3. Moltiplicazione di numeri relativi rappresentati in complemento a 2

Premessa.

Dati due numeri naturali x_r e y_r rappresentati in binario naturale su n cifre, il loro prodotto è rappresentabile su $2n$ cifre. Si ha infatti che, se $z_r=x_r\times y_r$:

$$0 \leq z_r \leq (2^n-1)(2^n-1) = 2^{2n} - 2^{n+1} + 1 < 2^{2n} \forall n > 0 \tag{3.1}$$

La rappresentazione Z_R di z_r in binario naturale (su $2n$ cifre) si può ottenere dalle rappresentazioni X_R e Y_R in binario naturale di x_r e y_r (su n cifre) eseguendo l'algoritmo di moltiplicazione "di shift e somma" che si suppone noto.

Moltiplicazione di numeri relativi rappresentati in complemento a 2

Siano ora x_r e y_r due numeri naturali rappresentati in binario naturale su n cifre che sono ottenuti come trasformazione attraverso complemento a due di due numeri relativi x e y. Ci poniamo il problema di determinare la rappresentazione in complemento a 2 del numero $z=xy$. Determiniamo intanto il minimo numero di cifre binarie m necessarie a rappresentare, secondo trasformazione in complemento, il numero z.

Ricordiamo che, dato il numero di bit m disponibili per la rappresentazione in complemento a 2, l'intervallo di valori h rappresentabili risulta:

$$-2^{m-1} \leq h \leq 2^{m-1} - 1 \tag{3.2}$$

La relazione precedente può essere scritta, in maniera alternativa, nel modo seguente:

$$-2^{m-1} \leq h < 2^{m-1} \tag{3.3}$$

Poiché sia ha:

$$-2^{n-1} \leq x \leq 2^{n-1} - 1 \quad ; \quad -2^{n-1} \leq y \leq 2^{n-1} - 1 \tag{3.4}$$

Si conclude che il prodotto $z=xy$ può assumere valori solo nell'intervallo:

$$-2^{n-1}(2^{n-1}-1) \leq z \leq 2^{2(n-1)} \tag{3.5}$$

Tenuto conto delle relazioni precedenti, si ha che z può essere rappresentato su m cifre solo se i possibili valori di z appartengono all'intervallo di valori rappresentabili con m cifre. Tenuto conto della 3.2 e 3.5, per quanto riguarda la disequazione più a destra deve essere

$$2^{2(n-1)} \leq 2^{m-1} - 1 \rightarrow 2^{2(n-1)} < 2^{m-1} \rightarrow m-1 > 2n-2 \rightarrow m \geq 2n \tag{3.6}$$

Per quanto riguarda la disequazione più a sinistra, invece, deve essere:

$$-2^{m-1} \leq -2^{n-1}(2^{n-1}-1) \rightarrow 2^{2(n-1)} - 2^{n-1} \leq 2^{m-1} \tag{3.7}$$

Come è evidente, la condizione 3.7 è certamente soddisfatta (per n>0) se la 3.6 è soddisfatta.

Possiamo quindi concludere che la rappresentazione in complemento del prodotto di due numeri rappresentati in complemento a 2 su n cifre, richiede almeno m=2n cifre.

Verifichiamo ora che l'algoritmo di moltiplicazione applicato alle rappresentazioni X_r e Y_r (in complemento a due su n cifre) dei numeri relativi x e y, **non produce** in generale la rappresentazione in complemento a 2 su $2n$ cifre del prodotto z.

Distinguiamo i seguenti casi:

1. almeno uno dei due fattori (x o y) è 0.
2. entrambi i fattori sono positivi
3. entrambi i fattori sono negativi
4. un fattore è positivo e l'altro è negativo.

Caso 1: In questo caso si verifica immediatamente che eseguendo l'algoritmo di moltiplicazione sulle rappresentazioni X_R e Y_R si ottiene la rappresentazione corretta di z=0.

