{i=0}^{N-1} l^{(i)}(t)l^{(i)}(s) \tag{7}$$

Where the functions $\{l^{(i)}\}_{i=0}^{N-1}$ are defined as follows:

$$l^{(i)}(t) = \begin{cases} (t-t_i)\mathbb{1}_{(t_i,t_{i+1}]}(t) + \delta\mathbb{1}_{(t_{i+1},t_N]}(t), & i=0,\dots,N-2 \\ (t-t_{N-1})\mathbb{1}_{(t_{N-1},t_N]}, & i=N-1 \end{cases} \tag{8}$$

i.e. for $i=0,\dots,N-2$ we have:

$$l^{(i)}(t) = \begin{cases} 0, & t\le t_i \\ t-t_i, & t_i<t\le t_{i+1} \\ \delta, & t>t_{i+1} \end{cases} \tag{9}$$

while for for $i=N-1$ we have:

$$l^{(i)}(t) = l^{(N-1)}(t) = \begin{cases} 0, & t\le t_{N-1} \\ t-t_i, & t>t_{N-1} \end{cases} \tag{10}$$

*note: we are working only with times in the interval $[0,T]$ here.*

We now introduce the following information operators $\mathcal{I}_s := (I_1(s),\dots,I_J(s))^T$ where:

$$I_i(s)g := \int_\Omega \phi_i(x)g(x,s)\mathrm{d}x \tag{11}$$

To update our belief in the distribution of $u$ we will condition on the following events: $\mathcal{I}_{t_i}\mathcal{L}u = \mathcal{I}_{t_i}f =: F^{(i)}$ sequentially for $i=1,\dots,N$. Letting $\tilde{A}_t := \mathcal{I}_t\mathcal{L}$ we seek, for a fixed realisation of $f$ (and hence of the $\{F^{(i)}\}$), the following conditional distributions:

$$u|\{\tilde{A}_{t_1}u = F^{(1)},\dots,\tilde{A}_{t_p}u = F^{(p)}, f\}\sim\mathcal{N}(m_p,V_p) \tag{12}$$

for $p\in\{1,\dots,N\}$. We make the following claim:

**Proposition 0.1.** With the prior specified as above we have that $m_p$ and $V_p$ are given as follows:

$$m_p(x,t) := \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{p} l^{(i-1)}(t)\Phi(x)^*\boldsymbol{c}^{(i)} \tag{13}$$

$$(V_p g)(x,t) := \int_\Omega \int_0^T k^{(p)}_{x,t,y,s} g(y,s)\mathrm{d}s\mathrm{d}y \tag{14}$$

$$\boldsymbol{c}^{(i)} := Q^{-1}\left[F^{(i)} - A\boldsymbol{\gamma}_{i-1}\right] \text{ for } i = 1,\dots,p \tag{15}$$

$$k^{(p)}_{x,t,y,s} := \sum_{j=1}^{J} \lambda_j \phi_j(x)\phi_j(y)k^{(p)}(t,s) \tag{16}$$

$$k^{(p)}(t,s) := \sum_{i=p}^{N-1} l^{(i)}(t)l^{(i)}(s) \tag{17}$$

and where the $\{\boldsymbol{\gamma}_i\}_{i=0}^p$ are defined recursively by:

$$\boldsymbol{\gamma}_0 := \boldsymbol{\gamma}, \tag{18}$$

$$\boldsymbol{\gamma}_i := Q^{-1}\left[M\boldsymbol{\gamma}_{i-1} + \delta F^{(i)}\right] \text{ for } i \geq 1 \tag{19}$$

The matrices $M$ and $A$ are the Galerkin Mass and Stiffness matrices respectively, i.e. $M_{ij} := \int_\Omega \phi_i(x)\phi_j(x)\mathrm{d}x$ and $A_{ij} := \int_\Omega a(x)\nabla\phi_i(x)\cdot\nabla\phi_j(x)\mathrm{d}x$. The matrix $Q := (M + \delta A)$. Further, we have that evaluating the conditional mean $m_p$ at time $t_p$ yields the following:

$$m_p(x,t_p) = \Phi(x)^*\boldsymbol{\gamma}_p \tag{20}$$

Thus, we can see that this choice of prior yields (for a fixed realisation of $f$) what the classical Backward-Euler Galerkin method yields.

*Proof:* We proceed via proof by induction.

For $p = 1$ it follows that the distribution of $u|\{\tilde{A}_\delta u = F^{(1)}, f\}$ is Gaussian $\mathcal{N}(m_1, V_1)$ by considering the following joint distribution:

$$\begin{pmatrix} u \\ \tilde{A}_\delta u \end{pmatrix} = \begin{pmatrix} I \\ \tilde{A}_\delta \end{pmatrix} u \sim \mathcal{N}\left(\begin{pmatrix} m_o \\ \tilde{A}_\delta m_0 \end{pmatrix}, \begin{pmatrix} V_0 & V_0\tilde{A}_\delta^* \\ \tilde{A}_\delta V_0 & \tilde{A}_\delta V_0 \tilde{A}_\delta^* \end{pmatrix}\right)$$

It follows that the conditional distribution is Gaussian and the mean and covariance are given by:

$$m_1 = m_0 + V_0\tilde{A}_\delta^*(\tilde{A}_\delta V_0 \tilde{A}_\delta^*)^{-1}(F^{(1)} - \tilde{A}_\delta m_0) \tag{21}$$

$$V_1 = V_0 - V_0\tilde{A}_\delta^*(\tilde{A}_\delta V_0 \tilde{A}_\delta^*)^{-1}\tilde{A}_\delta V_0 \tag{22}$$

To make progress we must now start computing various terms needed for our mean and covariance update rules. We start with $V_0\tilde{A}_\delta^*$. We have:

$$V_0\tilde{A}_\delta^* = V_0\mathcal{L}^*(I_1(\delta)^*,\dots,I_J(\delta)^*)$$

We can thus see that we need to be able to compute terms of the form $V_0\mathcal{L}^*I_i(\delta)^* = V_0(I_i(\delta)\mathcal{L})^*$. Now since the operator $I_i(\delta)\mathcal{L}$ takes in a function on $\Omega\times[0,T]$ and outputs a real number its adjoint should take in a real number and output a function on $\Omega\times[0,T]$. This adjoint should satisfy the following relation:

$$\alpha(I_i(\delta)\mathcal{L}g) = \int_\Omega \int_0^T ((I_i(\delta)\mathcal{L})^*\alpha)(x,t)g(x,t)\mathrm{d}t\mathrm{d}x \;\forall g,\; \forall\alpha\in\mathbb{R} \tag{23}$$

