Solutions

1. Consider the following continuous-time signal

$$x_a(t) = \cos(16\pi t) + \sin(8\pi t) \quad -\infty < t < \infty$$

sampled at a sampling rate of $F_s = 1/T = 16\,\text{Hz}$.

(a) (5 points) Find and sketch the sampled signal, $x(n) = x_a(t)|_{t=nT}$.
Find N_p, the fundamental period of the signal in terms of samples.

Solution:

$$\begin{aligned} x(n) &= [\cos(16\pi t) + \sin(8\pi t)]|_{t=\frac{n}{16}} \\ &= \boxed{\cos(\pi n) + \sin(\frac{\pi}{2} n)} \\ &= [\dots \ 1 \ -2 \ \underline{1} \ 0 \ 1 \ -2 \ 1 \ 0 \ \dots] \end{aligned}$$

See Fig. 5
The fundamental period is $N_p = 4$.

(b) (10 points) Find the N_p-point DFT of $x(n)$, starting at $n = 0$.

Solution: 4-DFT of $\cos(\pi n)$ is:

$$[1 \ -1 \ 1 \ -1] \longleftrightarrow [0 \ 0 \ 4 \ 0]$$

4-DFT of $\sin(\frac{\pi}{2} n)$ is:

$$[0 \ 1 \ 0 \ -1] \longleftrightarrow [0 \ -j2 \ 0 \ j2]$$

Therefore, 4-DFT of $x(n)$ is:

$$[0 \ 0 \ 4 \ 0] + [0 \ -j2 \ 0 \ j2] = \boxed{[0 \ -j2 \ 4 \ j2]}$$

(c) (15 points) In Computer Assignment 1, we learned about the concept of zero-padding, a method to improve the frequency resolution by concatenating zeros to the end of the input sequence, and thus taking the transform of a longer sequence. However, inserting the zeros in between the samples seems to be an intuitively better method. Let

$$x_c(n) = [\dots \ 0 \ x(-1) \ 0 \ \underline{x(0)} \ 0 \ x(1) \ 0 \ x(2) \ \dots]$$

Find the $2N_p-$point DFT of $x_c(n)$, starting at $n = 0$.

Solution:

$$x_{c\_\cos}(n) = \begin{bmatrix}1 & 0 & -1 & 0 & 1 & 0 & -1 & 0\end{bmatrix} \longleftrightarrow X_{c\_\cos}(k) = \begin{bmatrix}0 & 0 & 4 & 0 & 0 & 0 & 4 & 0\end{bmatrix}$$

$$\begin{aligned} x_{c\_\sin}(n) &= \begin{bmatrix}0 & 0 & 1 & 0 & 0 & 0 & -1 & 0\end{bmatrix} \\ &= \begin{bmatrix}0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\end{bmatrix} + \begin{bmatrix}0 & 0 & 0 & 0 & 0 & 0 & -1 & 0\end{bmatrix} \end{aligned}$$

$$\longleftrightarrow X_{c\_\sin}(k) = \begin{bmatrix}1 & -\mathrm{j} & -1 & \mathrm{j} & 1 & -\mathrm{j} & -1 & \mathrm{j}\end{bmatrix} - \begin{bmatrix}1 & \mathrm{j} & -1 & -\mathrm{j} & 1 & \mathrm{j} & -1 & -\mathrm{j}\end{bmatrix}$$

Since $x_c(n) = x_{c\_\cos}(n) + x_{c\_\sin}(n)$, $X_c(k) = X_{c\_\cos}(k) + X_{c\_\sin}(k)$.

$$\boxed{X_c(k) = \begin{bmatrix}0 & -j2 & 4 & j2 & 0 & -j2 & 4 & j2\end{bmatrix}}$$

(d) (10 points) In general, the resulting signal from the modified zero padding process described in (c) can be expressed analytically as

$$x_D(n) = \begin{cases} x(\frac{n}{2}), & n \text{ is even} \\ 0 & \text{otherwise} \end{cases}$$

Let $X(\mathrm{e}^{j\omega})$ be the DTFT of $x(n)$. Derive $X_D(\mathrm{e}^{j\omega})$, the DTFT of $x_D(n)$, in terms of $X(\mathrm{e}^{j\omega})$.

