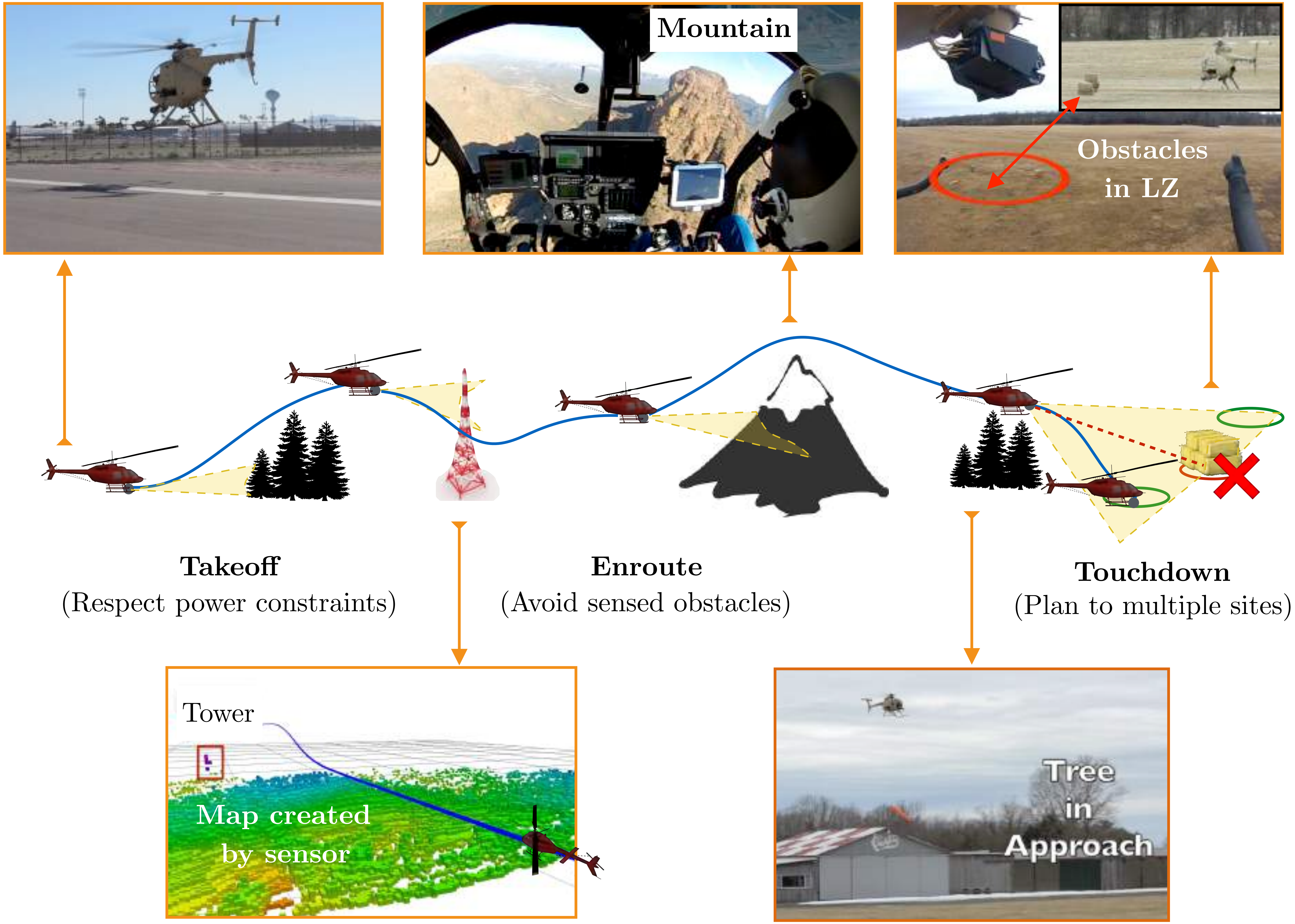
Hard MDPs and how to solve them

Sanjiban Choudhury

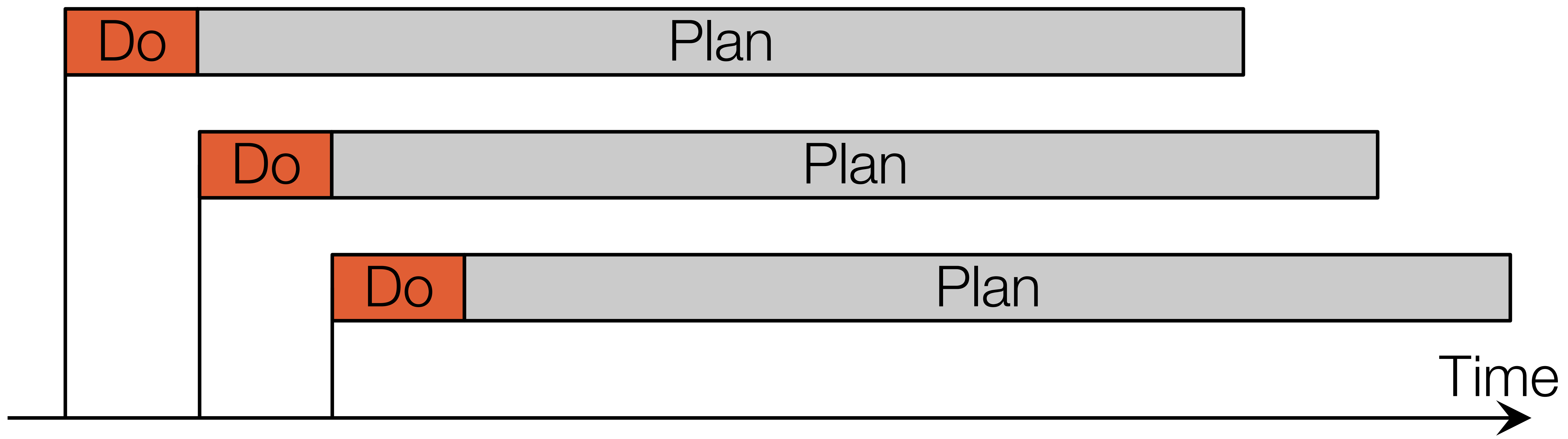
Cornell Bowers CIS Computer Science


Nirvana!
POMDP
Non-convex /
Non-differentiable
Constraints
Long-Horizons

Long Horizons


Mountain
Obstacles
in LZ
Takeoff
(Respect power constraints)
Enroute
(Avoid sensed obstacles)
Touchdown
(Plan to multiple sites)
Tower
Map created
by sensor
Tree
in
Approach

Receding Horizon Control (also called MPC!)



