Burgers equation

Lets look at the optimal control with PDE constraint problem, where the equation is the burgers equation:

$$\begin{aligned}
u_t + uu_x - \nu u_{xx} &= 0 \text{ for } (x,t) \in \Omega \times (0,\infty) \\
u(x,t) &= h(x,t) \text{ for } (x,t) \in \partial\Omega \times (0,\infty) \\
u(x,0) &= g(x) \text{ for } x \in \Omega
\end{aligned}$$

Here $\Omega = (a,b)$. The functional is on the form:

$$J(u(g),g) = \int_0^T \int_\Omega u(x,t)^2 dxdt$$

Here we want to minimize J with respect to the initial condition g. If we differentiate J with respect to g, we get:

$$\begin{aligned}
\hat{J}'(g)(s) &= \langle u'(g)^* J_u, s \rangle \\
&= \langle -E_g p, s \rangle
\end{aligned}$$

where p is the solution of the adjoint equation:

$$E_u^* p = J_u$$

By E I mean burgers equation, which means that:

$$\begin{aligned}
E_u &= \frac{\partial}{\partial t} + u_x + u\frac{\partial}{\partial x} - \nu\Delta + \delta_{t=0} + \delta_{\partial\Omega} \\
E_g &= -\delta_{t=0} \\
E_u^* &= -\frac{\partial}{\partial t} - u\frac{\partial}{\partial x} - \nu\Delta + \delta_{t=T} + (1+h)\delta_{\partial\Omega} \\
J_u &= 2u
\end{aligned}$$

The adjoint equation would then look like:

$$\begin{aligned}
-p_t - up_x - \nu p_{xx} &= 2u \text{ for } (x,t) \in \Omega \times (0,\infty) \\
p(x,t) &= 0 \text{ for } (x,t) \in \partial\Omega \times (0,\infty) \\
p(x,T) &= 0 \text{ for } x \in \Omega
\end{aligned}$$

This would mean that the gradient of $\hat{J}$ is:

$$\hat{J}'(g)(s) = \langle p(\cdot,0), s \rangle$$

Two time intervals
Divide the time interval $[0,T]$ into $[0,T_1]$ and $[T_1,T_2]$, with $T_2 = T$. Then solve the equation separately on the two intervals for functions u^1 and u^2.

The only difference is that u^2 has its own initial condition at $t = T_1$, that we call λ. To solve the problem, we now have to add a penalty term to the functional. The new functional looks as the following:

$$J(u(g), g, \lambda) = \int_0^T \int_\Omega u(x,t)^2 dxdt + \frac{\mu}{2}\int_\Omega (u^1(x,T_1) - \lambda(x))^2 dx$$

We now have a new unknown λ, which is the initial condition of the second time interval. We now want the gradient of the new functional:

$$\begin{aligned}\langle \hat{J}'(g,\lambda), (s,l)\rangle &= \langle \frac{\partial u(g,\lambda)}{\partial(g,\lambda)}^* J_g(u(g,\lambda), g, \lambda), (s,l)\rangle + \langle J_g + J_\lambda, (s,l)\rangle \\ &= \langle -(E_g + E_\lambda)p, (s,l)\rangle + \langle J_g + J_\lambda, (s,l)\rangle\end{aligned}$$

Again p is the solution of the adjoint equation, and as for u, p is separated into two equations. To derive the gradient lets divide E into E^1 and E^2:

$$\begin{aligned}E^1 &= u_t^1 + u^1 u_x^1 - \nu u_{xx}^1 + \delta_{t=0}(u^1 - g) + \delta_{\partial\Omega}(u^1 - h) \\ E^2 &= u_t^2 + u^2 u_x^2 - \nu u_{xx}^2 + \delta_{t=T_1}(u^2 - \lambda) + \delta_{\partial\Omega}(u^2 - h)\end{aligned}$$

If we differentiate E we get the following:

$$\begin{aligned}E_u^i &= \frac{\partial}{\partial t} + u_x^i + u^i \frac{\partial}{\partial x} - \nu\Delta + \delta_{t=T_{i-1}} + \delta_{\partial\Omega} \\ E_g^1 &= -\delta_{t=0} \\ E_\lambda^2 &= -\delta_{t=T_1}\end{aligned}$$

Now let us differentiate J:

$$\begin{aligned}J_g &= 0 \\ J_u &= 2u + \mu(u^1 - \lambda)\delta_{t=T_1} \\ \langle J_\lambda, l\rangle &= -\mu\int_\Omega (u^1(x,T_1) - \lambda(x))l(x)dx\end{aligned}$$

We need the adjoint of E_u^i, but this is almost the same as above, so instead, we can write down the two adjoint equations:
$i = 2$:

$$\begin{aligned}-p_t^2 - u^2 p_x^2 - \nu p_{xx}^2 &= 2u^2 \text{ for } (x,t) \in \Omega \times (T_1, T_2) \\ p^2(x,t) &= 0 \text{ for } (x,t) \in \partial\Omega \times (T_1, T_2) \\ p^2(x, T_2) &= 0 \text{ for } x \in \Omega\end{aligned}$$

