Linear Quadratic Regulator:

The Analytic MDP

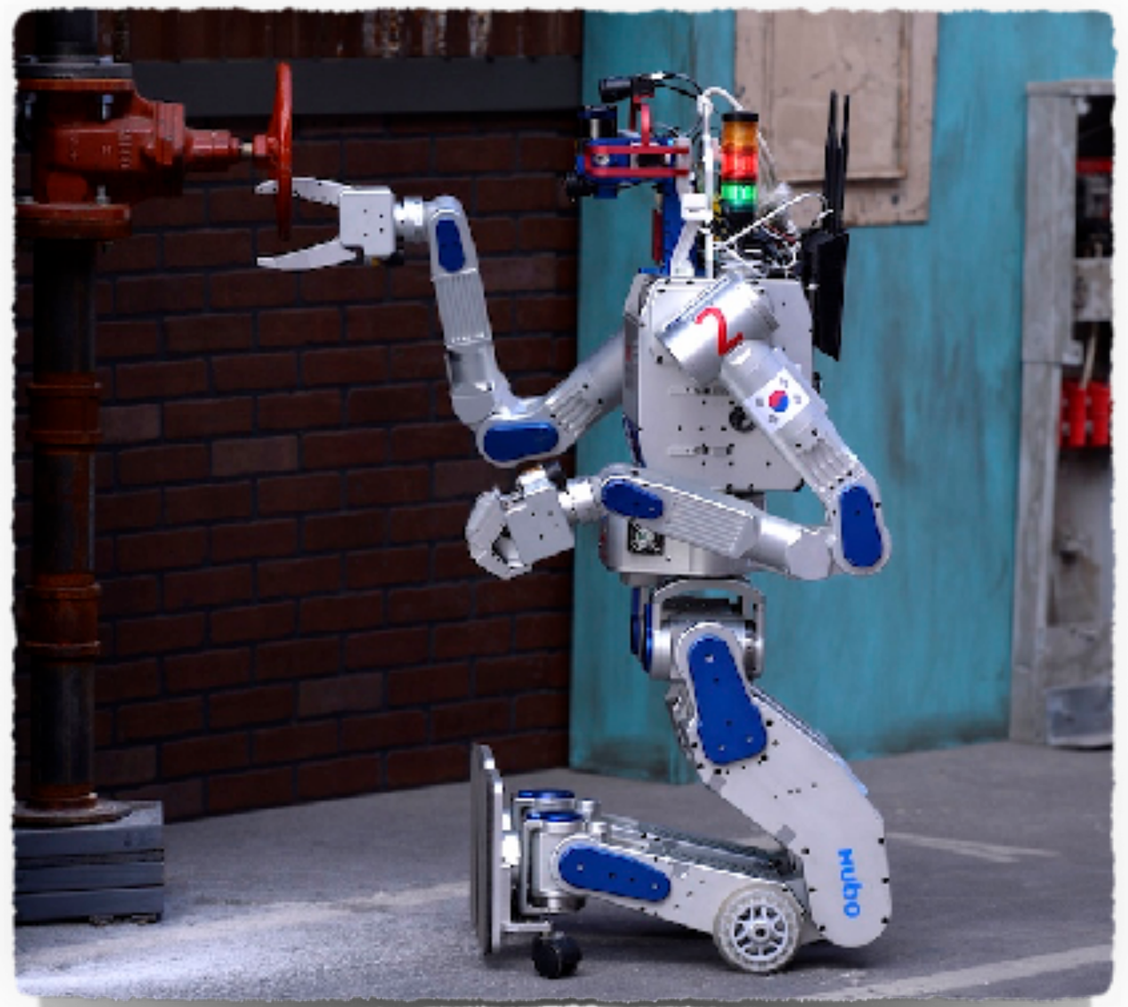
Sanjiban Choudhury

Cornell Bowers CIS
Computer Science

Announcements

1. No office hours for Sanjiban on Thursday this week :(

It's time to bring in the robots!




BostonDynamics

Activity!

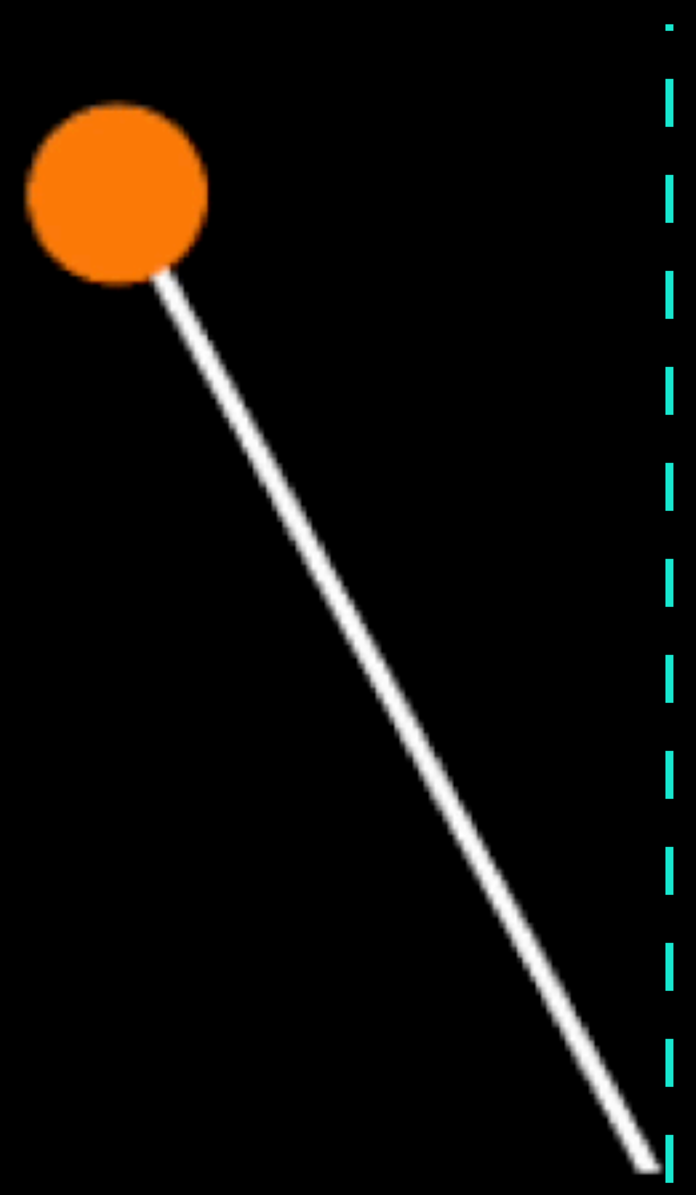
Think-Pair-Share

Think (30 sec): How do we model the Atlas backflip as a Markov Decision Problem <S, A, C, T>?

