1.

(a) $P(d_k|\theta_1+\theta_2) = \prod_{i=1}^{M}\left[\lambda_1 P(W_i=X_i|\theta_1)+\lambda_2 P(W_i=X_i|\theta_2)\right]$

$\Rightarrow$ log-likelihood $= L(\theta) = \sum_{i|X_i=0} \log(\lambda_1 P(W_i=0|\theta_1)+\lambda_2 P(W_i=0|\theta_2)) + \sum_{i|X_i=1} \log(\lambda_1 P(W_i=1|\theta_1)+\lambda_2 P(W_i=1|\theta_2))$

(b) Introduce $Z_i = \begin{cases} 1, & \text{if word } w_i \text{ comes from } \theta_1 \\ 0, & \text{otherwise} \end{cases}$

$\Rightarrow L_c(\theta) = \sum_{i=1}^{M}\left[Z_i \log(\lambda_1 P(W_i=X_i|\theta_1)) + (1-Z_i)\log(\lambda_2 P(W_i=X_i|\theta_2))\right]$

$\Rightarrow$ E-step : $Q(\theta;\theta^{(n)}) = E_H\left[L_c(\theta)|X,\theta^{(n)}\right]$, with $P(Z_i=1|X,\theta^{(n)}) = \dfrac{\lambda_1^{(n)}P(W_i=X_i|\theta_1)}{\lambda_1^{(n)}P(W_i=X_i|\theta_1)+\lambda_2^{(n)}P(W_i=X_i|\theta_2)}$, $P(Z=0|X,\theta^{(n)}) = 1-P(Z=1|X,\theta^{(n)})$

$P_0(Z_i=1|X,\theta^{(n)}) = \dfrac{\lambda_1^{(n)}P(W_i=0|\theta_1)}{\lambda_1^{(n)}P(W_i=0|\theta_1)+\lambda_2^{(n)}P(W_i=0|\theta_2)}$, $P_0(Z_i=0|X,\theta^{(n)}) = 1-P_0(Z_i=1|X,\theta^{(n)})$

$\Rightarrow P_1(Z_i=1|X,\theta^{(n)}) = \dfrac{\lambda_1^{(n)}P(W_i=1|\theta_1)}{\lambda_1^{(n)}P(W_i=1|\theta_1)+\lambda_2^{(n)}P(W_i=1|\theta_2)}$, $P_1(Z_i=0|X_1,\theta^{(n)}) = 1-P_1(Z_i=1|X,\theta^{(n)})$

$$= \sum_{i|X_i=0}\left[\frac{\lambda_1^{(n)}P(W_i=0|\theta_1)}{\lambda_1^{(n)}P(W_i=0|\theta_1)+\lambda_2^{(n)}P(W_i=0|\theta_2)}\log(\lambda_1 P(W_i=0|\theta_1)) + \frac{\lambda_2^{(n)}P(W_i=0|\theta_2)}{\lambda_1^{(n)}P(W_i=0|\theta_1)+\lambda_2^{(n)}P(W_i=0|\theta_2)}\log(\lambda_2 P(W_i=0|\theta_2))\right] + \sum_{i|X_i=1}\left[\frac{\lambda_1^{(n)}P(W_i=1|\theta_1)}{\lambda_1^{(n)}P(W_i=1|\theta_1)+\lambda_2^{(n)}P(W_i=1|\theta_2)}\log(\lambda_1 P(W_i=1|\theta_1)) + \frac{\lambda_2^{(n)}P(W_i=1|\theta_1)}{\lambda_1^{(n)}P(W_i=1|\theta_1)+\lambda_2^{(n)}P(W_i=1|\theta_2)}\log(\lambda_2 P(W_i=1|\theta_2))\right]$$

