Homework #6 ECE 6560 Marcus Pereira

1.> Derive the Euler-Lagrange Equation from Directional Derivative

$$E(f) = \int_0^1 L(f(x), \dot{f}(x), x)\,dx$$

* If we perturb $f(x)$ by $g(x)$, then $f'(x)$ is also perturbed and that is given by $\frac{\partial}{\partial x}(f(x)+\varepsilon g(x)) = \dot{f}(x)+\varepsilon\dot{g}(x)$

$$\frac{\partial E(f)}{\partial g} = \lim_{\varepsilon\to 0} \frac{E(f+\varepsilon g)-E(f)}{\varepsilon} = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\int_0^1 \left[L(f(x)+\varepsilon g(x), \dot{f}(x)+\varepsilon\dot{g}(x), x) - L(f(x),\dot{f}(x),x)\right]dx\right]$$

Using Taylor's expansion on the first integrand gives,

$$\int_0^1 L(f(x)+\varepsilon g(x), \dot{f}(x)+\varepsilon\dot{g}(x), x)\,dx = \int_0^1 \left\{L(f(x),\dot{f}(x),x) + \frac{\partial L}{\partial f}\varepsilon g(x) + \theta[(\varepsilon g)^2] + \frac{\partial L}{\partial \dot{f}}(\varepsilon\dot{g}(x)) + \theta[(\varepsilon\dot{g})^2]\right\}dx$$

$$\therefore \frac{\partial E(f)}{\partial g} = \lim_{\varepsilon\to 0}\int_0^1 \frac{1}{\varepsilon}\left[\cancel{L(f(x),\dot{f}(x),x)} + \frac{\partial L}{\partial f}\varepsilon g(x) + \frac{\partial L}{\partial \dot{f}}(\varepsilon\dot{g}(x)) + \theta[(\varepsilon g(x))^2] + \theta[(\varepsilon\dot{g}(x))^2] - \cancel{L(f(x),\dot{f}(x),x)}\right]dx$$

The partial terms are constant w.r.t limit $\varepsilon\to 0$ and the $\theta(\varepsilon g)^2/\varepsilon$, $\frac{\theta(\varepsilon\dot{g})^2}{\varepsilon}$ are higher order terms which go to zero in the limit, giving the following

$$\frac{\partial E(f)}{\partial g} = \int_0^1 \left(\frac{\partial L}{\partial f}g(x) + \frac{\partial L}{\partial \dot{f}}\dot{g}(x)\right)dx$$

Using integration by parts, $\int_0^1 \frac{\partial L}{\partial \dot{f}}\dot{g}(x)\,dx = \frac{\partial L}{\partial \dot{f}}g(x)\Big|_0^1 - \int_0^1 \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial \dot{f}}\right)g(x)\,dx$

The first term goes to zero for 2 cases:-

i) periodic functions

ii) $g(x)$ is zero at the boundaries (i.e no perturbation at $x=1$ and at $x=0$)

$$\therefore \frac{\partial E(f)}{\partial g} = \int_0^1\left[\frac{\partial L}{\partial f}g(x) - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial \dot{f}}\right)g(x)\right]dx = \int_0^1 g(x)\left[\frac{\partial L}{\partial f} - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial \dot{f}}\right)\right]dx$$

The above directional derivative vanishes if $\frac{\partial L}{\partial f} - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial \dot{f}}\right) = 0$

This condition for local minimum/maximum gives the Euler-Lagrange equation

$$\frac{\partial L}{\partial f} - \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial \dot{f}}\right) = 0$$

2) Finite dimensional versus infinite-dimensional gradient

$$\hat{E}(f_0, f_1, \dots, f_N) = \sum_{n=0}^{N-1} L\left(f_n, \frac{f_{n+1}-f_n}{\Delta x}, x_n\right)\Delta x$$

$$\nabla\hat{E} = \left[\frac{\partial \hat{E}}{\partial f_0}, \frac{\partial \hat{E}}{\partial f_1}, \dots, \frac{\partial \hat{E}}{\partial f_{N-1}}\right]^T$$ is the gradient vector

