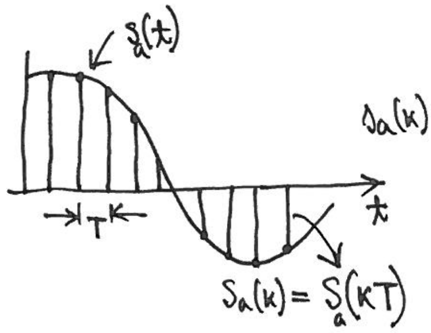
Quantization

- In nature, we mostly observe continuous analog signals. For example, speech, image, video.
- Using a sampler, we first sample the continuous signal to a discrete signal. But samples of a discrete signal has infinite resolution (analog).
- For digital processing, we must need a finite bit budget. Or, a finite resolution representation. So, <u>quantization</u> comes into picture.

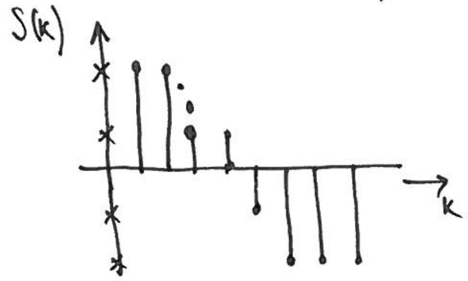


$s_a(t)$: a continuous time domain signal.

$s_a(k) = s_a(kT)$: sampled at each time interval T.

$s(k) = Q\left[s_a(kT)\right] = Q\left(s_a(k)\right)$: If b bits are available, then $s(k)$ can take one out of 2^b different discrete values which are in the range of highest and lowest values of $s_a(k)$.

Let us take $b=2$, and so only 4 values allowed.



$q(k) = x(k) - \hat{x}(k)$ = quantization noise.

$\hat{x}(k) = Q\left[x(k)\right]$ = quantized value of $x(k)$.

Mean square error $(MSE) = \left(\sum_{\forall k} q^2(k)\right)\frac{1}{\{\# k\}} = \mathcal{E}\{q^2(k)\}$

Signal-to-quantization-noise-ratio $(SQNR) = \dfrac{\sigma_x^2}{\sigma_q^2} = \dfrac{\mathcal{E}\{x^2\}}{\mathcal{E}\{q^2\}}$

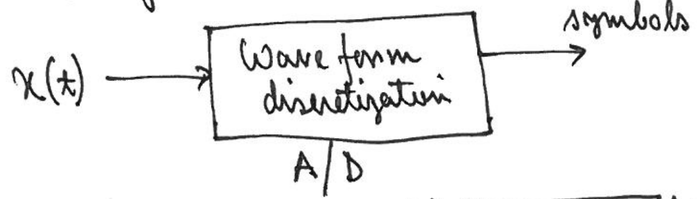
(assume both $x(k)$ and $q(k)$ are zero mean signals).

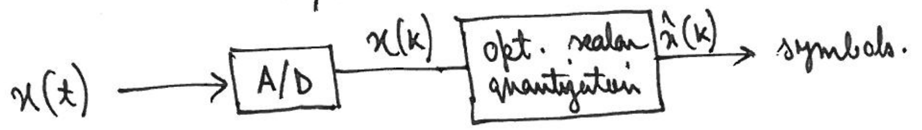
What we want? ⓐ σ_q^2 as minimum as possible given bit budget.

<u>other way</u> ⓑ Bit budget as minimum as possible given a maximum allowable σ_q^2.

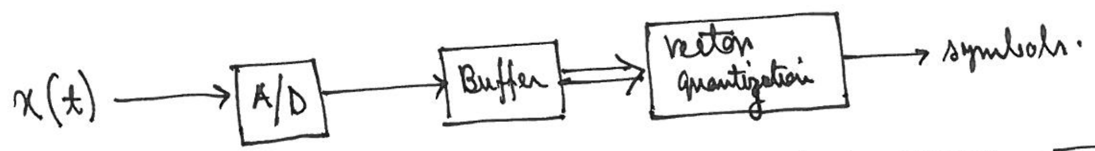
- An understanding of channel symbols (source codes over channel) Explain