$i = 1$:

$$\begin{aligned}-p_t^1 - u^1 p_x^1 - \nu p_{xx}^1 &= 2u^1 \text{ for } (x,t) \in \Omega \times (0, T_1) \\ p^1(x,t) &= 0 \text{ for } (x,t) \in \partial\Omega \times (0, T_1) \\ p^1(x, T_1) &= \mu(u^1(x,T_1) - \lambda(x)) \text{ for } x \in \Omega\end{aligned}$$

This gives us the following gradient:

$$\begin{aligned}
\langle \hat{J}'(g,\lambda),(s,l)\rangle &= \langle -(E_g+E_\lambda)p,(s,l)\rangle + \langle J_g + J_\lambda,(s,l)\rangle \\
&= \int_\Omega p^1(x,0)s(x)dx + \int_\Omega p^2(x,T_1)l(x)dx - \mu \int_\Omega (u^1(x,T_1)-\lambda(x))l(x)dx \\
&= \int_\Omega p^1(x,0)s(x)dx + \int_\Omega (p^2(x,T_1)-p^1(x,T_1))l(x)dx
\end{aligned}$$

m time intervals
Divide $[0,T]$ into m intervals $[T_{i-1},T_i]$, where $0=T_0<T_1<\cdots<T_m=T$. We then solve burgers equation for $\{u^i(x)\}_{i=1}^m$ on each interval with initial conditions $\{\lambda_i(x)\}_{i=0}^{m-1}$, where $\lambda_0(x)=g(x)$. As for the two interval case, we need to add penalty terms to the functional, to solve the optimization problem. The penalized functional looks like this:

$$J(u(g),g,\lambda) = \int_0^T \int_\Omega u(x,t)^2 dxdt + \frac{\mu}{2}\sum_{i=1}^{m-1}\int_\Omega (u^i(x,T_i)-\lambda_i(x))^2 dx$$

The state equation on each interval looks like:

$$\begin{aligned}
u^i_t + u^i u^i_x - \nu u^i_{xx} &= 0 \text{ for } (x,t)\in\Omega\times(T_{i-1},T_i) \\
u^i(x,t) &= h(x,t) \text{ for } (x,t)\in\partial\Omega\times(T_{i-1},T_i) \\
u^i(x,0) &= \lambda_{i-1}(x) \text{ for } x\in\Omega
\end{aligned}$$

If write this equation as en operator E^i, we get:

$$E^i = u^i_t + u^i u^i_x - \nu u^i_{xx} + \delta_{t=T_{i-1}}(u^i-\lambda_{i-1}) + \delta_{\partial\Omega}(u^i-h)$$

Now lets look at the gradient. The gradient expression for m intervals is the same as for 2 intervals, however λ is now a vector, i.e. $\lambda=(\lambda_1,...,\lambda_{m-1})$. First the gradient:

$$\langle \hat{J}'(g,\lambda),(s,l)\rangle = \langle -(E_g+E_\lambda)p,(s,l)\rangle + \langle J_g+J_\lambda,(s,l)\rangle$$

Again p is the solutions of the adjoint equations $(E^i_u)^*p^i = J_{u^i}$. Lets state the values of the different components in the gradient:

$$\begin{aligned}
E^i_u &= \frac{\partial}{\partial t} + u^i_x + u^i\frac{\partial}{\partial x} - \nu\Delta + \delta_{t=T_{i-1}} + \delta_{\partial\Omega} \\
E^1_g &= -\delta_{t=0} \\
E^i_{\lambda_i} &= -\delta_{t=T_i} \; i\neq 1 \\
J_g &= 0 \\
J_u &= 2u + \mu\sum_{i=1}^{m-1}(u^i-\lambda_i)\delta_{t=T_i} \\
\langle J_{\lambda_i}, l\rangle &= -\mu\int_\Omega (u^i(x,T_i)-\lambda_i(x))l_i(x)dx
\end{aligned}$$

This gives us the following adjoint equations:

$$\begin{aligned}
-p^i_t - u^i p^i_x - \nu p^i_{xx} &= 2u^i \text{ for } (x,t) \in \Omega \times (T_{i-1}, T_i) \\
p^i(x,t) &= 0 \text{ for } (x,t) \in \partial\Omega \times (T_{i-1}, T_i) \\
p^m(x,T) &= 0 \text{ for } x \in \Omega \\
p^i(x,T_i) &= \mu(u^i(x,T_i) - \lambda_i(x)) \text{ for } x \in \Omega \text{ and } i \neq m
\end{aligned}$$

We can also show that the gradient looks like the following:

$$\begin{aligned}
\langle \hat{J}'(g,\lambda),(s,l)\rangle &= \langle -(E_g + E_\lambda)p,(s,l)\rangle + \langle J_g + J_\lambda,(s,l)\rangle \\
&= \int_\Omega p^1(x,0)s(x)dx + \sum_{i=1}^{m-1} \int_\Omega (p^{i+1}(x,T_i) - p^i(x,T_i))l_i(x)dx
\end{aligned}$$