Step 1: Solve optimization problem to a horizon

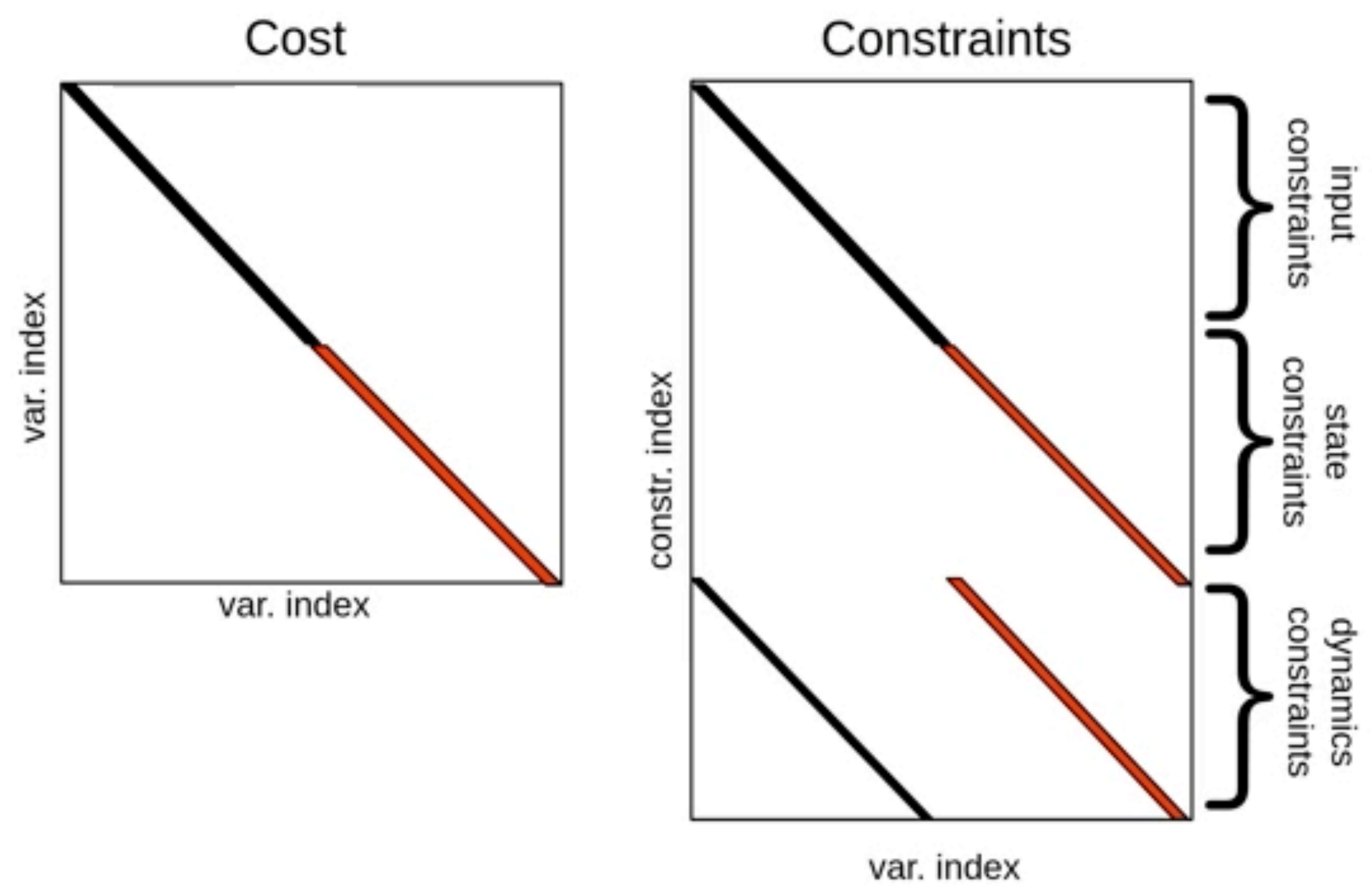
Step 2: Execute the first control

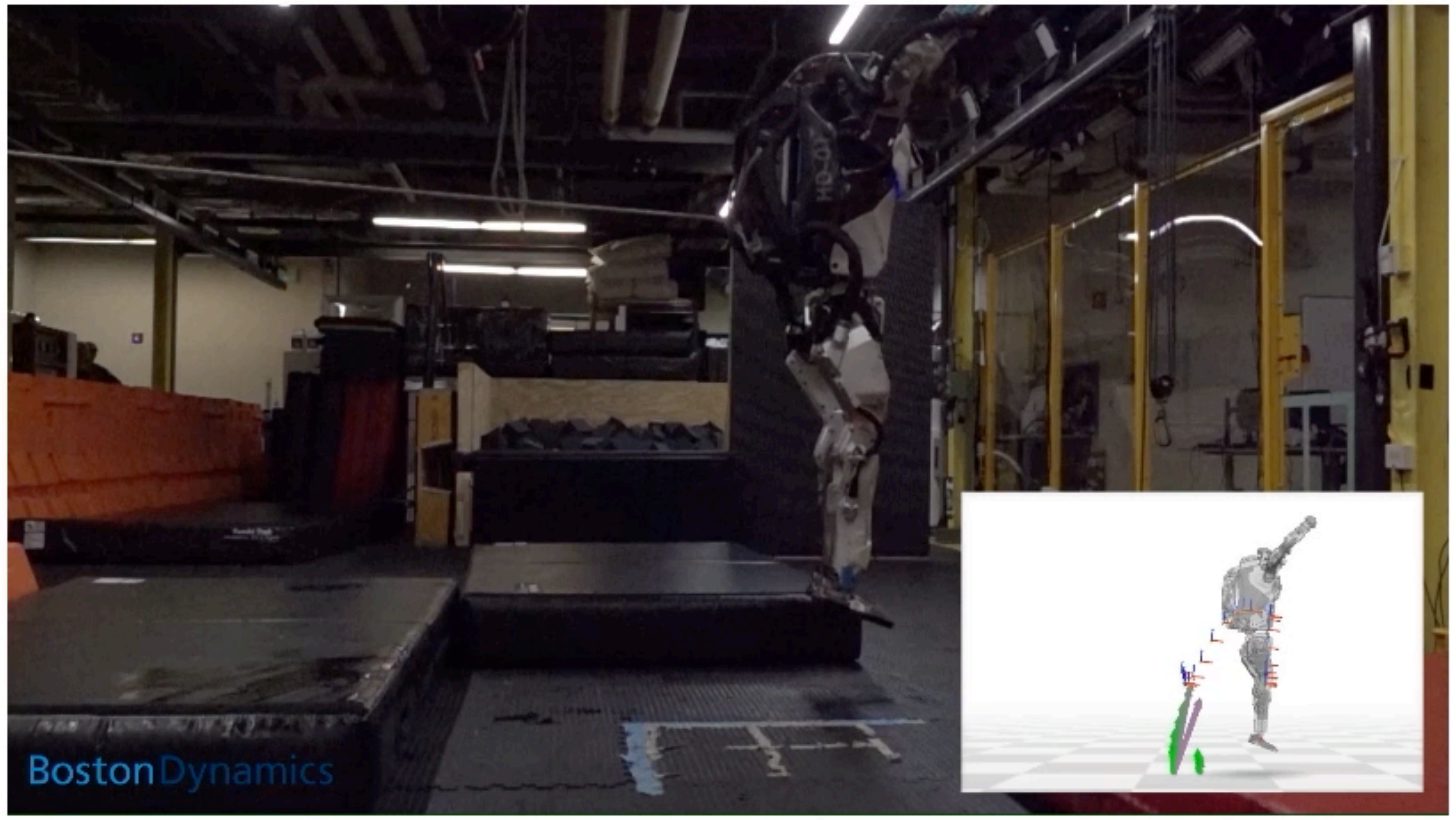
Step 3: Repeat!