Caso 2: In questo caso si ha x_r=x; y_r=y; L'algoritmo di moltiplicazione produce la rappresentazione corretta (su $2n$ cifre) del numero positivo $z_p= x_r \times y_r$ e poiché la trasformazione di $z=x\times y$ in complemento a 2 su $2n$ cifre è proprio $z_r=x\times y= x_r\times y_r$, si ha che la rappresentazione Z_R ottenuta mediante l'algoritmo di prodotto è la rappresentazione corretta (in complemento a 2 su 2n cifre) del prodotto z.

Caso 3: Supponiamo che sia y<0. In questo caso si ha: x_r=x; $y_r=2^n+y$. L'algoritmo di moltiplicazione produce la rappresentazione corretta (su $2n$ cifre) del numero positivo $z_p= x_r\times y_r=x\times(2^n+y)$. Tuttavia, poiché la rappresentazione corretta in complemento su $2n$ cifre del numero (negativo) xy è la rappresentazione in binario naturale del numero $z_r=2^{2n}+xy$, è evidente che la rappresentazione X_P del risultato dell'operazione di moltiplicazione non è la rappresentazione del prodotto xy.

Caso 4: In questo caso si ha che $x_r=2^n+x$ e $y_r=2^n+y$. L'algoritmo di moltiplicazione produce la rappresentazione corretta (su $2n$ cifre) del numero positivo $z_p= x_r\times y_r=(2^n+x)\times(2^n+y)=2^{2n}+(x+y)2^n+xy$. Anche in questo caso, tuttavia, poiché la rappresentazione corretta in complemento su $2n$ cifre del numero (positivo) xy è la rappresentazione in binario naturale del numero z_r=xy, è evidente che la rappresentazione Z_P del risultato dell'operazione di moltiplicazione non è la rappresentazione del prodotto xy.

Supponiamo ora di restringerci all'insieme dei numeri x e y tali che sia possibile rappresentare correttamente il loro prodotto su sole n cifre, ovvero che risulti:

$$-2^{n-1} \leq xy < 2^{n-1} \tag{3.8}$$

Se siamo in questa condizione particolare, è immediato osservare dalle relazioni precedenti che il numero $z_r=z_p \bmod 2^n$ è la corretta trasformazione del prodotto in complemento a 2^n e che pertanto le n cifre meno significative del prodotto Z_P costituiscono la corretta rappresentazione del prodotto in complemento a 2 su n cifre (ovviamente, come detto in premessa, solo nel caso in cui il prodotto sia rappresentabile su sole n cifre).

Questa osservazione ci consente di definire un semplice algoritmo per la determinazione della rappresentazione del prodotto di numeri rappresentati in complemento a due su q cifre. Infatti:

- dati due numeri relativi x e y rappresentabili su q cifre, si può estenderne la rappresentazione su n=$2q$ cifre;
- a questo punto è certo che il prodotto $x\times y$ è rappresentabile su n cifre;
- si esegue l'algoritmo di prodotto fra le rappresentazioni in complemento X_R e Y_R di x e y su n=$2q$ cifre ottenendo un vettore di $2n$=$4q$ cifre;
- Gli n bit meno significativi della rappresentazione X_P (su $2n$ cifre) costituiscono la rappresentazione del prodotto $z=x\times y$ su n=$2q$ cifre.

4. Rappresentazione e operazioni su numeri "con la virgola" mediante numeri interi (il caso della notazione 1.n)

L'uso della virgola per la separazione fra unità e cifre decimali nella consueta notazione decimale consente di estendere la rappresentazione dai numeri interi a un insieme di numeri razionali che, spesso, costituiscono una approssimazione di numeri reali. In presenza della virgola, la regola di interpretazione in decimale della stringa di cifre X_D $\{X_{n-1}\ X_{n-2}\ldots X_i, X_{i-1},\ldots, X_0\}$ è la seguente:

$$x = \sum_{k=0}^{n-1} X_k \times 10^{k-i} = 10^{-i} \times \sum_{k=0}^{n-1} X_k \times 10^k \qquad (4.1)$$

dove i è il numero di cifre a destra della virgola.

In sostanza, quando scriviamo il numero 15,46 intendiamo riferirci al numero razionale 1546/100.