Solution:

$$X_D(\mathrm{e}^{j\omega}) = \sum_{n=-\infty,\text{even}}^{\infty} x(\frac{n}{2})\mathrm{e}^{-jn\omega}$$

Let $n' = n/2$. Since n is even, n' will always be an integer.

$$X_D(\mathrm{e}^{j\omega}) = \sum_{n'=-\infty}^{\infty} x(n')\mathrm{e}^{-j2n'\omega} = \sum_{n'=-\infty}^{\infty} x(n')\mathrm{e}^{-j(2\omega)n'} = \boxed{X(\mathrm{e}^{j2\omega})}$$

(e) (10 points) Use the result you got from part (d) to derive $X_D(k)$, the $2N_p-$point DFT of $x_D(n)$. Check if your derivation confirms with your answer in part (c).

Solution:

$$X_D(k) = X_D(\mathrm{e}^{\mathrm{j}\omega})\Big|_{\omega=\frac{2k\pi}{2N_p}} = X(\mathrm{e}^{\mathrm{j}2\omega})\Big|_{\omega=\frac{2k\pi}{2N_p}}$$

Let $\omega' = 2\omega = \frac{2k\pi}{N_p}$. Therefore,

$$X_D(k) = X(\mathrm{e}^{j\omega'})\Big|_{\omega'=\frac{k\pi}{2N_p}} = \boxed{X(k), \quad k = 0, 1, ..., 2N_p - 1}$$

From (c) we see that $X_c(k)$ is $X(k)$ in part(b) with double length, which confirms with our derivation.

2. Suppose you have a signal $x(n)$ consisting of 16384 data points. The magnitude of $X(k) = \text{DFT}\{x(n)\}$ is shown in Fig. 1. The STFT is then performed on $x(n)$ with a Hamming window of length $N_w = 512$ and 50% overlap. The magnitudes and phases of the modulation curves along two of the peaks, $\text{DFT}\{X(n, k_1)\}$ and $\text{DFT}\{X(n, k_2)\}$, are shown in Figs. 2a, 2b, 3a, and 3b, respectively.

   (a) (30 points) Sketch the spectrogram of $x(n)$.

   (b) (10 points) Sketch $x(n)$.

   Hint: The magnitude and phase of 64−point DFT of

$$f(n) = \begin{cases} 1 & 0 \le n \le 31 \\ 0 & 32 \le n \le 63 \end{cases}$$

   are shown in Figs. 4a and 4b. You can assume that Figs.2a, 2b and Figs. 4a, 4b are essentially the same.

