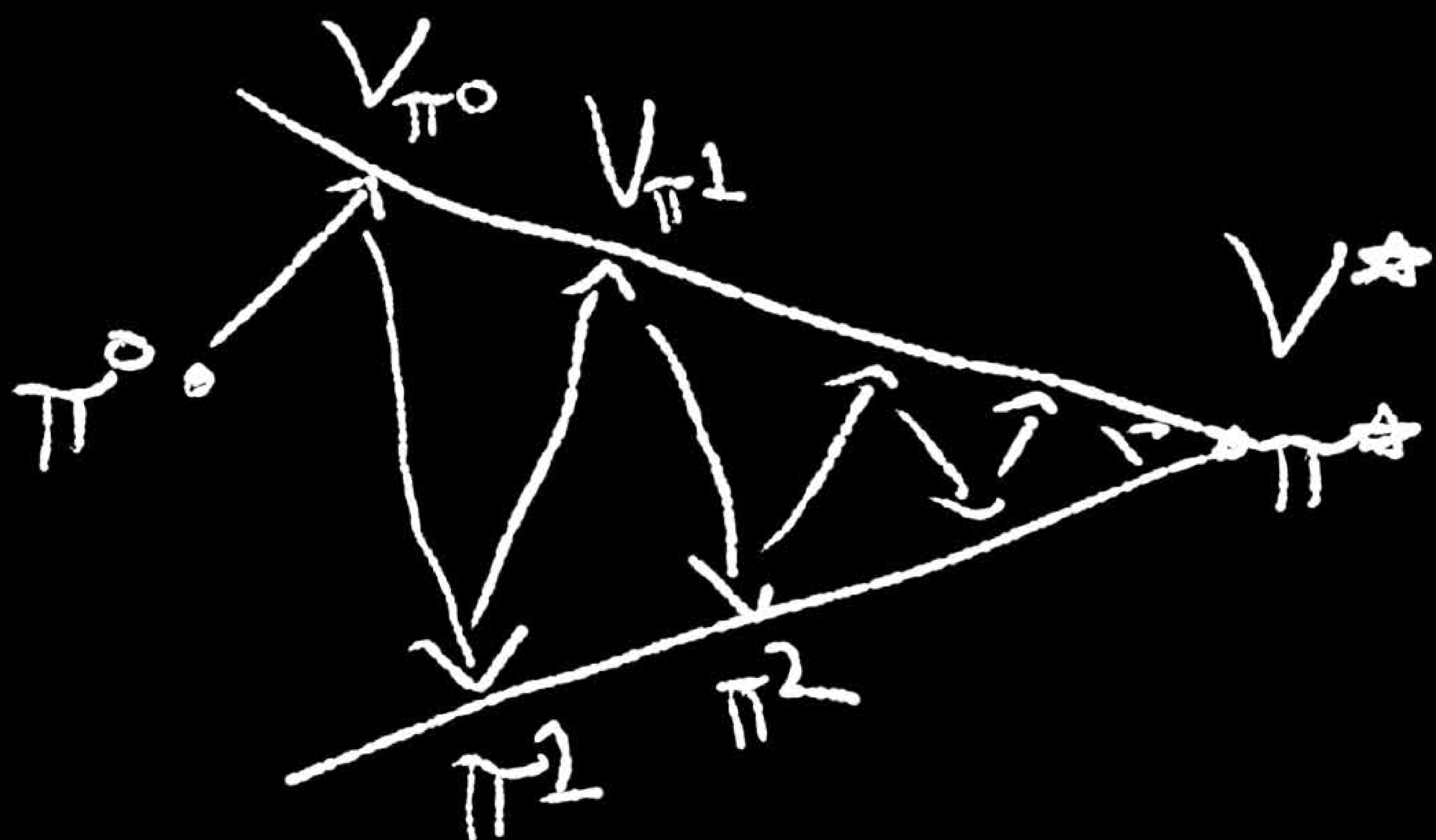
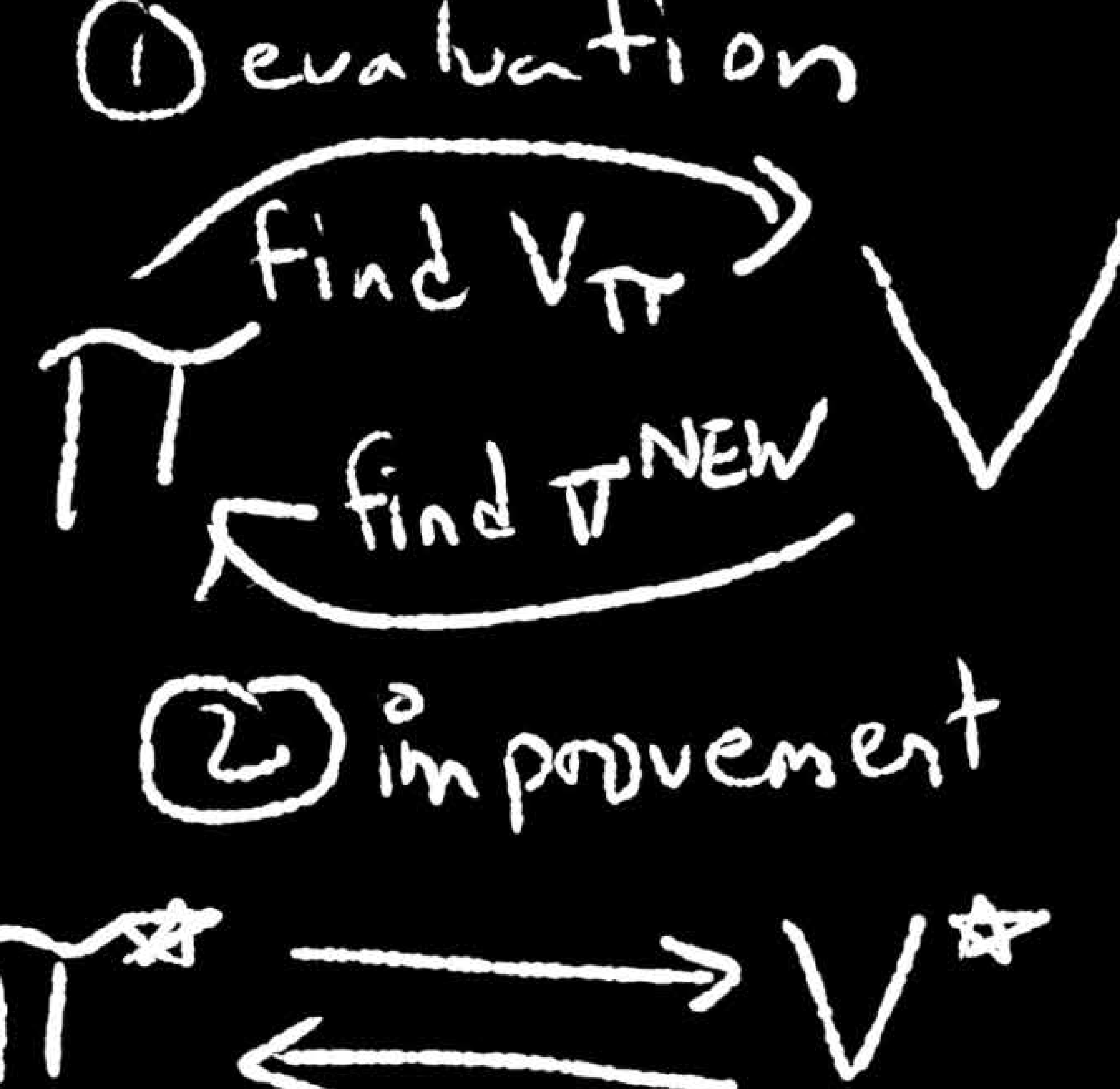
# III A Policy Iteration (PI) Algos + variants

