Problem 1

$$y(t) = x(t) + w(t), \quad t \in \{0, 1, \dots, n-1\}$$

$$x(t) = \sum_{k=1}^{N} a_k \cos\left(2\pi f_k t + \varphi_k\right), \quad f_i \neq f_j, \forall i \neq j$$

$$w(t) \sim \mathcal{N}(0, \sigma^2)$$

Suppose that there are altogether n data points and we assume $n \gg 1$.

Define $\alpha_k := a_k \cos(\varphi_k), \beta_k = -a_k \sin(\varphi_k)$. Then $a_k = \sqrt{\alpha_k^2 + \beta_k^2}$ and $\varphi_k = -\arctan^{-1}\left(\frac{\beta_k}{\alpha_k}\right)$.

Then the data can be written as

$$y(t) = \sum_{k=1}^{N} \alpha_k \cos(2\pi f_k t) + \beta_k \sin(2\pi f_k t)$$

Then in vector forms the observation equation can be written as

$$\vec{y} = \begin{bmatrix} \cos(2\pi f_1 \cdot 0) & \dots & \cos(2\pi f_N \cdot 0) & \sin(2\pi f_1 \cdot 0) & \dots & \sin(2\pi f_N \cdot 0) \\ \cos(2\pi f_1 \cdot 1) & \dots & \cos(2\pi f_N \cdot 1) & \sin(2\pi f_1 \cdot 1) & \dots & \sin(2\pi f_N \cdot 1) \\ \cos(2\pi f_1 \cdot (n-1)) & \dots & \cos(2\pi f_N \cdot (n-1)) & \sin(2\pi f_1 \cdot (n-1)) & \dots & \sin(2\pi f_N \cdot (n-1)) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \dots \\ \alpha_N \\ \beta_1 \\ \dots \\ \beta_N \end{bmatrix} + \vec{w}$$

Denote $c_i = [\cos(2\pi f_i 0), \cos(2\pi f_i 1), \dots, \cos(2\pi f_i (n-1))]^\top$,
$s_i = [\sin(2\pi f_i 0), \sin(2\pi f_i 1), \dots, \sin(2\pi f_i (n-1))]^\top$, $\vec{\gamma} = [\alpha_1, \dots \alpha_N, \beta_1, \dots, \beta_N]^\top$, and
$\vec{f} = [f_1, f_2, \dots, f_N]^\top$, we can simplify the expression as

$$\vec{y} = H(\vec{f})\vec{\gamma} + \vec{w}$$

Due to that the noise possesses gaussian pdf, with some simple calculation we discover that our optimization problem takes the following equivalent forms:

$$\begin{aligned} \underset{\{f_k, \varphi_k, a_k\}}{\text{maximize}} \ \ln p_{\mathrm{y}}(\vec{y}; \{f_k, \varphi_k, a_k\}) &= \underset{\vec{f}, \vec{\gamma}}{\text{minimize}} \ \left\| \vec{y} - H(\vec{f})\vec{\gamma} \right\|_2^2 \\ &= \underset{\vec{f}}{\text{minimize}} \underbrace{\underset{\vec{\gamma}}{\text{minimize}} \ \left\| \vec{y} - H(\vec{f})\vec{\gamma} \right\|_2^2}_{\text{Standard Least-Square Problem}} \\ &= \underset{\vec{f}}{\text{minimize}} \left\| \vec{y} - H(\vec{f}) \left(\underbrace{\left(H(\vec{f})^\top H(\vec{f}) \right)^{-1} H(\vec{f})^\top y}_{\text{Optimal } \gamma \text{ given } \vec{f}} \right) \right\|_2^2 \\ &= \underset{\vec{f}}{\text{maximize}} \ y^\top H(\vec{f}) \left(H(\vec{f})^\top H(\vec{f}) \right)^{-1} H(\vec{f})^\top y \end{aligned}$$

With the fact that MLE estimator is invariant under non-linear transforms, we conclude:

$$
\begin{aligned}
\hat{f}_{MLE}(\vec{y}) =& \underset{\vec{f}}{\operatorname{argmax}}\ y^\top H(\vec{f})\left(\frac{1}{n}H(\vec{f})^\top H(\vec{f})\right)^{-1} H(\vec{f})^\top y \\
=& \underset{\vec{f}}{\operatorname{argmax}}\ \left\| H(\vec{f})^\top y \right\|_{\left(\frac{1}{n}H(\vec{f})^\top H(\vec{f})\right)^{-1}} \\
\begin{bmatrix} \hat{\alpha}_{MLE} \\ \hat{\beta}_{MLE} \end{bmatrix} =& \hat{\gamma}_{MLE} = \gamma(\hat{f}_{MLE}) = \left(H(\hat{f}_{MLE})^\top H(\hat{f}_{MLE})\right)^{-1} H^\top(\hat{f}_{MLE})\vec{y} \\
\begin{bmatrix} \hat{a}_{k,MLE} \\ \hat{\varphi}_{k,MLE} \end{bmatrix} =& \begin{bmatrix} \sqrt{\hat{\alpha}^2_{k,MLE} + \hat{\beta}^2_{k,MLE}} \\ -\arctan^{-1}\left(\frac{\hat{\beta}_{k,MLE}}{\hat{\alpha}_{k,MLE}}\right) \end{bmatrix}
\end{aligned}
$$