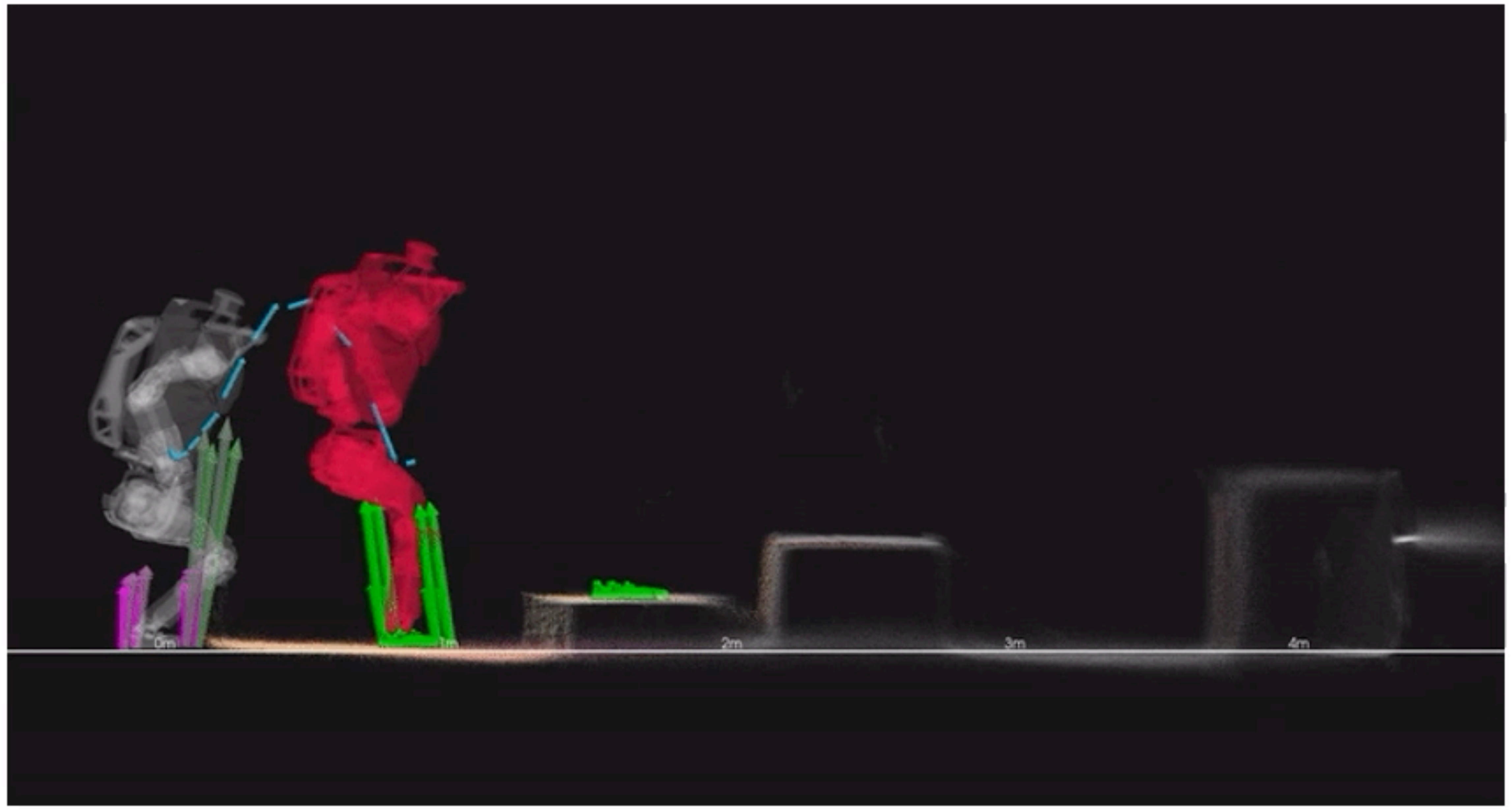
Constraints

Model-Predictive Control

- Continuously optimizes trajectory subject to nonlinear momentum dynamics
- Solve for future kinematic configurations
- Leverages optimized code and problem structure for speed







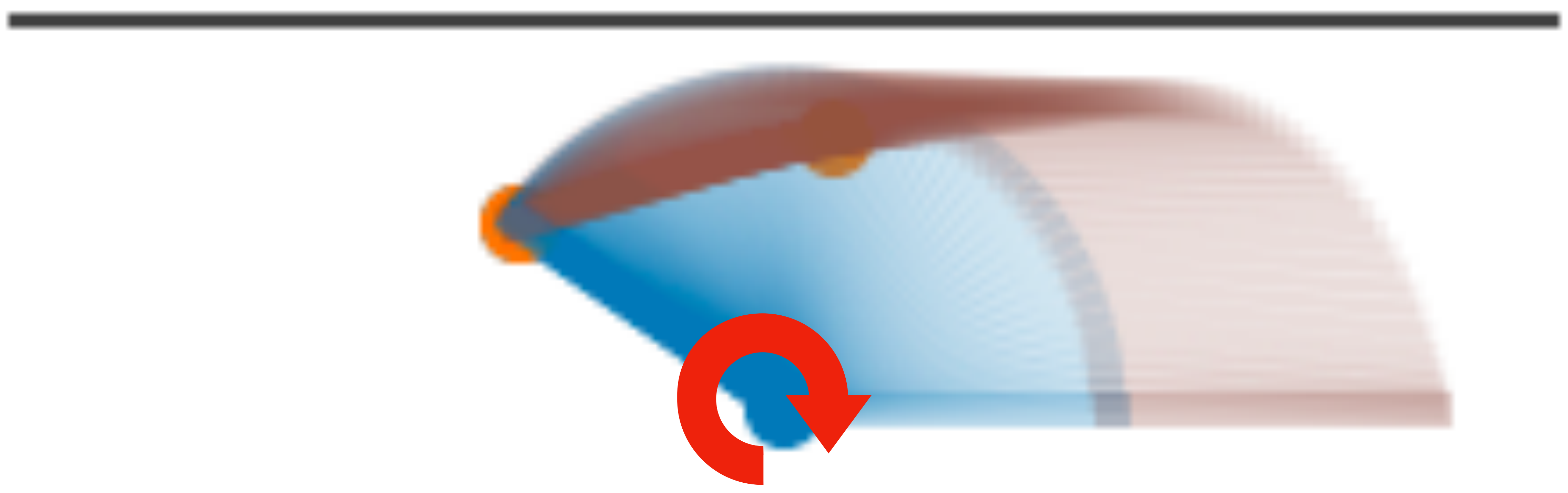
Activity!

Brainstorm!

We want to move our n-link manipulator from A to B but satisfy two constraints

#1: Don't exceed torque limit

#2: Don't hit wall

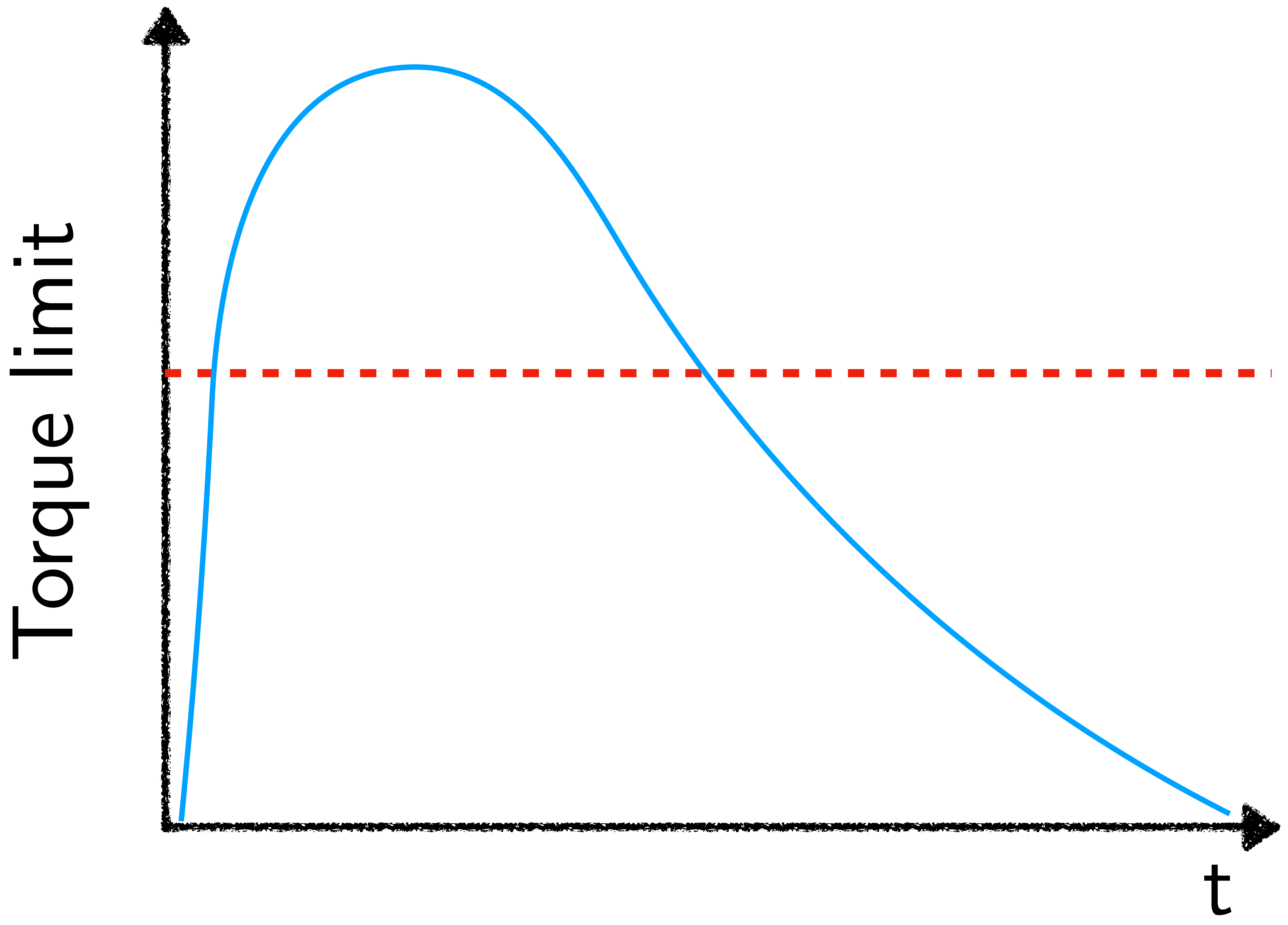


How do we hack iLQR to solve #1? #2?