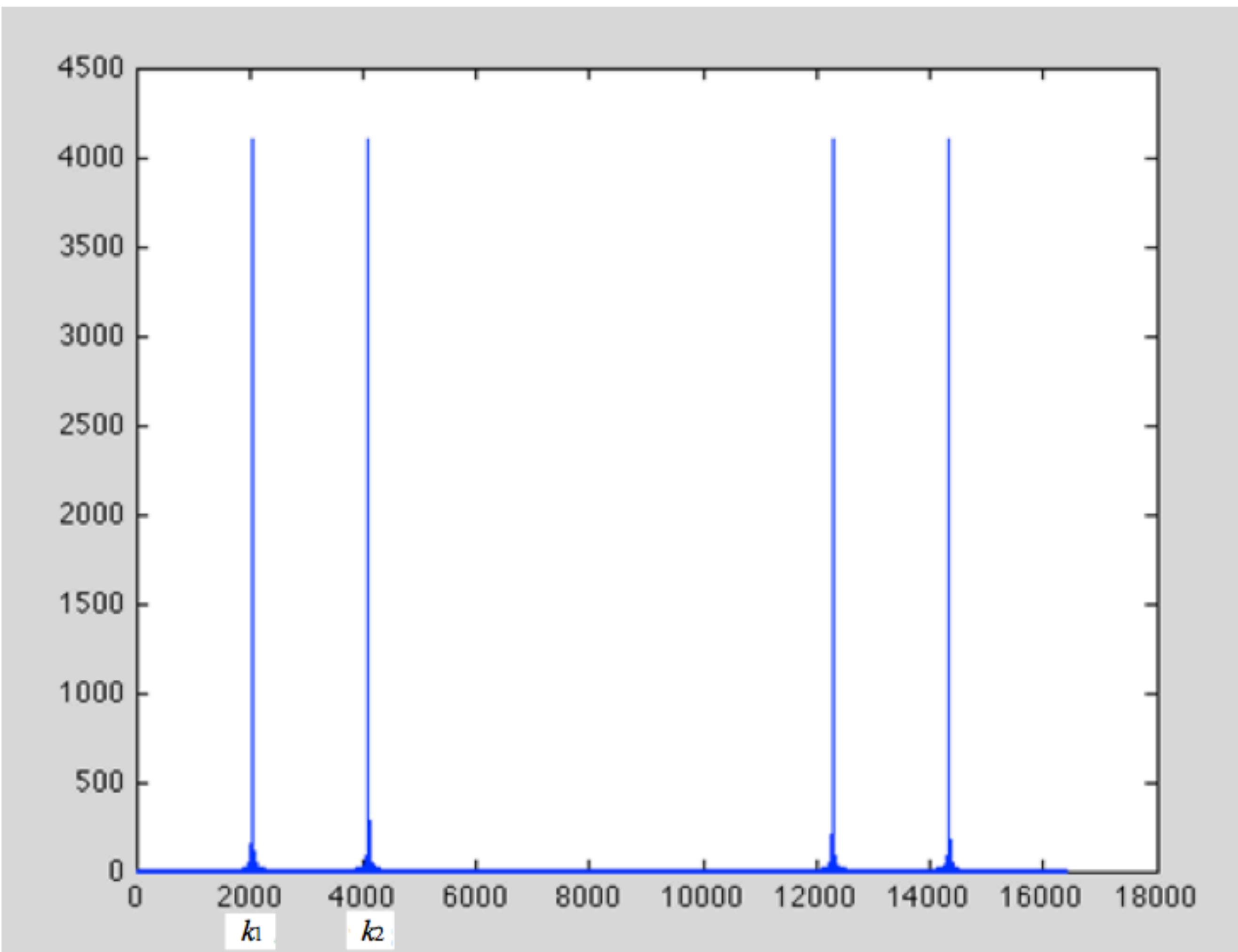
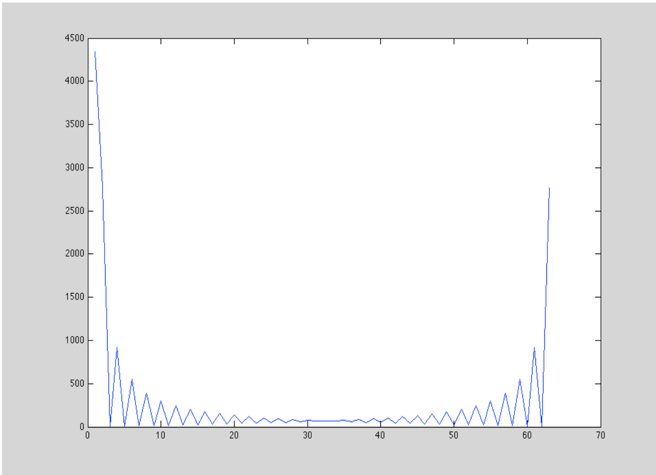
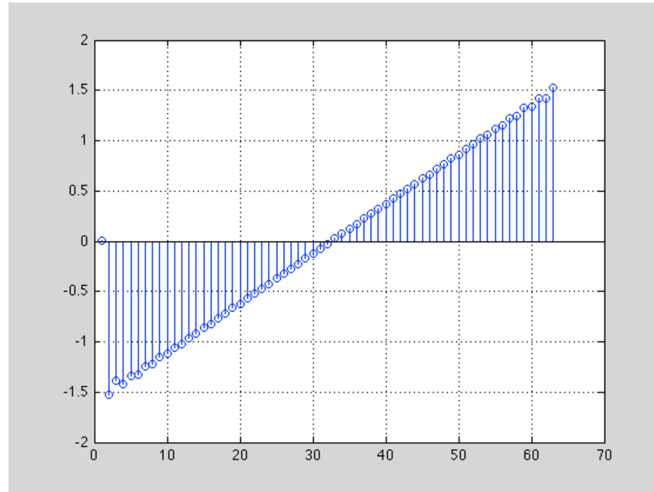