In what follows we will calculate the MLE for the frequency under the assumption that $n \gg 1$ and $f_k \in (0, \frac{1}{2})$.

$$
\frac{1}{n}H(\vec{f})^\top H(\vec{f}) = \frac{1}{n}\begin{bmatrix} (c_i^\top c_j)_{N\times N} & (c_i^\top s_j)_{N\times N} \\ (s_i^\top c_j)_{N\times N} & (s_i^\top s_j)_{N\times N} \end{bmatrix}
$$

Due to the fact that as $n \gg 1$. the following relations are easily obtained by trigonometry identities along with the assumption that $f_k \in (0, \frac{1}{2})$, see appendix:

$$
\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}c_i^\top c_j =& \frac{1}{2}\delta_{i,j} \\
\lim_{n\to\infty} \frac{1}{n}s_i^\top s_j =& \frac{1}{2}\delta_{i,j} \\
\lim_{n\to\infty} \frac{1}{n}c_i^\top s_j =& 0
\end{aligned}
$$

Using these relations we immediately conclude that

$$
\frac{1}{n}H(\vec{f})^\top H(\vec{f}) \approx \frac{1}{2}\mathbb{I}_{2N\times 2N} \qquad \text{when n} \gg 1.
$$

which then implies

$$
\begin{aligned}
\hat{f}_{MLE}(\vec{y}) =& \underset{\vec{f}}{\operatorname{argmax}}\ 2\cdot \left\| H(\vec{f})^\top y \right\|_2^2 = \underset{\vec{f}}{\operatorname{argmax}}\ 2\cdot \sum_{k=1}^{N}\left(\sum_{t=0}^{n-1} y(t)\cos(2\pi f_k t)\right)^2 + \left(\sum_{t=0}^{n-1} y(t)\sin(2\pi f_k t)\right)^2 \\
=& \underset{\vec{f}}{\operatorname{argmax}}\ 2\sum_{k=1}^{N}(\mathrm{Re}^2 + \mathrm{Im}^2)\left(\sum_{t=0}^{n-1} y(t)e^{j2\pi f_k t}\right) = \underset{\vec{f}}{\operatorname{argmax}}\ 2\sum_{k=1}^{N}\left|\sum_{t=0}^{n-1} y(t)e^{j2\pi f_k t}\right|^2 \\
=& \underset{\vec{f}}{\operatorname{argmax}}\ 2\sum_{k=1}^{N} |\mathrm{FFT}\{\mathrm{y}\}|^2(f_k)
\end{aligned}
$$

Under the assumption that $f_i \neq f_j$ whenever $i \neq j$, our derivations above clearly manifests that the MLE estimators for the frequencies are the top N frequencies of the power spectrum of the sampled data $\{y[t]\}_{t=0}^{n-1}$.

Consequently, if we denote the top-N frequencies as $\{\hat{f}_k\}_{k=1}^{N}$, then we obtain

$$
\begin{aligned}
\begin{bmatrix} \hat{\alpha}_{MLE} \\ \hat{\beta}_{MLE} \end{bmatrix} =& \hat{\gamma}_{MLE} = \left(H(\hat{f}_{MLE})^\top H(\hat{f}_{MLE})\right)^{-1} H^\top(\hat{f}_{MLE})\vec{y} \\
=& \frac{2}{N}\left[\dots \ \textstyle\sum_{t=0}^{n-1} y(t)\cos\left(2\pi \hat{f}_k t\right) \ \dots \ \textstyle\sum_{t=0}^{n-1} y(t)\sin\left(2\pi \hat{f}_k t\right) \ \dots\right]^\top
\end{aligned}
$$

which indicates that the MLE estimates for the k-th amplitude and initial phase are exactly the amplitude and inverse phase of the Fourier transform of the data y at the k-th top power frequency.

$$
\hat{a}_{k,MLE} = \sqrt{\hat{\alpha}^2_{k,MLE} + \hat{\beta}^2_{k,MLE}} = \sqrt{\mathrm{Re}^2(\vec{y}^\top e^{j2\pi \hat{f}_k \vec{t}}) + \mathrm{Im}^2(\vec{y}^\top e^{j2\pi \hat{f}_k \vec{t}})} = \left|\sum_{t=0}^{n-1} y(t)e^{j2\pi \hat{f}_k t}\right| = |\mathrm{FFT}\{y\}|(\hat{f}_k)
$$

$$\hat{\varphi}_{k,MLE} = -\arctan^{-1}\left(\frac{\hat{\beta}_{k,MLE}}{\hat{\alpha}_{k,MLE}}\right) = -\arctan^{-1}\left(\frac{\mathrm{Im}}{\mathrm{Re}}\left(y^{\top}e^{j2\pi\hat{f}_k\vec{t}}\right)\right) = -\,\mathrm{Angle}\left(\mathrm{FFT}\{y\}(\hat{f}_k)\right)$$