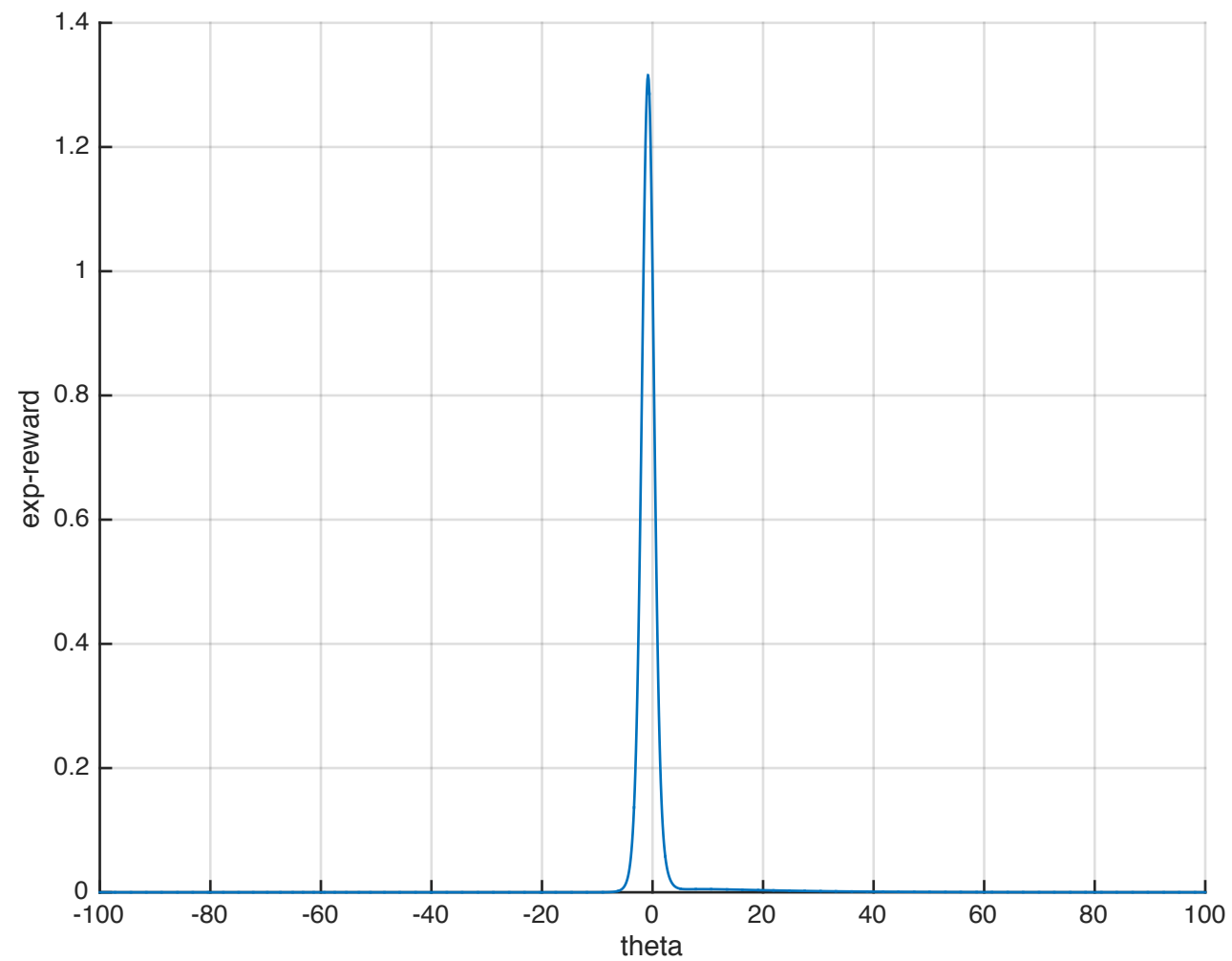
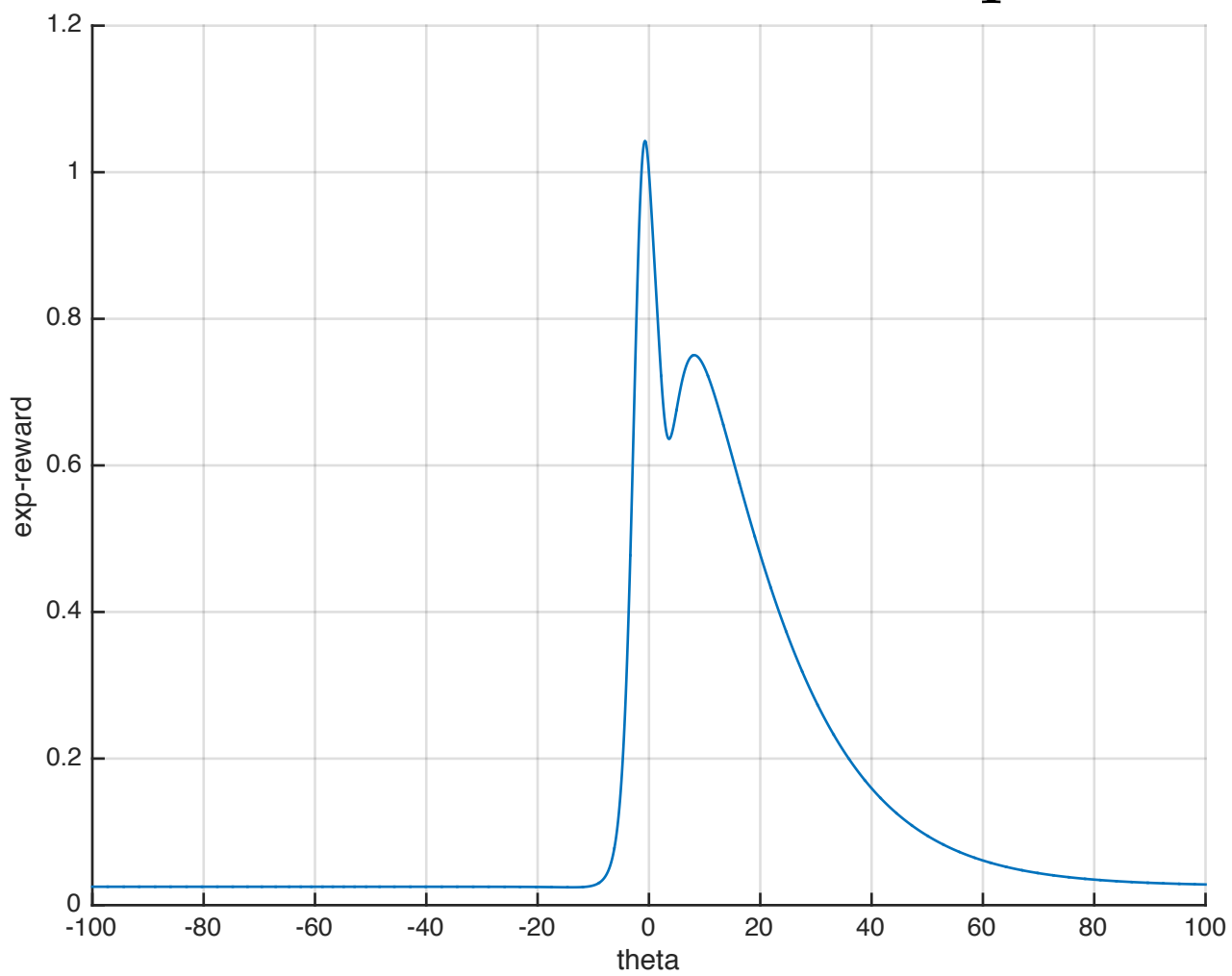