Before we discussed
① Policy Evaluation ② Policy Improvement.




Schematic of Policy Iter. Algo.

# Algo (Policy Iteration (PI) Algo)

0) Initialize admissible $\pi^0$. Set $m=0$

1) Policy Eval: Set $V_\pi(x_k)=0 \ \forall x_k \in X$, $\pi \leftarrow \pi^m$

for $j=0,1,\ldots$

$$\longrightarrow V_\pi^{j+1}(x_k) = c(x_k, u_k) + \gamma \cdot V_\pi^j(x_{k+1}) \quad \forall x_k \in X$$

where $u_k = \pi(x_k)$, $x_{k+1} = f(x_k, u_k)$

2) Policy Improvement: Set $V_{\pi^{OLD}} \leftarrow \underbrace{V_{\pi}^{j+1}}_{\text{from 1)}}$

$$\pi^{NEW} = \arg\min_{\pi(\cdot)} \left\{ c(x_k, \pi(x_k)) + \gamma \cdot V_{\pi^{OLD}}(x_{k+1}) \right\}$$

where $x_{k+1} = f(x_k, \pi(x_k))$

Set $\pi^{m+1} = \pi^{NEW}$,

Go to step 1) + set $m \leftarrow m+1$.

Remarks: The policy eval step 1) is computationally if $j \to \infty$, and we must perform each iteration $\forall\, x_k \in X$.

- for now consider $X$ as a discrete state space.

Q: Can we truncate $j$ to $j = M$ iters?

A: YES! "Generalized Policy Iter (GPI)"

Q: Can we truncate to $M = 1$ iteration?

A: YES! "Value Iteration (VI)"

Textbooks: Sutton & Barto, Bertsekas.

Ref: [Howard 1960], [Puterman 1978] for theory.