Consider $\frac{\partial \hat{E}}{\partial f_n}$. There are 2 terms that will contribute to this derivative which are the $(n-1)^{th}$ and the n^{th} terms as other terms don't depend on f_n and will be zero.

$$\therefore \frac{\partial \hat{E}}{\partial f_n} = \left[\frac{\partial}{\partial f_n} L\left(f_{n-1}, \frac{f_n - f_{n-1}}{\Delta x}, x_{n-1}\right) + \frac{\partial}{\partial f_n} L\left(f_n, \frac{f_{n+1}-f_n}{\Delta x}, x_n\right)\right]\Delta x$$

$$= \Delta x\left[\cancelto{0}{L_1^{n-1}\frac{\partial}{\partial f_n}(f_{n-1})} + L_2^{n-1}\frac{\partial}{\partial f_n}\left(\frac{f_n - f_{n-1}}{\Delta x}\right) + L_1^{n}\frac{\partial}{\partial f_n}(f_n) + L_2^{n}\frac{\partial}{\partial f_n}\left(\frac{f_{n+1}-f_n}{\Delta x}\right)\right]$$

where, L_1^n : partial derivative of L w.r.t the first input for n^{th} term

L_2^n : partial derivative of L w.r.t the second input for n^{th} term

$$\therefore \frac{\partial \hat{E}}{\partial f_n} = \Delta x\left[L_1^n(1) + L_2^n\left(\frac{-1}{\Delta x}\right) + L_2^{n-1}\left(\frac{1}{\Delta x}\right)\right]$$

$$= \Delta x\left[L_1\left(f_n, \frac{f_{n+1}-f_n}{\Delta x}, x_n\right) - \frac{1}{\Delta x}L_2\left(f_n, \frac{f_{n+1}-f_n}{\Delta x}, x_n\right) + \frac{1}{\Delta x}L_2\left(f_{n-1}, \frac{f_n - f_{n-1}}{\Delta x}, x_{n-1}\right)\right]$$

$$= \Delta x\left[L_1\left(f_n, \frac{f_{n-1}-f_n}{\Delta x}, x_n\right) - \frac{L_2\left(f_n, \frac{f_{n+1}-f_n}{\Delta x}, x_n\right) - L_2\left(f_{n-1}, \frac{f_n - f_{n-1}}{\Delta x}, x_{n-1}\right)}{\Delta x}\right]$$

The only term that will be different is the first term as there is no $(n-1)^{th}$ term for $n=0$

$$\therefore \frac{\partial \hat{E}}{\partial f_0} = \frac{\partial}{\partial f_0} L\left(f_0, \frac{f_1 - f_0}{\Delta x}, x_0\right)\Delta x = \left(L_1^0 \frac{\partial}{f_0}(f_0) + L_2^0\frac{\partial}{\partial f_0}\left(\frac{f_1 - f_0}{\Delta x}\right)\right)\Delta x$$

$$= \left(L_1^0 - \frac{1}{\Delta x}L_2^0\right)\Delta x$$

$$\therefore \frac{\partial \hat{E}}{\partial f_0} = \Delta x\left[L_1\left(f_0, \frac{f_1 - f_0}{\Delta x}, x_0\right) - \frac{1}{\Delta x}L_2\left(f_0, \frac{f_1 - f_0}{\Delta x}, x_0\right)\right]$$

$$= \Delta x\left[L_1\left(f_0, \frac{f_1 - f_0}{\Delta x}, x_0\right) - \frac{L_2\left(f_0, \frac{f_1 - f_0}{\Delta x}, x_0\right) - L_2\left(f_{-1}, \frac{f_0 - f_{-1}}{\Delta x}, x_{-1}\right)}{\Delta x}\right]$$

