Here we are considering all gradients wrt to θ the parameter for the policy

$$\nabla v_\pi(s) = \nabla[\sum_a \pi(a|s) q_\pi(s,a)]$$

(This is how it is defined)

$$= \sum_a [\nabla\pi(a|s)\,\mathrm{q}_\pi(s,a)] + \sum_a [\pi(a|s)\nabla[\sum_{s',r} p(s',r|s,a)\gamma(r + \mathrm{v}_\pi(s'))]]$$

We define $\phi(s) = \sum_a \nabla\pi(a|s)\,\mathrm{q}_\pi(s,a)$, its like a policy grad for each action and value

$$= \phi(s) + \gamma\sum_a \pi(a|s)\sum_{s'} p(s'|s,a)\nabla\,\mathrm{v}_\pi(s')$$

$$= \phi(s) + \gamma\sum_{s'} \rho_\pi(s\to s', k=1)\nabla\,\mathrm{v}_\pi(s')$$

Expanding on $v(s')$

$$= \phi(s) + \gamma\sum_{s'} \rho_\pi(s\to s', k=1)[\phi(s') + \sum_{a'} \pi(a'|s')\sum_{s''} p(s''|s',a')\gamma\nabla\,\mathrm{v}_\pi(s'')]$$

$$= \phi(s) + \gamma\sum_{s'} \rho_\pi(s\to s', k=1)\phi(s') + \gamma^2\sum_{s'} \rho_\pi(s\to s', k=1)\sum_{a'} \pi(a'|s')\sum_{s''} p(s''|s',a')\nabla\,\mathrm{v}_\pi(s'')$$

$$= \phi(s) + \gamma\sum_{s'} \rho_\pi(s\to s', k=1)\phi(s') + \gamma^2\sum_{s''}\sum_{s'} \rho_\pi(s\to s', k=1)\,\rho_\pi(s'\to s'', k=1)\nabla\,\mathrm{v}_\pi(s'')$$

$$= \phi(s) + \gamma\sum_{s'} \rho_\pi(s\to s', k=1)\phi(s') + \gamma^2\sum_{s''} \rho_\pi(s\to s'', k=2)\nabla\,\mathrm{v}_\pi(s'')$$

$$= \phi(s) + \gamma\sum_{s'} \rho_\pi(s\to s', k=1)\phi(s') + \gamma^2\sum_{s''} \rho_\pi(s\to s'', k=2)\phi(s'') + \gamma^3\sum_{s'''} \rho_\pi(s\to s''', k=3)\phi(s''') + \ldots + \gamma^n\sum_{s^n} \rho_\pi(s\to s^n, k=n)\phi(s^n) + \gamma^{n+1}\sum_{s^{n+1}} \rho_\pi(s\to s^{n+1}, k=n+1)\nabla\,\mathrm{v}_\pi(s^{n+1})$$

$$= \sum_{x\in\mathcal{S}}\sum_{k=0}^{\infty}\gamma^k\,\rho_\pi(s\to x, k)\phi(x)$$

$$= \sum_{x\in\mathcal{S}}\sum_{k=0}^{\infty}\gamma^k\,\rho_\pi(s\to x, k)\sum_a \nabla\pi(a|x)\,\mathrm{q}_\pi(x,a)$$

So we get the policy gradient Theorem for γ

Defining the $\rho(s\to s', k, \pi)$ function. It tells use about the probability of reaching from state s to state s' in k steps.

$\rho(s \to s, k=0) = 1$ Probability one
$\rho(s \to s', k=1, \pi) = \sum_a \pi(a|s) p(s'|s,a)$ One step transition probability following policy π

$\rho_\pi(s \to x, k = t+1) = \sum_{s'} \rho_\pi(s \to s', t)\, \rho_\pi(s' \to x, 1)$
Now we can extend the PG Theorem to get the objective update wrt to the initial value function gradient update

$$\nabla J(\theta) = \nabla \, \mathrm{v}_\pi(s_0)$$
$$= \sum_{x \in \mathcal{S}} \sum_{k=0}^{\infty} \gamma^k \, \rho_\pi(s \to x, k) \sum_a \nabla \pi(a|x) \, \mathrm{q}_\pi(x,a)$$