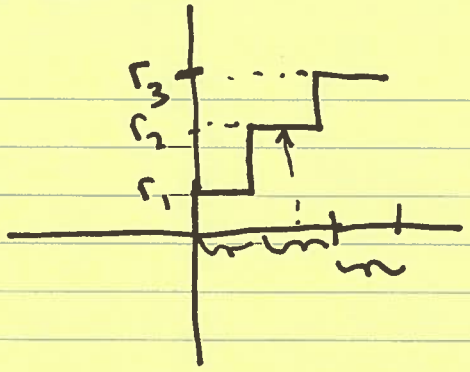
Lecture 5



Suppose $x = 3.5$ 4 bit converter

$$Q = \frac{R}{2^B} \qquad R = 10$$

3.5

N + M = B

IEEE Floating Point Standard

How to represent amplitude values of digital signals?

Let's assume signals are limited to a range of ± 1.

Suppose $x[5] = 1096.35$

$B = 16$ $R = 2^B = 65,534$

± 32767

$$\frac{1096.35}{32767} = 0.03345 \Rightarrow \times 32767 = [\text{int}]$$

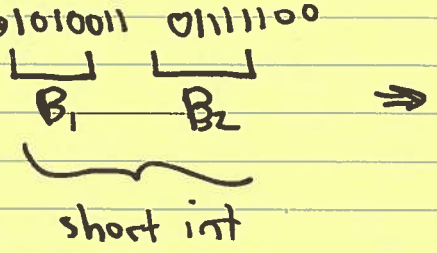
$$x[5] = a_0 2^0 + a_1 2^{-1} + a_2 2^{-2} + \dots$$

BCD: Binary coded decimal

16 bit integers $\Rightarrow \pm 32767 \Rightarrow [-1, 1]$
signed integer

short int i

audio files: 16-bit integers (raw)



two's complement 16-bit numbers

Data Types in C:

IEEE floats.
- float: 32 bit on 32 bit or 64 bit proc.
- double: 64 bit

long int: ~~32~~ 64 bits

unsigned short int: $[0, 2^{16}]$

short int: 16 bits $\longrightarrow$ [-32767, 32767]

int

byte "unsigned char" $\Rightarrow$ 8 bits

char