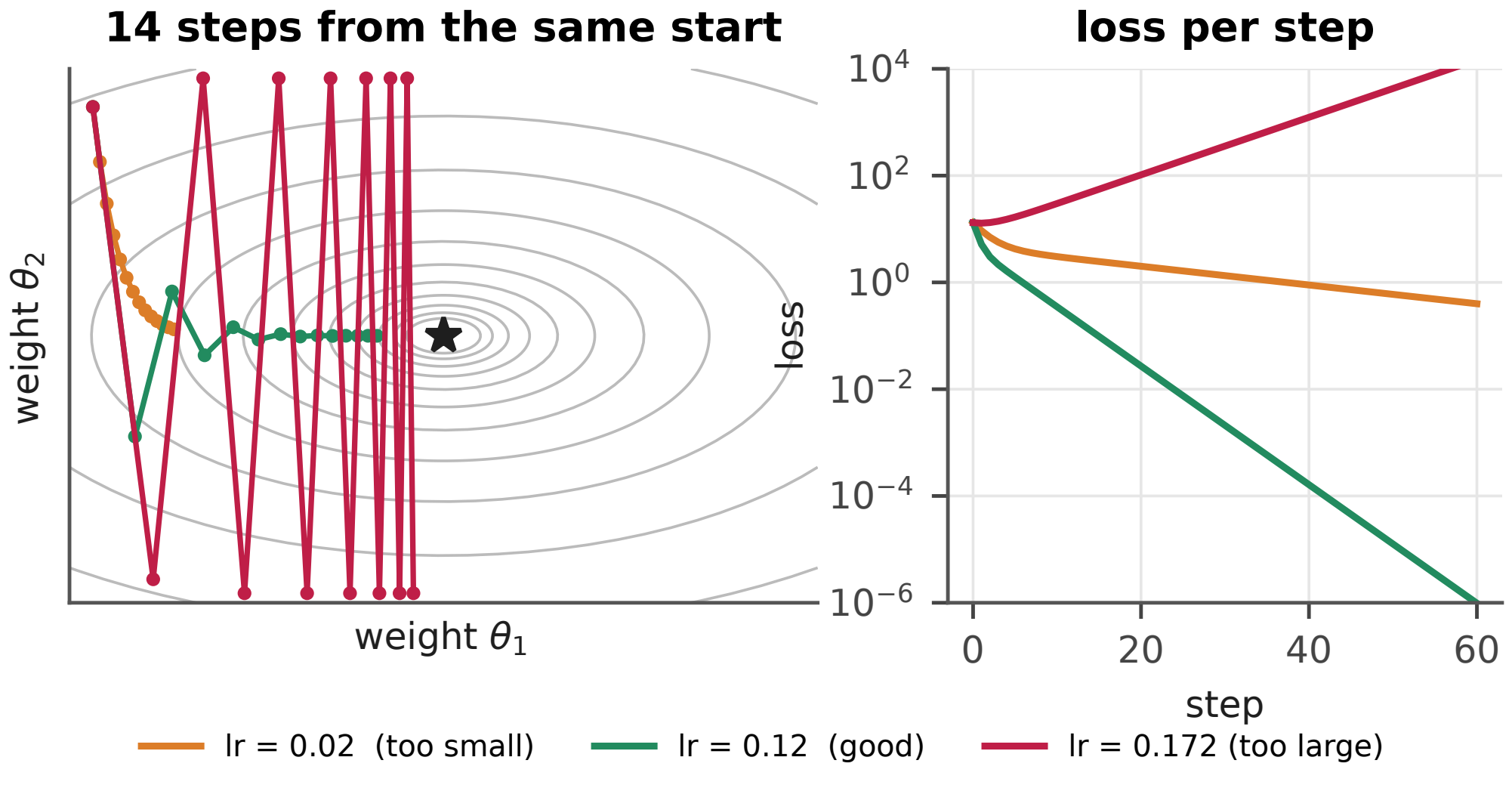

14 steps from the same start
loss per step
weight θ_2
weight θ_1
loss
10^4
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
0
20
40
60
step
lr = 0.02 (too small)
lr = 0.12 (good)
lr = 0.172 (too large)