Re-parameterization:

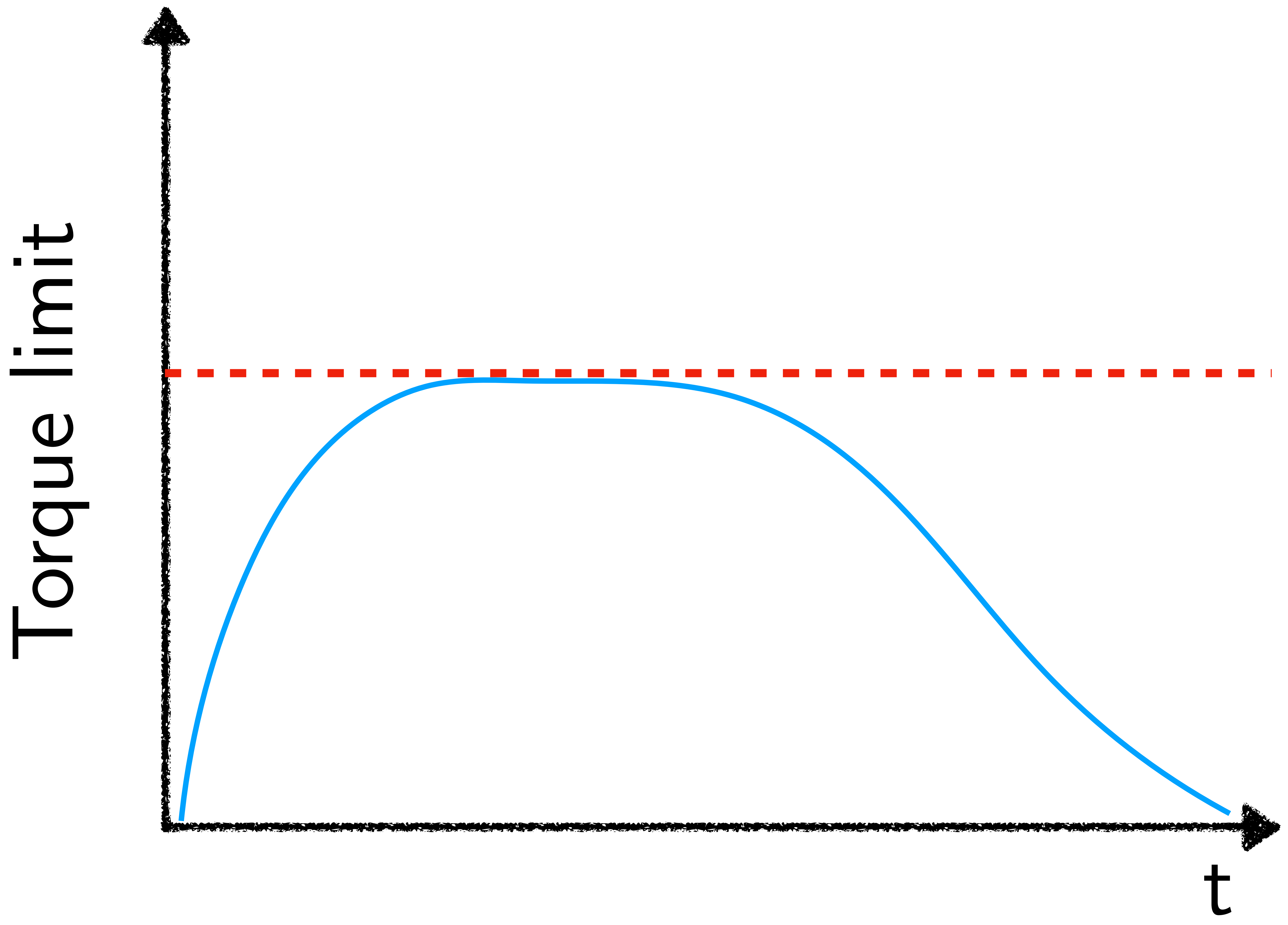
The quick 'n' easy way to solve constraints!

Example: Swing up using iLQR



How do we enforce a torque limit?

$$\tau_{min} \leq \tau \leq \tau_{max}$$



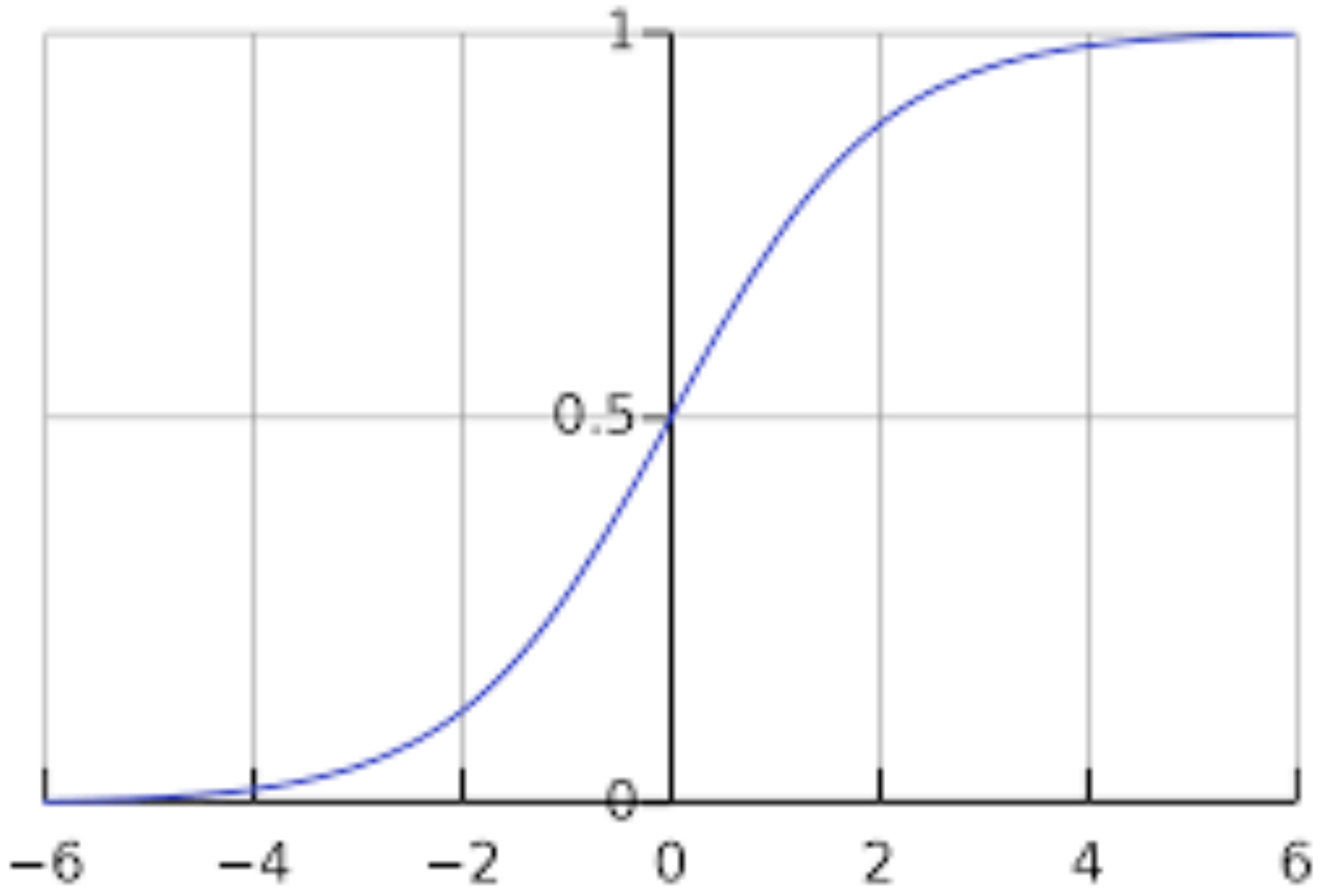
Idea: Reformulate the variables so the constraint must be satisfied

