Note that the stopping criterion is satisfied when the difference between successive value function estimates is sufficiently small. In particular, the loop terminates if the difference is less than θ for each state. And, the closer we want the final value function estimate to be to the optimal value function, the smaller we need to set the value of θ.

Feel free to play around with the value of θ in your implementation; note that in the case of the FrozenLake environment, values around `1e-8` seem to work reasonably well.

For those of you who are interested in *more rigorous* guidelines on how exactly to set the value of θ, you might be interested in perusing this paper, where you are encouraged to pay particular attention to Theorem 3.2. Their main result of interest can be summarized as follows:

> Let V^{final} denote the final value function estimate that is calculated by the algorithm. Then it can be shown that V^{final} differs from the optimal value function v_* by at most $\frac{2\theta\gamma}{1-\gamma}$. In other words, for each $s \in \mathcal{S}$,

> $\max_{s\in\mathcal{S}} |V^{\text{final}}(s) - v_*(s)| < \frac{2\theta\gamma}{1-\gamma}$.

Please use the next concept to complete **Part 6: Value Iteration** of `Dynamic_Programming.ipynb`. Remember to save your work!

If you'd like to reference the pseudocode while working on the notebook, you are encouraged to open this sheet in a new window.

Feel free to check your solution by looking at the corresponding section in `Dynamic_Programming_Solution.ipynb`.