La presenza della "virgola" come parte della rappresentazione è comoda nella scrittura, ma non è essenziale. Si ottiene lo stesso risultato se teniamo traccia, nello scrivere una stringa di cifre, del numero i che deve essere usato nella regola di interpretazione. Se i=0 la regola di interpretazione si riduce a quella relativa ai numeri interi. Ovviamente si possono rappresentare in questo modo solo numeri razionali riconducibili al rapporto fra un numero intero e una potenza della radice. Per esempio, il numero 8/3 non è in nessun caso rappresentabile esattamente visto che non è riducibile al rapporto fra un numero intero e una potenza di 10 (si noti che se usassimo invece una radice pari a 3 il numero sarebbe rappresentabile esattamente).

Capita spesso di avere un numero finito di cifre a disposizione per dover rappresentare una quantità fisica f che assume valori fra un valore minimo f_{min} e un valore massimo f_{maz} (si pensi a un segnale analogico in uscita da un sistema con una dinamica fissata). La quantità fisica può assumere valori solo positivi, solo negativi o positivi e negativi. In generale, scelto F=max($|f_{min}|$,$|f_{max}|$) si ha certamente che:

$$-F \leq f \leq F \qquad (4.2)$$

Ovvero è sempre possibile trovare un valore positivo F per il quale vale la 4.2.

Supponiamo per il momento di avere a disposizione la possibilità di rappresentare solo numeri interi relativi x_f compresi in un intervallo simmetrico che va da $-x_{Mf}$ a $+x_{Mf}$. Possiamo allora scegliere di rappresentare f (in generale con un errore) cercando il numero intero che meglio approssima f in base alla relazione:

$$\frac{x_f}{x_{Mf}} F \approx f \qquad (4.3)$$

E' facile intuire che tanto più grande è il massimo modulo x_{Mf} del numero rappresentabile, tanto minore è l'errore massimo che si può commette nell'approssimare un qualunque valore f con l'espressione alla sinistra della 4.3. Una volta che diamo per scontata la relazione alla sinistra della 4.3 e concordiamo sui valori x_{Mf} e F, f è univocamente individuata dal numero relativo x_f e il numero relativo x_f può essere rappresentato in complemento. Operazioni di somma fra due grandezze omogenee (con lo stesso valore di x_{Mf} e di F) si possono eseguire come somma in complemento a 2 e il risultato è corretto se la somma delle due grandezze rientra nell'intervallo "fisico" originario, ovvero se è rappresentabile.

Un caso più interessante è quello in cui sia necessario eseguire un prodotto fra una grandezza f e una grandezza g immaginando che anche per g si accetti una rappresentazione mediante un numero relativo x_g tale per cui si abbia:

$$\frac{x_g}{x_{Mg}} G \approx g \tag{4.4}$$

Sia ora $h=fg$ il risultato del prodotto delle due grandezze. Il valore massimo (in modulo) di h è $H=FG$. Per rappresentare h possiamo sfruttare il numero relativo x_h che si ottiene moltiplicando i numeri relativi x_f e x_g in accordo con la seguente relazione:

$$h = fg \approx \frac{x_h}{x_{Mh}} H = \frac{x_f x_g}{x_{Mf} x_{Mg}} FG \tag{4.5}$$

Il problema di questo approccio, però, sta nel fatto che il numero intero relativo x_h, qualunque sia la rappresentazione che scegliamo, ha bisogno di un numero maggiore di cifre per essere rappresentato e questo costituisce un problema, specialmente se le moltiplicazioni da fare sono molte, visto che ad ogni moltiplicazione corrisponde un aumento delle cifre necessarie per la rappresentazione del prodotto. Per rendercene conto supponiamo di scegliere, come rappresentazione per i numeri interi relativi x_f e x_g la rappresentazione in complemento a due. Ricordiamo che questo significa trasformare prima i numeri relativi x_f e x_g in numeri naturali x_{Rf} e x_{Rg} mediante l'operazione complemento usando come costante di complementazione una opportuna potenza di 2, per poter poi usare una rappresentazione canonica dei numeri trasformati. Supponiamo inoltre di avere il vincolo che il numero n di cifre disponibili per la rappresentazione di ogni numero sia n=8.

