

I start with a simple neural net with weights $W1, W2$ and biases $b1, b2$.

```
using Flux

W1 = rand(5,2)
b1 = rand(5)

W2 = rand(1, 5)
b2 = rand(1)

# set neural net parameters
ps = params(W1, b1, W2, b2)

# build a simple neural net
u(x, W1, b1, W2, b2) = sum(W2*tanh.(W1*x .+ b1) .+ b2)
```

Using Flux.jl I can differentiate u with respect to x

```
u_x = (x, W1, b1, W2, b2) -> gradient(
    x -> u(x, W1, b1, W2, b2), x)[1]
```

Now, If I want, I can differentiate u_x with respect to any of the weights and biases with out any issue.

```
∇u = gradient(ps) do
    sum(u_x(x, W1, b1, W2, b2))
end
```

The issue arises if I try to differentiate u_x with respect to x before computing the gradient with respect to the weights and biases.

```
u_xx = (x, W1, b1, W2, b2) -> gradient(
    x -> sum(u_x(x, W1, b1, W2, b2)), x)[1]

∇ = gradient(ps) do
    sum(u_xx(x, W1, b1, W2, b2))
end
```