An Efficient Residue to Weighted Converter for a New Residue Number System

Alexander Skavantzos
Department of Electrical and Computer Engineering
Louisiana State University, Baton Rouge, LA 70803
Phone: (504) 388-5240; Fax: (504) 388-5200
e-mail: alex@ee.lsu.edu

Abstract

The Residue Number System (RNS) is an integer system appropriate for implementing fast digital signal processors since it can support parallel, carry-free, high-speed arithmetic. In this paper a new RNS system and an efficient implementation of its residue-to-weighted converter are presented. The new RNS is a balanced S-moduli system appropriate for large dynamic ranges. The new residue-to-binary converter is very fast and hardware-efficient and is based on a 1's complement multiperand adder adding operands of size only 80% of the size of the system's dynamic range.

1. Introduction and Background

The Residue Number System (RNS) [1] is an integer system capable of supporting parallel, carry-free, high-speed arithmetic. The system also offers some useful properties for error detection, error correction and fault tolerance in digital systems. Important areas of application of the RNS include:

1. Digital Signal Processing (DSP) intensive computations such as digital filtering, convolutions, correlations and DFT and FFT computations [2]-[13].
2. Direct Digital Frequency Synthesis [14].

Recent work in RNS arithmetic has resulted in the development of the Quadratic Residue Number System (QRNS) [15]-[16], [7], [17]-[18] the Quadratic Like Residue Number System (QLRNS) [19], the Modified Quadratic Residue Number System (MQRNS) [20] and the generalization of the above Quadratic Residue Systems, the Polynomial Residue Number System (PRNS) [21]-[22], [9]-[10]. All these systems can support complex DSP operations using minimum computational complexity and maximum parallelism.

The basis for an RNS is a set of relatively prime integers

\[ S = \{m_1, m_2, \ldots, m_L\}, \text{ where } (m_i, m_j) = 1 \text{ for } i \neq j \]  

with \((m_i, m_j)\) indicating the greatest common divisor of \(m_i\) and \(m_j\). The set \(S\) is called the moduli set, while the dynamic range of the system is defined by the product \(M\) of all the moduli \(m_i\) in the set \(S\). Any integer \(X\) belonging to \(Z_M = \{0, 1, 2, \ldots, M-1\}\) has a unique RNS representation given by

\[ X \rightarrow^{\text{RNS}} (X_1, X_2, \ldots, X_L) \]  

where

\[ X_i = \langle x \rangle_{m_i}, \; i = 1, 2, \ldots, L \]  

while \(\langle x \rangle_m\) denotes the operation \(x \mod m\). If the integers \(X\) and \(Y\) have RNS representations \((X_1, X_2, \ldots, X_L)\) and \((Y_1, Y_2, \ldots, Y_L)\) respectively, then the RNS representation of \(W = X \otimes Y\) (where \(\otimes\) denotes addition, subtraction or multiplication) is given by

\[ W \rightarrow^{\text{RNS}} (W_1, \ldots, W_L), W_i = \langle X_i \otimes Y_i \rangle_{m_i}, \; i = 1, \ldots, L \]  

Equation (4) demonstrates the parallel, carry-free nature of the RNS. It must be mentioned that in order to ensure fast internal RNS processing, the moduli \(m_1, m_2, \ldots, m_L\) should be as small as possible. The reconstruction of \(X\) from its residues \((X_1, X_2, \ldots, X_L)\) is based on the Chinese Remainder Theorem (CRT) [1] shown by equation (5)

\[ X = \left\lfloor \sum_{i=1}^{L} \left( \frac{X_i N_i}{M_i} \right) M \right\rfloor \]  

where

\[ M = \prod_{i=1}^{L} m_i \]  

\[ M_i = \frac{M}{m_i}, N_i = \left( \frac{M_i}{m_i} \right)_{m_i}, \; i = 1, 2, \ldots, L \]
The notation \( \left( M_i^{-1} \right)_{m_i} \) in equation (7) denotes the multiplicative inverse of \( M_i \) modulo \( m_i \). Another way for converting the RNS representation \( \left( X_1, X_2, \ldots, X_L \right) \) into its weighted form \( X \) is by using the Mixed Radix Conversion (MRC) formula [1] shown by equation (8)

\[
X = X'_1 + m_1 X'_2 + m_1 m_2 X'_3 + \ldots + m_1 m_2 \ldots m_{L-1} X'_L
\]

