1. (30 points) **Sampling from Hell:** There's a minor problem with the sampling theorem as usually presented (and we've been no exception). This problem will help you correct it.

(a) (10 points) A signal \( m(t) \) has bandwidth 20kHz. What minimum sampling rate \( f_s \) will allow \( m(t) \) to be recovered from the samples \( \{m(kf_s)\} \) where \( k \) is an integer?

**SOLUTION:** The Nyquist sampling theorem is usually stated as \( f_s \geq 2W \) where \( W \) is the bandwidth of the signal. So one would naturally say \( f_s \geq 40 \text{kHz} \).

(b) (10 points) For \( A \) some constant, we sample \( s(t) = A \sin(2\pi f_0 t) \) at times \( t \in \{k \frac{2f_0}{f_s}\} \). What are the sample values? Can \( s(t) \) be reconstructed from these samples?

**SOLUTION:** All the samples are zero, so there's no way to unambiguously reconstruct the signal.

(c) (10 points) Sketch the spectrum of

\[
q(t) = s(t) \sum_k \delta(t - \frac{k}{f_s})
\]

for \( s(t) = \sin(2\pi f_0 t) \) where \( f_s \) is the sampling rate. Identify the problem with the sampling theorem implied by the previous part and then rewrite the sampling theorem correctly.

**SOLUTION:**

\[
Q(f) = f_s S(f - kf_s) = \frac{f_s}{2f_0} \sum_k (\delta(f - f_0 - kf_s) - \delta(f + f_0 - kf_s))
\]

When \( f_s = 2f_0 \) we have

\[
Q(f) = \frac{f_0}{f} \sum_k \left( \delta(f - (2k + 1)f_0) - \delta(f + (1 - 2k)f_0) \right)
\]

or

\[
Q(f) = \frac{f_0}{f} \left( \sum_k \delta(f - (2k + 1)f_0) - \sum_k \delta(f - (2k - 1)f_0) \right) = 0
\]

That is, the spectra cancel out – because all their energy is right at the band edge where they will overlap if \( f_s = 2f_0 \). When we derived the Nyquist theorem, the picture we drew implied that the signal had zero energy at the band edge.

Therefore, the correct statement of Nyquist is \( f_s > 2W \) – i.e., the sampling rate is strictly greater than twice the bandwidth.
2. (30 points) Quantization from Hell:

The PDF of signal levels attained by a signal \( x(t) \) is

\[
f_X(x) = \frac{1}{4}\delta(x + 3) + \frac{1}{4}\delta(x + 1) + \frac{1}{4}\delta(x - 1) + \frac{1}{4}\delta(x - 3)
\]

(a) (10 points) Please derive an optimal 2-bit quantizer for the signal \( x(t) \). What is the expected error of your quantization function?

**SOLUTION:** The signal takes on exactly 4 values, so it’s obvious that \( q_0 = -3, q_1 = -1, q_2 = 1 \) and \( q_3 = 3 \). The bins almost don’t matter, but using Lloyd-Max we have \( x_k = \frac{q_k + q_{k+1}}{2} \) so, \( x_1 = -2, x_2 = 0, \) and \( x_3 = 2 \). The error is zero because the only four possible values of \( x(t) \) are perfectly represented.

(b) (20 points) Please derive an optimum 1-bit quantizer for \( x(t) \). What is the expected error of your quantization function?

**SOLUTION:** We have a single break point (two bins) and we need to satisfy Lloyd Max:

\[
\frac{q_k + q_{k+1}}{2} = x_k
\]

and

\[
q_k = E[X | X \in (x_{k-1}, x_k)]
\]

As always, we assume \( x_0 = -\infty \) and \( x_L = \infty \) (where \( L \) is the number of levels).

In this case, the conditional means are simple as a function of \( x_1 \).