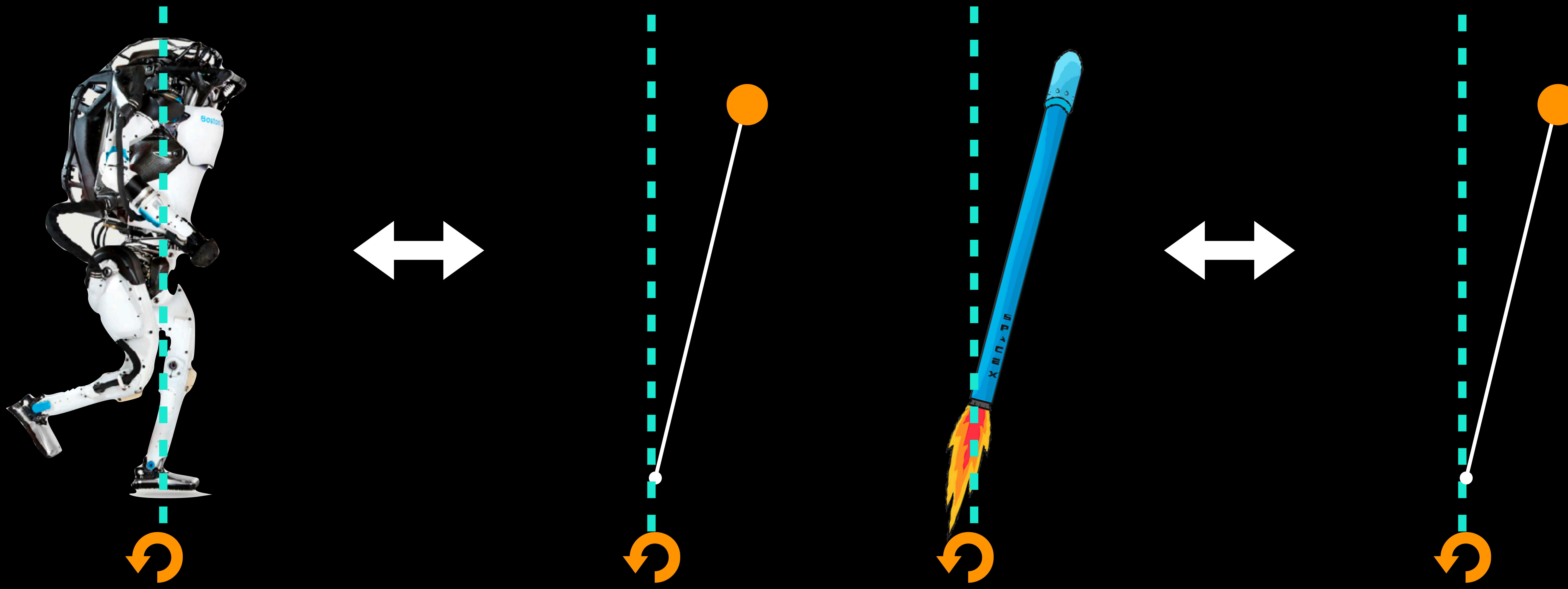
Pair: Find a partner

Share (45 sec): Partners exchange ideas

The Inverted Pendulum: A fundamental dynamics model

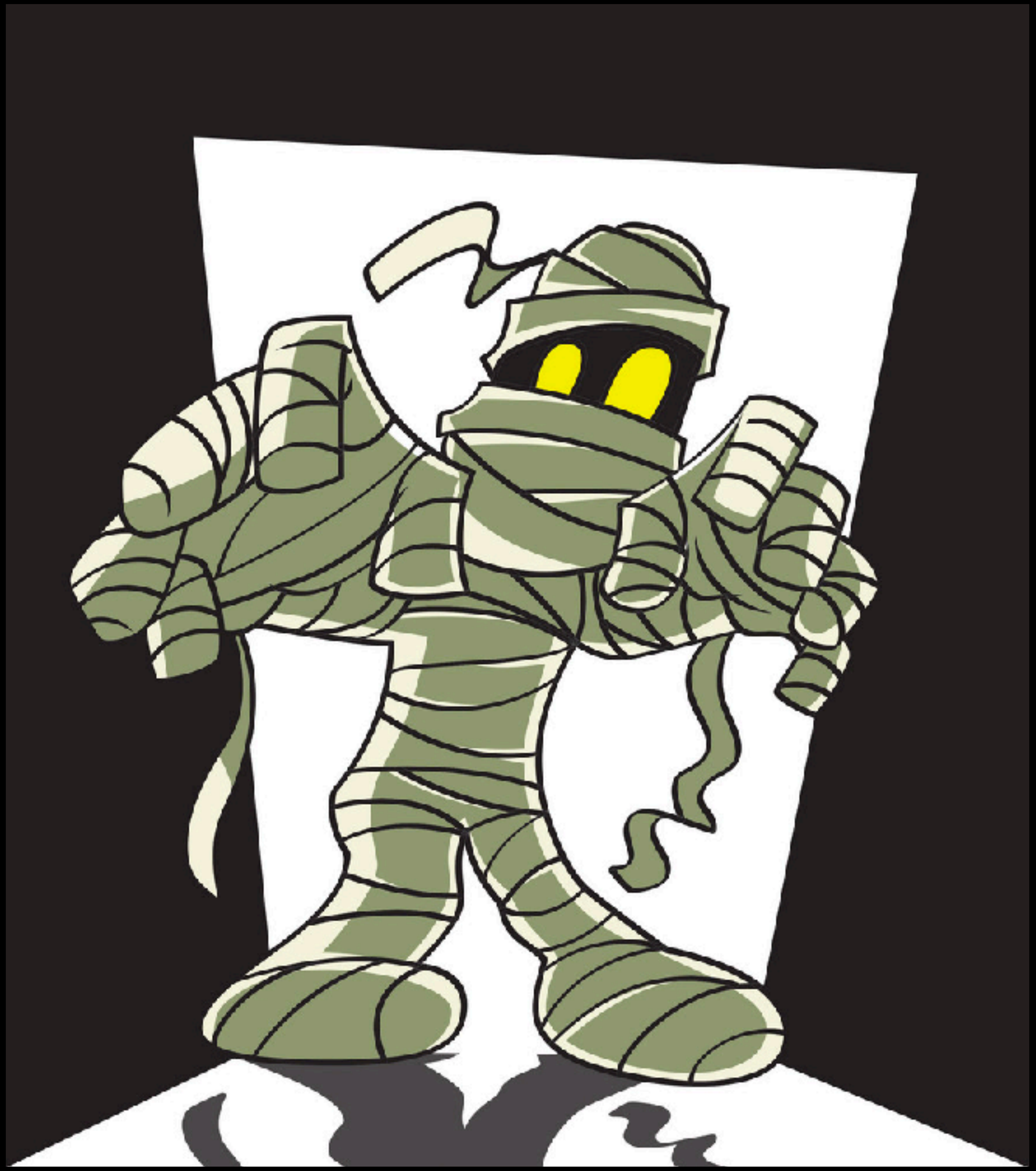


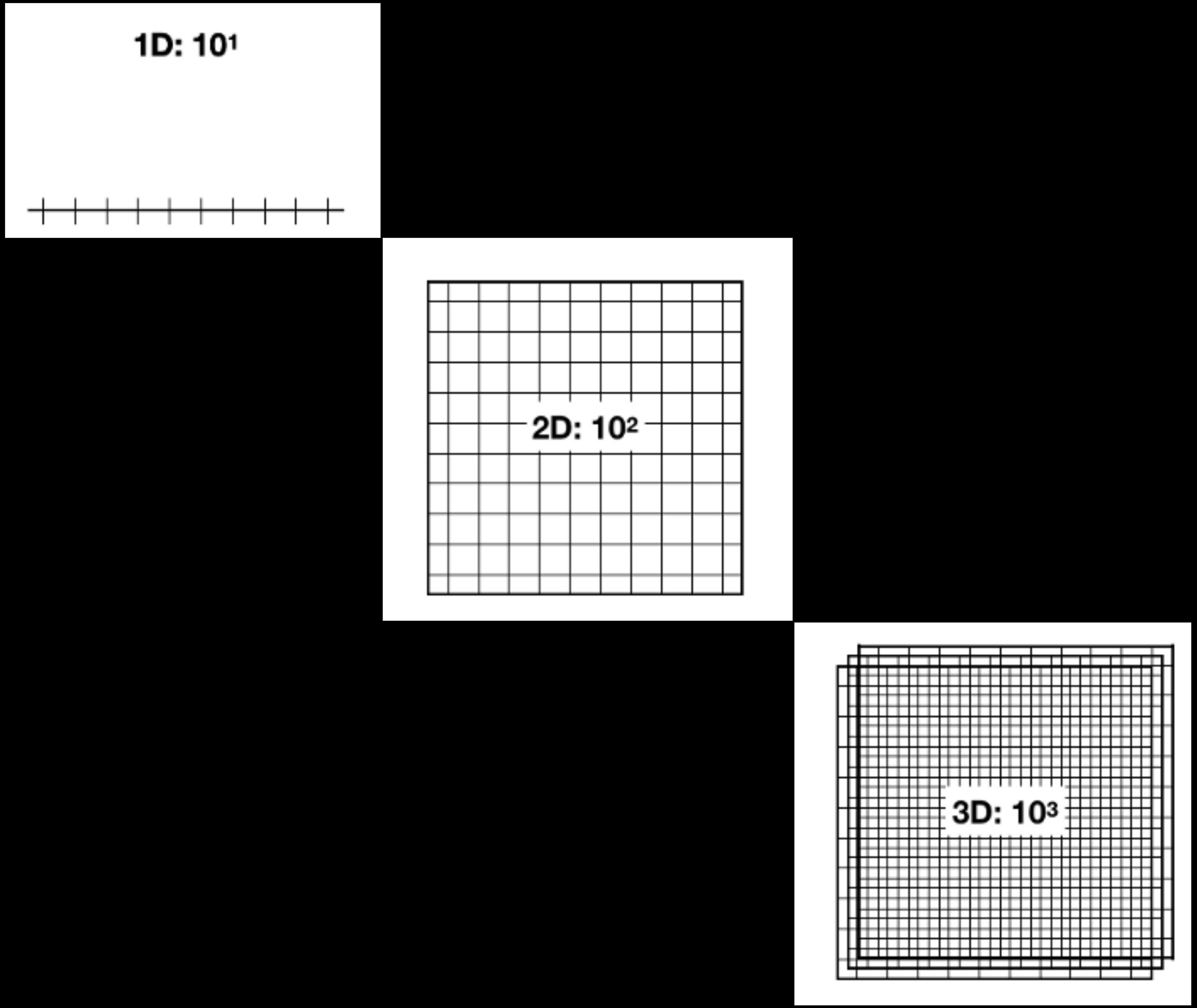
Humanoid balancing Rocket landing



Why not discretize the dynamics and apply value / policy iteration?

THE CURSE OF DIMENSIONALITY

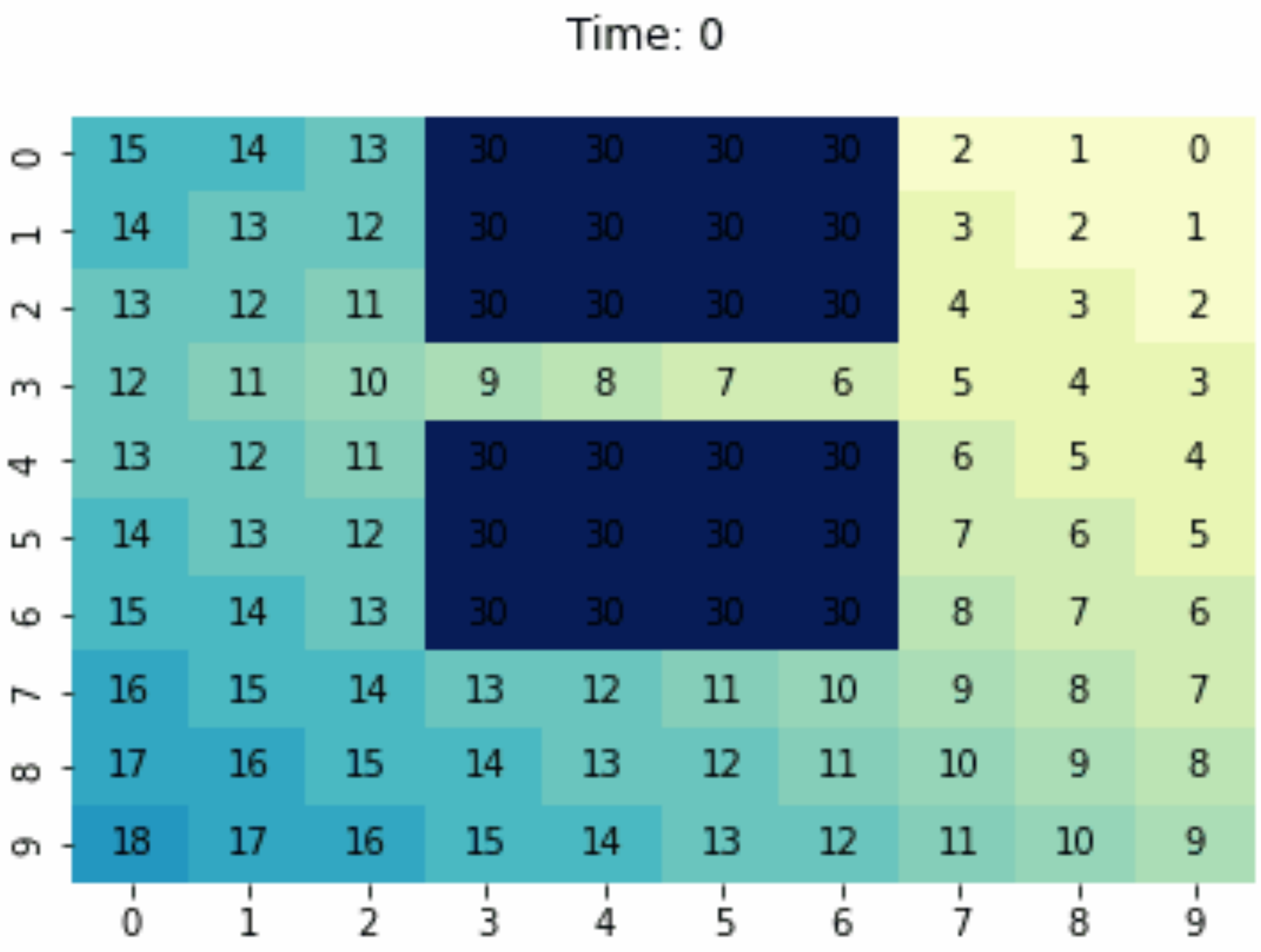




No Discretization!

