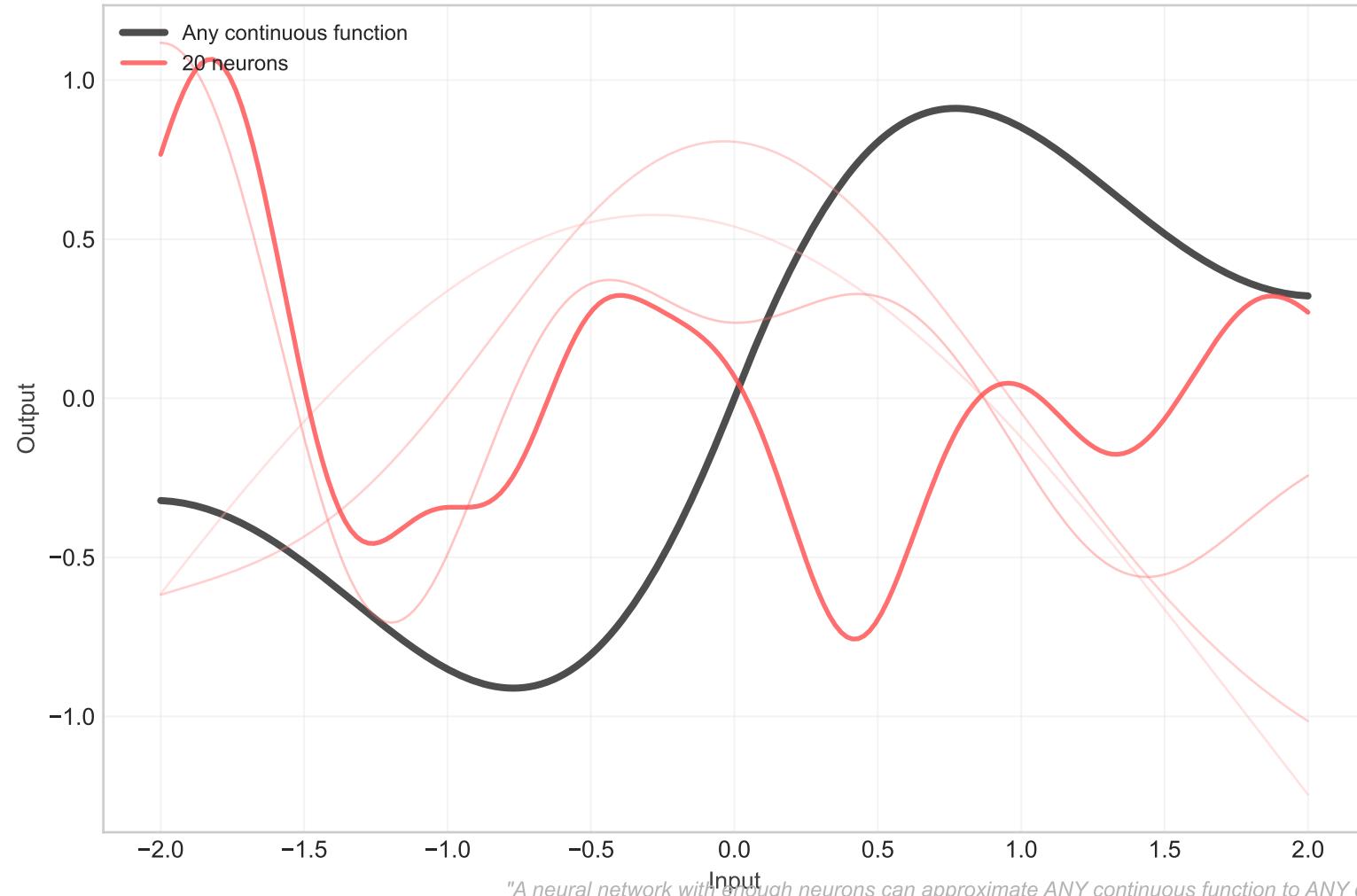


The Universal Approximation Theorem (Cybenko, 1989)

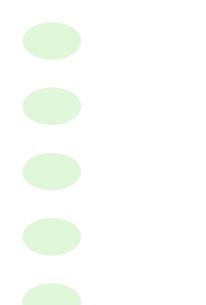
Universal Approximation: More Neurons = Better Fit



Network Width



Few



More



Many

Network Depth



More layers = More complex patterns