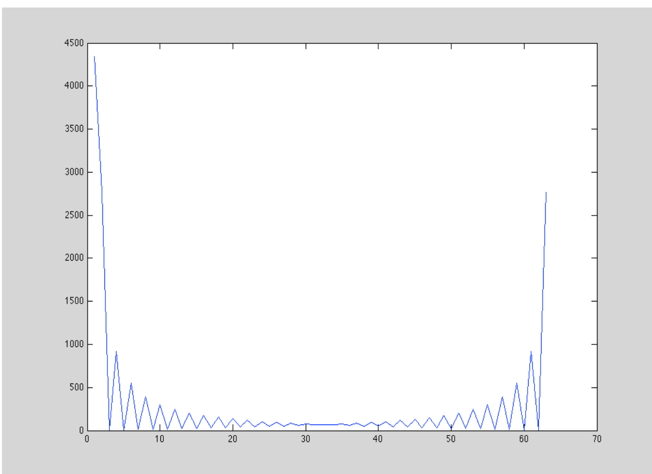
Figure 1: $|X(k)|$.



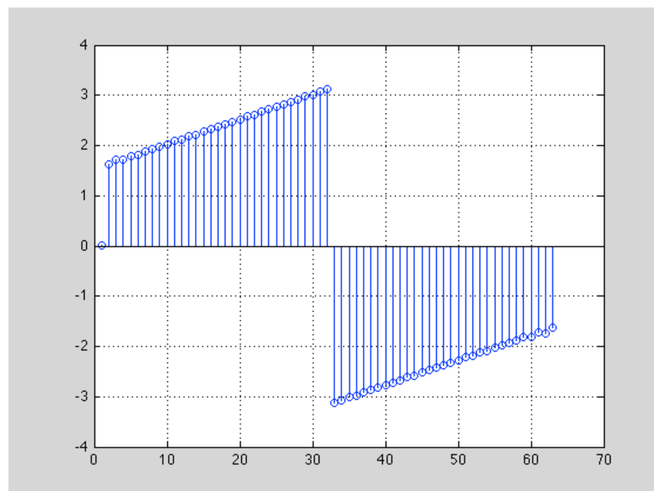
(a) Magnitude of DFT$\{X(n, k_1)\}$.



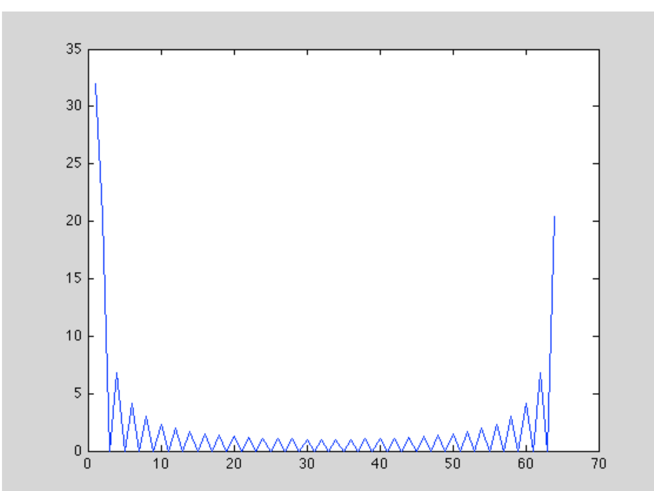
(b) Phase of DFT$\{X(n, k_1)\}$.