Can we analytically *represent* and *update* the value function?

$$V^*(s) = \min_a \left[c(s, a) + \gamma \mathbb{E}_{s' \sim \mathcal{T}(s,a)} V^*(s) \right]$$



$$V^*(s) = \min_a \left[c(s,a) + \gamma \mathbb{E}_{s' \sim \mathcal{T}(s,a)} V^*(s)\right]$$

Can represent analytically ...
(piecewise linear?)

But updating seems hard!

Can we analytically represent and *update* the value function?

Yes*

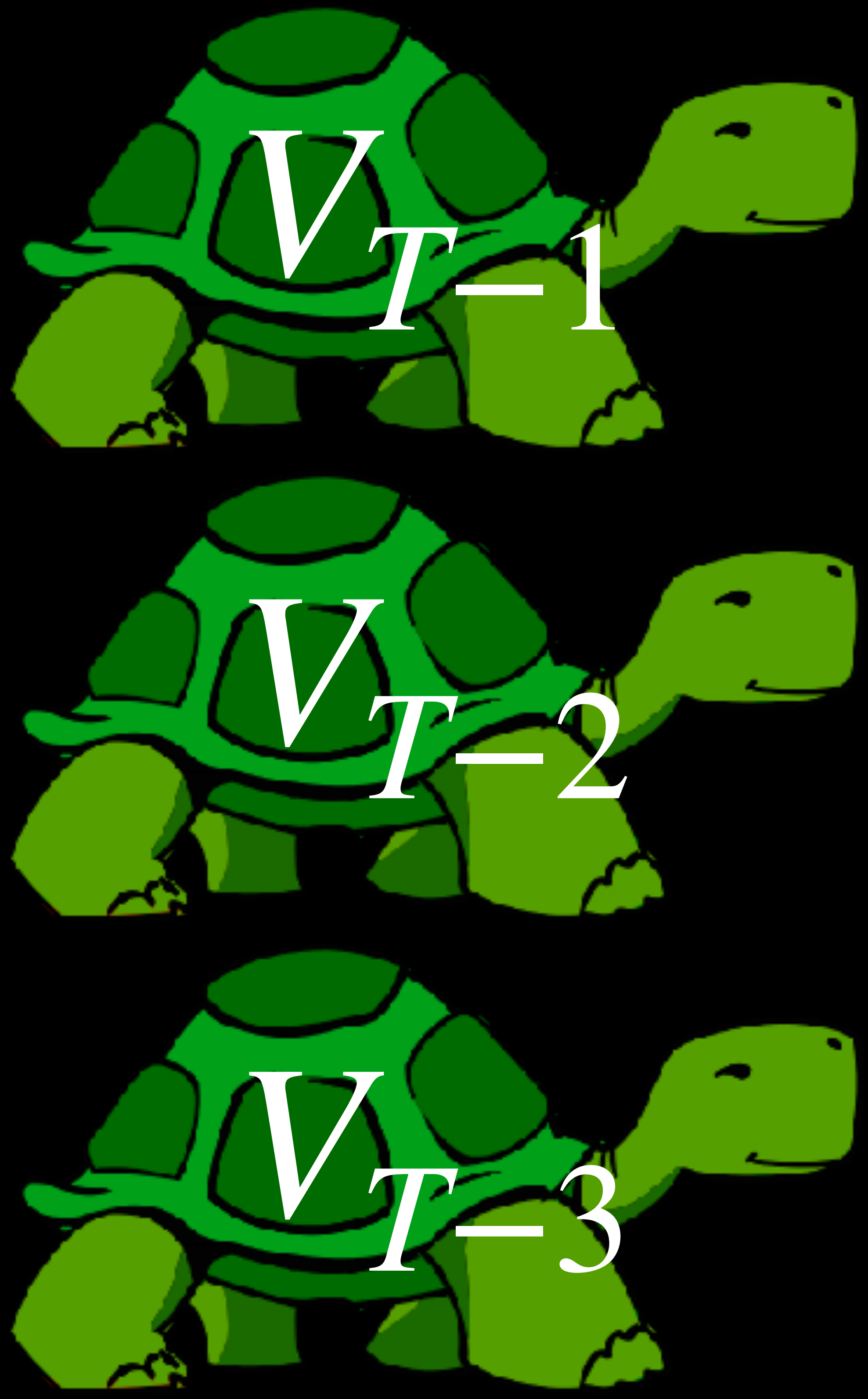
*linear dynamics, quadratic costs

Let’s formalize!

It’s quadratics all the way down!

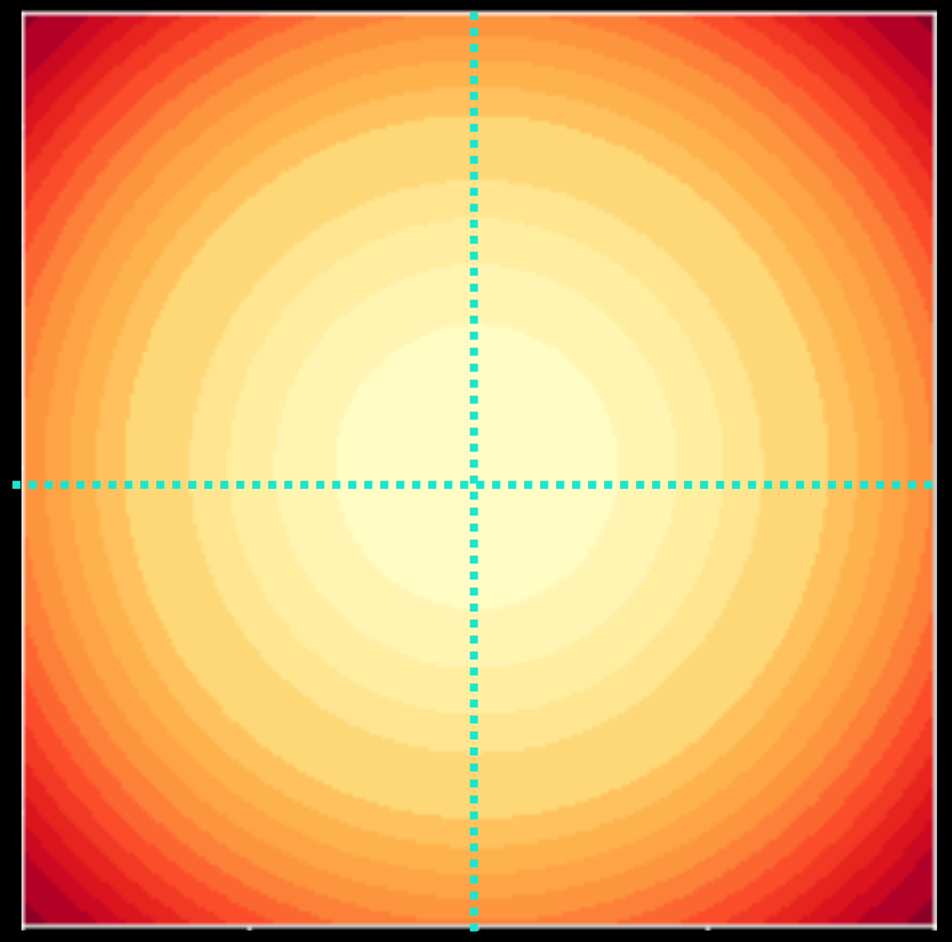
$$K_t = (R + B^T V_{t+1} B)^{-1} B^T V_{t+1} A$$