Con n=8 possono essere rappresentati tutti i numeri interi relativi da -2^{n-1} incluso (nel nostro caso -128) fino a $+(2^{n-1}-1)$ incluso (nel nostro caso +127). Sono possibili diverse scelte per i fattori di normalizzazione x_{Mf} e x_{Mg} ma una scelta particolarmente semplice è quella di scelgiere $x_{Mf}=x_{Mg}=2^{n-1}$ (nel nostro caso 128). Con questa scelta possiamo rappresentare i valori $-F$ e $-G$, ma no i valori $+F$ e $+G$ visto che il valore massimo dei rapporti x_f/x_{Mf} e x_g/x_{Mg} è sempre strettamente minore di uno. Si noti tuttavia che all'aumentare del numero n di cifre disponibili per la rappresentazione, il valore massimo rappresentabile si avvicina al valore massimo della grandezza fisica e che in ogni caso si dimensionano i sistemi di calcolo in modo da avere un po' di margine per evitare la saturazione (ovvero i valori scelti pe F e G sono di solito maggiori, di un certo fattore di sicurezza che dipende dalle applicazione, del massimo modulo dei valori assunto effettivamente dalle grandezze rappresentate.

Supponiamo ora che f sia il valore di una tensione istantanea con una dinamica che va da -9 a 9 V e che g sia il valore di una corrente con una dinamica che va da -14 mA a +14 mA. Scegliamo allora, per sicurezza, F=10 V e G=15 mA. Supponiamo ora di dover calcolare la potenza istantanea $h=fg$. L'operazione di trasformare una grandezza continua in una grandezza discretizzata è generalmente svolta dai convertitori digitali analogici che svolgono l'operazione di quantizzazione, ovvero associano ad intervalli di valori reali un numero intero relativo che rappresenta i valori dell'intervallo, in accordo con un qualche criterio di approssimazione dei valori (arrotondamento, troncamento ecc. ecc.). In questo esempio saremo noi a svolgere il ruolo dei "convertitori". Supponiamo che a un certo istante sia f=2.3 V e g=-5.1 mA. Un modo per arrivare a determinare i numeri relativi x_f e x_g per la rappresentazione è usare direttamente le 4.3 e 4.4 nella forma:

$$x_f \approx \frac{f}{F} x_{Mf} \qquad x_g \approx \frac{g}{G} x_{Mg} \tag{4.3}$$

Nel nostro caso abbiamo:

$$x_f \approx \frac{2.3}{10}128 \approx 29.440 \approx 29 \quad ; \qquad x_g \approx -\frac{5.1}{15}128 \approx -43.520 \approx -44 \tag{4.4}$$

Si noti che, ovviamente, siccome x_f e x_g devono essere interi relativi, il risultato delle operazioni nella 4.4 è stato arrotondato all'intero più vicino. Se si possono accettare errori statisticamente maggiori sul risultato finale, si può eseguire un troncamento invece dell'arrotondamento. Queste "scelte" fanno parte della calibrazione del convertitore AD usato, in una situazione reale, per ottenere x_f e x_g. Per procedere nei calcoli usando la rappresentazione in complemento a 2, dobbiamo prima trasformare i numeri e poi operare sulle rappresentazioni dei numeri ottenuti. Si ha:

$$\begin{aligned} x_{Rf} &= 29 \bmod 2^8 = 29 \quad \rightarrow \quad X_{Rf} = \{0001\ 1101\} \\ x_{Rg} &= -44 \bmod 2^8 = 212 \rightarrow \quad X_{Rg} = \{1101\ \ 0100\} \end{aligned} \tag{4.5}$$

Possiamo ora facilmente ottenere la rappresentazione in complemento a due (su $2n$=16 cifre) del numero x_h. Ovviamente, intendiamo riferirci alla possibilità di calcolare, come discusso nella sezione 3, il numero naturale x_{Rh} che è la trasformazione (modulo 2^{2n}) del numero x_h a partire dalle rappresentazioni dei numeri x_{Rf} e x_{Rg}. Si ha:

```
                         x'Rf
                                     xRf
                        0000000000011101  x

                        1111111111010100  =
                         x'Rg        xRg
-----------------------------------------
                        0000000000000000
                       0000000000000000
                      0000000000011101
                     0000000000000000
                    0000000000011101
                   0000000000000000
                  0000000000011101
                 0000000000011101
                0000000000011101
               0000000000011101
              0000000000011101
             0000000000011101
            0000000000011101
           0000000000011101
          0000000000011101
         0000000000011101
-----------------------------------------
       ----------------1111101100000100
         Non ci interessano        xRh
```