Using this we can now compute:

$$\begin{aligned}(V_0(I_i(\delta)\mathcal{L})^*\alpha)(x,s) &= \int_\Omega \int_0^T k^{(0)}_{x,s,y,w}((I_i(\delta)\mathcal{L})^*\alpha)(y,w)\mathrm{d}w\mathrm{d}y \\ &= \alpha(I_i(\delta)\mathcal{L}k^{(0)}_{x,s,\cdot,\cdot}) \\ &= \alpha\int_\Omega \phi_i(y)(\mathcal{L}k^{(0)}_{x,s,\cdot,\cdot})(y,\delta)\mathrm{d}y\end{aligned}$$

We now work out $(\mathcal{L}k^{(0)}_{x,s,\cdot,\cdot})(y,\delta)$ taking care to remember that $x, s$ are fixed and so $\mathcal{L}$ acts on the variables $y, \delta$:

$$\begin{aligned}(\mathcal{L}k^{(0)}_{x,s,\cdot,\cdot})(y,\delta) &= \partial_2 k^{(0)}(s,\delta)\sum_{j=1}^{J}\lambda_j\phi_j(x)\phi_j(y) - k^{(0)}(s,\delta)\nabla_y\cdot\left(a(y)\nabla_y\sum_{j=1}^{J}(\lambda_j\phi_j(x)\phi_j(y)\right)\\ &= \partial_2 k^{(0)}(s,\delta)\sum_{j=1}^{J}\lambda_j\phi_j(x)\phi_j(y) - k^{(0)}(s,\delta)\sum_{j=1}^{J}\lambda_j\phi_j(x)\nabla_y\cdot(a(y)\nabla_y\phi_j(y))\end{aligned}$$

So we can now compute:

$$\begin{aligned}(V_0(I_i(\delta)\mathcal{L})^*\alpha)(x,s) &= \alpha\int_\Omega\phi_i(y)\partial_2 k^{(0)}(s,\delta)\sum_{j=1}^{J}\lambda_j\phi_j(x)\phi_j(y)\mathrm{d}y - \alpha\int_\Omega\phi_i(y)k^{(0)}(s,\delta)\sum_{j=1}^{J}\lambda_j\phi_j(x)\nabla_y\cdot(a(y)\nabla_y\phi_j(y))\,\mathrm{d}y\\ &= \alpha\partial_2 k^{(0)}(s,\delta)\sum_{j=1}^{J}\lambda_j\phi_j(x)M_{ji} + \alpha k^{(0)}(s,\delta)\sum_{j=1}^{J}\lambda_j\phi_j(x)A_{ji}\end{aligned}$$

Using this result we can deduce that:

$$(V_0\tilde{A}^*_\delta\boldsymbol{v})(x,s) = \partial_2 k^{(0)}(s,\delta)\Phi(x)^*\Lambda M\boldsymbol{v} + k^{(0)}(s,\delta)\Phi(x)^*\Lambda A\boldsymbol{v} \tag{24}$$

for any $\boldsymbol{v}\in\mathbb{R}^J$, where $\Lambda = \mathrm{diag}\{\lambda_i\}_{i=1}^J$.

For our time kernel we can deduce:

$$\partial_1 k^{(0)}(t,s) = \sum_{i=0}^{N-1} l^{(i)\prime}(t)l^{(i)}(s) \tag{25}$$

$$\partial_2 k^{(0)}(t,s) = \sum_{i=0}^{N-1} l^{(i)}(t)l^{(i)\prime}(s) \tag{26}$$

$$\partial_1\partial_2 k^{(0)}(t,s) = \sum_{i=0}^{N-1} l^{(i)\prime}(t)l^{(i)\prime}(s) \tag{27}$$

$$\tag{28}$$

We also have:

$$l^{(i)\prime}(t) = \mathbb{1}_{(t_i,t_{i+1}]}(t) \tag{29}$$

for all $i = 0,\dots,N-1$. (*Note: this includes even the case of* $i = N-1$.) Noting that the kernel $k^{(0)}$ is symmetric we have:

$$\begin{aligned}k^{(0)}(s,\delta) &= k^{(0)}(\delta,s)\\ &= \sum_{i=0}^{N-1} l^{(i)}(\delta)l^{(i)}(s)\\ &= \sum_{i=0}^{N-1}\delta\delta_{i,0}l^{(i)}(s)\\ &= \delta l^{(0)}(s)\end{aligned}$$

and

$$\begin{aligned}\partial_2 k^{(0)}(s,\delta) &= \partial_1 k^{(0)}(\delta,s)\\ &= \sum_{i=0}^{N-1} l^{(i)\prime}(\delta)l^{(i)}(s)\\ &= \sum_{i=0}^{N-1}\delta_{i,0}l^{(i)}(s)\\ &= l^{(0)}(s)\end{aligned}$$

where we have used the following properties of the functions $\{l^{(i)}\}$ which can easily be shown:

$$l^{(i)}(t_j) = \delta \cdot \delta_{i,j-1} \tag{30}$$

$$l^{(i)\prime}(t_j) = \delta_{i,j-1} \tag{31}$$

for $i = 0, \dots, N-1$ and $j = 1, \dots, N$.

