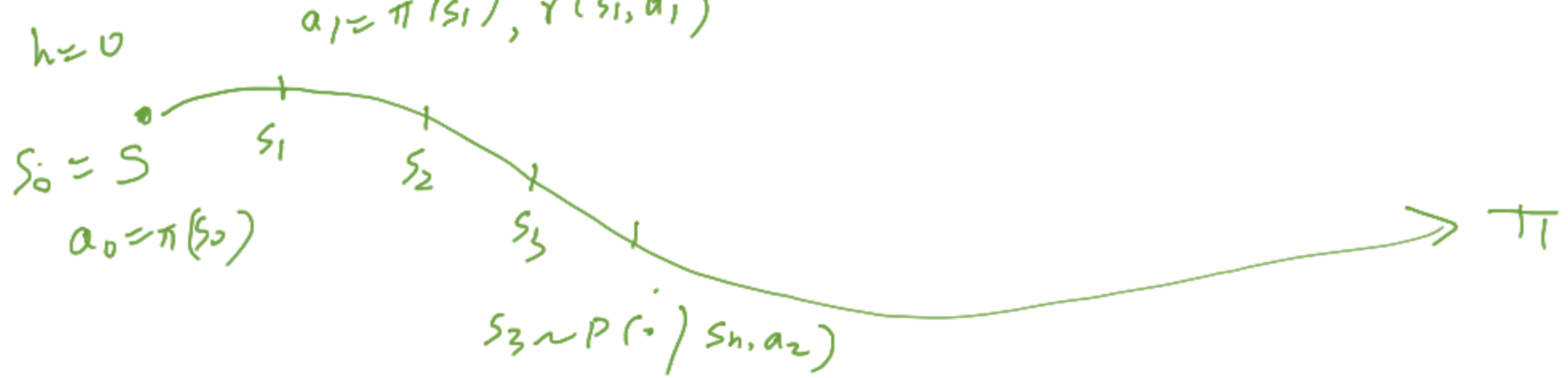
Hello !:

Value function:

Given policy $\pi: S \mapsto A$

$$V^{\pi}(s) = E\Big[r(s_0, a_0) + \gamma\, r(s_1, a_1) + \gamma^2 r(s_2, a_2) + \dots \;\Big|\; s_0 = s,\ a_h = \pi(s_h),\ s_{h+1} \sim P(\cdot \mid s_h, a_h) \Big]$$

