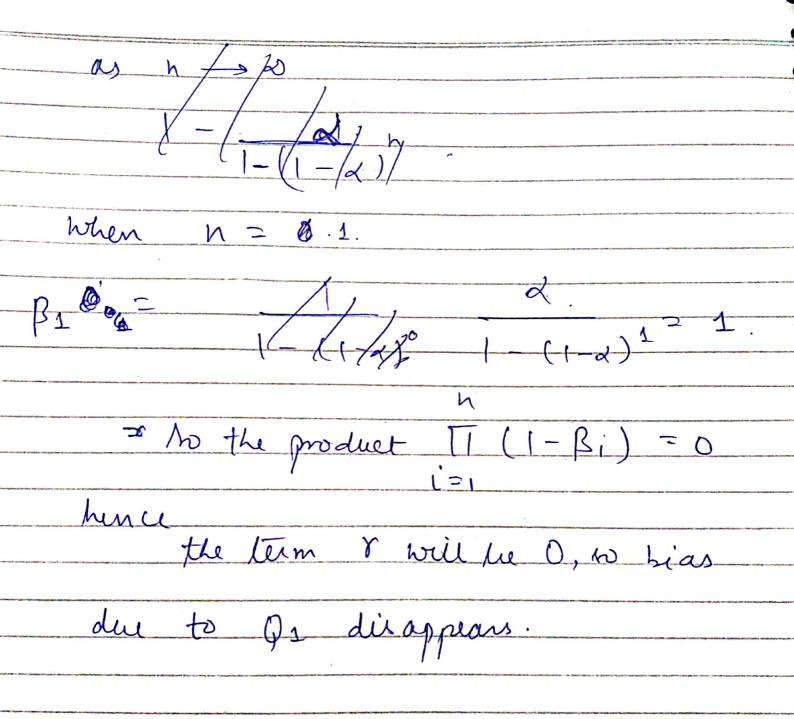
Exercise 2.7

$$B_n = Q$$
 $Q_0 = Q$
 $Q_0 = Q$
 $Q_1 = Q_1 + Q_2 + Q_3 + Q_4 + Q_4 + Q_5 +$

Bn Rn + (1-Bn). Qm = Bn Pn + (1-Bn) Bn-1 Rn-1 + (1-Bn-1), Qn+ = BARN + (1-BA) BAARNA + (1-BA)-(1-= Burn + (1-Bu). Fn-1 Fn-1 Fn-1 Fn-2 Pn-2 + (1-Bu-2) Que = Bn Rn + (1-Bn). Bn-18n+ + (1-Bn) Bn-28n-2 1-Bn). (1-Bn-2) (1-Bn-2) Qn-



Ex. 2.6

The early part of the graph Mows spike heraux it is exploring desting that time and sometime it will take the most optimal tep and sometime it won't. Rest once, it gets a stable estimate, the 90 of este optimal actions grows.

To avoid spikes, we can discourage exploring and encourage the system to do more of exploitation.