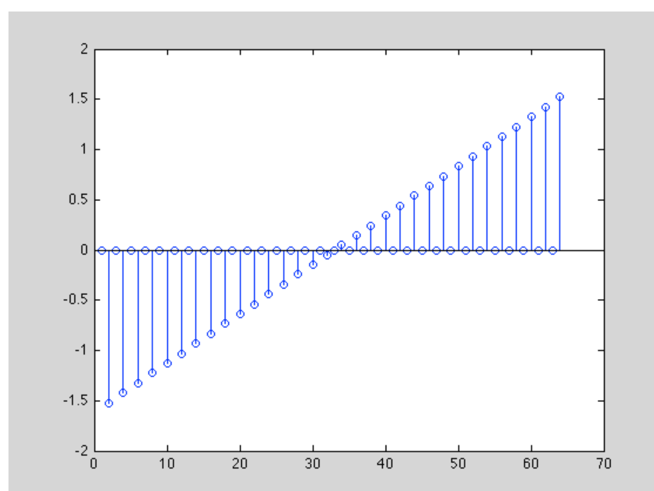
(a) Magnitude of DFT$\{X(n, k_2)\}$.



(b) Phase of DFT$\{X(n, k_2)\}$.



(a) Magnitude of DFT$\{f(n)\}$.



(b) Phase of DFT$\{f(n)\}$.

Solution:

(a) From figure 1 we see clearly that there are only two different main frequencies at $\frac{\pi}{2}$ and $\frac{\pi}{4}$ (some small side lobes around the two frequencies are due to windowing effect). But we can't see where these two frequencies are located with respect to time in the spectrogram. From fig. 2,3,4,5 we see that $X(n, k_1)$ and$X(n, k_2)$ are just two "shifted" rectangular windows. We can think about this as two windowed cosines located at different locations with respect to time. Fig. 6 and 7 indicates the location of first window (from the hint we can assume that fig. 2 and 3 are same as Fig. 6 and 7). By observing fig. 3 and 5, we can see that the phase difference between two windows is . The amount of n-shift can then be found from the property of DFT.

$$
\begin{aligned}
x(n - n_0) &\longleftrightarrow X(k)\mathrm{e}^{-j\frac{2\pi}{N}n_0 k} \\
\frac{2\pi}{N}n_0 &= \pi \\
n_0 &= \frac{N}{2}
\end{aligned}
$$

The location of the second window is then found.

The spectrogram should look like Fig. 6.

(b) From (a) we know that $x(n)$ is a discrete version of the plot in Fig. 7

3. (a) (5 points) Explain in two or three sentences why we need to define a new type of transform, the Short-Time Fourier Transform, for analyzing speech signals.

Solution: The speech signal is time-varying and the DFT alone does not reveal the transitions of the spectral content. The STFT decomposes the speech signal into windowed segments, and the resulting signal can be consider stationary for short enough windows. The standard DFT can therefore be performed to extract useful spectral information.

(b) (10 points) Consider the following continuous-time signal

$$x(t) = \begin{cases} \cos(2000\pi t) & -\infty < t \leq 1.0 \\ \cos(1000\pi t^2) & 1.0 < t < \infty \end{cases}$$

sampled at a sampling rate of $F_s = 8000\,\text{Hz}$. The STFT with a hamming window of length $N_w = 128$ and 50% overlap, defined as

$$X(n, e^{j\omega}) = \sum_{m=-\infty}^{\infty} x(m) w_h(64n - m) e^{-j\omega m}$$

is then performed on the sampled signal. Sketch $|X(64, e^{j\omega})|$ for $0 \leq \omega \leq \pi$.

Solution: Consider Fig. 8:
After discretizing the signal we get:

$$x(m) = \begin{cases} \cos(\frac{\pi}{4} m) & -\infty < m \leq 8000 \\ \cos(\frac{\pi}{64000} m^2) & 8000 < m < \infty \end{cases}$$

$n = 64$ corresponds to the window $(4096 - 64) < m \leq 4096$, $x(m) = \cos(\frac{\pi}{4} m)$. The main lobe bandwidth is

$$\frac{8\pi}{N_w} = \frac{8\pi}{128} = \frac{\pi}{16}.$$

See Fig. 9.