where, $L_2(f_{-1}, f_0 - f_{-1}, x_{-1}) = 0$ for fictitious point at $n=-1$

$$\therefore \nabla \hat{E} = \Delta x \begin{bmatrix} L_1\left(f_0, \frac{f_1 - f_0}{\Delta x}, x_0\right) - \dfrac{L_2\left(f_0, \frac{f_1 - f_0}{\Delta x}, x_0\right) - L_2\left(f_{-1}, \frac{f_0 - f_{-1}}{\Delta x}, x_{-1}\right)}{\Delta x} \\ \vdots \\ L_1\left(f_n, \frac{f_{n+1} - f_n}{\Delta x}, x_n\right) - \dfrac{L_2\left(f_n, \frac{f_{n+1} - f_n}{\Delta x}, x_n\right) - L_2\left(f_{n-1}, \frac{f_n - f_{n-1}}{\Delta x}, x_{n-1}\right)}{\Delta x} \\ \vdots \\ L_1\left(f_{N-1}, \frac{f_N - f_{N-1}}{\Delta x}, x_{N-1}\right) - \dfrac{L_2\left(f_{N-1}, \frac{f_N - f_{N-1}}{\Delta x}, x_{N-1}\right) - L_2\left(f_{N-1}, \frac{f_N - f_{N-1}}{\Delta x}, x_{N-1}\right)}{\Delta x} \end{bmatrix}$$

where, f_{-1} and f_N are fictitious points at the start and end of the sequence

We can see that as $N \to \infty$, we get $\frac{L_2^n - L_2^{n-1}}{\Delta x} \to \frac{\partial L_{\dot{f}}}{\partial x}\Big|_{f(x_n)}$ as $\left(\Delta x = \frac{1}{N}\right) \to 0$

where, $\dot{f}$ satisfies $\lim_{\Delta x \to 0} \frac{f_{n+1} - f_n}{\Delta x} = \frac{\partial f(x)}{\partial x}\Big|_{f(x_n)}$

Since $\Delta x = \frac{1}{N}$, we can see that as $N \to \infty$, we get the terms in the vector to converge to $L_f - \frac{\partial}{\partial x} L_{\dot{f}}$ evaluated at each point n upto a constant factor of $\frac{1}{N}$.

As $N \to \infty$, the "error" of the fictitious points at the 2 ends f_{-1} and f_N is negligible and can be ignored.

⟨3⟩ Gradient descent diffusion equation

$$E = \frac{1}{2}\iint e^{(I_x^2+I_y^2)}\,dx\,dy$$

Using the total variation gradient descent result derived in class we have for the above case,

$$c(\|\nabla I\|) = c\left(\sqrt{I_x^2+I_y^2}\right) = c(y) = e^{y^2}$$

$$\therefore\ \dot{c} = e^{y^2}\,2y = 2yc$$

$$\therefore\ I_t = \nabla\cdot\left(\frac{\dot{c}(\|\nabla I\|)}{\|\nabla I\|}\nabla I\right)$$

$$= \nabla\cdot\left(\frac{2\cancel{\|\nabla I\|}c}{\cancel{\|\nabla I\|}}\nabla I\right) \quad \text{as } y = \|\nabla I\| = \sqrt{I_x^2+I_y^2}$$

$$= \nabla\cdot\left(2c\,\nabla I\right)$$

$$= 2\left[\frac{\partial}{\partial x}\left(e^{I_x^2+I_y^2} I_x\right) + \frac{\partial}{\partial y}\left(e^{I_x^2+I_y^2} I_y\right)\right]$$

$$= 2\left[e^{I_y^2}\frac{\partial}{\partial x}\left(e^{I_x^2} I_x\right) + e^{I_x^2}\frac{\partial}{\partial y}\left(e^{I_y^2} I_y\right)\right]$$

$$= 2\left[e^{I_y^2}\left(e^{I_x^2} I_{xx} + I_x e^{I_x^2}\,2I_x I_{xx}\right) + e^{I_x^2}\left(e^{I_y^2} I_{yy} + I_y e^{I_y^2}\,2I_y I_{yy}\right)\right]$$

$$= 2e^{I_y^2}e^{I_x^2} I_{xx}\left[1 + 2(I_x)^2\right] + 2e^{I_x^2}e^{I_y^2} I_{yy}\left[1 + 2(I_y)^2\right]$$

$$\boxed{I_t = 2e^{(I_x^2+I_y^2)}\left[I_{xx} + 2(I_x)^2 I_{xx} + I_{yy} + 2(I_y)^2 I_{yy}\right]}$$