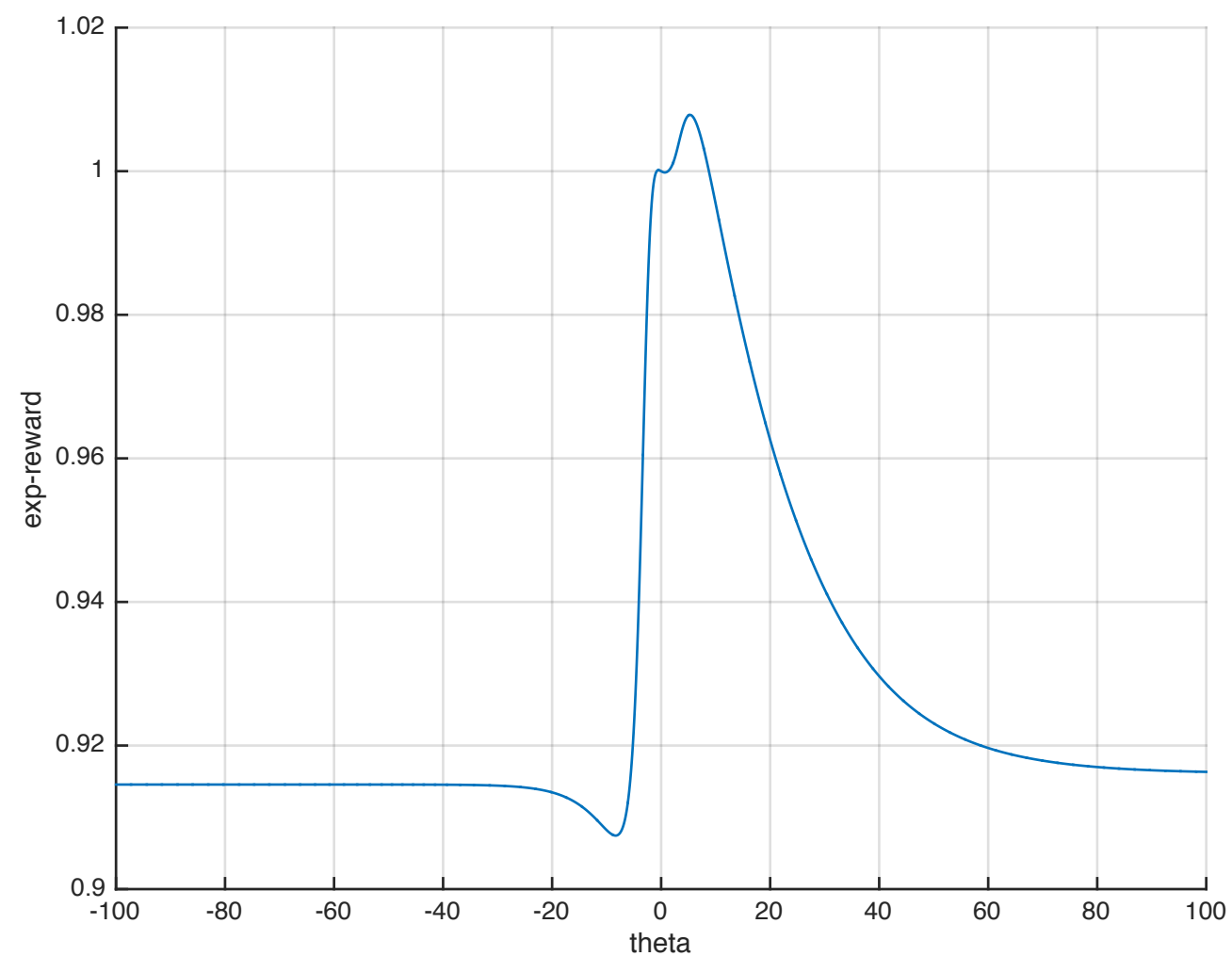
true reward landscape



$\tau = 0.2$



$\tau = 0.5$



$\tau = 10$