$a_1 = \pi(s_1),\ r(s_1, a_1)$



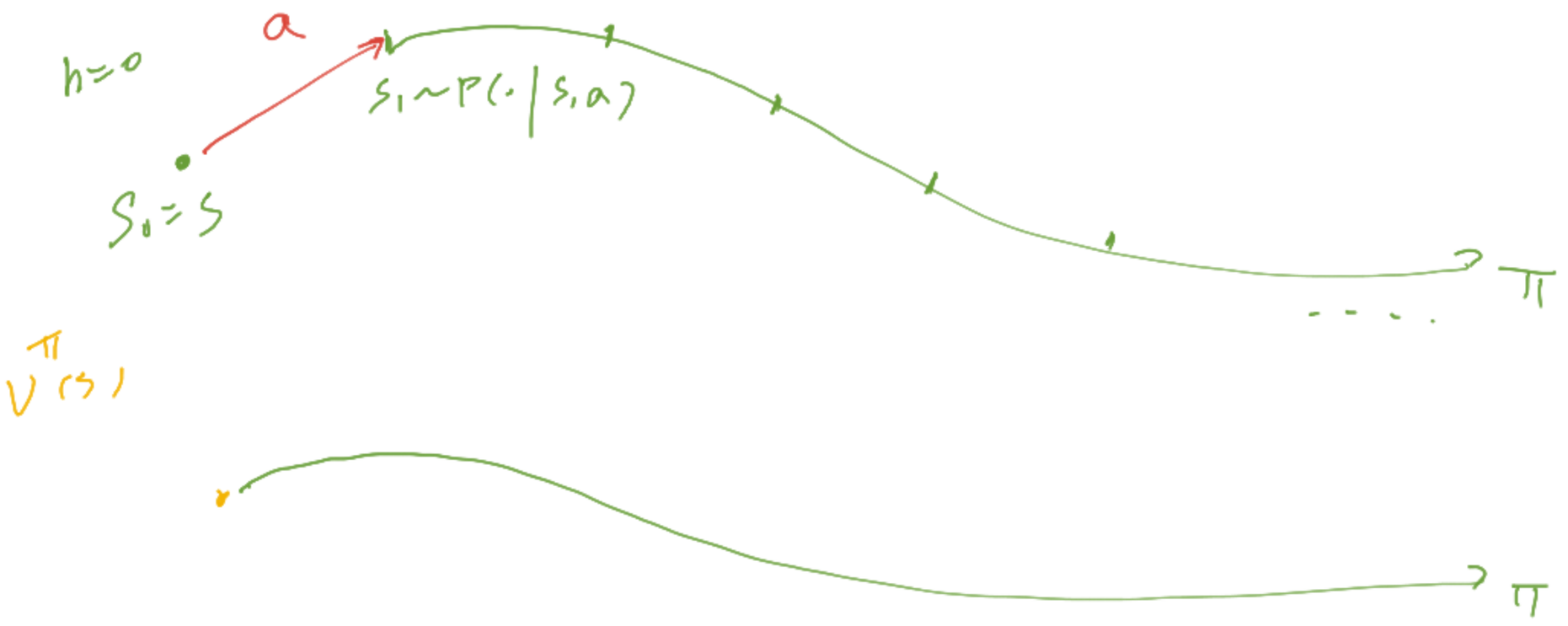
Given policy $\pi: S \mapsto A$ $(\pi \in \{\pi: S \mapsto A\})$

$\#\, A^S$ policies

Define

$$Q^{\pi}(s,a) = E\left[r(s_0,a_0) + \gamma\, r(s_1,a_1) + \gamma^2 r(s_2,a_2) \cdots \;\middle|\; \begin{array}{l} s_0 = s \\ a_0 = a \\ a_h = \pi(s_h) \\ s_{h+1} \sim P(\cdot \mid s_h, a_h) \end{array} \right]$$

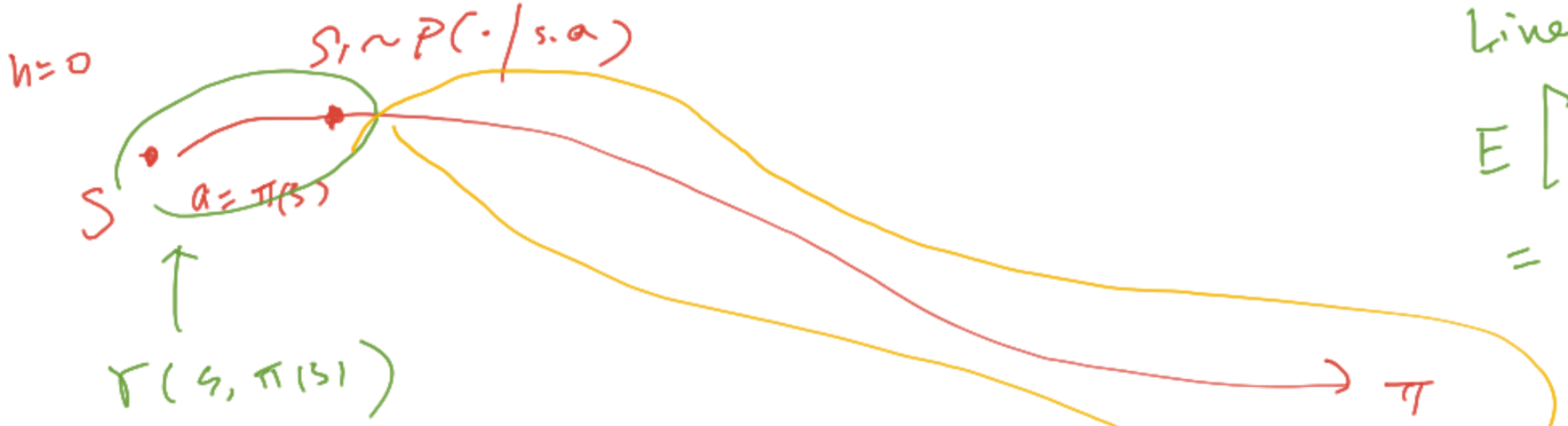
A



Bellman Equation:

(Value function)

$$V^{\pi}(s) = r(s, \pi(s)) + \gamma \, \mathbb{E}_{s' \sim P(\cdot|s,\pi(s))} \left[V^{\pi}(s') \right]$$



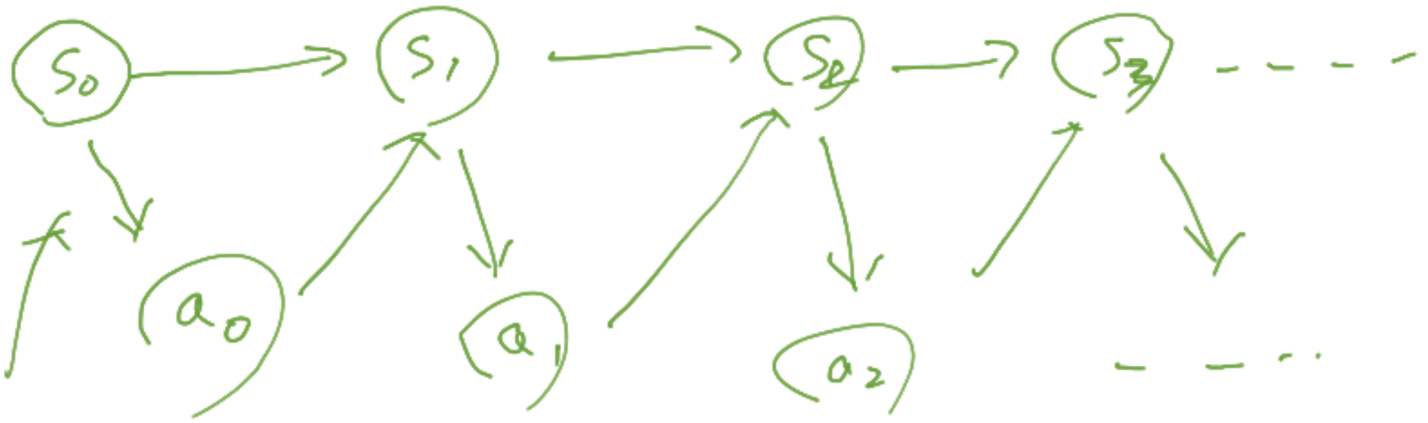
Linearity of Exp:

$$\mathbb{E}[a+b] = \mathbb{E}[a] + \mathbb{E}[b]$$

$$\mapsto V^{\pi}(s) = \mathbb{E}\left[r(s_0, a_0) + \gamma \, r(s_1, a_1) + \dots \mid s_0 = s,\ a_n = \pi(s_n),\ s_{n+1} \sim P(\cdot|s_n, a_n) \right]$$

$$= r(s, \pi(s)) + \mathbb{E}\left[\gamma \, r(s_1, a_1) + \dots \mid s_1 \sim P(\cdot|s, \pi(s)),\ s_{n+1} \sim P(\cdot|s_n, a_n),\ a_n = \pi(s_n) \right]$$

$$= r(s, \pi(s)) + \gamma \, \mathbb{E}\left[r(s_1, a_1) + \gamma \, r(s_2, a_2) + \dots \mid \dots \right]$$



$a_0 = \pi(s_0)$ $\quad s_1 \sim P(\cdot \mid s_0, a_0)$

$r + r^2 + r^3 + \cdots$

$r = 0.99$

$= \frac{r}{1-r} \approx 100$