We can now simplify (24) to:

$$\begin{aligned}(V_0 \tilde{A}_\delta^* \boldsymbol{v})(x,s) &= l^{(0)}(s)\Phi(x)^*\Lambda M\boldsymbol{v} + \delta l^{(0)}(s)\Phi(x)^*\Lambda A\boldsymbol{v} \\ &= l^{(0)}(s)\Phi(x)^*\Lambda(M+\delta A)\boldsymbol{v} \\ &= l^{(0)}(s)\Phi(x)^*\Lambda Q\boldsymbol{v}\end{aligned} \tag{32}$$

We now move onto computing:

$$\begin{aligned}\tilde{A}_\delta V_0 \tilde{A}_\delta^* &= \mathcal{I}_\delta \mathcal{L} V_0 \mathcal{L}^* \mathcal{I}_\delta^* \\ &= \begin{pmatrix} I_1(\delta) \\ \vdots \\ I_J(\delta)\end{pmatrix} \mathcal{L} V_0 \mathcal{L}^* \begin{pmatrix} I_1(\delta)^* & \dots & I_J(\delta)^*\end{pmatrix}\end{aligned}$$

This operator has $ij$-$th$ entry which is given by:

$$\begin{aligned}(\tilde{A}_\delta V_0 \tilde{A}_\delta^*)_{ij}\alpha &= I_i(\delta)\mathcal{L}V_0\mathcal{L}^* I_j(\delta)^*\alpha \\ &= \int_\Omega \phi_i(x)\left[(\mathcal{L}V_0(I_j(\delta)\mathcal{L})^*\alpha)(x,\delta)\right]\mathrm{d}x \\ &= \int_\Omega \phi_i(x)\Bigg[\alpha\partial_1\partial_2 k^{(0)}(\delta,\delta)\sum_{l=1}^J \lambda_l\phi_l(x)M_{lj} + \alpha\partial_1 k^{(0)}(\delta,\delta)\sum_{l=1}^J \lambda_l\phi_l(x)A_{lj} \\ &- \alpha\partial_2 k^{(0)}(\delta,\delta)\sum_{l=1}^J \lambda_l M_{lj}\nabla\cdot(a(x)\nabla\phi_l(x)) - \alpha k^{(0)}(\delta,\delta)\sum_{l=1}^J \lambda_l A_{lj}\nabla\cdot(a(x)\nabla\phi_l(x))\Bigg]\mathrm{d}x \\ &= \alpha\partial_1\partial_2 k^{(0)}(\delta,\delta)\sum_{l=1}^J \lambda_l M_{il}M_{lj} + \alpha\partial_1 k^{(0)}(\delta,\delta)\sum_{l=1}^J \lambda_l M_{il}A_{lj} \\ &+ \alpha\partial_2 k^{(0)}(\delta,\delta)\sum_{l=1}^J \lambda_l M_{lj}A_{il} + \alpha k^{(0)}(\delta,\delta)\sum_{l=1}^J \lambda_l A_{il}A_{lj} \\ &= \alpha\partial_1\partial_2 k^{(0)}(\delta,\delta)(M\Lambda M)_{ij} + \alpha\partial_1 k^{(0)}(\delta,\delta)(M\Lambda A)_{ij} + \alpha\partial_2 k^{(0)}(\delta,\delta)(A\Lambda M)_{ij} + \alpha k^{(0)}(\delta,\delta)(A\Lambda A)_{ij}\end{aligned}$$

We can thus conclude that $\tilde{A}_\delta V_0 \tilde{A}_\delta^*$ is the $J \times J$ matrix given by:

$$\tilde{A}_\delta V_0 \tilde{A}_\delta^* = \partial_1\partial_2 k^{(0)}(\delta,\delta)M\Lambda M + \partial_1 k^{(0)}(\delta,\delta)M\Lambda A + \partial_2 k^{(0)}(\delta,\delta)A\Lambda M + k^{(0)}(\delta,\delta)A\Lambda A \tag{33}$$

Using our time kernel and its derivatives we can easily deduce:

$$\begin{aligned}k^{(0)}(\delta,\delta) &= \delta^2 \\ \partial_1 k^{(0)}(\delta,\delta) &= \delta \\ \partial_2 k^{(0)}(\delta,\delta) &= \delta \\ \partial_1\partial_2 k^{(0)}(\delta,\delta) &= 1\end{aligned}$$

And thus:

$$\begin{aligned}\tilde{A}_\delta V_0 \tilde{A}_\delta^* &= M\Lambda M + \delta M\Lambda A + \delta A\Lambda M + \delta^2 A\Lambda A \\ &= M\Lambda(M+\delta A) + \delta A\Lambda(M+\delta A) \\ &= (M\Lambda + \delta A\Lambda)(M+\delta A) \\ &= (M+\delta A)\Lambda(M+\delta A) = Q\Lambda Q\end{aligned}$$

We can now make progress with the mean update equation. We first work out the following term using (32):

$$\begin{aligned}(V_0\tilde{A}_\delta^*(\tilde{A}_\delta V_0\tilde{A}_\delta^*)^{-1}F^{(1)})(x,s) &= l^{(0)}(s)\Phi(x)^*\Lambda Q(Q\Lambda Q)^{-1}F^{(1)} \\ &= l^{(0)}(s)\Phi(x)^*\Lambda QQ^{-1}\Lambda^{-1}Q^{-1}F^{(1)} \\ &= l^{(0)}(s)\Phi(x)^*Q^{-1}F^{(1)}\end{aligned}$$

For the other term involving $m_0$ in the mean update equation we must first work out $\tilde{A}_\delta m_0 = \mathcal{I}_\delta\mathcal{L}m_0$. To do this we compute:

$$\begin{aligned}(\mathcal{L}m_0)(x,t) &= \partial_t m_0 - \nabla\cdot(a(x)\nabla m_0(x,t)) \\ &= -\nabla(a(x)\nabla\Phi(x)^*\boldsymbol{\gamma}) \\ &= -\sum_{j=1}^{J}\gamma_j\nabla\cdot(a(x)\nabla\phi_j(x))\end{aligned}$$

Thus, the *i-th* entry of $\tilde{A}_\delta m_0$ can be computed as:

$$\begin{aligned}(\tilde{A}_\delta m_0)_i &= I_i(\delta)\mathcal{L}m_0 \\ &= \int_\Omega \phi_i(x)\left(-\sum_{j=1}^{J}\gamma_j\nabla\cdot(a(x)\nabla\phi_j(x))\right)\mathrm{d}x \\ &= \sum_{j=1}^{J}\gamma_j A_{ij} = (A\boldsymbol{\gamma})_i\end{aligned}$$

So $\tilde{A}_\delta m_0 = A\boldsymbol{\gamma}$. We can now compute the second term involving $m_0$ as follows (again using (32)):

$$\begin{aligned}((V_0\tilde{A}_\delta^*(\tilde{A}_\delta V_0\tilde{A}_\delta^*)^{-1}\tilde{A}_\delta)m_0)(x,s) &= (V_0\tilde{A}_\delta^*(Q\Lambda Q)^{-1}A\boldsymbol{\gamma})(x,s) \\ &= l^{(0)}(s)\Phi(x)^*\Lambda QQ^{-1}\Lambda^{-1}Q^{-1}A\boldsymbol{\gamma} \\ &= l^{(0)}(s)\Phi(x)^*Q^{-1}A\boldsymbol{\gamma}\end{aligned}$$