(8)

where \( X'_1, X'_2, \ldots, X'_L \) are the mixed radix digits of \( X \).

In this paper a very efficient new residue-to-weighted conversion technique is proposed. The technique is based on combining the CRT and MRC techniques and relies on a final adder of size smaller than the size of the system’s dynamic range. Section 2 offers the new decoding technique and the class of RNS systems which are appropriate for the new technique. A new 5-moduli RNS system and the efficient implementation of its residue-to-weighted converter are presented in section 3. Finally, section 4 offers conclusions.

2. RNS Systems with Efficient Residue to Weighted Conversion

Consider an L-moduli RNS system based on the moduli set \( S \) of (1) and consider converting the residue form \( \left( X_1, X_2, \ldots, X_L \right) \) into its weighted form \( X \) by using the Chinese Remainder Theorem (CRT) of (5). An implementation of the CRT equation (5) can be based on a multioperand adder \( M^* \) (\( M^* \) is the system’s dynamic range given by (6)). Such a \( M \) multioperand adder can be efficiently implemented by a mod \( M \) Carry Save Adder (CSA) tree and a mod \( M \) Carry Propagate Adder (CPA). Let \( N \) be the size of the dynamic range \( M \) in terms of number of bits. Then

\[
N = \left\lceil \log_2 M \right\rceil
\]

(9)
The following observations are in place:

1. The size \( N \) directly affects the speed and cost of the mod \( M \) CRT multioperand adder of (5). The larger the size \( N \) is, the higher the cost of the mod \( M \) CSA tree becomes. Also, the larger the size \( N \) is, the higher the cost and the propagation delay for the final mod \( M \) CPA becomes.

2. The form of the number \( M \) (the dynamic range) affects the speed and hardware complexity of the mod \( M \) CRT multioperand adder.

A new class of RNS systems appropriate for very efficient residue-to-weighted conversion is now presented. The RNS-to-weighted converters for these new systems rely on a hardware-efficient multioperand modulo adder of size smaller than the system’s dynamic range.

Consider an RNS system based on the moduli set \( S \) of (1) where one of the moduli (say the modulus \( m_i \)) is a power of two or

\[
m_i = 2^i
\]

(10)
The conversion of the RNS representation \( \left( X_1, X_2, \ldots, X_L \right) \) into its weighted form \( X \) will now take place by using a combination of the CRT and MRC techniques as follows:

Combine the channels mod \( m_2 \), mod \( m_3 \), \ldots, mod \( m_L \) using the CRT approach. The set used for this CRT is

\[
S^* = \left\{ m_2, m_3, \ldots, m_L \right\}
\]

(11)
and the performed CRT computation is

\[
X'_2 = \left( \sum_{i=2}^{L} \left( X_i N_i^* \right)_{m_i} M_i^* \right)_{M^*}
\]

(12)
where

\[
M^* = \prod_{i=2}^{L} m_i
\]

(13)

\[
M_i^* = \frac{M^*}{m_i}, N_i^* = \left( \left( M_i^* \right)^{-1} \right)_{m_i}, i = 2, 3, \ldots, L
\]

(14)
Apply now the MRC formula on channels mod \( m_i \) and mod \( M^* \) (see (10), (13) for \( m_i \), \( M^* \)) to get

\[
X = X'_1 + m_1 X'_2
\]

(15)
where

\[
X'_1 = X_1
\]