- \( x_1 < -3 \): \( q_0 \) undefined, \( q_1 = 0 \)
- \( x_1 \in (-3, -1) \): \( q_0 = -3, q_1 = 1 \)
- \( x_1 \in (-1, 1) \): \( q_0 = -2, q_1 = 2 \)
- \( x_1 \in (1, 3) \): \( q_0 = -1, q_1 = 3 \)
- \( x_1 > 3 \): \( q_0 = 0, q_1 \) undefined

The only choice consistent with the other Lloyd-Max condition \( \frac{q_k + q_{k+1}}{2} = x_k \) is \( x_1 \in (-1, 1) \) when we choose \( x_1 = 0 \) – which is the intuitively obvious choice for such a symmetric signal PDF.
3. (40 points) Phase Locked Loop from Hell:
An unlabeled block diagram of a phase-locked loop is shown in FIGURE 1. $\dot{\hat{\theta}}(t)$ is the derivative of $\hat{\theta}(t)$, an estimate of $\theta(t)$.

(a) (10 points) What is the block labeled A?
SOLUTION: Voltage controlled oscillator. Its input is integrated and used as the argument to the output sinusoid – in this case cosine.

(b) (10 points) What is the block labeled B?
SOLUTION: B is a low pass filter.

(c) (10 points) What is the block labeled C?
SOLUTION: C is a high gain inverting amplifier.

(d) (10 points) Explain the operation of the phase locked loop in detail.
SOLUTION:
Qualitatively: (3pts): The basic idea is a negative feedback loop which drives $e(t)$ toward zero and leaves $\dot{\hat{\theta}}(t) \approx \theta(t)$.
Quantitatively: (7pts): The output of the high gain amplifier is $-ce(t)$ which leads to
$$-ce(t) = -c\theta(t) + c\dot{\hat{\theta}}(t) = \dot{\hat{\theta}}(t)$$

or
$$-\frac{1}{c}\dot{\hat{\theta}}(t) + \hat{\theta}(t) = \theta(t)$$

With $c$ negative, the left hand differential equation is stable so transient $\hat{\theta}(t)$ die away. Since $c$ is large we have
$$\dot{\hat{\theta}}(t) \approx \theta(t)$$

which is what we want in a demodulator!

4. (50 points) Cora and Marty, the Demon Squirrel from Hell: Cora the Communications Engineer needs to guard against Marty, the demon squirrel from hell. Marty pops up at position $X_n$ and sets fire to whatever’s there. Cora has to put the fire out, but her hose is heavy and if it’s pointed in the wrong direction, she won’t douse the flames in time. So, Cora needs to predict where Marty will be. It should come as no surprise that Cora will use a linear predictor structure for this purpose.

Unbeknownst to Cora, Marty is a creature of habit whose firing positions follow
$$X_n = \frac{1}{2}X_{n-1} + G_n$$

where the $\{G_n\}$ are i.i.d. zero mean Gaussian random variable with unit variance. Cora will use a simple one-step estimate $\hat{X}_n = wX_{n-1}$ you need to provide her with the $w$ which minimizes $\epsilon^2 = E[(X_n - \hat{X}_n)^2]$.

(a) (10 points) Please derive an expression for Marty’s position based on all past inputs $G_k$ and his initial position $X_0$. Show that in the limit of large $n$, Marty’s initial position is irrelevant.
SOLUTION: Two ways to do this – using the result from linear system theory (which you probably do not remember) or by directly calculating

\[
X_n = 0.5X_{n-1} + G_n \\
= 0.5 (0.5X_{n-2} + G_{n-1}) + G_n \\
= 0.5 (0.5 (0.5X_{n-3} + G_{n-2}) + G_{n-1}) + G_n
\]

which we see to be

\[
X_n = \sum_{k=0}^{n-1} (0.5)^k G_{n-k} + (0.5)^n X_0
\]

(b) (10 points) \(X_n\) is a sequence of random variables owing to the random inputs \(G_n\). For very large \(n\), what is \(E[X_n]\), what is \(E[X_n^2]\), what is \(f_{X_n}(x)\)?

