2.f

To find λ_{MAP}, we start by detailing its expression:

$$\lambda_{MAP} = \arg\max_{\lambda} p(\lambda|k) =$$

From Bayes rule, we get:

$$= \arg\max_{\lambda} \frac{p(k|\lambda)p(\lambda)}{p(k)} = \arg\max_{\lambda} \frac{p(k|\lambda)p(\lambda)}{\int p(k|\lambda)p(\lambda)\partial\lambda}$$

As the question states, it can be hard to compute the integral in the denominator of the above expression. But because we are working with maximum a posteriori estimation, we are using the arg max operator on λ. This means we are looking through all the possible values of λ, and all of them will have the same common denominator, so we can simply ignore the integral and work on finding the λ that maximizes $p(k|\lambda)p(\lambda)$ without any effect on the estimator found. We then continue our computations:

$$\lambda_{MAP} = \arg\max_{\lambda} p(k|\lambda)p(\lambda) =$$

Using the log trick, we get:

$$= \arg\max_{\lambda} \log(p(k|\lambda)p(\lambda)) = \arg\max_{\lambda} \log p(k|\lambda) + \log p(\lambda)$$

Since we already have the log-likelihood, we just need to compute the expression for the prior $\log p(\lambda)$ before continuing with our λ_{MAP} computations. Considering the given prior distribution, we have:

$$\log p(\lambda) = \log p(\lambda|\alpha_1, \alpha_2) = \log \frac{\alpha_2^{\alpha_1}}{\Gamma(\alpha_1)} \lambda^{\alpha_1 - 1} e^{-\alpha_2 \lambda} =$$

$$= \alpha_1 \log \alpha_2 - \log \Gamma(\alpha_1) + (\alpha_1 - 1) \log \lambda - \alpha_2 \lambda =$$

Expanding the Gamma function $\Gamma(x) = (x-1)!$:

$$= \alpha_1 \log \alpha_2 - (\sum_{i=1}^{\alpha_1 - 1} \log(\alpha_1 - i)) + (\alpha_1 - 1) \log \lambda - \alpha_2 \lambda$$

Going back to the original equation $\arg\max_\lambda \log p(k|\lambda) + \log p(\lambda)$, we get:

$$\arg\max_\lambda \log p(k|\lambda) + \log p(\lambda) = \arg\max_\lambda \sum_{i=1}^{N} k_i \log \lambda - \sum_{i=1}^{N} \log(k_i!) - N\lambda + \alpha_1 \log \alpha_2 - (\sum_{i=1}^{\alpha_1 - 1} \log(\alpha_1 - i)) + (\alpha_1 - 1) \log \lambda - \alpha_2 \lambda =$$

$$= \arg\max_\lambda \log \lambda \sum_{i=1}^{N} k_i - \sum_{i=1}^{N} \log(k_i!) + \alpha_1 \log \alpha_2 - (\sum_{i=1}^{\alpha_1} \log(\alpha_1 - i)) + (\alpha_1 - 1) \log \lambda - (N + \alpha_2)\lambda =$$

$$= \arg\max_\lambda ((\sum_{i=1}^{N} k_i) + \alpha_1 - 1) \log \lambda - (N + \alpha_2)\lambda - \sum_{i=1}^{N} \log(k_i!) + \alpha_1 \log \alpha_2 - (\sum_{i=1}^{\alpha_1 - 1} \log(\alpha_1 - 1))$$

But to get to the value that maximizes the above expression w.r.t λ, we would take its derivative and equal it to zero, which would zero out all the terms that do not depend on λ. Therefore, we can rewrite the expression with only the terms that depend on λ without compromising the maximum a posteriori estimation:

$$\lambda_{MAP} = \arg\max_\lambda ((\sum_{i=1}^{N} k_i) + \alpha_1 - 1) \log \lambda - (N + \alpha_2)\lambda$$

Exactly what was expected.