$$V_t = Q + K_t^T R K_t + (A + BK_t)^T V_{t+1} (A + BK_t)$$



The LQR Algorithm

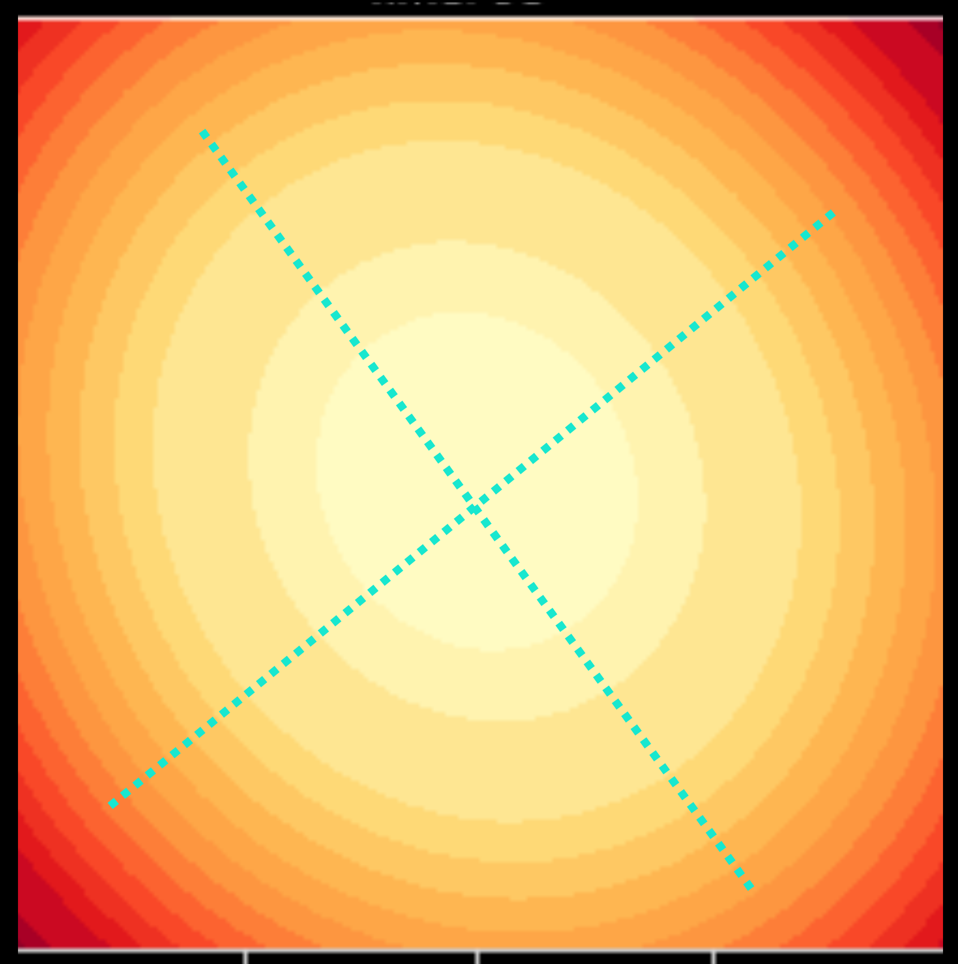
Initialize $V_T = Q$



For t = T … 1

Compute gain matrix

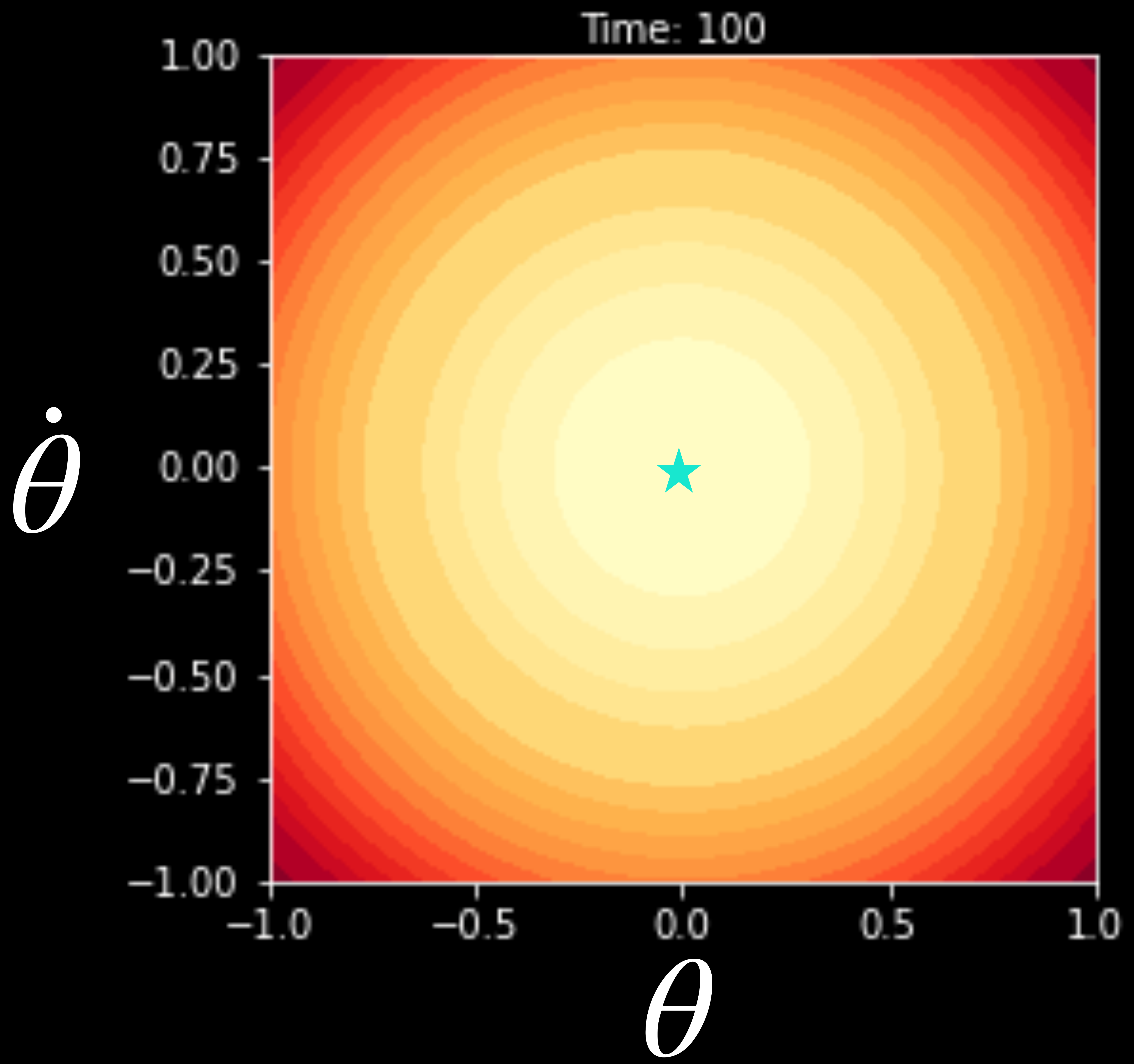
$$K_t = (R + B^T V_{t+1} B)^{-1} B^T V_{t+1} A$$



Update value

$$V_t = Q + K_t^T R K_t + (A + BK_t)^T V_{t+1} (A + BK_t)$$

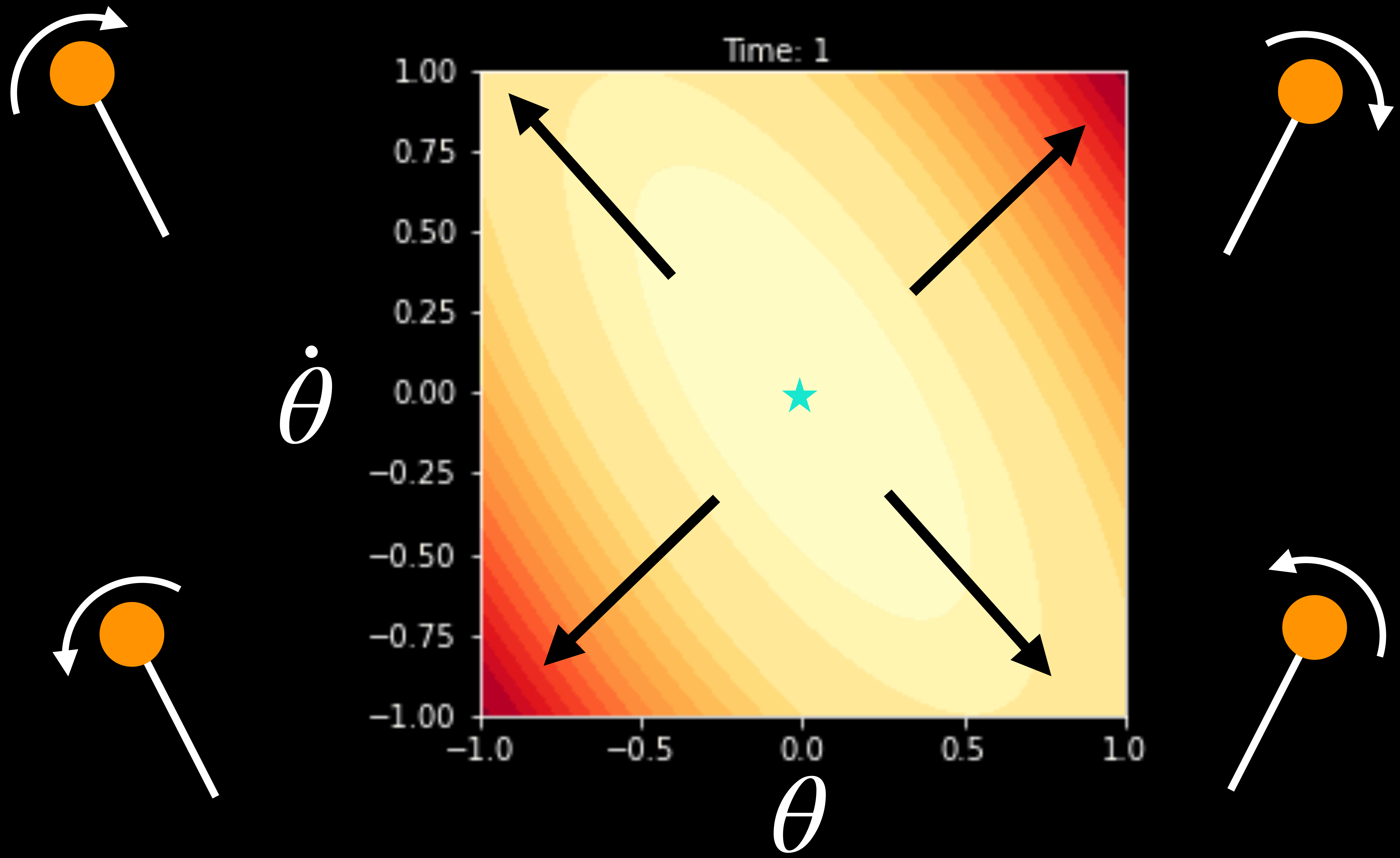
Value Iteration for Inverted Pendulum



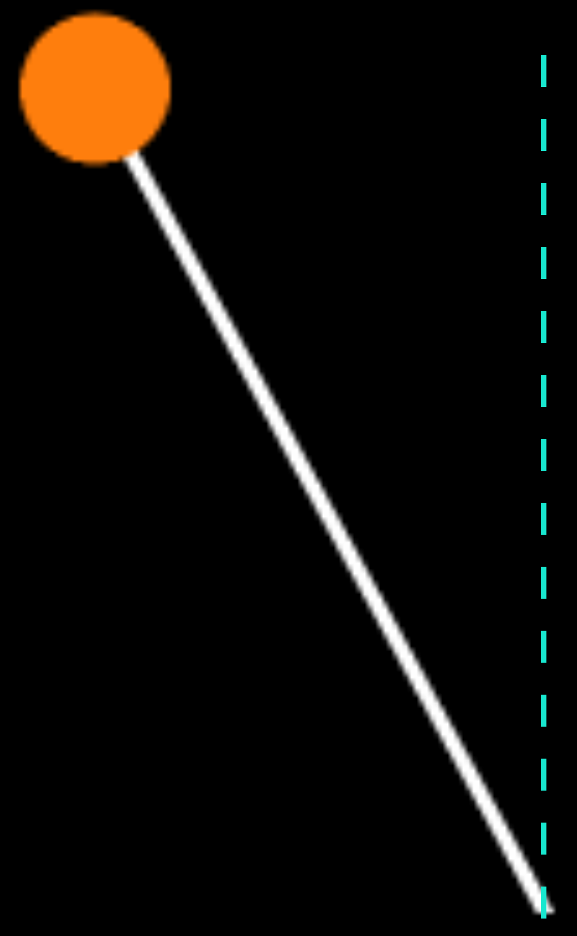
Value converges when system is stabilizable