Using (21) we can now compute:

$$\begin{aligned}m_1(x,s) &= \Phi(x)^*\boldsymbol{\gamma} + l^{(0)}(s)\Phi(x)^*Q^{-1}F^{(1)} - l^{(0)}(s)\Phi(x)^*Q^{-1}A\boldsymbol{\gamma} \\ &= \Phi(x)^*\boldsymbol{\gamma} + l^{(0)}(s)\Phi(x)^*Q^{-1}\left[F^{(1)} - A\boldsymbol{\gamma}\right] \\ &= \Phi(x)^*\boldsymbol{\gamma} + l^{(0)}(s)\Phi(x)^*\boldsymbol{c}^{(1)}\end{aligned}$$

From this we can note:

$$\begin{aligned}m_1(x,t_1) &= \Phi(x)^*\boldsymbol{\gamma} + \delta\Phi(x)^*\boldsymbol{c}^{(1)} \\ &= \Phi(x)^*\boldsymbol{\gamma} + \delta\Phi(x)^*Q^{-1}\left[F^{(1)} - A\boldsymbol{\gamma}\right] \\ &= \Phi(x)^*Q^{-1}\left[Q\boldsymbol{\gamma} + \delta F^{(1)} - \delta A\boldsymbol{\gamma}\right] \\ &= \Phi(x)^*Q^{-1}\left[M\boldsymbol{\gamma} + \delta F^{(1)}\right] \\ &= \Phi(x)^*\boldsymbol{\gamma}_1\end{aligned}$$

We now move on to computing the covariance $V_1$. To do this we must first work out $\tilde{A}_\delta V_0$. Computing this involves determining how $I_j(\delta)\mathcal{L}V_0$ acts on functions $g$ for $j = 1,\ldots,J$. We have:

$$I_j(\delta)\mathcal{L}V_0 g = \int_\Omega \phi_j(x)(\mathcal{L}V_0 g)(x,\delta)\mathrm{d}x$$

Now recalling that $V_0 g(x,\delta) = \int_\Omega \int_0^T k^{(0)}_{x,\delta,y,s} g(y,s) \mathrm{d}s\mathrm{d}y$ we deduce:

$$
\begin{aligned}
(\mathcal{L}V_0 g)(x,\delta) &= \int_\Omega \int_0^T (\mathcal{L}k^{(0)}_{\cdot,\cdot,y,s})(x,\delta) g(y,s)\mathrm{d}s\mathrm{d}y \\
&= \int_\Omega \int_0^T \left( \partial_1 k^{(0)}(\delta,s) \sum_{i=1}^J \lambda_i \phi_i(x)\phi_i(y) - k^{(0)}(\delta,s) \sum_{i=1}^J \lambda_i \nabla_x \cdot (a(x)\nabla_x \phi_i(x))\phi_i(y) \right) g(y,s)\mathrm{d}s\mathrm{d}y
\end{aligned}
$$

We can now perform the integration to obtain:

$$
\begin{aligned}
I_j(\delta)\mathcal{L}V_0 g &= \int_\Omega \phi_j(x) \left( \int_\Omega \int_0^T \left( \partial_1 k^{(0)}(\delta,s) \sum_{i=1}^J \lambda_i \phi_i(x)\phi_i(y) - k^{(0)}(\delta,s) \sum_{i=1}^J \lambda_i \nabla_x \cdot (a(x)\nabla_x \phi_i(x))\phi_i(y) \right) g(y,s)\mathrm{d}s\mathrm{d}y \right) \mathrm{d}x \\
&= \int_\Omega \int_0^T \left( \partial_1 k^{(0)}(\delta,s) \sum_{i=1}^J \lambda_i M_{ij} \phi_i(y) g(y,s) + k^{(0)}(\delta,s) \sum_{i=1}^J \lambda_i A_{ij} \phi_i(y) g(y,s) \right) \mathrm{d}s\mathrm{d}y \\
&= \sum_{i=1}^J \lambda_i \int_\Omega \int_0^T (\partial_1 k^{(0)}(\delta,s) M_{ij} + k^{(0)}(\delta,s) A_{ij})\phi_i(y) g(y,s)\mathrm{d}s\mathrm{d}y \\
&= \sum_{i=1}^J \lambda_i \left[ \int_0^T M_{ij} \partial_1 k^{(0)}(\delta,s)(I_i(s)g)\mathrm{d}s + \int_0^T A_{ij} k^{(0)}(\delta,s)(I_i(s)g)\mathrm{d}s \right] \\
&= \sum_{i=1}^J \lambda_i \left[ M_{ij} \left( \int_0^T \partial_1 k^{(0)}(\delta,s)(\mathcal{I}_s g)\mathrm{d}s \right)_i + A_{ij} \left( \int_0^T k^{(0)}(\delta,s)(\mathcal{I}_s g)\mathrm{d}s \right)_i \right] \\
&= \left( M\Lambda \int_0^T \partial_1 k^{(0)}(\delta,s)(\mathcal{I}_s g)\mathrm{d}s + A\Lambda \int_0^T k^{(0)}(\delta,s)(\mathcal{I}_s g)\mathrm{d}s \right)_j
\end{aligned}
$$

Thus we can deduce:

$$
\tilde{A}_\delta V_0 g = M\Lambda \int_0^T \partial_1 k^{(0)}(\delta,s)(\mathcal{I}_s g)\mathrm{d}s + A\Lambda \int_0^T k^{(0)}(\delta,s)(\mathcal{I}_s g)\mathrm{d}s \tag{34}
$$

We now utilise the specific form of the time kernel to simplify this to:

$$
\begin{aligned}
\tilde{A}_\delta V_0 g &= M\Lambda \int_0^T l^{(0)}(s)(\mathcal{I}_s g)\mathrm{d}s + \delta A\Lambda \int_0^T l^{(0)}(s)(\mathcal{I}_s g)\mathrm{d}s && (35) \\
&= (M + \delta A)\Lambda \int_0^T l^{(0)}(s)(\mathcal{I}_s g)\mathrm{d}s && (36) \\
&= Q\Lambda \boldsymbol{\nu}_g^{(0)} && (37)
\end{aligned}
$$

where $\boldsymbol{\nu}_g^{(i)} := \int_0^T l^{(i)}(s)(\mathcal{I}_s g)\mathrm{d}s$ for $i = 0, \dots, N-1$.

