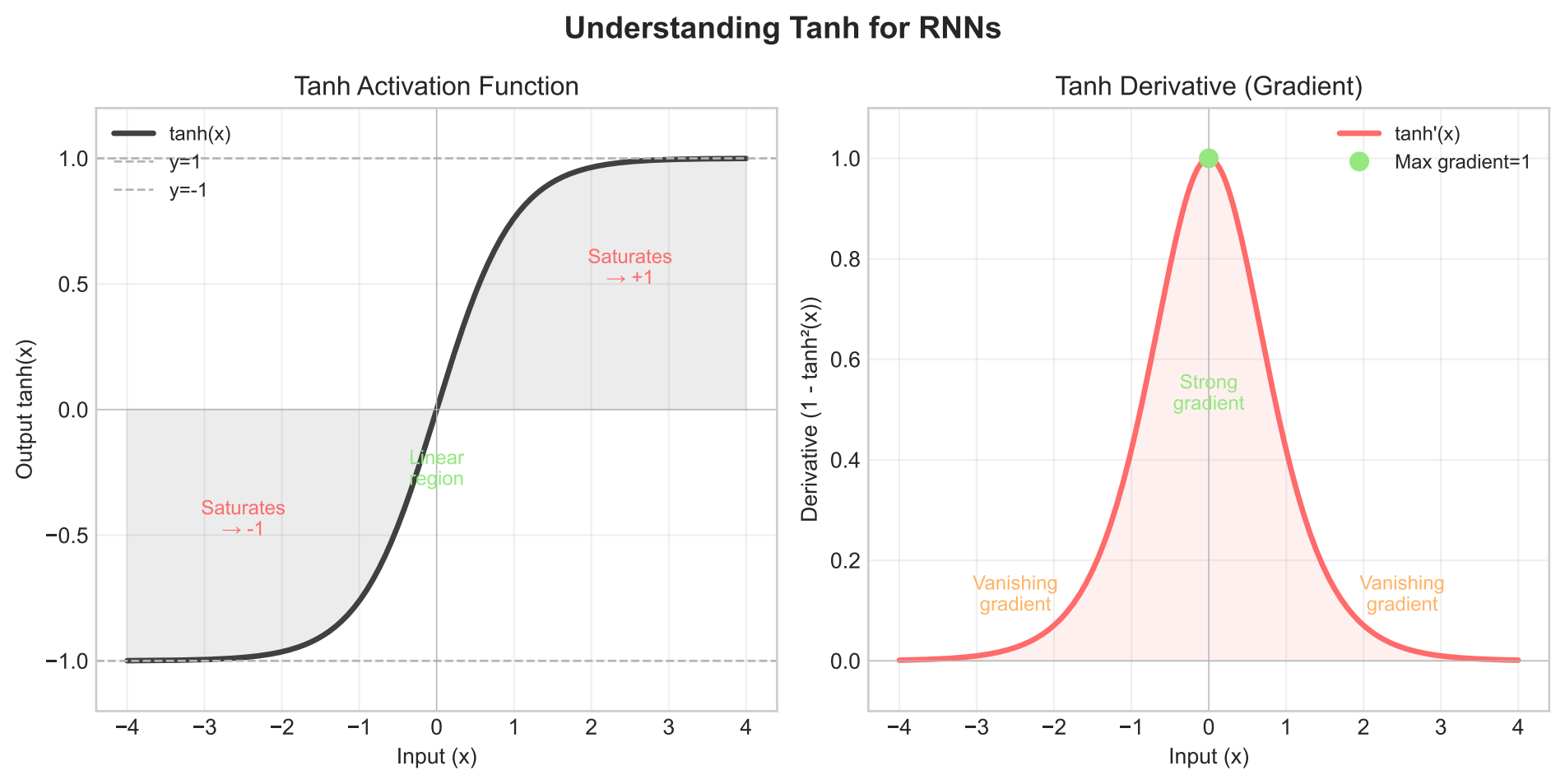
Understanding Tanh for RNNs
Tanh Activation Function
tanh(x)
y=1
y=-1
Saturates
→ +1
Linear
region
Saturates
→ -1
Output tanh(x)
Input (x)
Tanh Derivative (Gradient)
tanh'(x)
Max gradient=1
Strong
gradient
Vanishing
gradient
Vanishing
gradient
Derivative (1 - tanh²(x))
Input (x)