Quantization

A mathematical function, $y = f(x)$, has x as the independent variable and y as the dependent variable When this relation is plotted, we get the values of y on the $y - axis$ and that of x on the $x - axis$. In sampling, the independent variable is converted from continuous to discrete. Quantization handles the discretization of the dependent variables.

While sampling takes snapshots of the signal in steps of the sampling period, quantization defines discrete levels for which the signal magnitude is represented. For example, the magnitude of a sine wave lies between -1 and 1. The magnitude can therefore take any value within this range. If we define the number of quantization levels as 8, then it means we need only 8 values between -1 and 1, discarding the infinite number of values in between.

Method of Quantization

First, the number of quantization levels is defined. This is determined by the number of bits used in the quantizer. The number of bits would also determine the resolution of the quantization. The formula to calculate the number of levels (L) is:

$$L = 2^{\text{No of bits}}$$

This formula gives the number of combinations we can obtain with the number of bits we have. 2 bits gives 4 combinations of 0s and 1s. 3 bits gives 8 combinations of 0s and 1s.

Next we determine the length/range of the signal. This can also be obtained through the formula:

$$R = y_{max} - y_{min}$$

With our case of a sine wave, it's length would be $1 - (-1) = 2$.

The quantization levels are defined in steps. A step is calculated using the formula:

$$\Delta = \frac{R}{L} = \frac{(y_{max} - y_{min})}{2^n}$$

Therefore, with a 3-bit quantizer, we have 8 levels with a step size of $\frac{2}{8} = 0.25$.

Here's a table that shows the various binary representations for 3 bits and their corresponding quantization levels

Bits	Quantization levels
000	-4Δ = -1.00
001	-3Δ = -0.75
010	-2Δ = -0.50
011	-1Δ = -0.25
100	0Δ = 0.00
101	1Δ = 0.25
110	2Δ = 0.50
111	3Δ = 0.75

From the table, we notice that the maximum and minimum quantization levels are not symmetric about zero / non-uniform quantization. A method known as mid-rise quantization can be employed to shift the levels by half a step. (This would be shown at the end)

Next, is to map the various magnitudes of the signal to their corresponding quantization level. The quantization level corresponding to a signal value at x is given by:

$$i = \text{floor}\left(\frac{x - x_{\text{min}}}{\Delta}\right)$$

where i is the quantization index.

The quantized value is then given by:

$$x_q = x_{min} + \left(i + \frac{1}{2}\right)\Delta$$

$x - x_{min}$ shifts the signal upward, so that all values are positive. For each shift, there is a division by the step, $Delta$, to determine which quantization interval the signal lies in. The floor operation assigns the signal to the lower bound of that interval. The result is the quantization index, i, which indicates the position of the quantization level.

The value of the quantized signal is then calculated using the second equation. The addition of $\frac{1}{2}\Delta$ places the quantized value at the center of the interval, which corresponds to a mid-rise quantization.

With that, a new table with a symmetrical form is shown below:

Bits	Quantization Level
000	-0.875
001	-0.625
010	-0.375
011	-0.125
100	0.125
101	0.375
110	0.625
111	0.875