We can use this to now compute:

$$
\begin{aligned}
(V_0 \tilde{A}_\delta^* (\tilde{A}_\delta V_0 \tilde{A}_\delta^*)^{-1} \tilde{A}_\delta V_0 g)(x,t) &= (V_0 \tilde{A}_\delta^* (Q\Lambda Q)^{-1} Q\Lambda \boldsymbol{\nu}_g^{(0)})(x,t) \\
&= l^{(0)}(t)\Phi(x)^* \Lambda Q Q^{-1} \Lambda^{-1} Q^{-1} Q \Lambda \boldsymbol{\nu}_g^{(0)} \\
&= l^{(0)}(t)\Phi(x)^* \Lambda \boldsymbol{\nu}_g^{(0)}
\end{aligned}
$$

One can also easily show that the action of $V_0$ on functions can be rewritten as follows:

$$
(V_0 g)(x,t) = \sum_{i=0}^{N-1} l^{(i)}(t)\Phi(x)^* \Lambda \boldsymbol{\nu}_g^{(i)} \tag{38}
$$

Thus, we can conclude using (22) that:

$$\begin{aligned}(V_1 g)(x,t) &= \sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*\Lambda\boldsymbol{\nu}_g^{(i)} \\ &= \int_\Omega\int_0^T k_{x,t,y,s}^{(1)} g(y,s)\mathrm{d}s\mathrm{d}y \qquad (39)\end{aligned}$$

where:

$$k_{x,t,y,s}^{(1)} := \sum_{j=1}^{J}\lambda_j\phi_j(x)\phi_j(y)k^{(1)}(t,s) \qquad (40)$$

$$k^{(1)}(t,s) := \sum_{i=1}^{N-1} l^{(i)}(t)l^{(i)}(s) \qquad (41)$$

We thus see that the result holds for $p=1$. We now proceed with the inductive step. Assume the result holds for some $p<N$. As we did in the proof of the case $p=1$ we can easily deduce that

$$u|\{\tilde{A}_{t_1}u = F^{(1)},\ldots,\tilde{A}_{t_{p+1}}u = F^{(p+1)}, f\} \sim \mathcal{N}(m_{p+1}, V_{p+1}) \qquad (42)$$

where we have the update equations:

$$m_{p+1} = m_p + V_p\tilde{A}_{t_{p+1}}^*(\tilde{A}_{t_{p+1}}V_p\tilde{A}_{t_{p+1}}^*)^{-1}(F^{(p+1)} - \tilde{A}_{t_{p+1}}m_p) \qquad (43)$$

$$V_{p+1} = V_p - V_p\tilde{A}_{t_{p+1}}^*(\tilde{A}_{t_{p+1}}V_p\tilde{A}_{t_{p+1}}^*)^{-1}\tilde{A}_{t_{p+1}}V_p \qquad (44)$$

From the time kernel $k^{(p)}(t,s)$ we can work out:

$$\partial_1 k^{(p)}(t,s) = \sum_{i=p}^{N-1} l^{(i)\prime}(t)l^{(i)}(s) \qquad (45)$$

$$\partial_2 k^{(p)}(t,s) = \sum_{i=p}^{N-1} l^{(i)}(t)l^{(i)\prime}(s) \qquad (46)$$

$$\partial_1\partial_2 k^{(p)}(t,s) = \sum_{i=p}^{N-1} l^{(i)\prime}(t)l^{(i)\prime}(s) \qquad (47)$$

$$(48)$$

Using (30) and (31) we have

$$k^{(p)}(s,t_{p+1}) = \delta l^{(p)}(s) \qquad (49)$$

$$\partial_2 k^{(p)}(s,t_{p+1}) = l^{(p)}(s) \qquad (50)$$

These results will help with figuring out the analogue of (32). Since $V_p$ is of the same form as $V_0$ we can deduce:

$$\begin{aligned}(V_p\tilde{A}_{t_{p+1}}^*\boldsymbol{v})(x,s) &= \partial_2 k^{(p)}(s,t_{p+1})\Phi(x)^*\Lambda M\boldsymbol{v} + k^{(p)}(s,t_{p+1})\Phi(x)^*\Lambda A\boldsymbol{v} \qquad (51)\\ &= l^{(p)}(s)\Phi(x)^*\Lambda M\boldsymbol{v} + \delta l^{(p)}(s)\Phi(x)^*\Lambda A\boldsymbol{v} \qquad (52)\\ &= l^{(p)}(s)\Phi(x)^*\Lambda Q\boldsymbol{v} \qquad (53)\end{aligned}$$

Using (30) and (31) we can also conclude the following:

$$\begin{aligned}k^{(p)}(t_{p+1},t_{p+1}) &= \delta^2\\ \partial_1 k^{(p)}(t_{p+1},t_{p+1}) &= \delta\\ \partial_2 k^{(p)}(t_{p+1},t_{p+1}) &= \delta\\ \partial_1\partial_2 k^{(p)}(t_{p+1},t_{p+1}) &= 1\end{aligned}$$

Thus, just as in the proof for $p=1$ we have:

$$\begin{aligned}\tilde{A}_{t_{p+1}} V_p \tilde{A}^*_{t_{p+1}} &= \partial_1\partial_2 k^{(p)}(t_{p+1},t_{p+1}) M\Lambda M + \partial_1 k^{(p)}(t_{p+1},t_{p+1}) M\Lambda A + \partial_2 k^{(p)}(t_{p+1},t_{p+1}) A\Lambda M + k^{(p)}(\delta,\delta) A\Lambda A \\ &= Q\Lambda Q\end{aligned}$$

