$$R_j^{(n)} = \sum_i \mathbf{X}_j \frac{\partial L_i^{(n)}(\mathbf{X}, \mathbf{W})}{\partial \mathbf{X}_j} \frac{R_i^{(n-1)}}{L_i^{(n)}(\mathbf{X}, \mathbf{W})}$$

$$R_j^{(n)} = \mathbf{X}_j \left(\sum_i \frac{\partial L_i^{(n)}(\mathbf{X}, \mathbf{W})}{\partial \mathbf{X}_j} \frac{R_i^{(n-1)}}{L_i^{(n)}(\mathbf{X}, \mathbf{W})} \right)$$

$$\frac{R_i^{(n-1)}}{L_i^{(n)}(\mathbf{X}, \mathbf{W})} = \frac{R_i^{(n-1)}}{Z_1 + Z_2} = \frac{R}{Z_1 + Z_2}$$

sum $= 0$

for $i \in [1{,}1000] \cap \mathbb{N}$ {

compute $\frac{\partial L_i^{(n)}(\mathbf{X},\mathbf{W})}{\partial \mathbf{X}_j}$

multiply with $\frac{R_i^{(n-1)}}{L_i^{(n)}(\mathbf{X},\mathbf{W})}$

sum += product

}

return sum

$R_j^{(n)} =$ `x1 *torch.autograd.grad(Z1, x1, S1)[0]`