$$\tau_{min} \leq \tau \leq \tau_{max}$$

τ_{max}

τ

τ_{min}

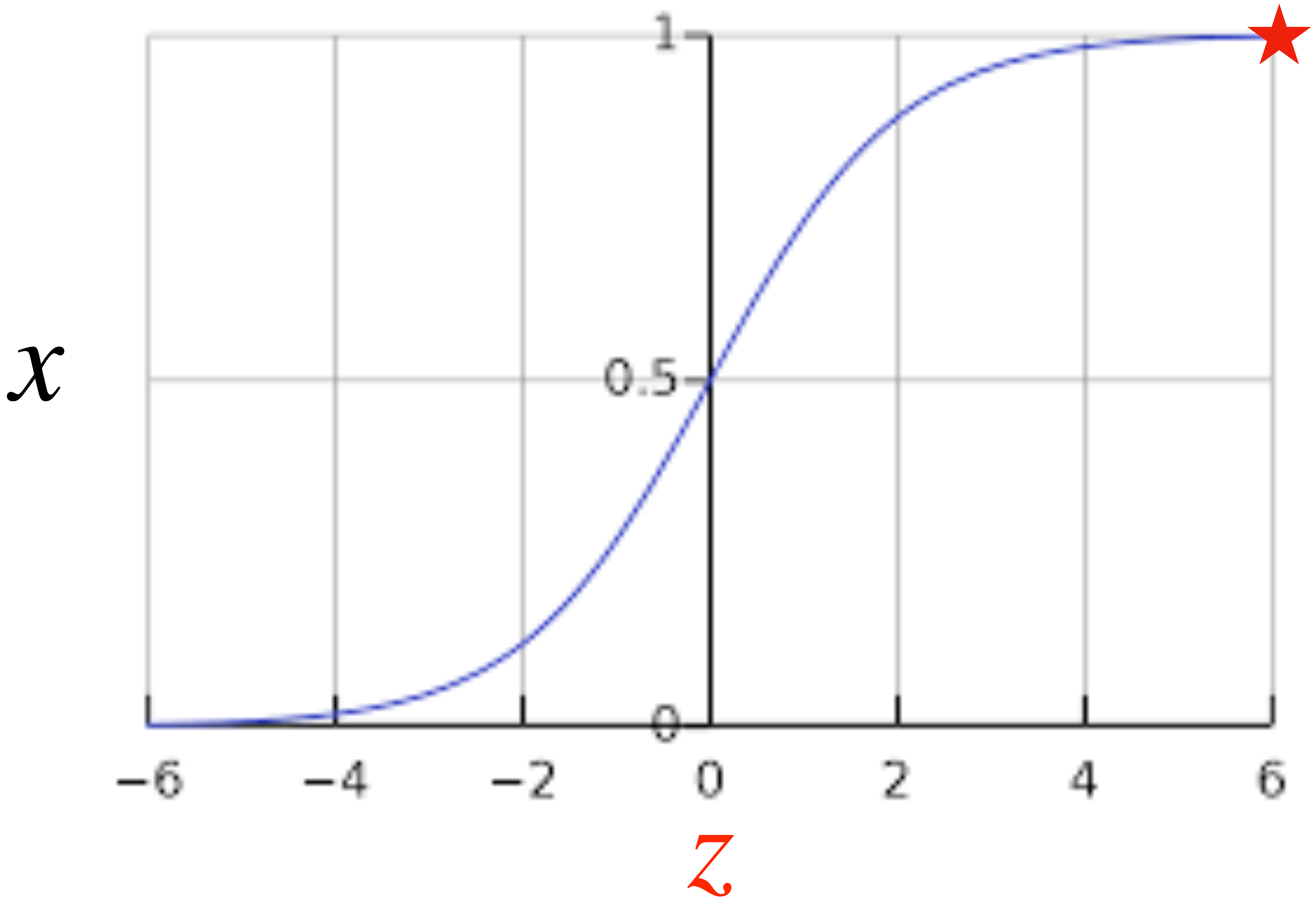


$-\infty$ z ∞

$$\tau = \text{Sigmoid}(z, \tau_{min}, \tau_{max})$$

… when does re-parameterization fail?

Failure: Stuck on the far side of the sigmoid

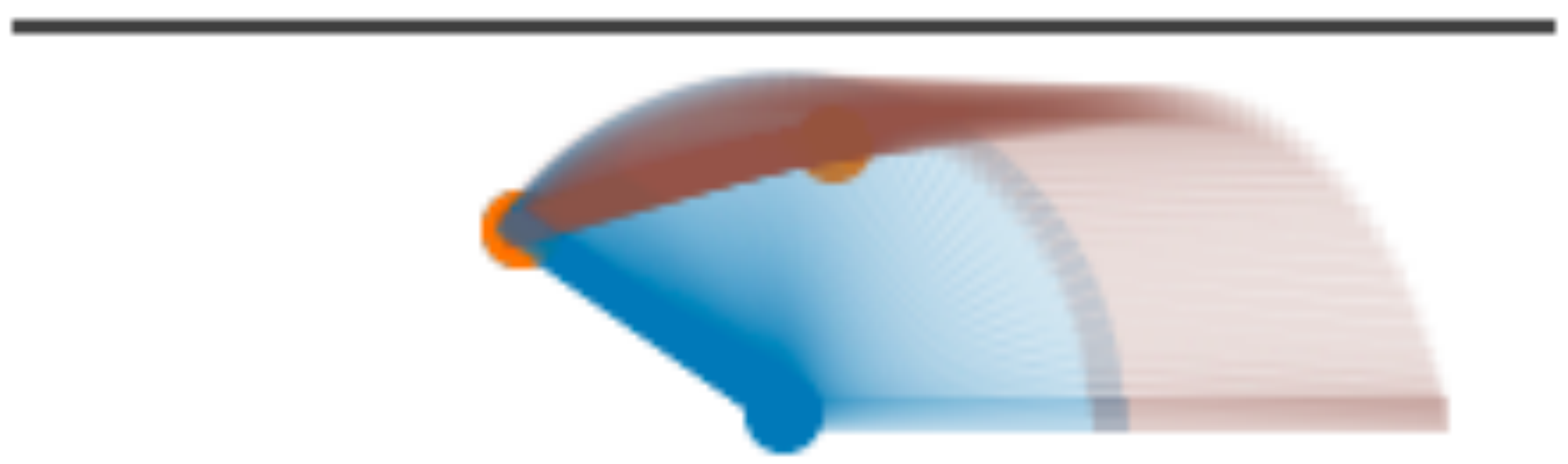


Let's say z is very high

What is $\dfrac{\partial x}{\partial z}$?

Failure 2: Constraints too complex to re-parameterize

Don't hit wall

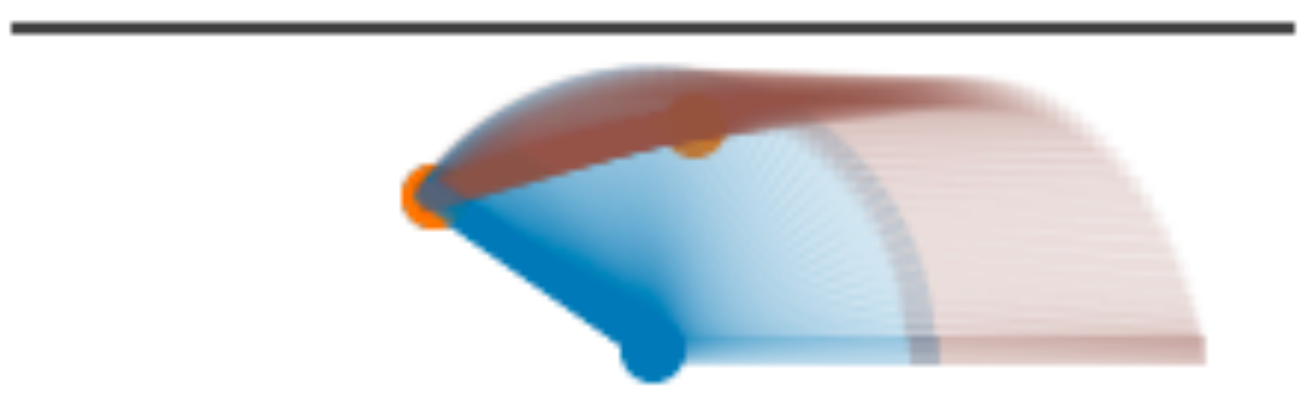


How do we handle more complex constraints?

$$\min_x f(x)$$

$$g(x) = 0$$

$$h(x) \leq 0$$



Hang on ....