We can now compute:

$$\begin{aligned}\left(V_p \tilde{A}^*_{t_{p+1}} (\tilde{A}_{t_{p+1}} V_p \tilde{A}^*_{t_{p+1}})^{-1} F^{(p+1)}\right)(x,s) &= l^{(p)}(s)\Phi(x)^*\Lambda Q(Q\Lambda Q)^{-1} F^{(p+1)} \\ &= l^{(p)}(s)\Phi(x)^*\Lambda Q Q^{-1}\Lambda^{-1} Q^{-1} F^{(p+1)} \\ &= l^{(p)}(s)\Phi(x)^* Q^{-1} F^{(p+1)}\end{aligned}$$

In order to finish with the mean update we must now compute $\tilde{A}_{t_{p+1}} m_p = \mathcal{I}_{t_{p+1}} \mathcal{L} m_p$. Recalling that

$$m_p(x,s) = \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{p} l^{(i-1)}(s)\Phi(x)^*\boldsymbol{c}^{(i)}$$

we can work out:

$$(\mathcal{L}m_p)(x,s) = \sum_{i=1}^{p} l^{(i-1)\prime}(s)\Phi(x)^*\boldsymbol{c}^{(i)} - \sum_{j=1}^{J}\left(\boldsymbol{\gamma} + \sum_{i=1}^{p} l^{(i-1)}(s)\boldsymbol{c}^{(i)}\right)_j \nabla\cdot(a(x)\nabla\phi_j(x))$$

Using this, together with (30) and (31), we can work out:

$$\begin{aligned}(\tilde{A}_{t_{p+1}} m_p)_k &= I_k(t_{p+1})\mathcal{L}m_p \\ &= \int_\Omega \phi_k(x)(\mathcal{L}m_p)(x,t_{p+1})\mathrm{d}x \\ &= -\sum_{j=1}^{J}\left(\boldsymbol{\gamma} + \sum_{i=1}^{p}\delta\boldsymbol{c}^{(i)}\right)_j \int_\Omega \phi_k(x)\nabla\cdot(a(x)\nabla\phi_j(x))\mathrm{d}x \\ &= \sum_{j=1}^{J}\left(\boldsymbol{\gamma} + \sum_{i=1}^{p}\delta\boldsymbol{c}^{(i)}\right)_j A_{kj} \\ &= \left[A\left(\boldsymbol{\gamma} + \sum_{i=1}^{p}\delta\boldsymbol{c}^{(i)}\right)\right]_k\end{aligned}$$

i.e. we have:

$$\tilde{A}_{t_{p+1}} m_p = A\left(\boldsymbol{\gamma} + \sum_{i=1}^{p}\delta\boldsymbol{c}^{(i)}\right) \tag{54}$$

We now claim that $\boldsymbol{\gamma} + \sum_{i=1}^{p}\delta\boldsymbol{c}^{(i)} = \boldsymbol{\gamma}_p$. This can be proven by induction. It is true for $p=1$ (see proof that $m_1(x,t_1) = \Phi(x)^*\boldsymbol{\gamma}_1$). Assume it holds for $p$. For $p+1$ we have:

$$\begin{aligned}\boldsymbol{\gamma} + \sum_{i=1}^{p+1}\delta\boldsymbol{c}^{(i)} &= \boldsymbol{\gamma} + \sum_{i=1}^{p}\delta\boldsymbol{c}^{(i)} + \delta\boldsymbol{c}^{(p+1)} \\ &= \boldsymbol{\gamma}_p + \delta\boldsymbol{c}^{(p+1)} \\ &= \boldsymbol{\gamma}_p + \delta Q^{-1}\left[F^{(p+1)} - A\boldsymbol{\gamma}_p\right] \\ &= Q^{-1}\left[(Q-\delta A)\boldsymbol{\gamma}_p + \delta F^{(p+1)}\right] \\ &= Q^{-1}\left[M\boldsymbol{\gamma}_p + \delta F^{(p+1)}\right] \\ &= \boldsymbol{\gamma}_{p+1}\end{aligned}$$

as claimed. Thus, $\tilde{A}_{t_{p+1}} m_p = A\boldsymbol{\gamma}_p$ and we can now compute:

$$\begin{aligned}\left(V_p \tilde{A}^*_{t_{p+1}} (\tilde{A}_{t_{p+1}} V_p \tilde{A}^*_{t_{p+1}})^{-1} \tilde{A}_{t_{p+1}} m_p\right)(x,s) &= l^{(p)}(s)\Phi(x)^*\Lambda Q(Q\Lambda Q)^{-1} A\boldsymbol{\gamma}_p \\ &= l^{(p)}(s)\Phi(x)^*\Lambda Q Q^{-1}\Lambda^{-1} Q^{-1} A\boldsymbol{\gamma}_p \\ &= l^{(p)}(s)\Phi(x)^* Q^{-1} A\boldsymbol{\gamma}_p\end{aligned}$$

It now follows from (43) that we have:

$$\begin{aligned}
m_{p+1}(x,s) &= \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{p} l^{(i-1)}(s)\Phi(x)^*\boldsymbol{c}^{(i)} + l^{(p)}(s)\Phi(x)^*Q^{-1}F^{(p+1)} - l^{(p)}(s)\Phi(x)^*Q^{-1}A\boldsymbol{\gamma}_p \\
&= \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{p} l^{(i-1)}(s)\Phi(x)^*\boldsymbol{c}^{(i)} + l^{(p)}(s)\Phi(x)^*Q^{-1}\left[F^{(p+1)} - A\boldsymbol{\gamma}_p\right] \\
&= \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{p} l^{(i-1)}(s)\Phi(x)^*\boldsymbol{c}^{(i)} + l^{(p)}(s)\Phi(x)^*\boldsymbol{c}^{(p+1)} \\
&= \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{p+1} l^{(i-1)}(s)\Phi(x)^*\boldsymbol{c}^{(i)}
\end{aligned}$$

Having obtained this formula we can now evaluate the mean at $s = t_{p+1}$ to deduce:

$$m_{p+1}(x,t_{p+1}) = \Phi(x)^*\left[\boldsymbol{\gamma} + \sum_{i=1}^{p+1}\delta\boldsymbol{c}^{(\boldsymbol{i})}\right] = \Phi(x)^*\boldsymbol{\gamma}_{p+1}$$

by the above claim.