(4.6)

Si ottiene quindi X_{Rh} =(1111 1101 0000 0100) ovvero x_{Rh}=64260 che, interpretato con riferimento a una costante di complementazione $2^{2n}=2^{16}=65536$ porta a x_h=64260-65536=-1276. Per risalire alla potenza istantanea h rappresentata da questo numero, possiamo usare la 4.5 e abbiamo:

$$h \approx \frac{x_h}{x_{Mh}} H = \frac{-1276}{128 \times 128} 10\ \mathrm{V}\ 15\ \mathrm{mA}\ \approx\ -11.68\ \mathrm{mW} \tag{4.7}$$

Questo risultato è vicino al valore "vero" di -2.3×5.1=11.73 mW. Se si aumenta il numero di cifre usate per la rappresentazione delle grandezze di partenza, si ottiene ovviamente una migliore approssimazione.

Il problema, però, è che ora la grandezza h è rappresentata su 16 cifre. Supponiamo, come spesso è il caso, che il numero di cifre per rappresentare qualunque grandezza nel nostro sistema sia fissata. Per esempio sia in ogni caso pari a 8. Ovviamente, così come abbiamo rappresentato tensioni e correnti su 8 cifre, dato un valore della potenza, è sicuramente possibile trovare una rappresentazione (al costo di una approssimazione) su sole 8 cifre, cercando il numero intero relativo x''_h per il quale si ha, con $x''_{Mh}=2^{n-1}$:

$$\frac{x''_h}{x''_{Mh}} \approx \frac{x_h}{x_{Mh}} \tag{4.8}$$

La stessa equazione 5.8 ci fornisce un modo per ottenere x''_h. Basta infatti cercare il numero intero relativo x''_h che meglio approssima l'espressione:

$$x''_h \approx \frac{x_h}{x_{Mh}}\ x''_{Mh} \tag{4.9}$$

Nel caso dell'esempio precedente, a partire a x_h=-1276 (si faccia attenzione che in questo contesto stiamo lavorando direttamente con i numeri e non con le loro rappresentazioni in complemento) si otterrebbe:

$$\frac{-1276}{128 \times 128} 128 = -9.96 => x''_h = -9\ (\text{tronc.})\ \text{o}\ x''_h = -10\ (\text{arrot.}) \tag{4.10}$$

Nella espressione precedente abbiamo evidenziato due possibili scelte che portano, come è evidente, a una differenza significativa nell'accuratezza della rappresentazione. In ogni caso, la 4.10 non ci serve a molto perché, mentre come esseri umani siamo in grado di gestire agevolmente i numeri con le rappresentazioni che abbiamo imparato ad usare nel corso degli anni, un sistema digitale lavora solo per le rappresentazioni e, in particolare, nel nostro caso con rappresentazioni in complemento a 2. La domanda che dobbiamo allora porci è se è possibile, partendo da x_{Rh}, una rappresentazione x''_{rH} senza bisogno di fare i conti sui "numeri veri" come abbiamo fatto noi. Intanto si osservi che, nella 4.9, tutte le volte che si deve trovare un valore su "meno cifre" si tratta di trovare l'intero x''_h che meglio approssima il numero $x_h/2^m$, con m maggiore di zero. Si noti che riusciamo a risolvere il problema per m=1, si può risolvere il problema per un genreico valore di m iterando il procedimento.

