First, let us show that for a policy $\pi$, if the initial and undiscounted state distributions, $\rho$ and $d_\pi$, are the same, then the undiscounted and discounted state distribution are completely equivalent. The proof is based on the definition of stationary distribution and an analytical form of discounted distribution.

*Proof.* Recall that the stationary distribution $d_\pi$ satisfies that

$$\sum_s d_\pi(s)\mathbb{P}_\pi(S' = s'|S = s) = d_\pi(s').$$

As known, the discounted stationary state distribution, $d_{\pi,\gamma}$ can be written as

$$d_{\pi,\gamma}(s) = (1-\gamma)\sum_{t\geq 0}\gamma^t\mathbb{P}_\pi(S_t = s).$$

Now let us consider that the initial distribution is the undiscounted distribution, that is $\mathbb{P}_\pi(S_0 = s) = d_\pi(s)$. Then, we can prove the statement by mathematical induction that for all $t > 0$, $\mathbb{P}_\pi(S_t = s) = d_\pi(s)$.

For $t = 0$, it is obvious.

For $t = k + 1$, let us assume the statement holds for $t = k$ such that $\mathbb{P}_\pi(S_k = \bar{s}) = d_\pi(\bar{s})$, for all states $\bar{s}$.

$$\begin{aligned}
\mathbb{P}_\pi(S_{k+1} = s) &= \sum_{\bar{s}}\mathbb{P}_\pi(S_k = \bar{s})\mathbb{P}_\pi(S' = s'|S = \bar{s}) \\
&= \sum_{\bar{s}} d_\pi(\bar{s})\mathbb{P}_\pi(S' = s'|S = \bar{s}) \\
&= d_\pi(s).
\end{aligned}$$

Thus, we can rewrite the discounted state distribution as

$$\begin{aligned}
d_{\pi,\gamma}(s) &= (1-\gamma)\sum_{t\geq 0}\gamma^t\mathbb{P}_\pi(S_t = s) \\
&= (1-\gamma)\sum_{t\geq 0}\gamma^t d_\pi(s) \\
&= d_\pi(s).
\end{aligned}$$

□

**Lemma 0.1.** *For any policy $\pi$, any positive constant $\epsilon$, and any MDP where the undiscounted state stationary distribution exists, if the total variation between the initial and undiscounted state distributions, $\rho$ and $d_\pi$ is small, that is,*

$$d_{TV}(\rho, d_\pi) \leq \epsilon,$$

*then the total variation between the discounted and undiscounted state distributions, $d_\pi$ and $d_{\pi,\gamma}$, is also small, that is,*

$$d_{TV}(d_{\pi,\gamma}, d_\pi) \leq \epsilon.$$

*Proof.* Recall that the total variation equals

$$d_{TV}(\rho, d_\pi) = \frac{1}{2}\sum_s |\rho(s) - d_\pi(s)|.$$

Next, let us compare the discounted and undiscounted state distributions.

$$\begin{aligned}
d_{TV}(d_{\pi,\gamma}, d_\pi) &= \frac{1}{2}\sum_{s'} |\sum_s (1-\gamma)\sum_{t\geq 0}\gamma^t \rho(s)\mathbb{P}_\pi(S_t = s'|S_0 = s) - \sum_s (1-\gamma)\sum_{t\geq 0}\gamma^t d_\pi(s)\mathbb{P}_\pi(S_t = s'|S_0 = s)| \\
&= \frac{1}{2}\sum_{s'}(1-\gamma)\sum_{t\geq 0}\gamma^t |\sum_s (\rho(s) - d_\pi(s))\mathbb{P}_\pi(S_t = s'|S_0 = s)| \\
&\leq \frac{1}{2}\sum_{s'}(1-\gamma)\sum_{t\geq 0}\gamma^t \sum_s \mathbb{P}_\pi(S_t = s'|S_0 = s)|\rho(s) - d_\pi(s)| \\
&= \frac{1}{2}(1-\gamma)\sum_{t\geq 0}\gamma^t \sum_s\sum_{s'} \mathbb{P}_\pi(S_t = s'|S_0 = s)|\rho(s) - d_\pi(s)| \\
&= \frac{1}{2}(1-\gamma)\sum_{t\geq 0}\gamma^t \sum_s |\rho(s) - d_\pi(s)| \\
&= (1-\gamma)\sum_{t\geq 0}\gamma^t d_{TV}(\rho, d_\pi) \\
&\leq \epsilon.
\end{aligned}$$

□