Why not put a really really really high cost for violating constraints?

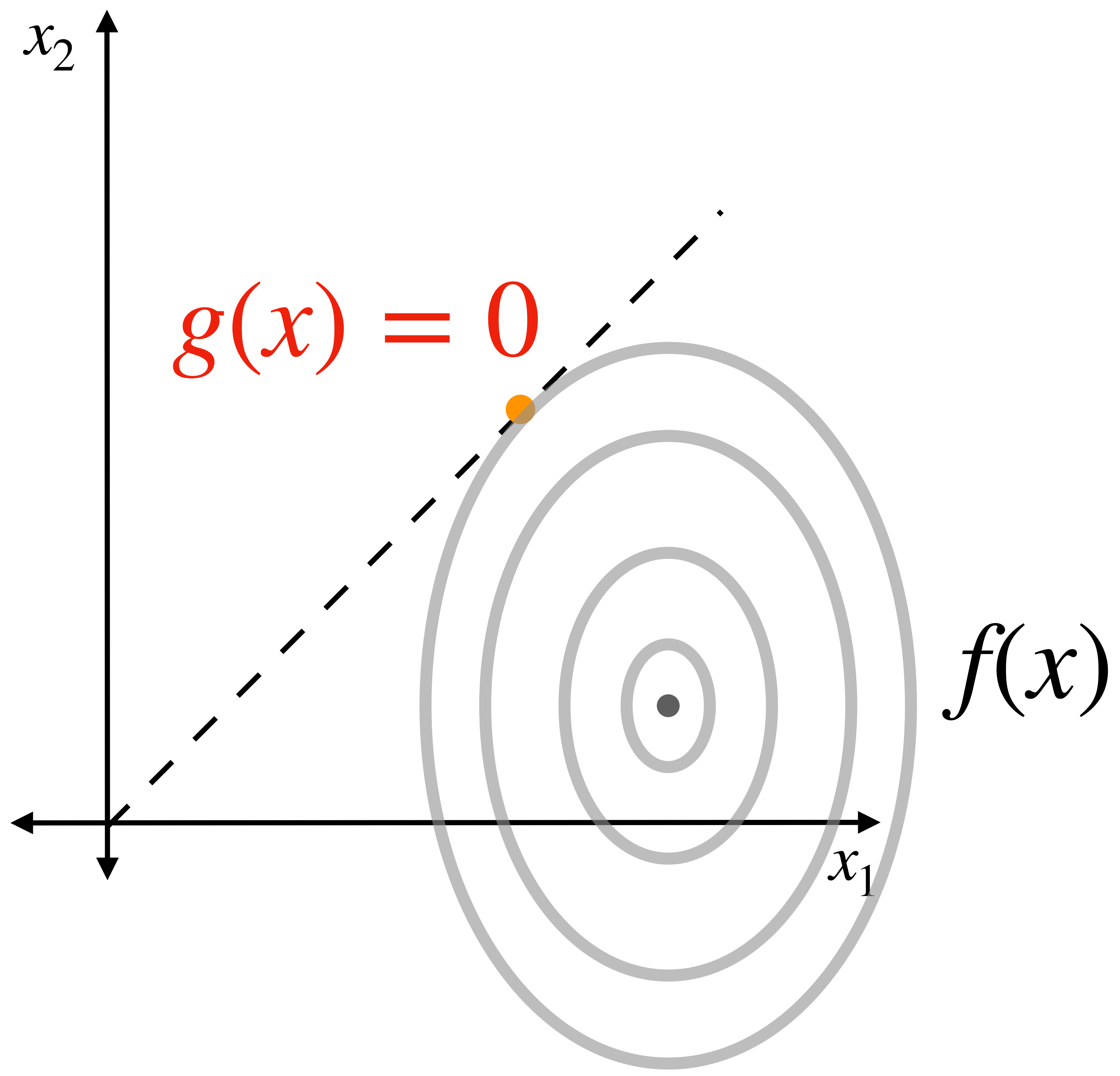
Penalty method

$$\min_x \quad f(x)$$
$$g(x) = 0$$

$$\min_x f(x) + \frac{\alpha}{2} g(x)^2$$

Seems easy to implement ... what could possibly go wrong?

What would be the gradient at the optimal value?