Il nostro problema si può porre nel modo seguente:

Dato x_{Rh} ottenuto come trasformazione di x_h modulo 2^n, come si ottiene da x_{Rh} il valore di x''_{Rh} che corrisponde a x''_h in mode che sia:

$$x''_h \approx \frac{x_h}{2} \tag{4.11}$$

Si ottiene un algoritmo molto semplice se, come regola di approssimazione per i valori x''_h si assume la seguente:

$$x''_h = \begin{cases} \dfrac{x_h}{2} & se\ x_h \text{ pari} \\ \dfrac{x_h - 1}{2} & se\ x_h \text{ dispari} \end{cases} \tag{4.12}$$

Si può infatti facilmente dimostrare (vedi sezione 5 più avanti) che questo si traduce, in termini di rappresentazione in cifre, con il fatto che la rappresentazione di x''_h si ottiene come sequenza delle n-1 cifre, da X_{n-1}, a X_1, usate per la rappresentazione di x_h su n cifre. Ovvero:

se

$$x_{Rh} = x_h\ mod\ 2^n\ \equiv (X_{n-1}, X_{n-2}, \dots, X_1, X_0) \tag{4.13}$$

e

$$x''_{Rh} = x''_h\ mod\ 2^{n-1}\ \equiv (X''_{n-2}, X''_{n-3}, \dots, X''_1, X''_0) \tag{4.14}$$

allora

$$X''_{n-2} = X_{n-1}, \dots, X''_0 = X_1 \tag{4.15}$$

In altre parole, nel contesto della rappresentazione di grandezze in termini dell'equazione 4.3, se si vuole ridurre il numero di cifre con il quale viene rappresentata la quantità f (e si è disposti a tollerare il conseguente errore di troncamento) è sufficiente eliminare dalla rappresentazione le m cifre più a destra, ricordando che per l'interpretazione della 4.3, che il termine al denominatore deve essere inteso come diviso per 2^m.

Con riferimento all'esempio precedente, si aveva:

X_{Rh}=(1111 1101 0000 0100); $C=2^{16}$, x_{rh}=64260; $x_{Mh}=2^{14}$=16384; h=−11.68 mW

Proviamo ora a ridurre la rappresentazione su un numero sempre minore di cifre:

X^{15}_{Rh}=(111 1110 1000 0010); $C=2^{15}$, x_{rh}=32130; $x_{Mh}=2^{13}$=8192; h=−11.68 mW

X^{14}_{Rh}=(11 1111 0100 0001); $C=2^{14}$, x_{rh}=16065; $x_{Mh}=2^{12}$=4096; h=−11.68 mW

X^{13}_{Rh}=(1 1111 1010 0000); $C=2^{13}$, x_{rh}=8032; $x_{Mh}=2^{11}$=2048; h=−11.71 mW

.

X^{9}_{Rh}=(1 1111 1010); $C=2^{9}$,x_{rh}=502; $x_{Mh}=2^{7}$=128; h=−11.71 mW

X^{8}_{Rh}=(1 1111 1010); $C=2^{8}$,x_{rh}=251; $x_{Mh}=2^{6}$=64; h=−11.71 mW

E così via.

Dagli esempi fatti si nota che se volessimo usare come costante x_{Mh} il valore 128, come abbiamo fatto quando lavoravamo con i numeri (e non con le rappresentazioni), anche per mantenere la consistenza con quanto fatto per i fattori del prodotto, saremmo costretti a mantenere il prodotto su 9 cifre. Alternativamente, se vogliamo ridurre la rappresentazione del prodotto su 8 cifre, siamo costretti a usare una costante x_{Mh} diversa da quella usata per i fattori e questo complica le cose nel caso di moltiplicazioni ripetute. Notiamo però che nel nostro caso particolare la rappresentazione su 9 cifre sarebbe riducibile, il che ci consentirebbe di usarne solo 8, ma se questo fosse possibile in generale, dovrebbe essere vero che sono sufficienti (2n-1) cifre per rappresentare il prodotto di due numeri in complemento a 2 su n cifre, cosa che avevamo dimostrato non essere vera.