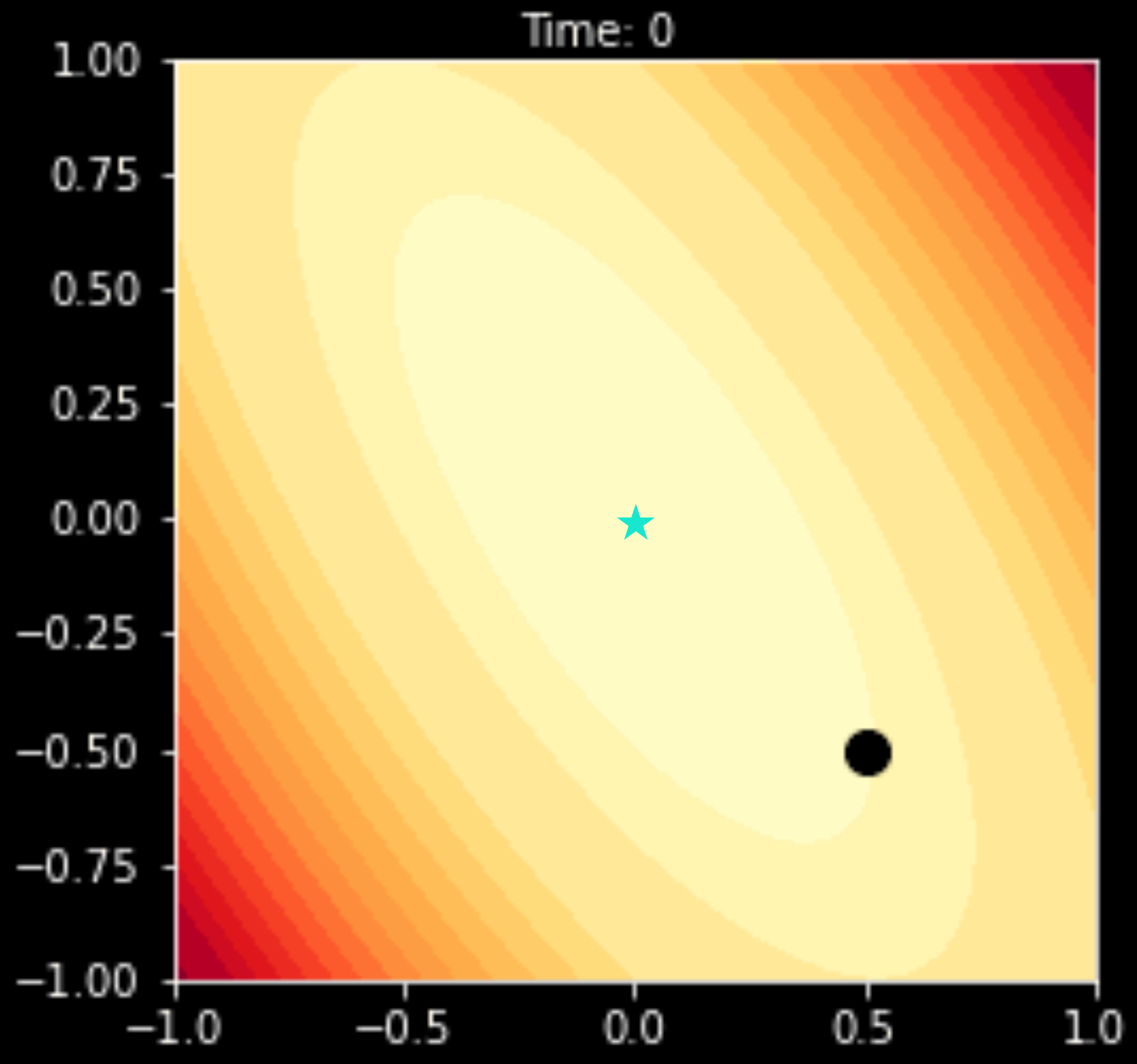
Can solve Ricatti equations for fixed point