(16)

\[
X'_2 = \left( m_1^{-1} \left( X_2 - X_1 \right) \right)_{M^*}
\]

(17)
Combining (17) and (12) yields

\[
X'_2 = \left( m_1^{-1} \left[ \sum_{i=2}^{L} \left( X_i N_i^* \right)_{m_i} M_i^* \right] - X_1 \right)_{M^*}
\]

(18)
Equation (18) dictates that \( X'_2 \) can be computed by a mod \( M^* \) multioperand adder of size \( N^* \) bits where \( N^* \) is
\[
N^* = \left\lceil \log_2 M^* \right\rceil = \left\lceil \log_2 \frac{M}{2^i} \right\rceil = N - i \tag{19}
\]

According to (15), (16) the desired \( X \) can be computed by
\[
X = X_i + m_i X_i^* \tag{20}
\]

Since \( m_i = 2^i \) (see (10)) and the residue \( X_i \) is an \( i \)-bit number, \( \langle X_i \rangle_{m_i} \), no computational hardware is needed to compute \( X \) according to equation (20). The desired \( X \) is just the result of concatenating \( X_i^* \) and \( X_i \) or
\[
X = X_i^*, X_i \tag{21}
\]

where comma (,) denotes concatenation. Thus, using the novel RNS decoding technique of (18) and (21) (which is based on combining CRT and MRC), the RNS-to-weighted conversion can rely on a multioperand modulo adder of size smaller than the size of the system's dynamic range. This is possible if one of the L moduli in the system is of form \( 2^i \). In this case, the size of the multioperand adder will be by \( i \) bits less than the size of the system's dynamic range (see (19)).

Our next concern is that the mod \( M^* \) multioperand adder of equation (18) be as fast and hardware-efficient as possible. This will of course depend on the form of the number \( M^* \) (see (13) for \( M^* \)). The best form of \( M^* \) is \( M^* = 2^s \). This is mathematically impossible, however, due to the fact that \( M^* = 2^s \) cannot be factored into pairwise relatively prime integers \( m_2, m_3, \ldots, m_L \). The second most attractive \( M^* \) is the form \( M^* = 2^n - 1 \) which is the choice considered by this paper.

3. A New RNS and its Converter Design

As seen in section 2, if a multimoduli RNS system relies on a moduli set with one modulus being of form \( 2^i \) and the product of the remaining moduli being of form \( M^* = 2^n - 1 \), then the RNS-to-weighted conversion can rely on an efficient adder of size smaller than the size of the system's dynamic range. This is possible due to the presented novel RNS decoding technique which is based on combining the CRT and MRC techniques. The simplest such RNS system is the 2-moduli system with \( m_1 = 2^n \) and \( M^* = m_2 = 2^n - 1 \). Another system of this category is the popular 3-moduli system with \( m_1 = 2^n \) and \( M^* = m_2 m_3 = 2^{2n} - 1 \) which implies \( m_2 = 2^n - 1 \) and \( m_3 = 2^n + 1 \) (23)-(27). Both the above mentioned systems rely on simple RNS-to-weighted conversion but are not appropriate for large dynamic ranges. This is due to the fact that in case of large dynamic ranges, large values of \( n \) are required resulting in apparent performance degradation of the system.

For large dynamic ranges, RNS systems with more than two or three moduli must be considered. A new 5-moduli RNS appropriate for efficient residue-to-weighted conversion is now presented. The new system is based on the set \( S \)
\[
S = \{ m_1, m_2, m_3, m_4, m_5 \}
\]
\[
= \{ 2^{n+1}, 2^n - 1, 2^n + 1, 2^n + 2^{\frac{n+1}{2}}, 2^n - 2^{\frac{n+1}{2}} + 1 \}, n \text{ is odd integer} \tag{22}
\]

Here
\[
m_1 = 2^{n+1} \tag{23}
\]
\[
M^* = \prod_{i=2}^{5} m_i = 2^{4n} - 1 \tag{24}
\]

The moduli \( m_1, m_2, \ldots, m_5 \) of the set \( S \) are pairwise relatively prime while the achieved dynamic range (in number of bits) is \( DR_S = 5n + 1 \) bits. Also, the set \( S \) implies balanced arithmetic since the sizes of the mod \( m_1, m_2, \ldots, m_5 \) processors are \( n + 1, n, n + 1 \) and \( n \) bits respectively.