Tuttavia, se si esamina più in dettaglio la situazione, si scopre che l'unico caso in cui sono necessarie 2n cifre è il caso in cui entrambi i fattori sono relativi al minimo numero rappresentabile, ovvero $x=y=-2^{n-1}$. In questo caso, infatti, si dovrebbe rappresentare il numero $z=xy=+2^{2(n-1)}$ che è più grande, di una sola unità, del massimo numero rappresentabile in complemento a 2 su 2n-1 cifre, ovvero proprio $2^{2(n-1)}-1$. Se si esclude questo caso, tutti i risultati del prodotto fra 2 numeri in complemento a due su n cifre sono rappresentabili su 2n-1 cifre, riducendo di 1 la rappresentazione del risultato su 2n cifre ottenuto dal moltiplicatore.

Ricordiamo che in questo momento ci stiamo occupando di come si possano usare numeri interi per rappresentare grandezze reali con segno. Se nella scelta dei valori massimi e minimo (le F, G o H di prima) riusciamo a garantire che il minimo valore negativo non sia mai assunto (contemporaneamente) dai fattori della moltiplicazione, allora per rappresentare il prodotto possiamo utilizzare solo 11 cifre e, con riferimento al problema specifico di ridurre il prodotto su un numero di cifre esattamente pari a quelle dei fattori, si possono usare come rappresentazione le *n* cifre che si ottengono escludendo la prima a sinistra (operazione di riduzione) e cancellando le n-1 più a destra. In questo modo, se le costanti di scala dei fattori (x_{Mf} e x_{Mg}) sono uguali fra loro, anche quella del prodotto ha lo stesso valore.

Con riferimento al nostro esempio, questo significa che possiamo usare come rappresentazione del prodotto la X^9_{Rh}, cancellando la prima cifra a destra e, ovviamente, interpretando il risultato come ottenuto per complementazione da $C=2^8$. In questo modo xMh rimane $2^{n-1}=128$, come nel caso dei singoli fattori.

Molte strutture di calcolo (specialmente nel contesto DSP) sono progettate per fornire in uscita le n cifre già correttamente allineate in uscita a un prodotto fra fattori su n cifre interpretati in accordo alla 4.3. Ci si riferisce a questo tipo di rappresentazione come rappresentazione 1.*q*, dove $q=n-1$. Nel caso di $n=16$, quindi, si parla di rappresentazione 1.15.

5. Rappresentazione del doppio e della metà di un numero nella rappresentazione in complemento a due.

Supponiamo che dato un numero x rappresentabile in complemento a 2 su N cifre ($C=2^N$), anche il doppio $y=2x$ sia rappresentabile. Si noti che il doppio di un numero si ottiene moltiplicando il numero per la radice della rappresentazione. Sia x non negativo. Si ha:

$$x_r = x; \quad y_r = 2x = 2x_r \tag{5.1}$$

In termini di rappresentazione:

$$x_r = \sum_{i=0}^{N-1} [X_i] 2^i \tag{5.2}$$

Si noti che se $2x$ deve essere rappresentabile questo significa necessariamente che x deve essere strettamente minore di 2^{N-2} e che quindi le cifre X_{N-1} e X_{N-2} devono essere entrambe "0". Dal confronto:

$$2x_r = \sum_{i=0}^{N-1} [X_i] 2^{i+1} = \sum_{i=0}^{N-1} [Y_i] 2^i = y_r \tag{5.3}$$

e tenuto conto della proprietà (enunciata e non dimostrata) che nel sistema di numerazione binario per i numeri naturali esiste una ed una sola rappresentazione per ciascun numero si ha:

$$Y_0 = 0; \quad Y_i = X_{i-1} \quad \text{per } 1 \le i < N \tag{5.4}$$

In altre parole, la rappresentazione del doppio per numeri non negativi si ottiene eseguendo l'operazione di shift a sinistra dalla stringa di bit (si ricordi che $X_{N-1}=X_{N-2}=0$), ponendo $Y_0=0$.

