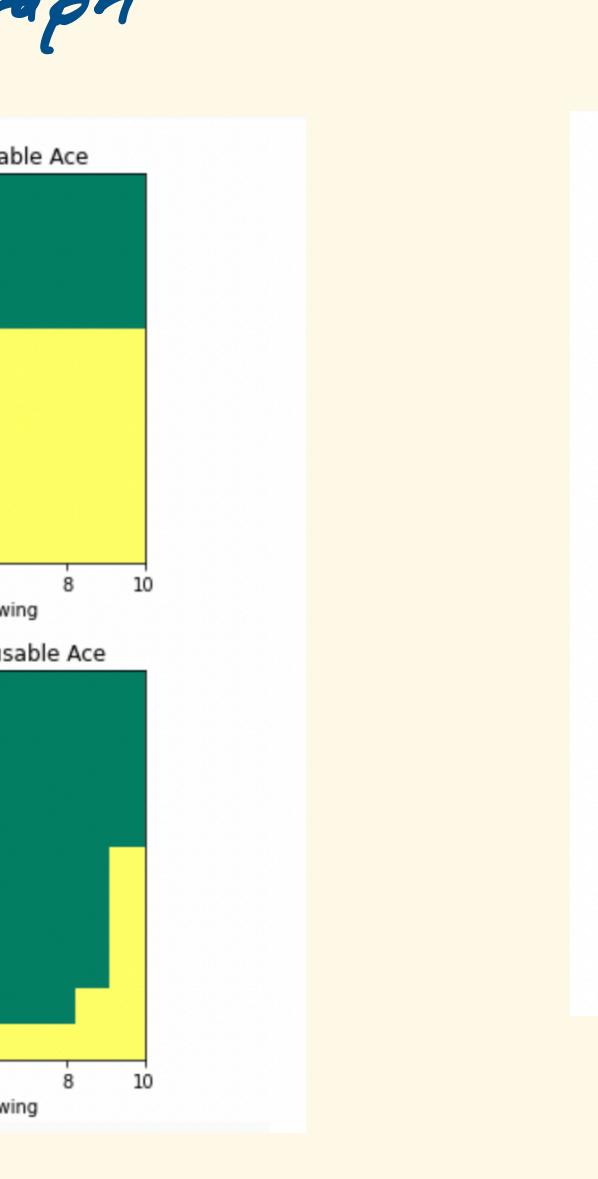


Ans 1.a)

In Example 5.1
the state is (p, d, a) & facing

this makes the state unique (p, d, a) ,
there can be same same (p, d) but
 $(p, d, a) \neq$ State would be unique for
each episode as it's keep on increasing
therefore Every visit MC would have no
effect.

1.b)



one episode =

$$G_t = R_t + \gamma G_{t+1}$$

$$G_0 = 1, G_1 = 2, G_2 = 3, G_3 = 4, G_4 = 5$$

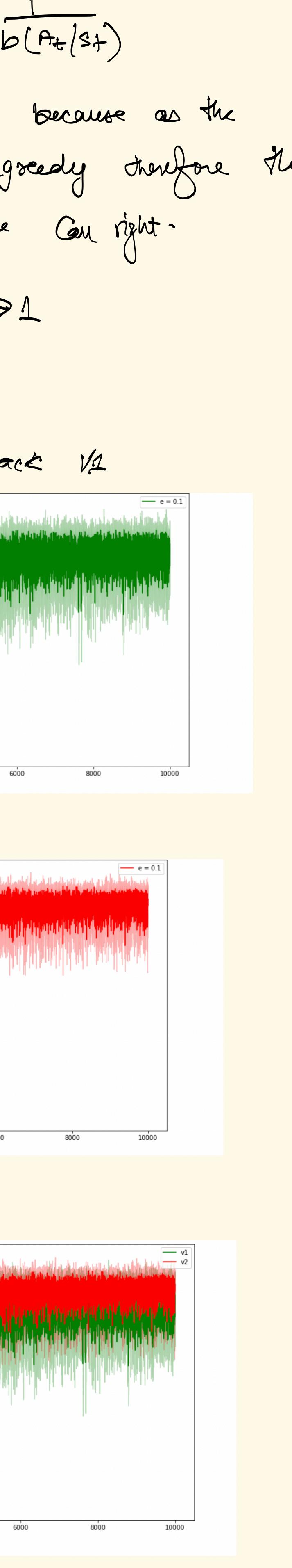
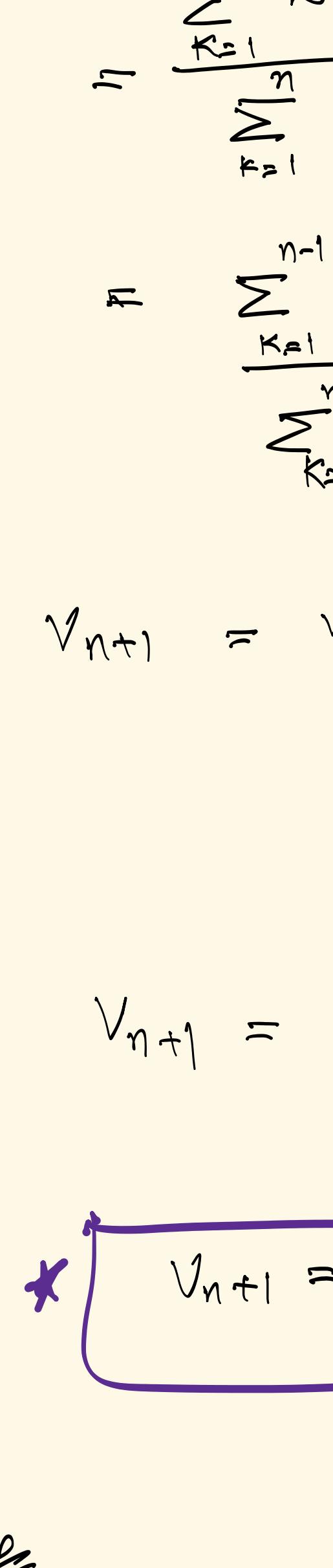
$$G_5 = 6, G_6 = 7, G_7 = 8, G_8 = 9, G_9 = 10$$

$$\text{first visit } G_0 = 10$$

$$\text{Every visit} = \frac{\sum_{i=0}^9 G_i}{10} = \frac{10 \times 11}{2 \times 10} = 5.5$$

Ans 2.a)

2.a Graph



Ans 2.b)

$\epsilon = 0$ setting demonstrate the importance of the policy in Monte-Carlo ES,

when $\epsilon = 0$ chooses the first greedy result & does not explore but this change with MC as in this it still explores

which in turn shows the importance of the MC as

$\therefore \epsilon = 0$ setting demonstrate the importance of doing exploring starts in Monte-Carlo ES

Ans 3

3.a)