M-step: $\dfrac{\partial Q}{\partial \lambda_1}=0$, $\lambda_2 = 1-\lambda_1 \Rightarrow$

$$\Rightarrow \sum_{i|X_i=0}\left[\frac{P_0(Z_i=1|X,\theta^{(n)})}{\lambda_1} - \frac{1-P_0(Z_i=1|X,\theta^{(n)})}{1-\lambda_1}\right] + \sum_{i|X_i=1}\left[\frac{P_1(Z_i=1|X,\theta^{(n)})}{\lambda_1} - \frac{1-P_1(Z_i=1|X,\theta^{(n)})}{1-\lambda_1}\right] = 0$$

$$\Rightarrow \sum_{i|X_i=0}\left[P_0(Z_i=1|X,\theta^{(n)})-\lambda_1\right] + \sum_{i|X_i=1}\left[P_1(Z_i=1|X,\theta^{(n)})-\lambda_1\right] = 0 \Rightarrow \lambda_1 = \frac{\sum_{i|X_i=0}P_0(Z_i=1|X,\theta^{(n)})+\sum_{i|X_i=1}P_1(Z_i=1|X,\theta^{(n)})}{M},\ \lambda_2 = 1-\lambda_1$$

In summary, E-step: calculate $P_0(Z_i=1|X,\theta^{(n)}) = \dfrac{\lambda_1^{(n)}P(W_i=0|\theta_1)}{\lambda_1^{(n)}P(W_i=0|\theta_1)+\lambda_2^{(n)}P(W_i=0|\theta_2)}$ and $P_1(Z_i=1|X,\theta^{(n)}) = \dfrac{\lambda_1^{(n)}P(W_i=1|\theta_1)}{\lambda_1^{(n)}P(W_i=1|\theta_1)+\lambda_2^{(n)}P(W_i=1|\theta_2)}$

Q-step : update λ_1, λ_2 as $\lambda_1 = \dfrac{\sum_{i|X_i=0}P_0(Z_i=1|X,\theta^{(n)})+\sum_{i|X_i=1}P_1(Z_i=1|X,\theta^{(n)})}{M}$, $\lambda_2 = 1-\lambda_1$

2.

Build transition matrix T where t_{ij} = Prob[state i to state j]

Q-step : update λ_1, λ_2 as $\lambda_1 = \dfrac{\sum_{i|x_i=0} P_0(z_i=1|X,\theta) + \sum_{i|x_i=1} P_1(z_i=1|X,\theta)}{M}$, $\lambda_2 = 1-\lambda_1$

2.

Build transition matrix T where t_{ij} = Prob[state i to state j]

$$T = \begin{bmatrix} 0.1 & 0.6 & 0.3 \\ 0.1 & 0.25 & 0.65 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

(a) It's aperiodic because we can find $k=1$ such that a state i go back to it self with path's length equals to multiple of k.

It's irreducible because every state can go to any state (strongly connected)

$\Rightarrow$ It's ergodic.

(b) For steady state π, ① $\sum_{i=1}^{3} \pi_i = 1$ ② $\pi = T^T \pi$

$$\Rightarrow \begin{cases} \overset{-0.9}{\cancel{0.1\pi_1}} + 0.1\pi_2 + 0.3\pi_3 = \cancel{\pi_1} \\ 0.6\pi_1 + \overset{-0.75}{\cancel{0.25}}\pi_2 + 0.5\pi_3 = \cancel{\pi_2} \\ 0.3\pi_1 + 0.65\pi_2 + \overset{-0.8}{\cancel{0.2}}\pi_3 = \cancel{\pi_3} \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

18

$\Rightarrow 2.05\pi_2 = 2.1\pi_3$

$\Rightarrow \pi_2 = \frac{42}{41}\pi_3$

$\Rightarrow \pi_1 = \frac{10}{9} \cdot \left(\frac{4.2}{41}\pi_3 + 0.3\pi_3\right)$

$= \frac{55}{123}\pi_3$

$\Rightarrow \pi_1 : \pi_2 : \pi_3 = 55 : 126 : 123$

$$\Rightarrow \pi = \begin{bmatrix} \frac{55}{304} \\ \frac{63}{152} \\ \frac{123}{304} \end{bmatrix}$$