Value Iteration for Inverted Pendulum

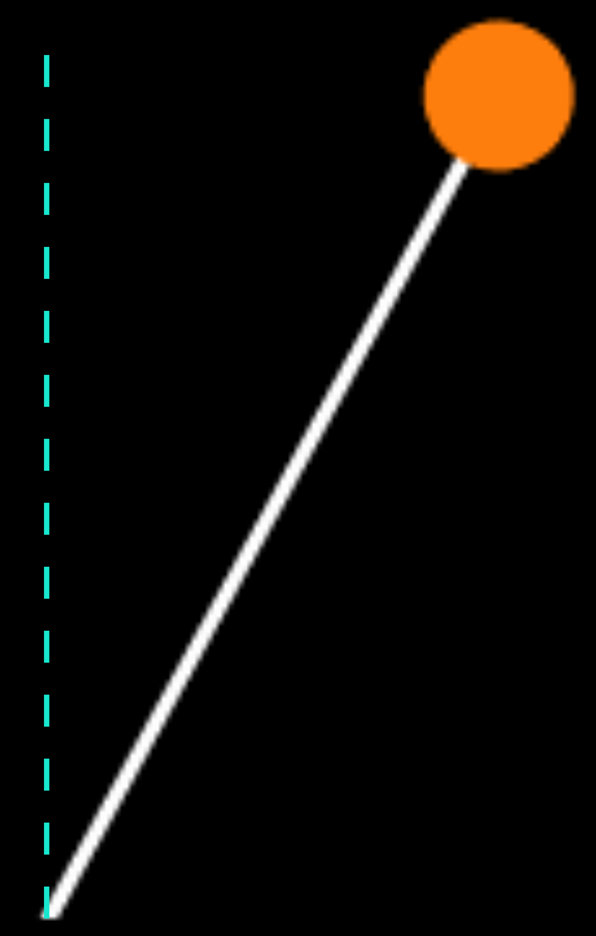


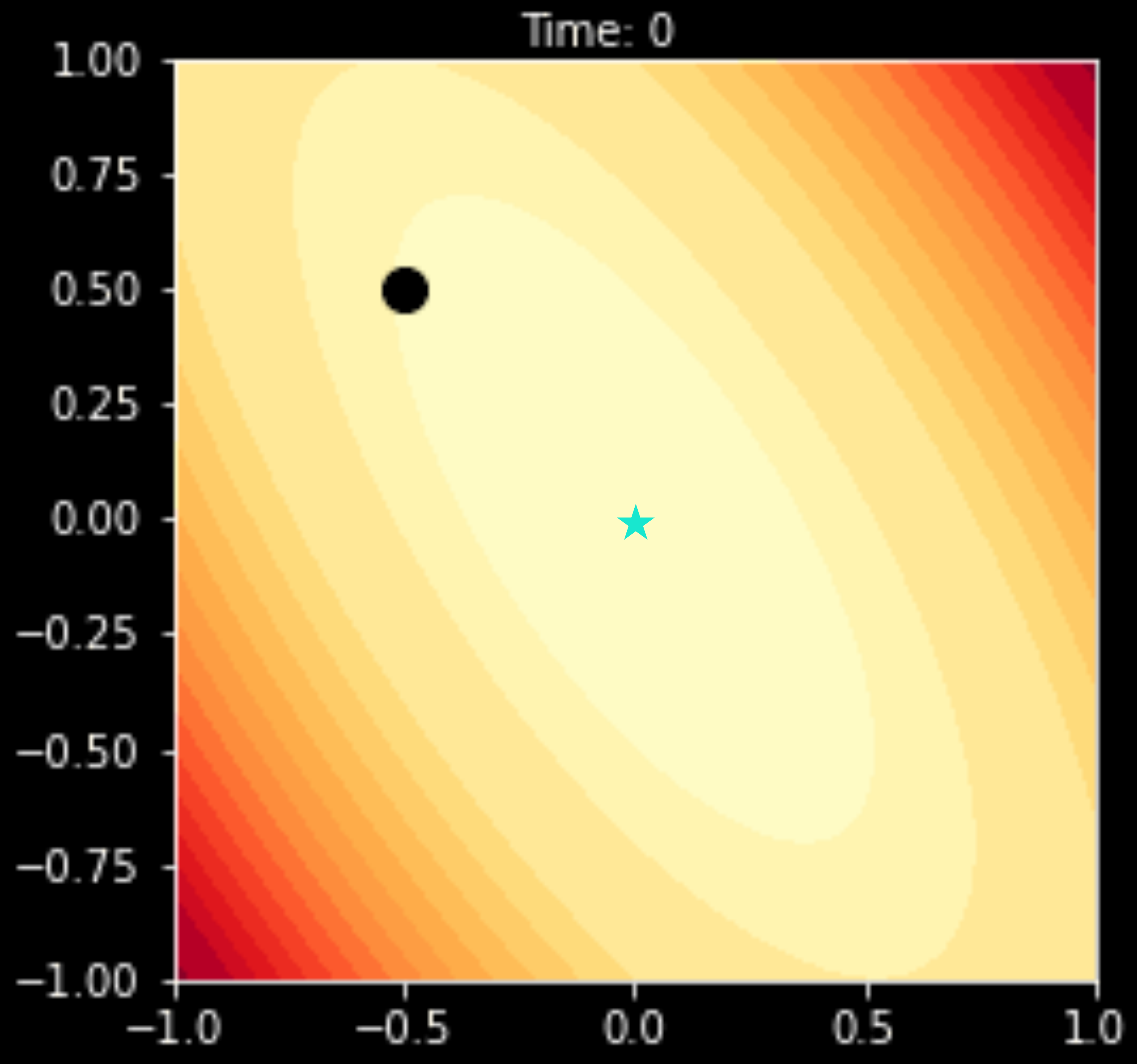
An Easy Starting Point



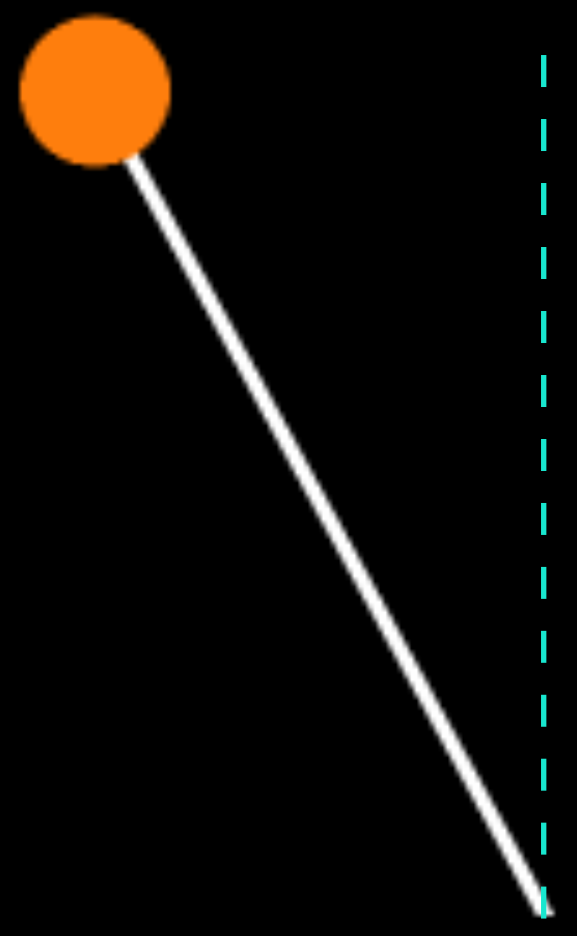


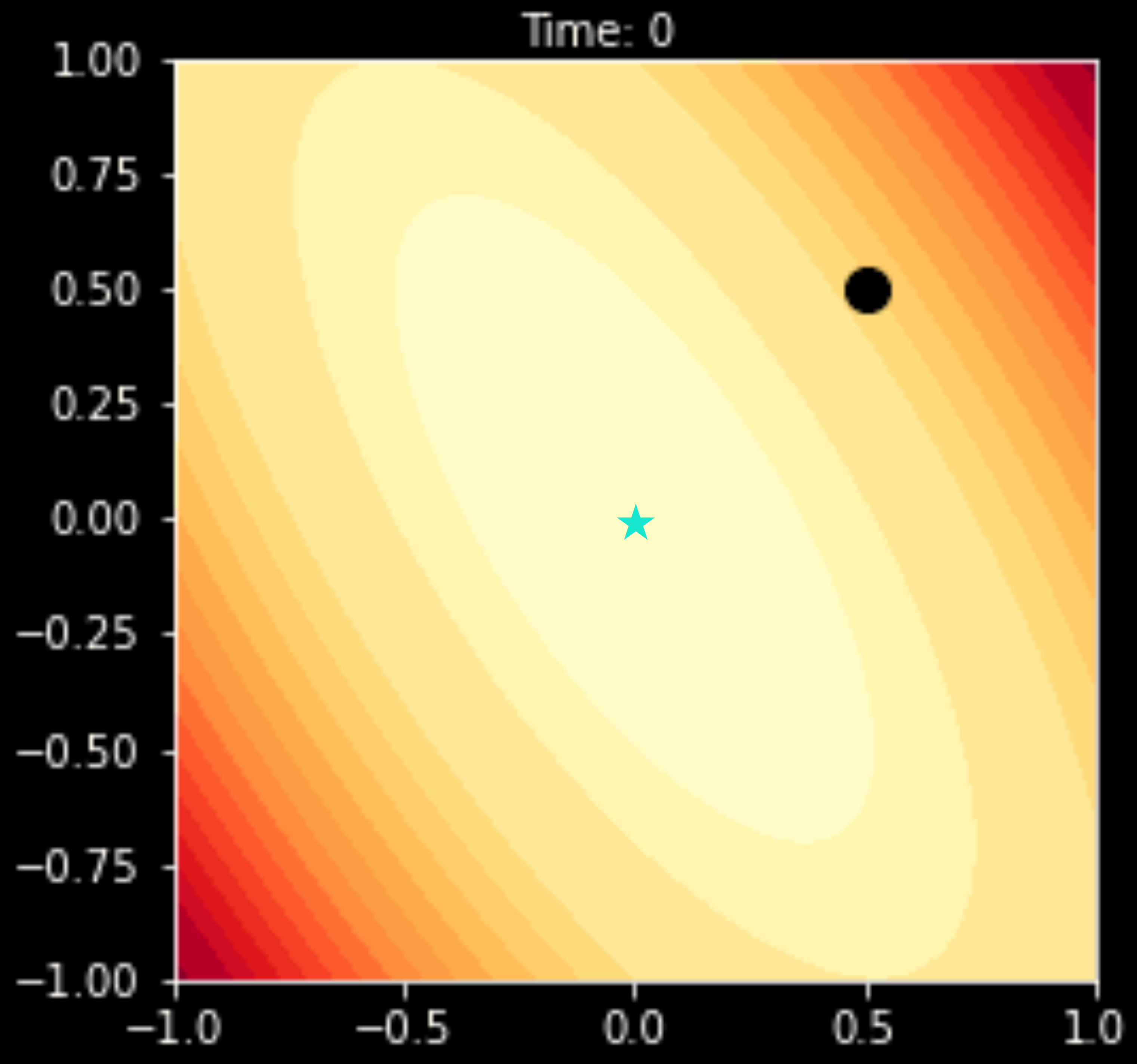
Another Easy Starting Point



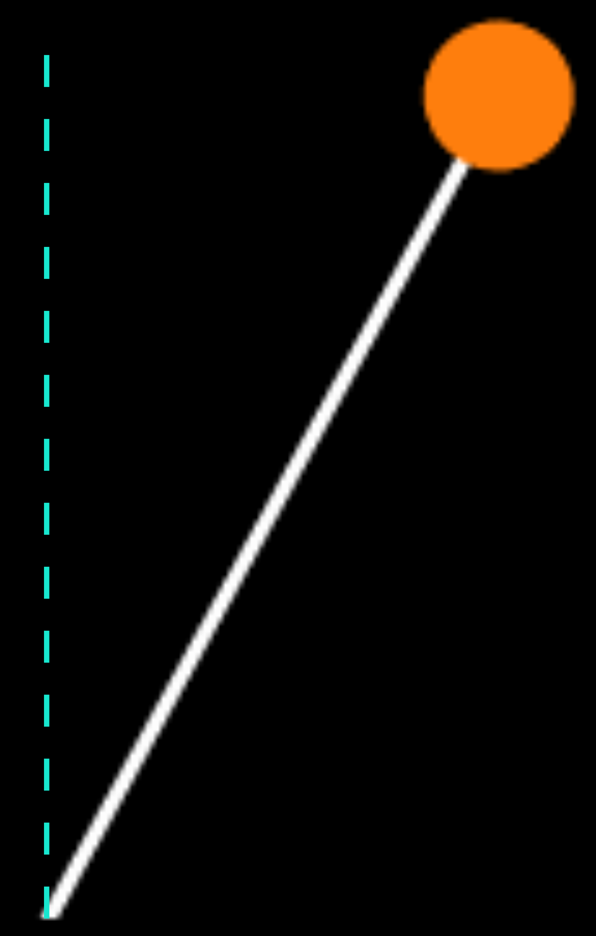


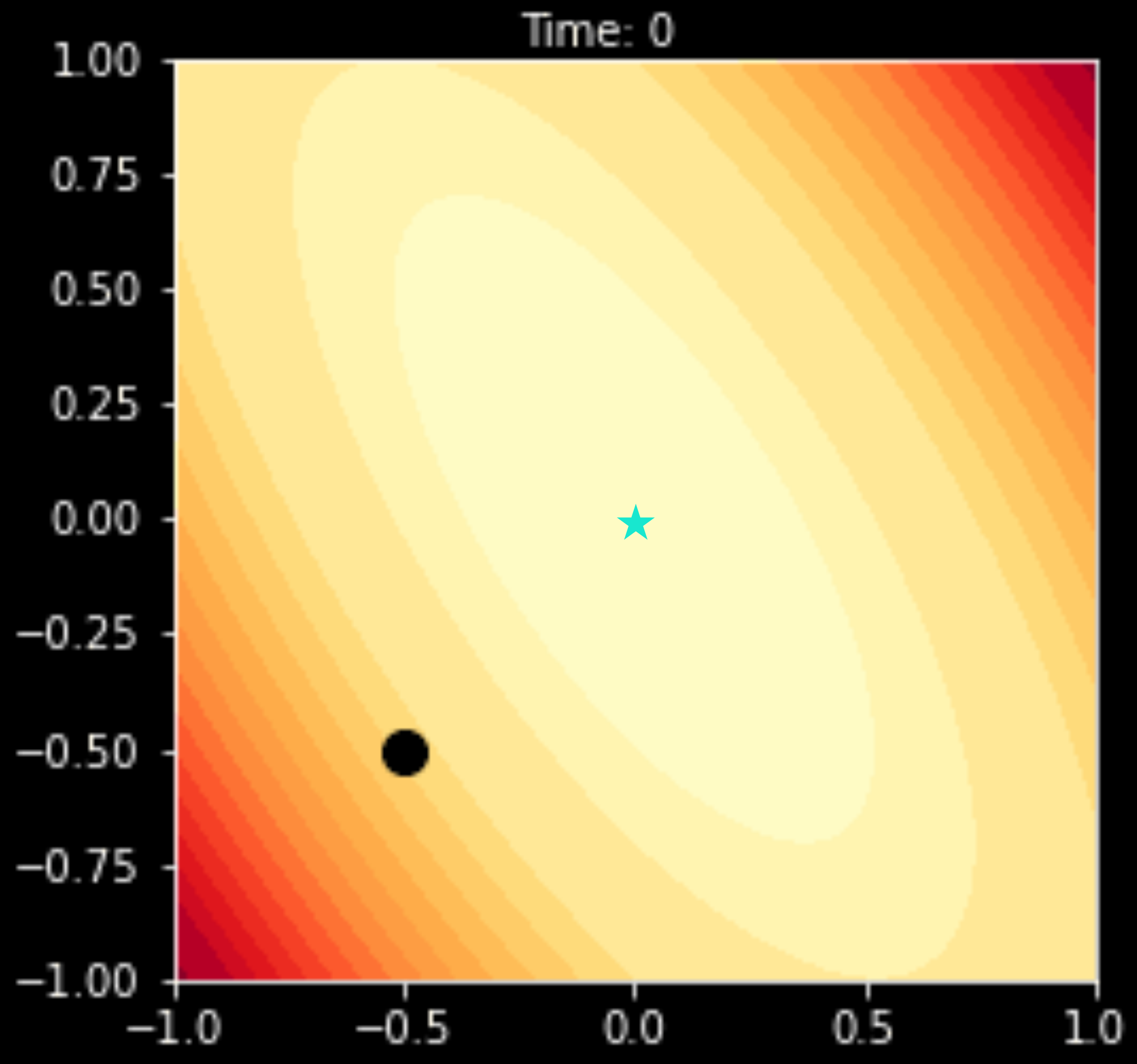
A Hard Starting Point





Another Hard Starting Point





When does LQR converge?

$$V = Q + K^T R K + (A + BK)^T V (A + BK)$$

$$K = (R + B^T V B)^{-1} B^T V A$$

When the closed loop system is stable, i.e.

Eigen values of (A+BK) are inside the unit circle on the complex plane

When does LQR converge?

$$V = Q + K^T R K + (A + BK)^T V (A + BK)$$

$$K = (R + B^T V B)^{-1} B^T V A$$

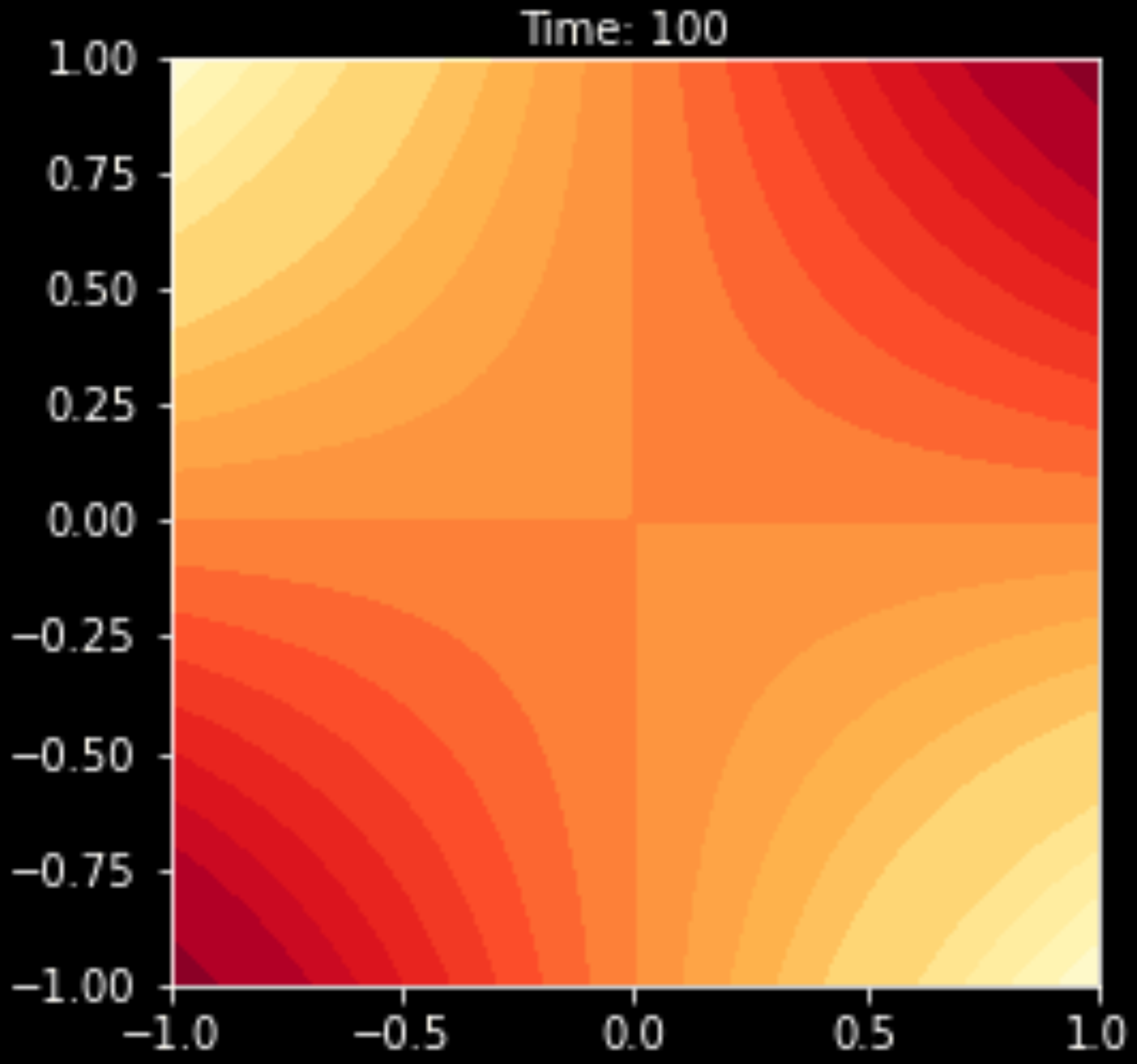
When the closed loop system is stable, i.e.
Eigen values of (A+BK) are inside the unit circle on the complex plane

How can we find the fixed point of this equation?

Discrete time algebraic ricatti equation (DARE)

https://en.wikipedia.org/wiki/Algebraic_Riccati_equation

What if Q is not PSD?