We now move on to the covariance update (44). The piece missing before we can proceed is the computation of $\tilde{A}_{t_{p+1}}V_p$. Following the proof of the case $p=1$ we can write:

$$\begin{aligned}
\tilde{A}_{t_{p+1}}V_p g &= M\Lambda\int_0^T \partial_1 k^{(p)}(t_{p+1},s)(\mathcal{I}_s g)\mathrm{d}s + A\Lambda\int_0^T k^{(p)}(t_{p+1},s)(\mathcal{I}_s g)\mathrm{d}s && (55)\\
&= M\Lambda\int_0^T l^{(p)}(s)(\mathcal{I}_s g)\mathrm{d}s + \delta A\Lambda\int_0^T l^{(p)}(s)(\mathcal{I}_s g)\mathrm{d}s && (56)\\
&= (M+\delta A)\Lambda\int_0^T l^{(p)}(s)(\mathcal{I}_s g)\mathrm{d}s && (57)\\
&= Q\Lambda\boldsymbol{\nu}_g^{(p)} && (58)
\end{aligned}$$

And thus, we can compute:

$$\begin{aligned}
(V_p\tilde{A}^*_{t_{p+1}}(\tilde{A}_{t_{p+1}}V_p\tilde{A}^*_{t_{p+1}})^{-1}\tilde{A}_{t_{p+1}}V_p g)(x,t) &= (V_p\tilde{A}^*_{t_{p+1}}(Q\Lambda Q)^{-1}Q\Lambda\boldsymbol{\nu}_g^{(p)})(x,t) \\
&= l^{(p)}(t)\Phi(x)^*\Lambda QQ^{-1}\Lambda^{-1}Q^{-1}Q\Lambda\boldsymbol{\nu}_g^{(p)} \\
&= l^{(p)}(t)\Phi(x)^*\Lambda\boldsymbol{\nu}_g^{(p)}
\end{aligned}$$

Thus, using (44) we can deduce:

$$\begin{aligned}
(V_{p+1}g)(x,t) &= \sum_{i=p+1}^{N-1} l^{(i)}(t)\Phi(x)^*\Lambda\boldsymbol{\nu}_g^{(i)} \\
&= \int_\Omega\int_0^T k^{(p+1)}_{x,t,y,s}g(y,s)\mathrm{d}s\mathrm{d}y
\end{aligned}$$

where:

$$\begin{aligned}
k^{(p+1)}_{x,t,y,s} &:= \sum_{j=1}^{J}\lambda_j\phi_j(x)\phi_j(y)k^{(p+1)}(t,s) \\
k^{(p+1)}(t,s) &:= \sum_{i=p+1}^{N-1} l^{(i)}(t)l^{(i)}(s)
\end{aligned}$$

Thus, the result is true for $p+1$. So by induction it is true for $p \in \{1,\ldots,N\}$. ■

Proposition 0.1 implies that conditioning at all $N$ time points will yield the following degenerate Gaussian distribution:

$$u|\{\tilde{A}_{t_1}u = F^{(1)},\ldots,\tilde{A}_{t_N}u = F^{(N)}, f\} \sim \mathcal{N}(m_N, 0) = \delta_{m_N} \tag{59}$$

since $V_N = 0$. This is a Dirac point mass located at the function:

$$m_N(x,t) = \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*\boldsymbol{c}^{(i)} \tag{60}$$

We now want to perform the marginalisation over the RHS noise term $f$. In order to do this it will help to rewrite $m_N$ as $m_N = Lf + c$ where $L$ is a bounded linear operator acting on $f$ and $c$ is a determinstic function (not depending on $f$). In order to rewrite $m_N$ in this form it will be useful to first prove the following result:

**Lemma 0.1.** The vectors $\{\boldsymbol{\gamma}_i\}$ satisfy the following:

$$\boldsymbol{\gamma}_i = (Q^{-1}M)^i\boldsymbol{\gamma} + \delta Q^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}F^{(j)} \quad \text{for } i \geq 1 \tag{61}$$

*Proof:* We proceed by induction. For $i = 1$ we have:

$$\begin{aligned}
\boldsymbol{\gamma}_1 &= Q^{-1}\left[M\boldsymbol{\gamma} + \delta F^{(1)}\right] \\
&= (Q^{-1}M)\boldsymbol{\gamma} + \delta Q^{-1}F^{(1)} \\
&= (Q^{-1}M)\boldsymbol{\gamma} + \delta Q^{-1}\sum_{j=1}^{1}(MQ^{-1})^{1-j}F^{(j)}
\end{aligned}$$

so the result is true for $i = 1$. Assume it is true for $i$. For $i+1$ we can use the recursive defintion of the $\{\boldsymbol{\gamma}_i\}$ to compute:

$$\begin{aligned}
\boldsymbol{\gamma}_{i+1} &= Q^{-1}\left[M\boldsymbol{\gamma}_i + \delta F^{(i+1)}\right] \\
&= Q^{-1}M\left[(Q^{-1}M)^i\boldsymbol{\gamma} + \delta Q^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}F^{(j)}\right] + \delta Q^{-1}F^{(i+1)} \\
&= (Q^{-1}M)^{i+1}\boldsymbol{\gamma} + \delta Q^{-1}\left[F^{(i+1)} + MQ^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}F^{(j)}\right] \\
&= (Q^{-1}M)^{i+1}\boldsymbol{\gamma} + \delta Q^{-1}\left[F^{(i+1)} + \sum_{j=1}^{i}(MQ^{-1})^{i+1-j}F^{(j)}\right] \\
&= (Q^{-1}M)^{i+1}\boldsymbol{\gamma} + \delta Q^{-1}\sum_{j=1}^{i+1}(MQ^{-1})^{(i+1)-j}F^{(j)}
\end{aligned}$$

so the result is true for $i+1$. Thus, the result holds for $i \geq 1$ by induction. ■