(c) (10 points) Sketch $|X(n, e^{j\frac{\pi}{4}})|$ for $0 \leq n \leq 251$.

Transition of the signal occurs at $m = 8000$, which corresponds to $n = \frac{8000}{64} + 1 = 126$

Solution: Consider the Fig. 10.

(d) (15 points) Sketch the spectrogram of the signal for $0 \leq \omega \leq \pi$ and $0 \leq n \leq 251$. You can ignore the side lobes.

Solution:

$$\omega = \begin{cases} \frac{\pi}{4} & m \leq 8000 \\ \frac{d}{dm}[\frac{\pi}{64000}m^2] = \frac{\pi}{32000}m & m > 8000 \end{cases}$$

the instant frequency ω increases linearly after $m = 8000$
$\omega = \frac{\pi}{4}$ at $m = 8000$ or $n = \frac{8000}{64} + 1 = 126$
$\omega = \frac{\pi}{2}$ at $m = 16000$ or $n = \frac{16000}{64} + 1 = 251$ Consider Fig. 11

4. Consider the signal:

$$s(n) = (-\alpha)^n \text{ for } n \geq 0$$

(a) (10 points) Find the first order prediction coefficient by setting up and solving the normal equation(s) of LPC below:

$$\sum_{n=-\infty}^{\infty} s(n)s(n-i) = \sum_{k=1}^{p} \alpha_k \sum_{n=-\infty}^{\infty} s(n-k)s(n-i)$$

(b) (10 points) Find the corresponding $E_{\min}$ and gain G from the minimum least-square error equation below:

$$E_{\min} = \sum_{n=-\infty}^{\infty} \left(s^2(n) - \sum_{k=1}^{p} \alpha_k s(n-k)s(n) \right)$$

(c) (5 points) If the energy of the source signal is normalized, i.e. $\sum_{n=-\infty}^{\infty} u^2(n) = 1$, explain from the least-square error function

$$E = \sum_{n=-\infty}^{\infty} (s(n) - \hat{s}(n))^2$$

why $E_{\min} = G^2$ when E achieves $E_{\min}$.

(d) (15 points) The current sample of the signal $s(n)$ can be perfectly predicted from the past sample as $s(n) = (-\alpha)s(n-1)$, which implies that the signal $s(n)$ can be perfectly predicted by the estimator $\hat{s}(n)$ with first order prediction. If so, should the minimum least-square error $E_{\min}$ be zero? If yes, explain why the all-pole model

$$V(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \alpha_1 z^{-1}}$$

with $G = 0$ makes sense. If No, explain why $E_{\min} \neq 0$ even though $\hat{s}(n)$ can perfectly predict $s(n)$

Solution:

(a) Let $i = 0$, then

$$\sum_{n=-\infty}^{\infty} s(n)s(n-i) = \sum_{n=-\infty}^{\infty} s^2(n) = \frac{1}{1-(-\alpha)^2}$$

$$\sum_{n=-\infty}^{\infty} s(n-k)s(n-i) = \sum_{n=-\infty}^{\infty} s(n-1)s(n) = \frac{-\alpha}{1-(-\alpha)^2}$$

Therefore,

$$\alpha_1 = \frac{-\alpha}{1-(-\alpha)^2} \frac{1-(-\alpha)^2}{1} = \boxed{-\alpha}$$

(b)

$$E_{\min} = \frac{1}{1-(-\alpha)^2} - (-\alpha)\frac{-\alpha}{1-(-\alpha)^2} = \boxed{1} = G^2 = G$$

(c)

$$s(n) - \hat{s}(n) = \sum_{k=1}^{p} a_k s(n-k)s(n) + Gu(n) - \sum_{k=1}^{p} a_k s(n-k)s(n) = Gu(n)$$

Therefore,

$$E_{\min} = \sum_{n=-\infty}^{\infty} G^2 u^2(n) = G^2 \sum_{n=-\infty}^{\infty} u^2(n) = G^2$$

(d) LPC can not predict the very first sample $s(0)$ out of nothing. So there must be a finite error between $\hat{s}(n)$ and $s(n)$. The missing term from the prediction is $(-\alpha)^0 = 1$, which corresponds to $Gu(n)$ with $G = 1$ and $u(n) = \delta(n)$.