$$x^T Q x \not\geq 0$$

$$Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

What if R is not positive definite?

$$u^T R u \not> 0$$

$$R = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Hint: Gain matrix update?

$$K_t = (R + B^T V_{t+1} B)^{-1} B^T V_{t+1} A$$

What about handling uncertainty?

Gaussian noise in dynamics ?

$$x_{t+1} \sim \mathcal{N}(Ax_t + Bu_t, \Sigma)$$

Some Trivia!

A productive year from Kalman

In 1960 three major papers were published by R. Kalman and coworkers...

1. One of these [Kalman and Bertram 1960], publicized the vital work of Lyapunov in the time-domain control of nonlinear systems.

2. The next [Kalman 1960a] discussed the optimal control of systems, providing the design equations for the linear quadratic regulator (LQR).

3. The third paper [Kalman 1960b] discussed optimal filtering and estimation theory, providing the design equations for the discrete Kalman filter.

(from http://www.uta.edu/utari/acs/history.htm)

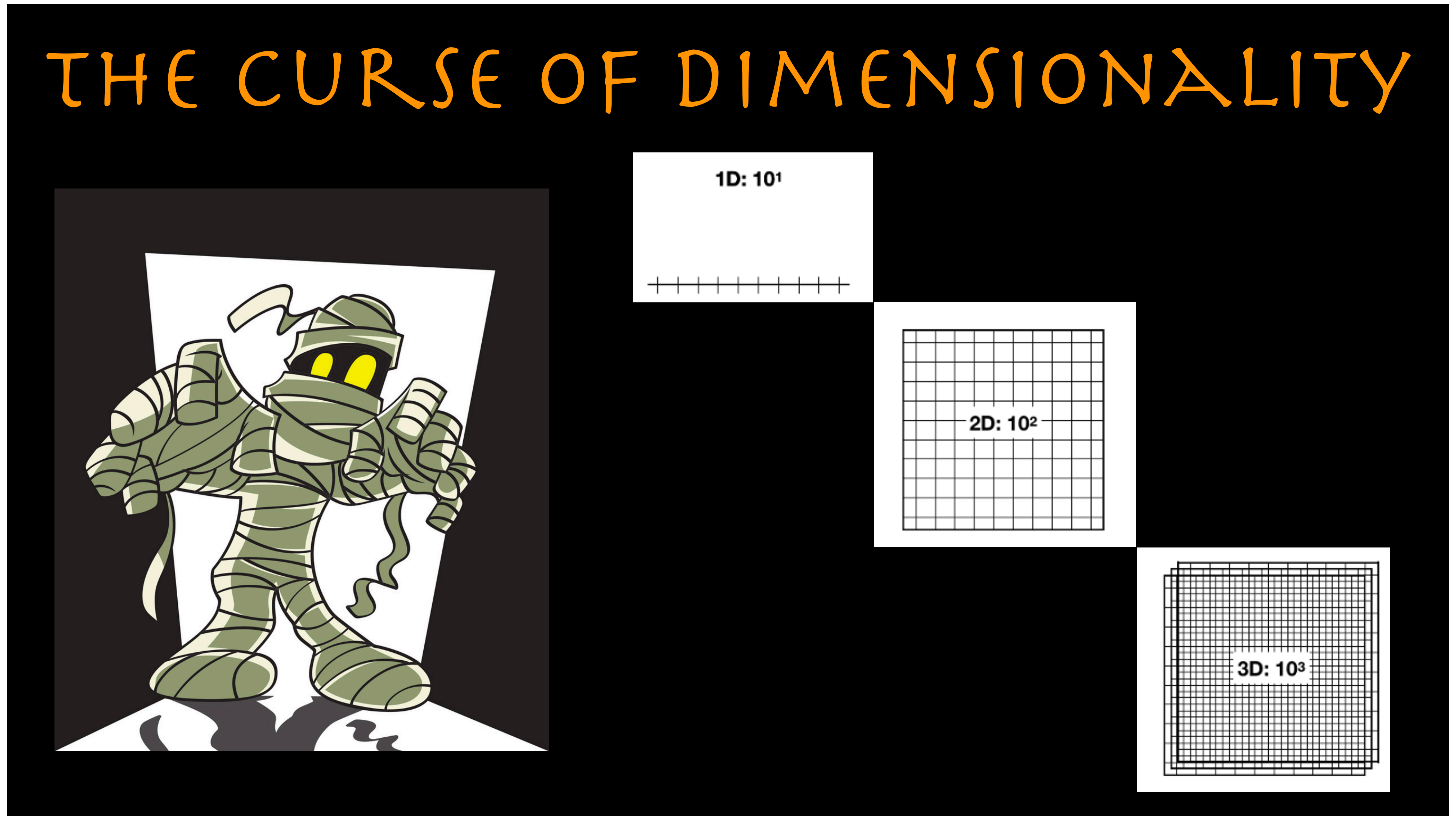
Trivia: Duality between control and estimation