Using Lemma 0.1 we can now rewrite $m_N$ as follows:

$$\begin{aligned}
m_N(x,t) &= \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*\boldsymbol{c}^{(i)} \\
&= \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}\left[F^{(i)} - A\boldsymbol{\gamma}_{i-1}\right] \\
&= \Phi(x)^*\boldsymbol{\gamma} + \sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}F^{(i)} - \sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}A\boldsymbol{\gamma}_{i-1} \\
&= \Phi(x)^*\boldsymbol{\gamma} + \left(\sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_i}\right) f - l^{(0)}(t)\Phi(x)^*Q^{-1}A\boldsymbol{\gamma} - \sum_{i=2}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}A\boldsymbol{\gamma}_{i-1} \\
&= \Phi(x)^*\left[I - l^{(0)}(t)Q^{-1}A\right]\boldsymbol{\gamma} + \left(\sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_i}\right) f - \sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}A\boldsymbol{\gamma}_i \\
&= \Phi(x)^*\left[I - l^{(0)}(t)Q^{-1}A\right]\boldsymbol{\gamma} + \left(\sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_i}\right) f \\
&- \sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}A\left((Q^{-1}M)^i\boldsymbol{\gamma} + \delta Q^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}F^{(j)}\right) \\
&= \Phi(x)^*\left[I - l^{(0)}(t)Q^{-1}A\right]\boldsymbol{\gamma} + \left(\sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_i}\right) f - \sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}A(Q^{-1}M)^i\boldsymbol{\gamma} \\
&- \delta\sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}AQ^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}F^{(j)} \\
&= \Phi(x)^*\left[I - l^{(0)}(t)Q^{-1}A - \sum_{i=1}^{N-1} l^{(i)}(t)Q^{-1}A(Q^{-1}M)^i\right]\boldsymbol{\gamma} \\
&+ \left[\sum_{i=1}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_i} - \delta\sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}AQ^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}\mathcal{I}_{t_j}\right] f \\
&= \Phi(x)^*\left[I - \sum_{i=0}^{N-1} l^{(i)}(t)Q^{-1}A(Q^{-1}M)^i\right]\boldsymbol{\gamma} \\
&+ \left[l^{(0)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_1} + \sum_{i=2}^{N} l^{(i-1)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_i} - \delta\sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}AQ^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}\mathcal{I}_{t_j}\right] f \\
&= \Phi(x)^*\left[I - \sum_{i=0}^{N-1} l^{(i)}(t)Q^{-1}A(Q^{-1}M)^i\right]\boldsymbol{\gamma} \\
&+ \left[l^{(0)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_1} + \sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_{i+1}} - \delta\sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}AQ^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}\mathcal{I}_{t_j}\right] f \\
&= \Phi(x)^*\left[I - \sum_{i=0}^{N-1} l^{(i)}(t)Q^{-1}A(Q^{-1}M)^i\right]\boldsymbol{\gamma} \\
&+ \left[l^{(0)}(t)\Phi(x)^*Q^{-1}\mathcal{I}_{t_1} + \sum_{i=1}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}\left(\mathcal{I}_{t_{i+1}} - \delta AQ^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}\mathcal{I}_{t_j}\right)\right] f \\
&= \Phi(x)^*\left[I - \sum_{i=0}^{N-1} l^{(i)}(t)Q^{-1}A(Q^{-1}M)^i\right]\boldsymbol{\gamma} + \left[\sum_{i=0}^{N-1} l^{(i)}(t)\Phi(x)^*Q^{-1}\left(\mathcal{I}_{t_{i+1}} - \delta AQ^{-1}\sum_{j=1}^{i}(MQ^{-1})^{i-j}\mathcal{I}_{t_j}\right)\right] f \\
&= c(x,t) + (Lf)(x,t)
\end{aligned}$$

*Note: we have used the convention that a sum from* $j=1$ *to* $j=0$ *is considered to be empty, i.e.* 0. In the above we have defined:

$$c(x,t) = \Phi(x)^* \left[ I - \sum_{i=0}^{N-1} l^{(i)}(t) Q^{-1} A (Q^{-1}M)^i \right] \gamma \tag{62}$$

$$(Lf)(x,t) = \Phi(x)^* \left[ \sum_{i=0}^{N-1} l^{(i)}(t) Q^{-1} \left( \mathcal{I}_{t_{i+1}} - \delta A Q^{-1} \sum_{j=1}^{i} (MQ^{-1})^{i-j} \mathcal{I}_{t_j} \right) \right] f \tag{63}$$

We will now marginalise over $f$ in order to obtain the averaged conditional distribution. To do this we will need the following Lemma (which we prove below):

**Lemma 0.2.** Let $f \sim \mathcal{N}(\bar{f}, K)$ where we assume that this Gaussian measure is on a Hilbert space of functions $\mathcal{H}_1 \subset \mathbb{R}^{\mathcal{X}}$. Suppose that for a fixed realisation of $f$ we have

$$y|f \sim \mathcal{N}(Lf + c, V)$$

where $L$ is a bounded linear operator from $\mathcal{H}_1$ to another Hilbert space $\mathcal{H}_2 \subset \mathbb{R}^{\mathcal{X}}$ (so $y$ lies in $\mathcal{H}_2$), $c$ is a deterministic function and the covariance opertor $V$ does not depend on $f$. Then marginalizing over $f$ yields:

$$y \sim \mathcal{N}(L\bar{f} + c, LKL^* + V)$$

as the averaged distribution of $y$.

*Proof:* The fact that $y|f \sim \mathcal{N}(Lf + c, V)$ is equivalent to saying that:

$$y = Lf + c + \tilde{y}$$

where $\tilde{y} \sim \mathcal{N}(0, V)$ is independent of $f$. Thus we have:

$$\begin{pmatrix} f \\ \tilde{y} \end{pmatrix} = \mathcal{N}\left( \begin{pmatrix} \bar{f} \\ 0 \end{pmatrix}, \begin{pmatrix} K & 0 \\ 0 & V \end{pmatrix} \right)$$

Since we can write:

$$y = \begin{pmatrix} L & 1 \end{pmatrix} \begin{pmatrix} f \\ \tilde{y} \end{pmatrix} + c$$

we deduce:

$$y \sim \mathcal{N}\left( L\bar{f} + c, \begin{pmatrix} L & 1 \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} L^* \\ 1 \end{pmatrix} \right) = \mathcal{N}(L\bar{f} + c, LKL^* + V)$$

as required. ■

We can now perform the marginalisation over $f$ noting that $V_N = 0$ does not depend on $f$ to obtain:

$$\int u|\{\tilde{A}_{t_1}u = F^{(1)}, \dots, \tilde{A}_{t_N}u = F^{(N)}, f\} \mathrm{d}f \sim \mathcal{N}(L\bar{f} + c, LKL^*) \tag{64}$$

with $L$ and $c$ given by equations (62) and (63) respectively. ***Note: We must still prove that*** $L$ ***is bounded and work out its adjoint*** $L^*$***.***