

Lecture 8 Quantization

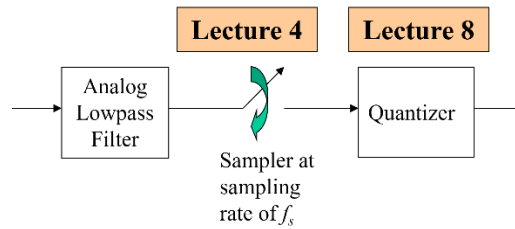
April 20, 2026

Notes by Jenna May

Quantization

- Interpretation of continuous quantity by finite set of discrete values
- Used in analog to digital conversion

Analog to Digital Conversion

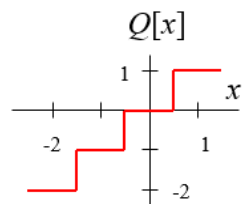


- The anti-aliasing analog continuous-time lowpass filter has stopband frequency less than $\frac{1}{2} f_s$ to reduce aliasing from sampling at a rate of f_s (*enforces sampling theorem*)
- Example tradeoffs in sampling rate and amplitude resolution for analog-to-digital converters a.k.a. digitizers are shown below. The sampling rate decreases down the rows, which means the sampling time increases. Having a longer sampling time allows more time to convert the analog amplitude into a digital approximation which leads to more bits in resolution.

Analog-to-Digital Converter Channels	Bits/sample	Sampling rate	Signal Bandwidth	
NI PXIe-5186	2	8	12.5 GSPS	5 GHz
NI PXI-5154	2	8	2 GSPS	1 GHz
NI PXIe-5122	1	14	100 MSPS	
NI PXI-5192	1	16	15 MSPS	
NI PXI-5192	1	24	500 KSPS	

Uniform Amplitude Quantization

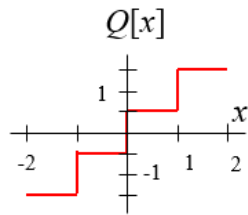
- Round to nearest integer
 - Can use two's complement representation {10, 11, 00, 01} or unsigned representation



$$\Delta = \frac{1 - (-2)}{2^2 - 1} = \frac{3}{3} = 1$$

- "Linear" region for the response is over x from -2 V to 1 V

- Rounding with offset
 - Average voltage is 0
 - No true 0 value in the output



- Step size Δ is 1 V computed as $(1.5 \text{ V} - (-1.5\text{V})) / 3 = 1 \text{ V}$.
- “Linear” region for the response is from -1.5 V to 1.5 V
- Input $x = 0 \text{ V}$, output is 0.5 V. Quantization error is 0.5V which is $\Delta/2$.
- Input $x = 0.9999 \text{ V}$, output is 0.5V. Quantization error is -0.4999 V.
- Quantization error: difference between output voltage and input voltage
- Over the “linear” region, the quantization error is from $-\Delta/2$ to $\Delta/2$
- If x is uniformly distributed, then the quantization error over the “linear” region is uniformly distributed from $-\Delta/2$ to $\Delta/2$ and has noise power σ^2 because the mean is zero.

Signal to Noise Ratio (SNR)

- SNR is the Signal to Quantization Noise Ratio (SQNR)
- Upper bound on SNR for audio for B bits: $\text{SNR} = 1.76 \text{ dB} + 6.02 \text{ dB/bit} * B$
- The upper bound on SNR is more accurate as the number of bits B increases
- CD audio uses 16 bits per sample: $\text{SNR} = 98.08 \text{ dB}$. In practice, audio CDs give 95 dB.

Quantization Noise

- The input to the quantizer includes desired signal plus thermal noise.
- If we use more bits than the SNR at the quantizer input would support, then the least significant bits are quantization the thermal noise and not the desired signal.
- If we use fewer bits than the SNR at the quantizer input would support, then we would lose resolution in the desired signal.
- The ideal setting is to set the number of bits in the quantizer so that the quantization noise power is equal to the thermal noise power.

How Many Quantization Bits Are Reliable? Slide 8-13.

- The input to the quantizer includes desired signal plus thermal noise, as discussed above in the Quantization Noise section
- Given the SNR at the input to the quantizer, we can work backwards to see how many bits B would represent the desired signal by using $\text{SNR} = 2 \text{ dB} + 6 B$.
- Technical specification for the [A/D converter on the STM ARM board](#) in lab says it has 94 dB SNR (15.3 effective bits) and 80 dB total harmonic distortion (13.0 effective bits).
 - 15.3 effective bits means that the least significant bit is a 16-bit quantization value is not accurate in representing the desired signal.