Due to the fact that \( m_1 \) and \( M^* \) are forms dictated by (23)-(24), the new RNS of set \( S \) (eq. (22)) can be decoded by using the novel technique of (18) and (21). Let the RNS representation of \( X \) be \( \langle X_1, X_2, \ldots, X_5 \rangle_X \). Then (18) becomes
\[
X_i^* = \langle m_1^{-1}(AM_1^* + BM_2^* + CM_3^* + DM_4^*) - X_i \rangle_{M^*} \tag{25}
\]

where
\[
M^* = 2^{4n} - 1 \tag{26}
\]
\[
m_1 = 2^{n+1} \tag{27}
\]
\[
\langle m_1^{-1} \rangle_{M^*} = 2^{3n-1} \tag{28}
\]
\[
A = \langle X_2 N_2 \rangle_{m_2} \tag{29}
\]
\[
B = \langle X_3 N_3 \rangle_{m_3} \tag{30}
\]
\[
C = \langle X_4 N_4 \rangle_{m_4} \tag{31}
\]
\[
D = \langle X_5 N_5 \rangle_{m_5} \tag{32}
\]
\[
M_2^* = (2^n + 1)(2^{2n} + 1) \tag{33}
\]
where $m_2, m_3, m_4, m_5$ are given by (22) while $N^*_{2}, N^*_{3}, N^*_{4}, N^*_{5}$ are given by (14). It must be mentioned that the computations of A, B, C, and D (equations (29) - (32)) are performed by the existing mod $m_i$ ($i = 2, ..., 5$) multipliers which belong to the RNS processing hardware. Since A, B, C, D, and $X_i$ belong to the rings of integers mod $(2^n - 1)$, mod $(2^n + 1)$, mod$(2^n + 2^{n+1} + 1)$, mod$(2^n - 2^{n+1} + 1)$, respectively, then these numbers are of lengths $n, n+1, n+1, n$ and $n+1$ bits respectively. Let the binary (bit-level) representations of A, B, C, D, and $X_i$ be

\[ A = (a_{n-1}a_{n-2}...a_1a_0) \]

\[ B = (b_{n-1}b_{n-2}...b_1b_0) \]

\[ C = (c_{n-1}c_{n-2}...c_1c_0) \]

\[ D = (d_{n-1}d_{n-2}...d_1d_0) \]

\[ X_i = (x_{n-1}x_{n-2}...x_1x_0) \]

Combining equations (26), (28), (33) - (41) together with (25) and using simple properties of arithmetic mod$(2^n - 1)$ results in

\[ X^*_2 = \sum_{i=1}^{11} Z_i \]

where $Z_1, Z_2, ..., Z_{11}$ are the following $4n$-bit vectors

\[ Z_1 = (a_0a_{n-1}a_{n-2}...a_1)^4 \]

\[ Z_2 = b_0(0)^{n-1} b_{n-1}...b_3b_0(0)^{n-1} b_{n-1}...b_1 \]

\[ Z_3 = \bar{b}_n\bar{b}_{n-1}...\bar{b}_3\bar{b}_0(1)^{n-1} \bar{b}_n\bar{b}_{n-1}...\bar{b}_1\bar{b}_0(1)^{n-1} \]

\[ Z_4 = c_{n-1}...c_0(0)^{n-n} c_n(c_{n-1}...c_1c_0)^2 \]

\[ Z_5 = c_{n-1}...c_0(1)^{n-n} c_n(c_{n-1}...c_1c_0)^2 \]

