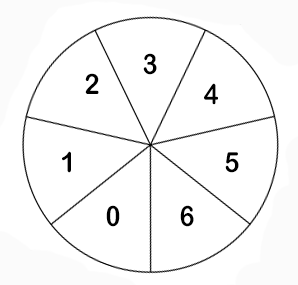
**Lab 5: Reinforcement Learning**

**Part II**

**Question 1 (theory):** In the report,

a) describe your choices of state and reward functions, and

b) describe **in your own words** the purpose of the different components in the Q-learning update that you implemented. In particular, what are the Q-values?



The orientation got a value between -π and + π that represent an infinity of values.  
That’s why we used the provided function « discretize » in order to split the amount of state into 7 different state :

The objectif is that the angle is as cloth as possible to 0 so our ship will always look at the top.   
Firstly, we’ve tried linear values like angle , so when the angle is cloth to 0, the reward should tend toward +∞ while it would be only to 1 when it get cloth to – π or + π. The problem was that difference of rewards was not mark enough so we used « stronger » modification to change it’s values : . In this way, as soon as Abs(angle) > 1, it get low values while the opposite got big rewards.

Here the equation calculating the new values for (state S for an action a)



Represent the reward of being on the current state

Represent the value of the next iteration on the new action a’. This values is multiplied by  **γ** that represent the importance given to the next step. Higher this value is and faster will tend to

Finally, **α** represent the learning rate, in other terms, higher this value is and faster will learn from the iteration.

**Question 2:** Try turning off exploration from the start before learning. What tends to happen? Explain why this happens in your report.

By turning off the exploration at the starting of the application, the ship is just falling without sarting any rocket. It’s a logical consequence because the is the only state in which the ship has been and so it’s the only state that have a reward > 0. Because the algorithm do not try any new case while the exploration is off, this action for this state stay the best move for the ship that the program « know »