Sravan Tumuluri
skt3rt
floatingpoint.pdf
9/23/13

Floating point number 1: -8.796875

Sign bit is 1 because it is negative

8.796875/(2^3) = 1.099609375

Exponent: 3 + 127 = 130 which is 1000 0010

Mantissa: 1.099609375 – 1 = 0.099609375

Mantissa in binary: 1/16 + 1/32 + 1/256 + 1/512 = 0001 1001 1000 0000 0000 000

Number in binary: 1100 0001 0000 1100 1100 0000 0000 0000

Little Endian Binary: 0000 0000 1100 0000 0000 1100 1100 0001

Little Endian Hex: 0x 00C00CC1

Floating Point Number 2: 0x00c01f40

Number in Little Endian Binary: 0000 0000 1100 0000 0001 1111 0010 0000

Number in Big Endian Binary: 0010 0000 0001 1111 1100 0000 0000 0000

Converting to float:

Sign is +, bit = 0.

Exponent: 0100 0000 = 64

Mantissa: 0011 1111 1000 0000 0000 000 = 0.4921875

Float = (1-2*0) * (1+0.4921875) *2^(-63) = 1.61783293 e-19