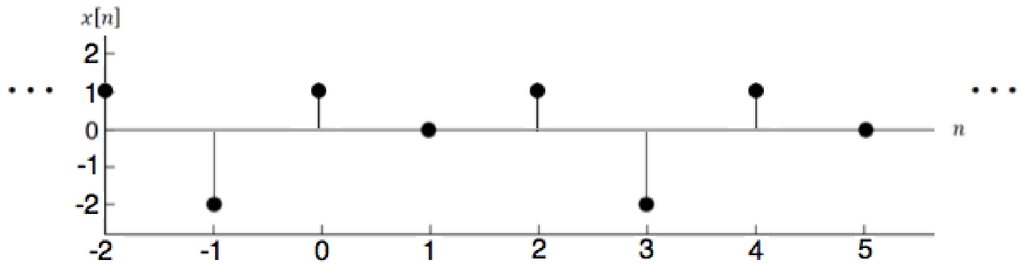


Figure 5: Problem 1a

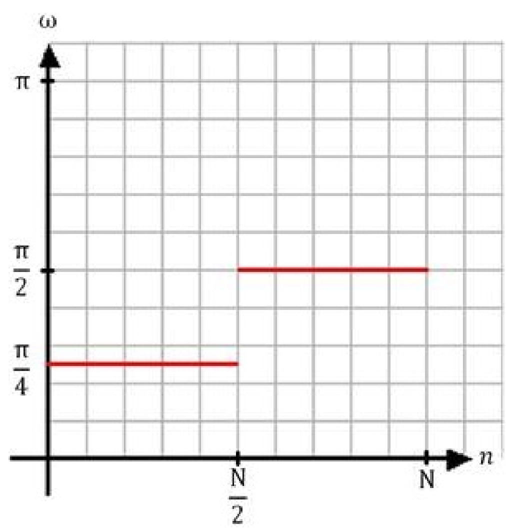


Figure 6: Problem 2a

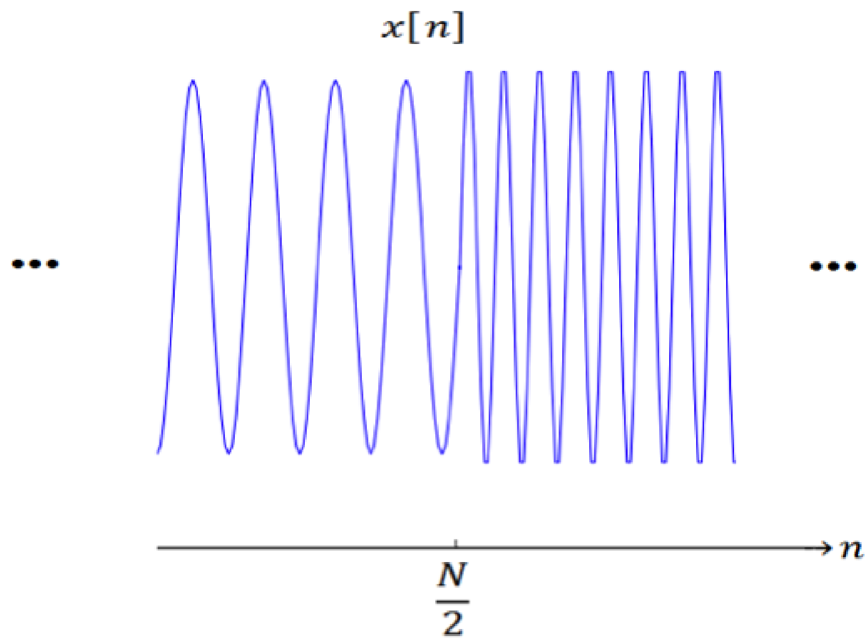


Figure 7: Problem 2b

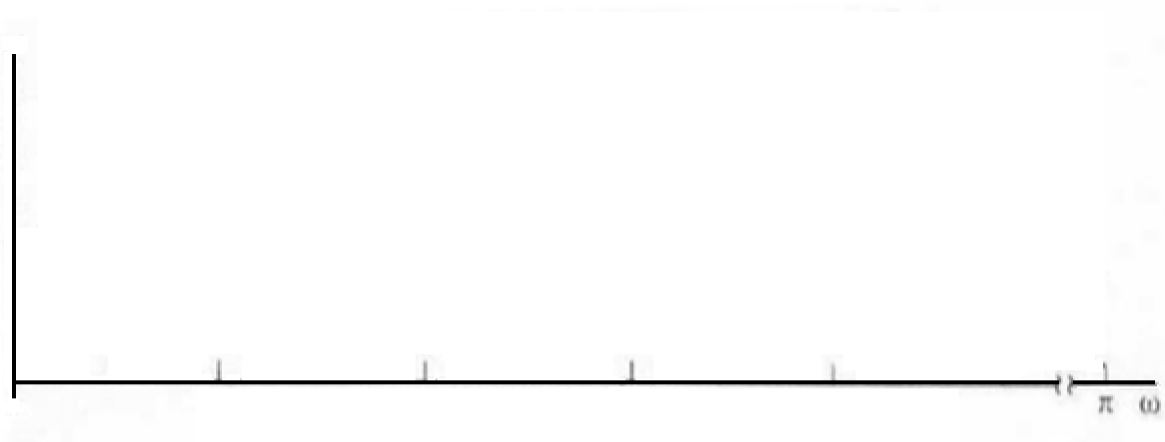