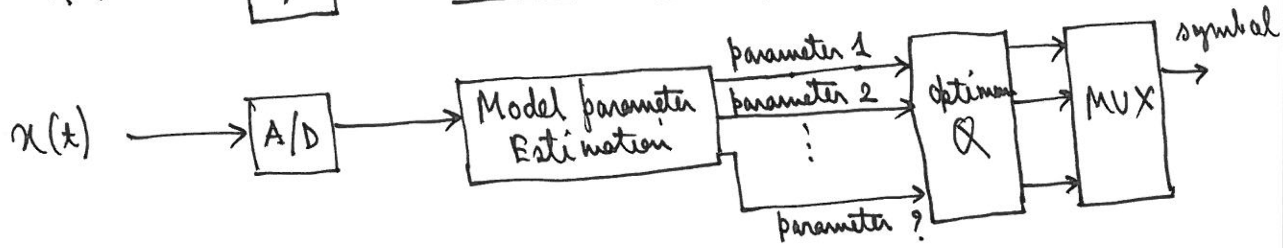
Some coding schemes:

①


②


Scalar quantization

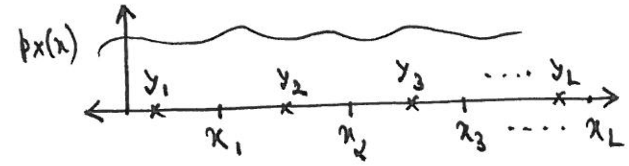
③


④


Scalar Quantization

- Mapping of x to a finite set of symbols. x is real valued. $-\alpha \leq x \leq +\alpha$

- Intervals of x: I_K, $1 \leq K \leq L$
  (finite no) I_K: $x_{K-1} \leq x \leq x_K$



- Reconstructed value for each interval I_K : y_K (real valued quantity)
  (Quantized)

$$y = Q(x) \quad : \text{quantization of } x \text{ to } y.$$

$$y \in \{y_K\}_{K=1}^{L}$$

$$q = x - Q(x) = x - y.$$

- Minimize σ_q^2 for a fixed L. $(L = 2^b)$

- $$\sigma_q^2 = \mathcal{E}\{q^2\} = \mathcal{E}\{(x-y)^2\}$$
$$= \int_{-\alpha}^{+\alpha} (x-y)^2 \, p_X(x) \, dx \qquad [p_X(x) : \text{pdf of } x]$$
$$= \sum_{K=1}^{L} \int_{x_{K-1}}^{x_K} (x-y_K)^2 \, p_X(x) \, dx$$

- **Optimal quantizer**: Solve $\{x_K, y_K \; ; \; 1 \leq K \leq L\}$ such that σ_q^2 is minimum

(4)

<u>Design of optimal quantizer</u>

- For most of the pdfs, there is no closed form solution for the ~~pdfs~~ parameters $\{x_k, y_k\}$ to design an optimal quantizer.

- Lloyd-Max Algo : Two steps (Iteration)

  (1) solve for $\{y_k\}$ given $\{x_k\}$.

  (2) solve for $\{x_k\}$ given $\{y_k\}$.

Theorem: $\{x_k^t, y_k^t, \; 1 \leq k \leq L\}$ leads to lower σ_q^2 than the case $\{x_k^{t-1}, y_k^{t-1}, \; 1 \leq k \leq L\}$ where t is the iteration index.

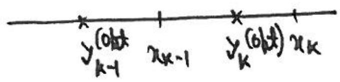
$$\sigma_q^2 = \sum_{k=1}^{L} \int_{x_{k-1}}^{x_k} (x - y_k)^2 \, p_x(x) \, dx.$$

Step 1 :
$$\frac{\partial \sigma_q^2}{\partial y_k} = 0 \quad , \; \forall k .$$

or,
$$\int_{x_{k-1}}^{x_k} 2(x - y_k)(-1) \, p_x(x) \, dx$$

or,
$$y_k^{(opt)} = \frac{\int_{x_{k-1}}^{x_k} x \, p_x(x) \, dx}{\int_{x_{k-1}}^{x_k} p_x(x) \, dx} \rightarrow \text{centroid} \quad \text{- - - (a)}$$

Step 2 : Given $\{y_k^{(opt)}\}_{k=1}^{L}$.



For $x > x_k$, $(x - y_k)^2 < (x - y_{k-1})^2$

or, $x^2 - 2x y_k + y_k^2 < x^2 - 2x y_{k-1} + y_{k-1}^2$

or, $2x y_k - 2x y_{k-1} + y_{k-1}^2 - y_k^2 > 0$

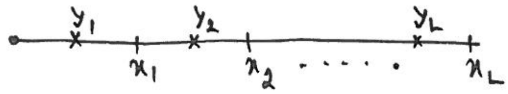
or, $2x(y_k - y_{k-1}) - (y_k + y_{k-1})(y_k - y_{k-1}) > 0 \quad [\because y_k - y_{k-1} > 0]$

or, $2x - (y_k + y_{k-1}) > 0$

or, $x > \frac{1}{2}(y_k + y_{k-1})$

or, $\boxed{x_{k-1}^{(opt)} = \frac{1}{2}(y_k + y_{k-1})}$ - - - - (b)

