The Pumping Lemma

Claim. Let L be a regular language. There exists a constant $p > 0$ such that, for any string $w \in L$ with $|w| \geq p$, w can be written $w = xyz$ such that the following three conditions hold:

1. $|y| > 0$
2. $|xy| \leq p$
3. $xy^i z \in L$ for any $i \geq 0$.

Proof. Since L is regular, there exists some DFA $M = (Q, \Sigma, \delta, q_0, F)$ with $L(M) = L$. Let $p = |Q|$, and let $w \in L$ have $|w| \geq p$.

Let w_i be the ith symbol of w for $i = 1, \ldots, |w|$; and let $r_i = \hat{\delta}(q_0, w_1 \ldots w_i)$. Then the sequence $(r_0, r_1, \ldots, r_p)$ has length $p+1$. Since $|Q| = p$, this implies that there exist k and l, with $0 \leq k < l \leq p$, for which $r_k = r_l$. That is, $w_1 \ldots w_p$ drives M through a cycle.

Let $x = w_1 \ldots w_k$, $y = w_{k+1} \ldots w_l$, and $z = w_{l+1} \ldots w_{|w|}$. Clearly $y > 0$. Also, $|xy| = l - k + k = l \leq p$. Furthermore, it is clear from the choice of y and the above discussion that $\hat{\delta}(q_0, xz) = \hat{\delta}(q_0, w)$ and that $\hat{\delta}(r_k, y) = r_k$. Hence, $\hat{\delta}(q_0, xy^i z) = \hat{\delta}(q_0, xz) = \hat{\delta}(q_0, w)$ for all $i \in N$. Since $w \in L(M)$, this implies $xy^i z \in L(M)$ for all $i \in N$ and we have the claim. ∎