

Problem 1:

a) Let X be $X = x_{n-1}x_{n-2}\dots x_1x_0$ ①.

The additive inverse of X is

$$-X = 2's \text{ compl. of } X = (\overline{x_{n-1}} \overline{x_{n-2}} \dots \overline{x_1} \overline{x_0}) + 1 \quad ②$$

$$\begin{array}{r} ①, ② \Rightarrow X + (-X) = \frac{x_{n-1}x_{n-2}\dots x_1x_0}{\overline{x_{n-1}} \overline{x_{n-2}} \dots \overline{x_1} \overline{x_0}} \\ +) \end{array}$$

$$\begin{array}{r} \\ \\ \hline 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{array}$$

$$\begin{array}{r} +) \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{r} \\ \\ \hline (1 & 0 & 0 & 0 & 0 & 0 & 0)_2 = 2^n \end{array}$$

or $\boxed{X + (-X) = 2^n}$ ③

On the other hand $\boxed{X + (-X) = 0}$ ④.

Eqs ③, ④ $\Rightarrow \boxed{2^n = 0}$ ⑤.

But the carry out has a weight factor of 2^n which is equivalent to zero (see ⑤).

Thus the carry out must be ignored.

b). For addition 2 to produce a carry out it must be $x_{n-1} = x_{n-2} = \dots = x_1 = x_0 = c = 1$. But this is impossible since the worst case addition 1 gives \longrightarrow next page \longrightarrow

$$(x+y)_{\text{wsgx}} = \begin{array}{r} 111\cdots 111 \\ +) 111\cdots 111 \\ \hline 1111\cdots 110 \end{array} \quad 3755, \text{ Test 1 (2)} \\ \text{Solutions, Spr. 04}$$

(c)

$$\boxed{000000} \boxed{1001} \boxed{0} \quad \text{Initialization}$$

$\downarrow 11010$ $\rightarrow 1 \times \text{multiplicand / double shift}$

$$\boxed{11010} \boxed{1001} \boxed{0}$$

$$\boxed{11110} \boxed{1010} \boxed{0}$$

$\downarrow 01100$ $\rightarrow -2 \times \text{multiplicand / double shift}$

$$\boxed{01010} \boxed{1010} \boxed{0}$$

$$\boxed{00010} \boxed{1010} \boxed{1}$$

$\rightarrow \text{product} = (00101010)_2 = (+42)_{10}$

(d) $A = A_1 \times 2^n + A_0 \quad (1)$

$A = B \times Q + R \quad (2)$

(1), (2) $\Rightarrow \boxed{A_1 \times 2^n + A_0 = B \times Q + R} \quad (3)$

Since $A_1 < B$, the maximum of left side of (3) is $A_1_{\text{wsgx}} \times 2^n + A_0_{\text{wsgx}} = (B-1) \times 2^n + 2^n - 1 = B \times 2^n - 2^n + 2^n - 1 = B \times 2^n - 1$ or

$\boxed{\text{wsgx} \cdot \text{left side of (3)} = B \times 2^n - 1} \quad (4).$

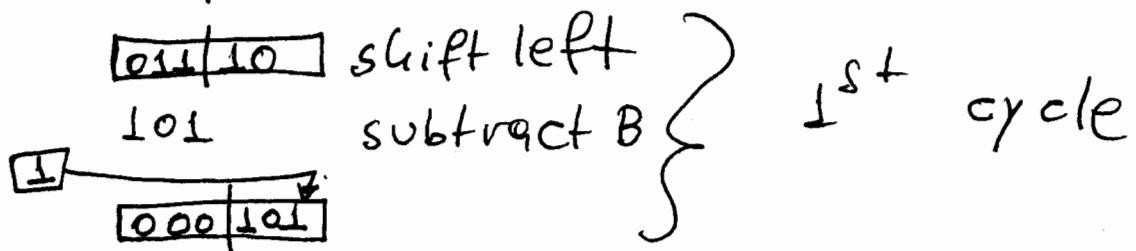
An n -bit quotient Q can achieve the same maximum of (4) for the right side of (3). Just see that if Q is n -bit long, then $(B \times Q + R)_{\text{wsgx}} = B \times Q_{\text{wsgx}} + R_{\text{wsgx}} = B \times (2^n - 1) + B - 1 = B \times 2^n - 1$.

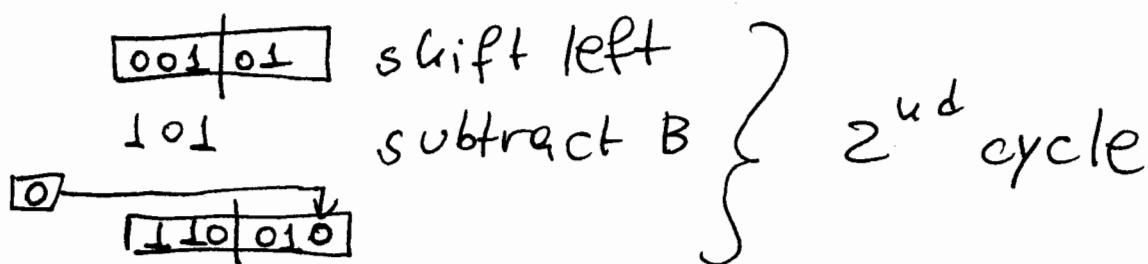
3755, Test 1 Solutions, Spr. 04.