~~Sequential~~ Lloyd-Max (LM) Algorithm



0. Initialize : choose $\{y_k^0\}$, $t = 1$
1. solve for $\{x_k^*\}_{k=1}^{L}$ using equation (b).
2. Solve for $\{y_k^*\}$, using equation (a).
3. Till convergence.

$$\boxed{\text{Optimal Quantization} \xrightarrow[\text{leads to}]{} \text{Non-uniform Quantization for an arbitrary pdf}}$$

Some properties for Opt. Quantization

1. Mean

$$E[q] = E\{x - Q(x)\} = \sum_{k=1}^{L} \int_{x_{k-1}}^{x_k} (x - y_k)\, p_X(x)\, dx$$

For opt. Q.

$$y_k = \frac{\int_{x_{k-1}}^{x_k} x\, p_X(x)\, dx}{\int_{x_{k-1}}^{x_k} p_X(x)\, dx}$$

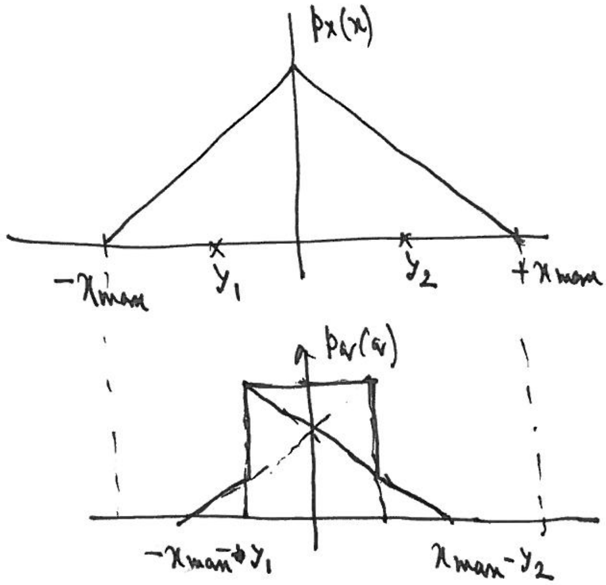
$$\therefore\ E[q] = \sum_{k=1}^{L} \left[\int_{x_{k-1}}^{x_k} x\, p_X(x)\, dx - \int_{x_{k-1}}^{x_k} x\, p_X(x)\, dx \right] = 0$$

$$\boxed{\text{Quantization noise } q \text{ is always zero mean.}}$$

2 PDF of q

$$
\begin{aligned}
q &= x - Q(x) \\
&= x - y_k \quad \text{when } x_{k-1} \leq x \leq x_k \quad (x \in I_k), \ \forall k.
\end{aligned}
$$

$$
\begin{aligned}
p_q(q) &= \sum_{k=1}^{L} P(x \in I_k) \; p_q(q = x - y_k \mid x \in I_k) \\
&= \sum_{k=1}^{L} P(\cancel{x \in I_k}) \cdot \frac{p_q(q = x - y_k, x \in I_k)}{P(\cancel{x \in I_k})} \\
&= \sum_{k=1}^{L} p_q(q = x - y_k, \ x \in I_k) \\
&= \sum_{k=1}^{L} p_x(x = q + y_k, \ x \in I_k)
\end{aligned}
$$

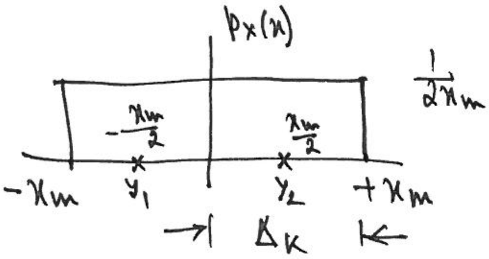


$y_2 = -y_1$

- As L increases, $p_q(q)$ becomes more narrow.
- Bounded $p_x(x) \Rightarrow$ bounded $p_q(q)$.
- In general $p_q(q)$ is not-uniform. But for large L, it may be approximated as uniform.
- If $p_x(x)$ is uniform, then $p_q(q)$ is uniform.