Supponiamo or che x sia negativo. Si ha:

$$x_r = C + x = 2^N + x; \quad y_r = C + 2x = 2^N + 2x = 2(2^{N-1} + x) = 2x'_r \tag{5.5}$$

Si noti che se il doppio di x (negativo) deve essere rappresentabile, significa che:

$$y = 2x \ge -2^{N-1} \;\rightarrow\; x \ge -2^{N-2} \;\rightarrow x_r = 2^N + x \ge 2^N - 2^{N-2} = 2^{N-1} + 2^{N-2} \tag{5.6}$$

Dall'equazione precedente si deduce che $X_{N-1}= X_{N-2}=1$. Questo significa che x può essere rappresentato su N-1 cifre. Questo significa anche che la quantità x'_r in 5.5 è la trasformazione in complemento del numero x su N-1 cifre. Sappiamo anche che la rappresentazione di x'_r si ottiene semplicemente ignorando la cifra X_{N-1} dalla rappresentazione di x su N cifre. Infine, sempre dalla 5.5, otteniamo che y_r è il numero naturale che si ottiene moltiplicando per 2 il numero naturale x'_r, la cui rappresentazione su N cifre in binario naturale si ottiene come in Eq. 5.4.

In definitiva, se il doppio di x è rappresentabile, la sua rappresentazione si ottiene a partire dalla rappresentazione di x ignorando la cifra X_{N-1} , facendo uno shift a sinistra di una posizione di tutte le cifre rimanenti e ponendo l'ultima cifra a 0.

Per quanto riguarda la possibilità di rappresentare la metà di un numero intero, si deve intanto osservare che la metà di un numero intero è un numero intero solo nel caso di numeri pari (incluso lo zero). Limitiamoci al caso di numeri pari. Se x è pari si ha che anche x_r è pari (indipendentemente dal fatto che x sia positivo, negativo o nullo) e che si ha quindi X_0=0. Con x pari, y=x/2 è un numero intero. Per x non negativo si ha:

$$x_r = x; \quad y_r = x/2 = x_r/2 \tag{5.7}$$

Dalla relazione (si ricordi che X_0=0):

$$x_r/2 = \sum_{i=i}^{N-1} [X_i] 2^{i-1} = \sum_{i=0}^{N-1} [Y_i] 2^i = y_r \tag{5.8}$$

La relazione 5.8 è soddisfatta se:

$$Y_{N-1} = 0; \quad Y_i = X_{i+1} \quad \text{per } 0 \leq i < N-1 \quad (5.9)$$

Questa operazione corrisponde a ignorare lo zero più a destra nella rappresentazione di x_r, spostare di un posto verso destra tutte le cifre rimanenti e porre Y_{N-1} a 0. Si noti che, avendo a che fare con un numero non negativo, siamo certi che X_{N-1} è certamente nullo.

Per x negativo (pari) si ha:

$$x_r = C + x = 2^N + x; \quad y_r = C + \frac{x}{2} = 2^N + \frac{x}{2} = 2^{N-1} + \frac{(2^N + x)}{2} = 2^{N-1} + \frac{x_r}{2} \tag{5.10}$$

Ricordando che X_0 è nullo, si ottiene che per ottener la rappresentazione di y_r basta ignorare la cifra X_0, fare uno shift a destra delle cifre rimanenti e porre Y_{N-1} per tenere conto della quantità 2^{N-1} nella 5.10. Tenuto conto del fatto che se abbiamo a che fare con un numero negativo deve essere necessariamente X_{N-1}=1, possiamo ridurre l’algoritmo per ottenere la rappresentazione della metà di un numero pari in complemento a due alle relazioni seguenti (si ricordi che X_0=0):

$$Y_{N-1} = X_{N-1}; \quad Y_i = X_{i+1} \quad \text{per } 0 \leq i < N-1 \tag{5.11}$$

La 5.11 corrisponde all’operazione nota come shift aritmetico a destra.

Nel caso in cui il numero sia dispari, possiamo rappresentare la metà di uno dei due interi più vicini, accettando di commettere un errore. E’ facile dimostrare che dal punto di vista delle operazioni da compiere conviene sempre prendere in considerazione il numero pari immediatamente inferiore, perché in termini di rappresentazione questo coincide con la rappresentazione del numero di partenza, salvo porre la cifra più a sinistra a 0. Usando per un momento la rappresentazione in modulo e segno per parlare dei numeri coinvolti, questo significa che nel caso di 7, per esempio, rappresenteremo 3 (la metà di 6), mentre nel caso di -7 rappresenteremo -4 (la metà di -8).