R. Kalman "A new approach to linear filtering and prediction problems." (1960)

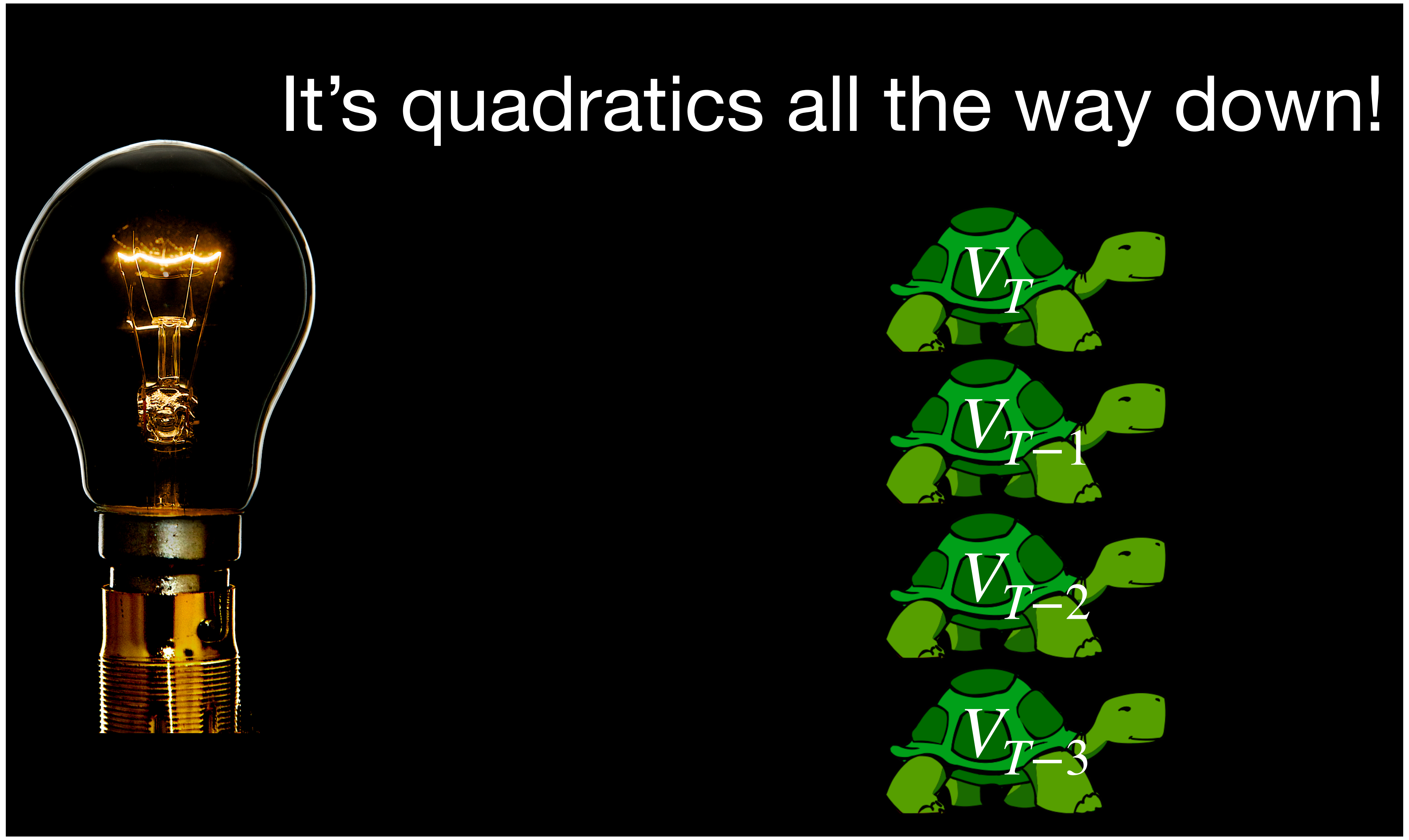
linear-quadratic regulator	**Kalman-Bucy filter**	
V	Σ	(state variance)
A	A^{T}	(dynamics)
B	H^{T}	(dynamics noise)
R	DD^{T}	(measurement)
Q	CC^{T}	(motion noise)
t	$t_f - t$	

(Table from E.Todorov "General duality between optimal control and estimation", CDC, 2008)

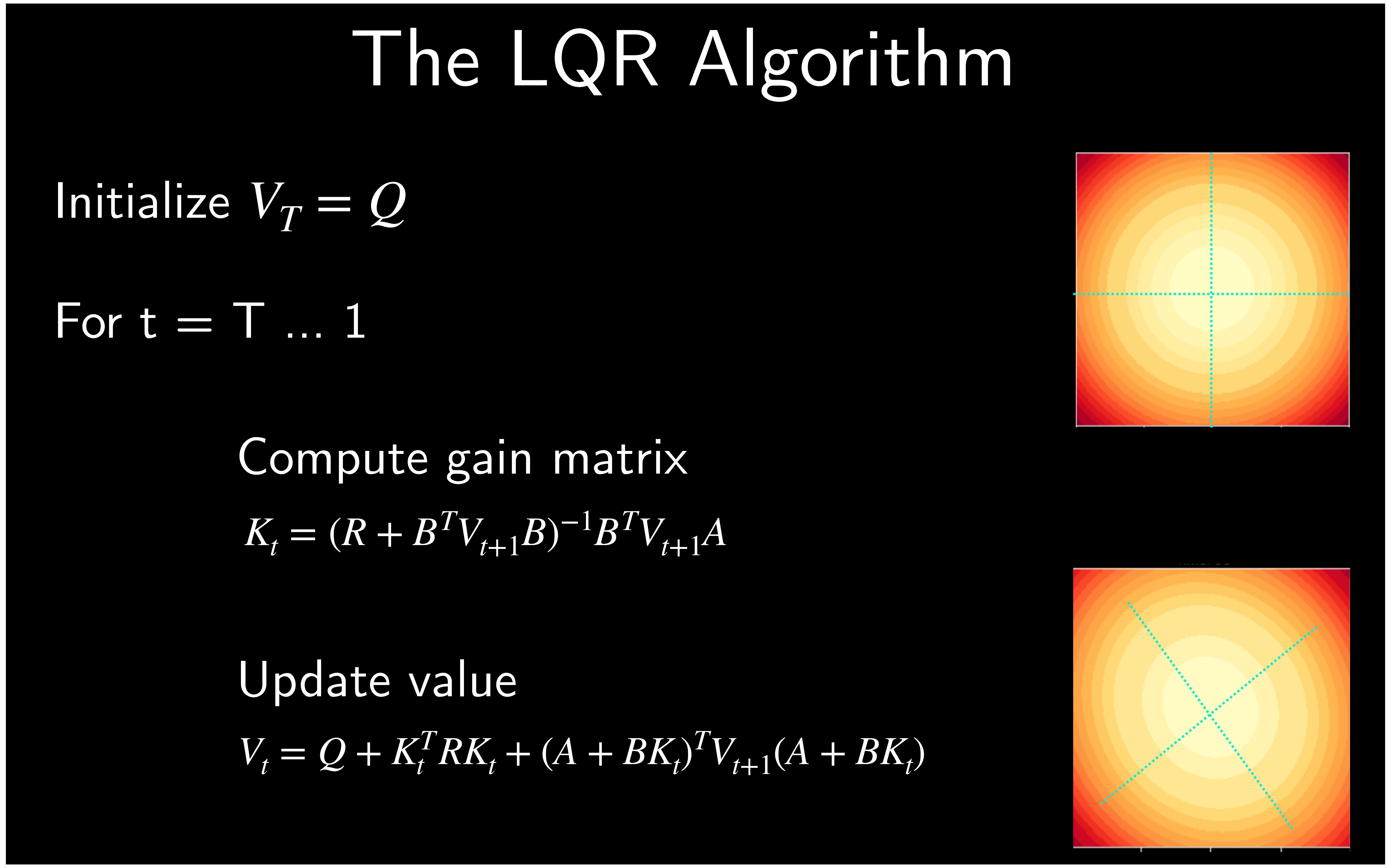
tl;dr



THE CURSE OF DIMENSIONALITY
1D: 10¹
2D: 10²
3D: 10³



It's quadratics all the way down!
V_T
V_{T-1}
V_{T-2}
V_{T-3}



The LQR Algorithm
Initialize $V_T = Q$
For t = T … 1
Compute gain matrix
$K_t = (R + B^T V_{t+1} B)^{-1} B^T V_{t+1} A$
Update value
$V_t = Q + K_t^T R K_t + (A + BK_t)^T V_{t+1} (A + BK_t)$