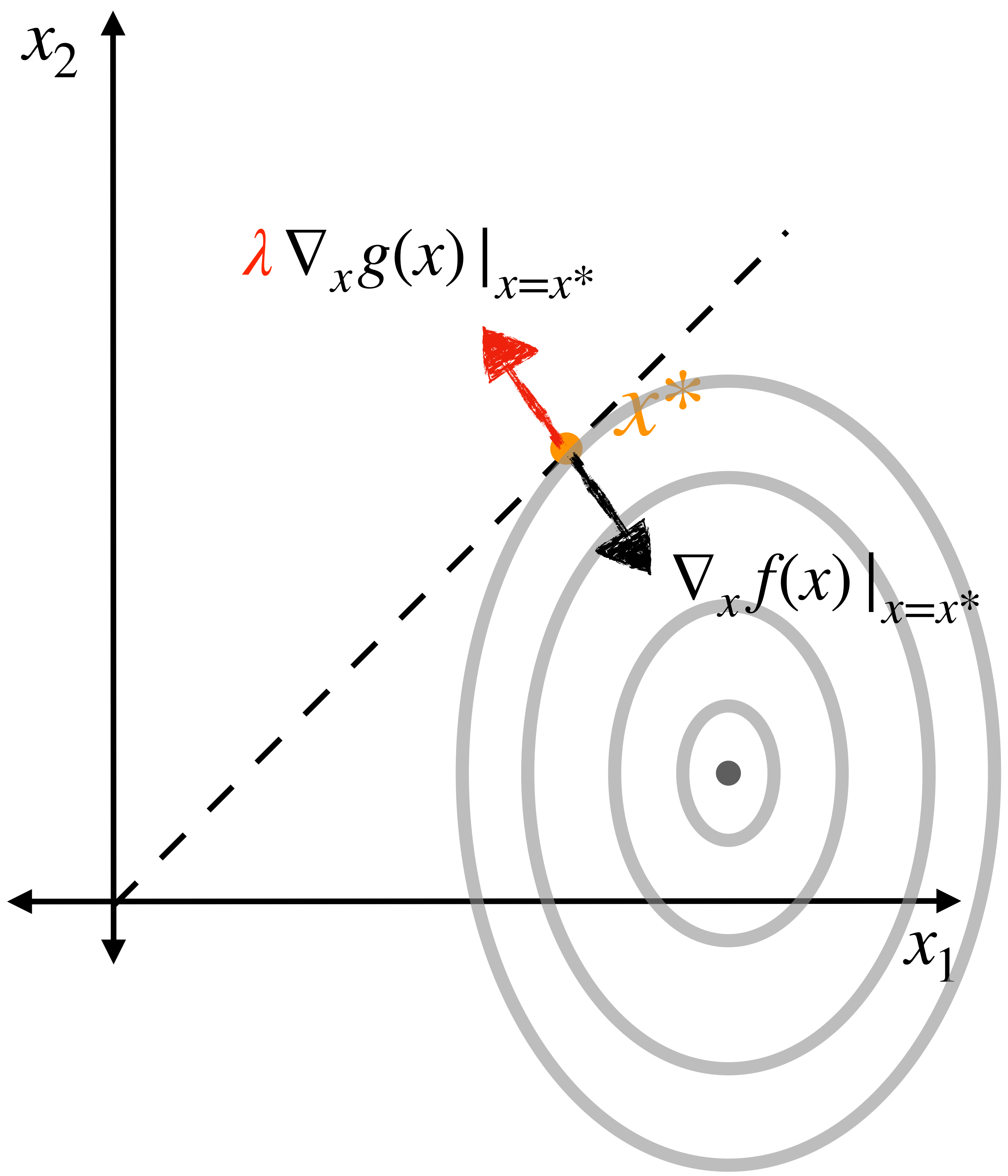


$$\min_x \quad f(x)$$

$$g(x) = 0$$

$$\min_x f(x) + \frac{\alpha}{2} g(x)^2$$

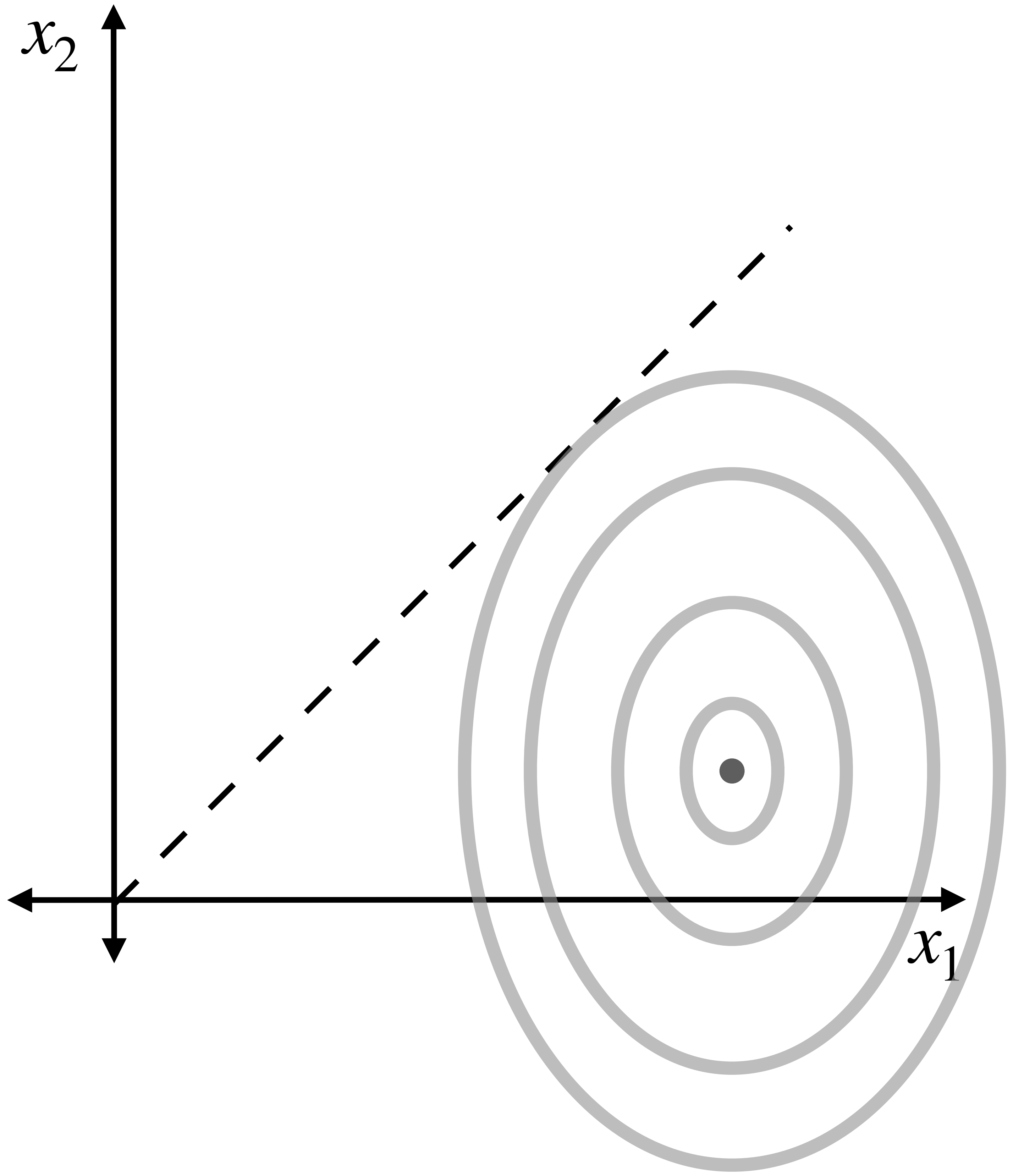
Lagrange's key insight



V1: A statement on the gradient

$$\nabla_x f(x)|_{x=x^*} = \lambda \nabla_x g(x)|_{x=x^*}$$

Lagrange's key insight



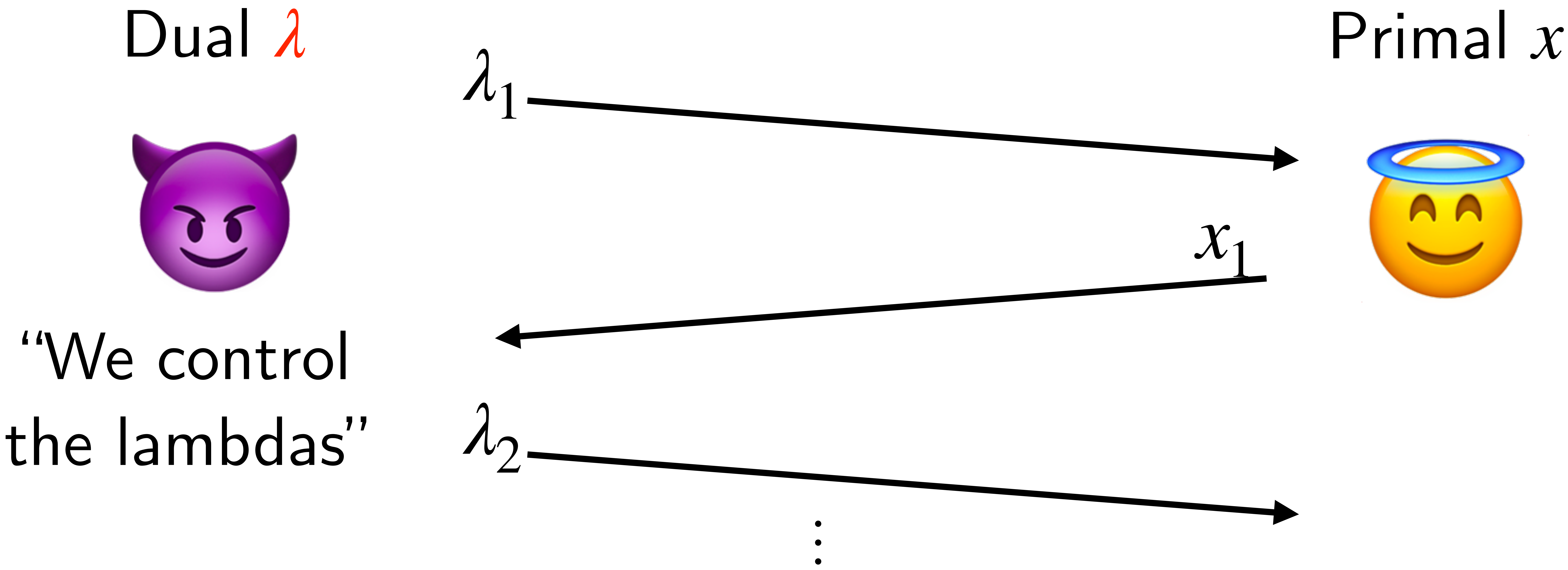
V2: A game!

$$\max_{\lambda} \min_{x} f(x) - \lambda^T g(x)$$

Lagrange Multipliers

We have seen such games before!

$$\min_{x} \max_{\lambda} f(x) - \lambda^T g(x)$$



Stably change λ

Follow the Regularized Leader!