Figure 8: Problem 3b

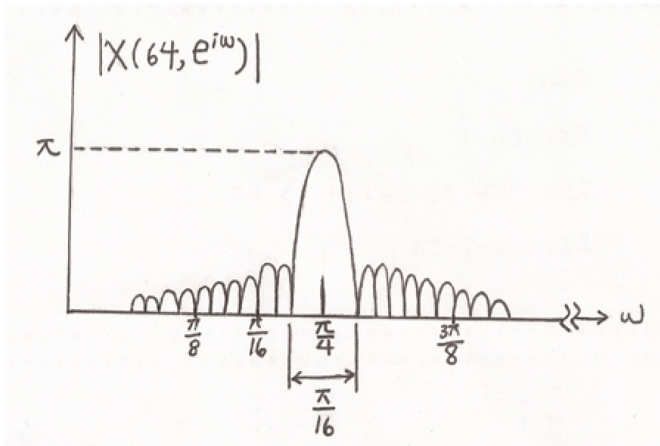


Figure 9: Problem 3b

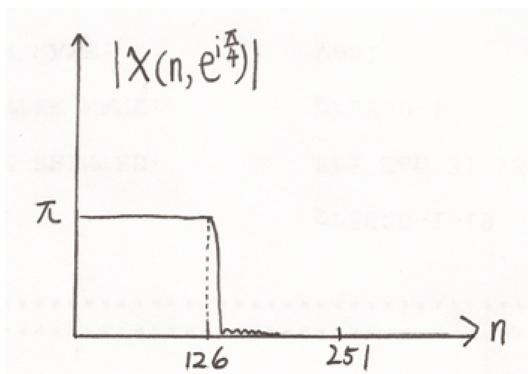


Figure 10: Problem 3c

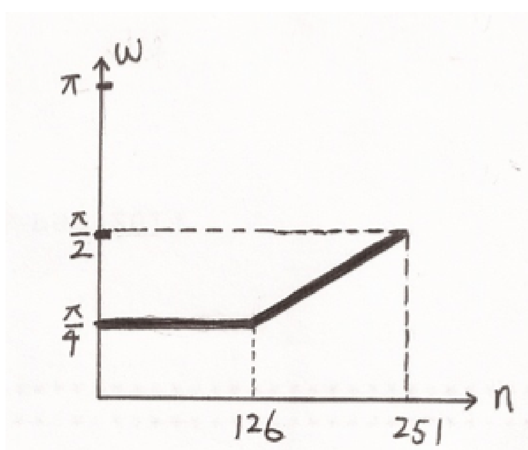


Figure 11: Problem 3d