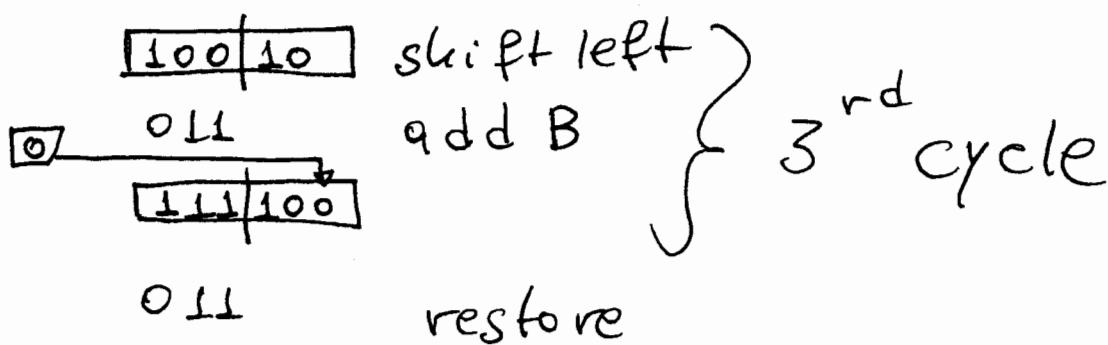
(3)

Therefore, the length of n bits is sufficient for the quotient and overflow does not occur.

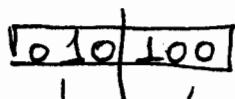
e)  Initialization

 shift left }
101 subtract B } 1st cycle

 shift left }
101 subtract B } 2nd cycle

 shift left }
011 add B } 3rd cycle

011 restore



$$R = (010)_2 = 2 \quad \leftarrow \quad \rightarrow Q = (100)_2 = 4$$

f). (i) \rightarrow True ; (ii) \rightarrow False ; (iii) \rightarrow False.

g). (i) $Q < 0$; (ii) $R > 0$.

Problem 2:

(a) The range of the fraction is $0.5 \leq f \leq 1 - 2^{-17}$

The range of the exponent is $-2^7 \leq e \leq 2^7 - 1$
 or $-128 \leq e \leq 127$. Thus the positive floating point dynamic range is:

$$0.5 \times 2^{-128} \leq A^+ \leq (1 - 2^{-17}) \times 2^{127}$$

The negative floating point dynamic range is:
 $-(1 - 2^{-17}) \times 2^{127} \leq A^- \leq -0.5 \times 2^{-128}$

b) Align/Adjust

$$e_1 - e_2 = e_1 + 2^3 \text{ compl. of } e_2 = 1001$$

$$\begin{array}{r} +) 0110 \\ \hline 01111 \end{array}$$

$\hookrightarrow c=0$ means $e_1 - e_2 < 0$

or $e_2 > e_1$. Then $e_2 - e_1 = 2^3 \text{ compl. of } (1111) =$
 $= (0001)_2 = (1)_{10}$. Thus N_1 :

s_1	e_2	f_1'
0	1010	01010

Subtract fractions: Here a true subtraction takes place. $f_1' - f_2 = f_1' + 2^3 \text{ compl. of } f_2 =$
 $= 01010$

$$\begin{array}{r} +) 01111 \\ \hline 011001 \end{array}$$

$\hookrightarrow c=0$ means negative result. Thus \Rightarrow

EE 3755, Test 1 Solutions, Spr. 04 ⑤

$N_3 = N_1 - N_2$ will be a negative number
with fraction = 2's compl. of $(11001) = (00111)$

or N_3 : $\boxed{1 \mid 1010 \mid 00111}$

Postnormalizing we get N_3 : $\boxed{1 \mid 1000 \mid 11100}$

Exponent underflow did not occur since

$$e_3 = (1000)_2 = (8)_{10} \in [0 \text{ } 15].$$

C). Consider the integers A and B where

$$A = q_{n-1} q_{n-2} \dots q_1 q_0 \underbrace{000 \dots 000}_{n \text{ zeroes}}.$$

$$B = b_{n-1} b_{n-2} \dots b_1 b_0.$$

Divide A by B to get an n-bit quotient Q
and an n-bit remainder R. Let Q' be

$$Q' = \begin{cases} Q & \text{if } \frac{R}{B} < 1/2 \\ Q+1 & \text{if } \frac{R}{B} \geq 1/2 \end{cases}$$

Let Q' be represented in binary as

$$Q' = q'_{n-1} q'_{n-2} \dots q'_1 q'_0. \text{ Then}$$

$$f_3 = \frac{f_1}{f_2} \cong \bullet q'_{n-1} q'_{n-2} \dots q'_1 q'_0$$

EE 3755, Test 1 Solutions, Spr. 04 (6).

(d). The dynamic range of 4-bit biased exponents is $DR = [0 \ 15]$. Here $f_1 > f_2$ and thus alignment of dividend is needed. Thus, the exponent of the quotient is $e_1 + 1 - e_2 + bi_{QS} = 5 + 1 - 14 + 8 = 0 \in [0 \ 15]$. Therefore, neither exponent overflow nor exponent underflow occurs.

(e). The dynamic range of 4-bit biased exponents is $DR = [0 \ 15]$. Here $f_1 \times f_2 = (0.1000)_2 \times (0.1000)_2 = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 2^{-2} = (0.0100)_2$. Postnormalization will be needed and the exponent of the product will be $e_1 + e_2 - bi_{QS} - 1 = 11 + 13 - 8 - 1 = 15 \notin [0 \ 15]$. Therefore, either exponent overflow or exponent underflow occurs.

(f). Here $e_{biased} = (1011)_2 = 11$; $bi_{QS} = 2^3 = 8$. Therefore $e_{unbiased} = e_{biased} - bi_{QS} = 11 - 8 = 3$. Thus, A value $= -(0.111100)_2 \times 2^3 = -(111.100)_2 = -(7.5)_{10}$.