SOLUTION: \(X_n\) is a sum of Gaussians so \(X_n\) is Gaussian. The term in \(X_0\) goes to zero geometrically in \(n\), so for large \(n\) we can ignore it. \(E[X_n]\) is therefore zero because all the \(G_k\) have mean zero. \(E[X_n^2]\) is the sum of the squares of the individual terms in the sum because they are zero mean and independent. The mean square of each term \((0.5)^k G_{n-k}\) is \((0.25)^k\) since the \(G_k\) are zero mean and unit variance. Thus, in the limit of large \(n\), the total variance is \(\sum_{k=0}^{\infty} (0.25)^k = \frac{4}{3}\). So \(X_n \sim \mathcal{N}(0, \frac{4}{3})\)

(c) (10 points) Find an expression for \(E[X_n X_{n-\ell}]\) the correlation function of the random sequence \(X_n\). Show that \(E[X_n X_{n-\ell}] = R_X(\ell)\) does not depend on \(n\) for \(n\) large.

SOLUTION: We know that

\[
X_n = (0.5)\ell X_{n-\ell} + \sum_{k=0}^{\ell} (0.5)^k G_{n-k}
\]

so

\[
E[X_n X_{n-\ell}] = E \left[ \left( (0.5)\ell X_{n-\ell} + \sum_{k=0}^{\ell} (0.5)^k G_{n-k} \right) X_{n-\ell} \right] = (0.5)^\ell E[X_{n-\ell}^2]
\]

because \(X_{n-\ell}\) is independent of the subsequent \(G_{n-k}\), \(k = 0, 1, \ldots, \ell\). Since \(E[X_n^2] = E[X_{n-\ell}^2] = \frac{4}{3}\), we then have

\[
R_X(\ell) = (0.5)^\ell \frac{4}{3}
\]

which obviously does not depend on \(n\).

(d) (10 points) Cora can form a good estimate of the correlation function \(R_X(\ell)\). What \(w\) minimizes \(\varepsilon^2\)? Is it the same as or different from \(\frac{1}{2}\), the coefficient from Marty’s equation of motion? What is the resultant error, \(\varepsilon^2\).

SOLUTION:

\[
\varepsilon^2 = E \left[ (X_n - wX_{n-1})^2 \right]
\]

Taking the derivative w.r.t. \(w\) and setting to zero gives us

\[
R_X(1) = wR_X(0)
\]

Taking the second derivative (just to be sure about convexity) gives us \(R_X(0) > 0\) so the optimization is indeed convex. Finally we evaluate \(w\) as

\[
w = \frac{R_X(1)}{R_X(0)} = \frac{(0.5)^{\frac{1}{2}}}{\frac{4}{3}} = \frac{1}{2}
\]
which is the same as Marty's coefficient – quite a relief since we’d expect our estimates to mirror reality!

As for the error:

\[ \varepsilon^2 = E \left[ (X_n - 0.5X_{n-1})^2 \right] = E[X_n^2] + E[X_{n-1}^2]/4 - E[X_nX_{n-1}] = \frac{5}{3} - \frac{2}{3} = 1 \]

(e) (10 points) What are the weights \( w_1 \) and \( w_2 \) if Cora uses a two-step predictor \( \hat{X}_n = w_1X_{n-1} + w_2X_{n-2} \)? What is the error, \( \varepsilon^2 \)?

**SOLUTION:** You can use the formulae or start from scratch as above. Either way, we end up with two equations in two unknowns.

\[
\begin{align*}
R_X(1) &= R_X(0)w_1 + R_X(1)w_2 \\
R_X(2) &= R_X(1)w_1 + R_X(0)w_2
\end{align*}
\]

We could invert this implicit matrix equation or just do the usual elimination to obtain

\[
R_X(0)R_X(1) - R_X(1)R_X(2) = (R_X^2(0) - R_X^2(1))w_1
\]

so that

\[
w_1 = R_X(1) \frac{R_X(0) - R_X(2)}{R_X(0) - R_X(1)} = \frac{4143}{3234} \cdot \frac{16}{9} = \frac{1}{2}
\]

which means

\[
w_2 = \frac{1}{R_X(1)} \left[ R_X(1) - R_X(0) \frac{1}{2} \right] = 0
\]

which is exactly the same estimator as before and has the same error as before.

We could have guessed this, by the way. The equations of motion explicitly state that if you’re given \( X_{n-1} \), the next step depends ONLY on \( X_{n-1} \) and \( G_n \). So it makes sense to ignore and previous values of position other than \( X_{n-1} \).