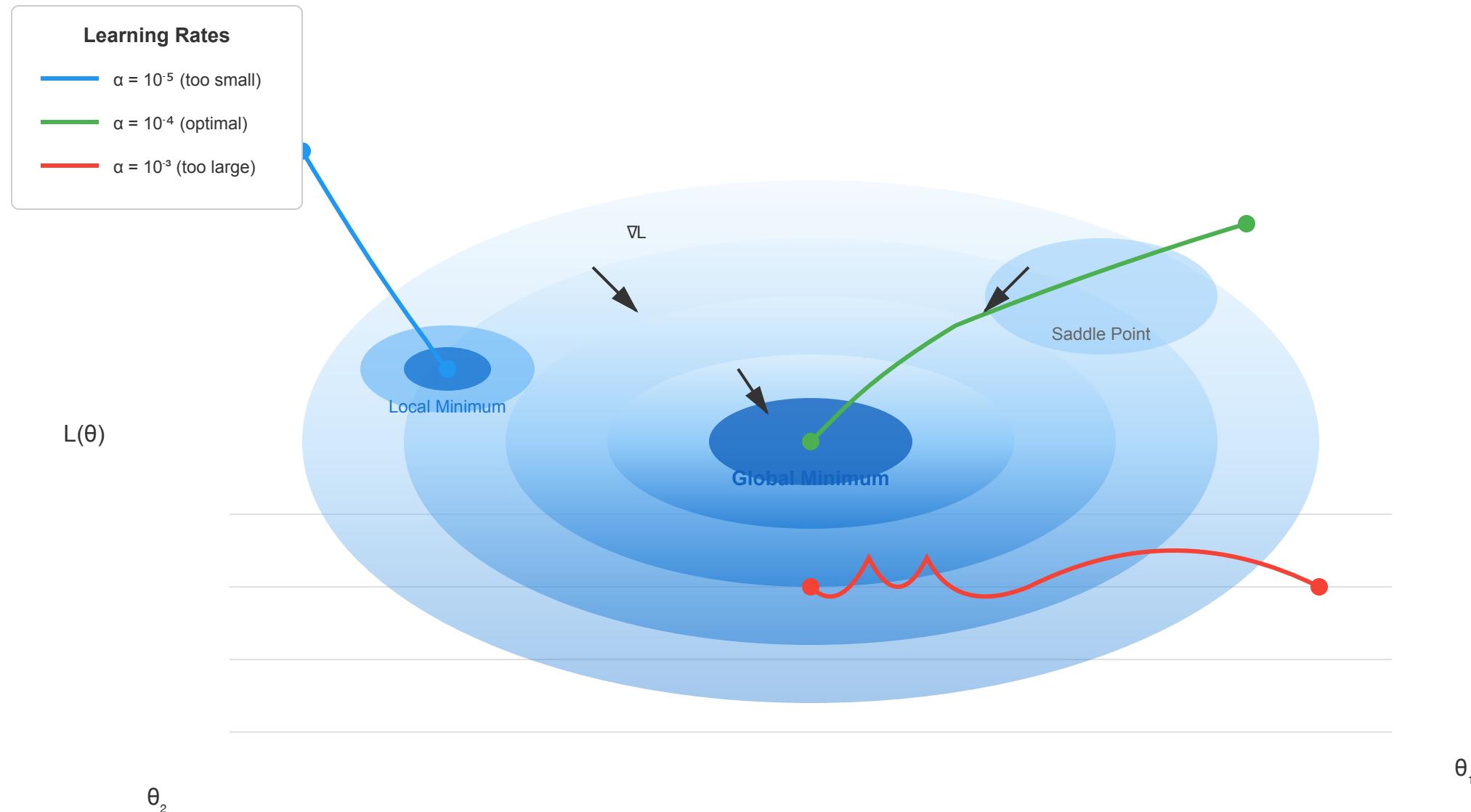


Loss Landscape with Gradient Descent Trajectories



Gradient descent follows ∇L downhill. Learning rate α controls step size: too small is slow, too large oscillates.