Symmetric Uniform PDF :

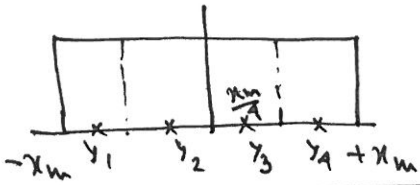
1 bit



$$\Delta_K = \frac{x_m - (-x_m)}{2} = \frac{2x_m}{2}$$

$$\sigma_q^2 = \int_{-x_m}^{0} \left(x - \left(\frac{-x_m}{2}\right)\right) p_x(x)\, dx + \int_{0}^{x_m} \left(x - \frac{x_m}{2}\right) p_x(x)\, dx.$$

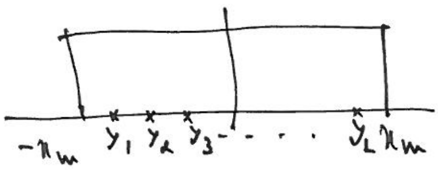
2 bit



$$\Delta_K = \frac{2x_m}{4}$$

$\boxed{\text{H.W. Can we find } \sigma_q^2 = ?}$

$L = 2^b$ bit



Now, $\sigma_x^2 = \frac{x_m^2}{3}$.

H.W.

check: $SQNR = L^2 = 2^{2R}$

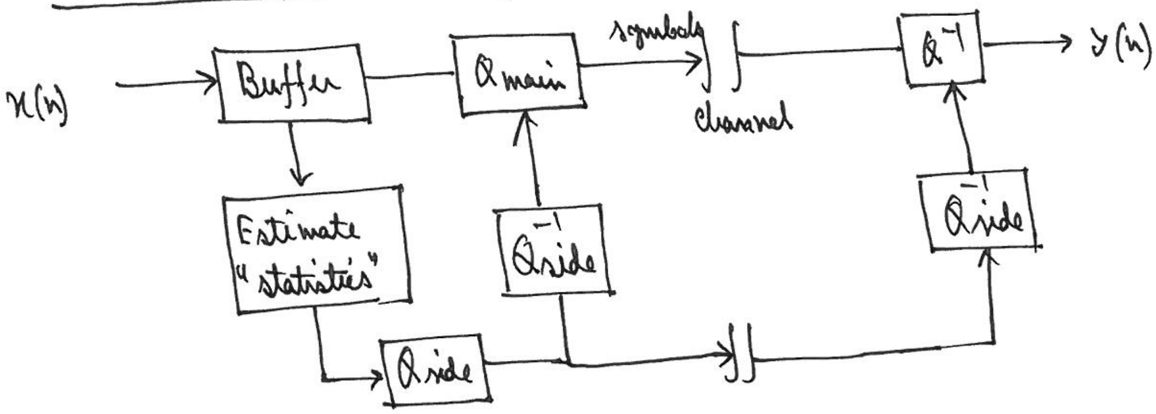
$$SQNR\big|_{dB} = 10 \log_{10} L^2 = 10 \log_{10} 2^{2R} = 6.02R \text{ dB}$$

I think this result should come. Please Check it. (H.W.).

Adaptive Quantization

- Suppose $p_x(x)$ changing with "time"
- But all our design tricks are ~~mainly~~ mainly for stationary pdf.

// Forward Adaptation (AQF)



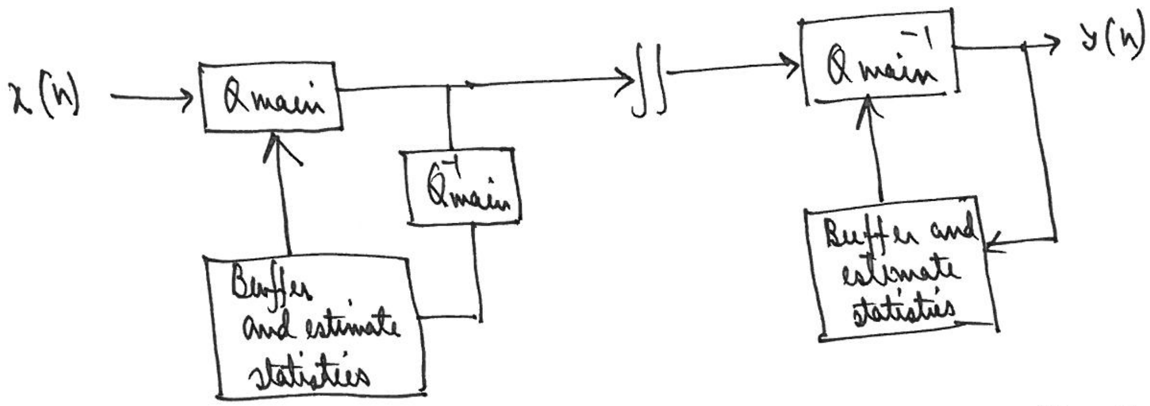
Total Rate = Rmain + Rside

Buffer delay of N samples.

N ⟶ number of learning samples.

Rside $\propto \frac{1}{N}$

// Backward Adaptation (AQB)



Receiver is more complex than AQF

Delay: NO (one sample)

Statistics is approximated by quantized samples : To have synchrony between Tx and Rx

N is not a problem.