Specific FTRL: Gradient Descent

Augmented Lagrangian

$$\min_x \max_{\lambda} f(x) - \lambda^T g(x)$$

For t = 1 … T

😈 Update λ_t

$$\lambda_{t+1} = \lambda_t - \eta g(x_t)$$

😇 Update x_t

$$x_{t+1} = \arg\min_x f(x) - \lambda_{t+1}^T g(x)$$
$$= \arg\min_x f(x) - \lambda_t^T g(x) + \eta g(x)^2$$

… and more

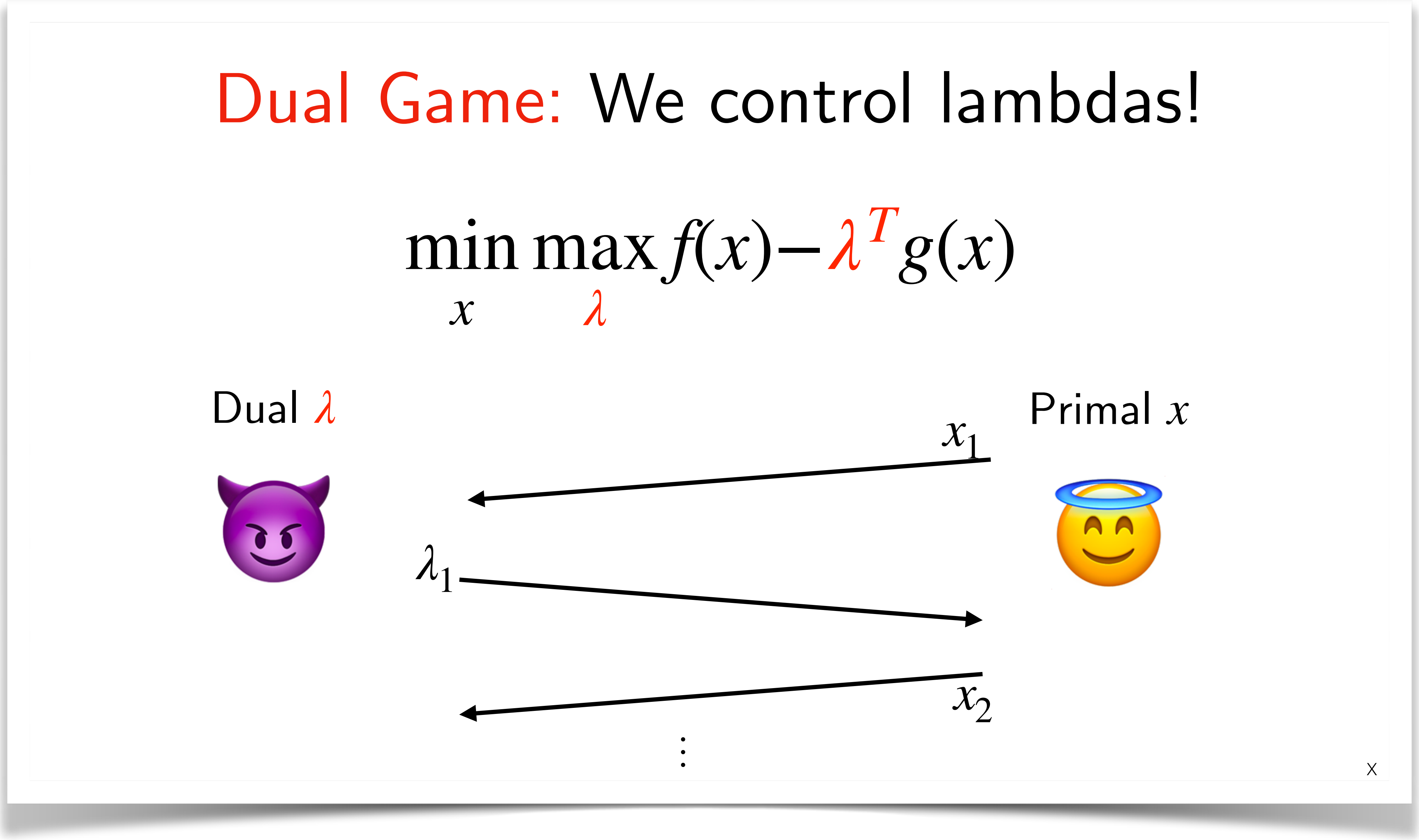
… and more

Non-convex / Non-differentiable

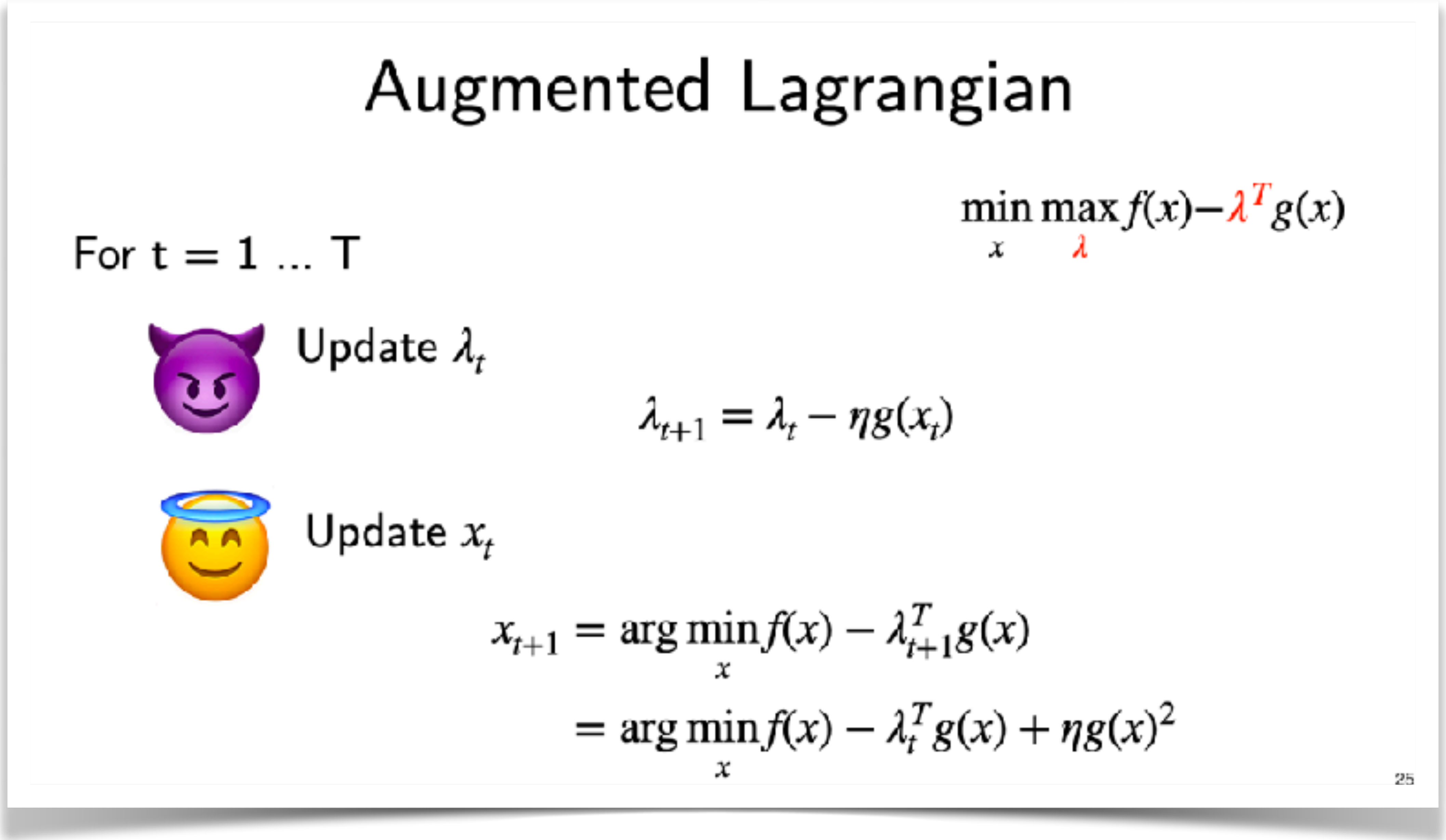
Partial Observability

What if the MDP is not known?

tl;dr


Dual Game: We control lambdas!
$\min_x \max_\lambda f(x) - \lambda^T g(x)$
Dual λ
Primal x
x_1
λ_1
x_2
⋮
x


Dual player is too
aggressive …


Augmented Lagrangian
$\min_x \max_\lambda f(x) - \lambda^T g(x)$
For t = 1 … T
Update λ_t
$\lambda_{t+1} = \lambda_t - \eta g(x_t)$
Update x_t
$x_{t+1} = \arg\min_x f(x) - \lambda_{t+1}^T g(x)$
$= \arg\min_x f(x) - \lambda_t^T g(x) + \eta g(x)^2$
25