\[ Z_6 = 1(c_n)^{2n} (1)^{2n-1} \]

\[ Z_7 = (0)^{n-1} d_{n-1}d_{n-2}...d_1d_0(0)^{n-1} \]

\[ Z_8 = \bar{d}_n\bar{d}_{n-1}...\bar{d}_1\bar{d}_0(1)^{n-n} \bar{d}_n\bar{d}_{n-1}...\bar{d}_1\bar{d}_0(1)^{n-n} \]

\[ Z_9 = d_0d_{n-1}...d_1d_0(1)^{n-n} d_0d_{n-1}...d_1d_0(1)^{n-n} \]

\[ Z_{10} = d_{n-1}d_{n-2}...d_1d_0(1)^{n-n} d_{n-1}d_{n-2}...d_1d_0(1)^{n-n} \]

\[ Z_{11} = x_nx_{n-1}...x_1x_0(1)^{n-n} \]

In the equations (43) - (53) the notation $(0)^n$ indicates a string of n zeros, $(1)^n$ indicates a string of n ones while $(abcd)^k$ indicates a 4k-bit vector where the 4-bit string $abcd$ is repeated k times. For example $(1)^3(ab)^2(0)^4$ means 111babab0000. In order to improve the readability of the paper, derivations for the expressions of $Z_1, Z_2, ..., Z_{11}$ (equations (43) - (53)) are not provided here.

Figure 1 shows an efficient implementation of the new RNS-to-binary converter for the new 5-moduli system of (22). The converter implements equation (42) using the carry save adder (CSA) approach. The converter consists of a mod $(2^{4n} - 1)$ CSA tree and a mod $(2^{4n} - 1)$ carry propagate adder (CPA). Each mod $(2^{4n} - 1)$ CSA in the tree consists of 4n full adders (FAs). The outputs of each mod $(2^{4n} - 1)$ CSA are the 4n-bit summation vector $S_i$ and the 4n-bit vector $C_i$, where $C_i$ is the left-rotated by one bit carry-out vector (end-around carry). The presence of zeros and ones in the vectors of (43) - (53) implies that some of the full adders (FAs) in the CSAs can be replaced by simpler gates (AND, OR, XOR etc.) costing less and having smaller propagation delay than a FA. The replication of certain
binary strings in the vectors of (43) - (53) implies that some CSAs will rely on fewer than 4n FAs or simplified FAs. The mod \((2^{4n} - 1)^\) CPA is a 1's complement carry propagate adder of size 4n bits. An area-time efficient design of such an adder is provided by reference [28]. The result of the 5-moduli RNS-to-weighted conversion is the concatenation of the 4n-bit vector \(X_2^n\) and the \((n+1)\)-bit residue \(X_1\) (see figure 1 and equation (21)). The cost and delay for the new RNS-to-weighted converter are

\[
\text{Converter-cost} < 36nC_1 + C_2 \quad (54)
\]

\[
\text{Converter-delay} = 5D_1 + D_2 \quad (55)
\]

where \(C_1\) and \(D_1\) are the cost and delay for a full adder (FA) while \(C_2\) and \(D_2\) are the cost and delay for a mod \(\left(2^{4n} - 1\right)\) carry propagate adder (CPA). The parameter \(n\) is the number of bits per moduli channel; (the system has a dynamic range of \(5n+1\) bits).

4. Conclusion

In this paper a class of multimoduli RNS systems appropriate for very efficient residue-to-weighted binary conversion has been presented. Specific emphasis was placed on a new balanced 5-moduli and its converter design. The presented L-moduli RNS class relies on one modulus being of the form \(m_l = 2^l\) and the product of the remaining L-1 moduli being of form \(M^* = 2^a - 1\). The proposed residue-to-weighted conversion technique is based on combining the Chinese Remainder Theorem (CRT) and the Mixed Radix Conversion (MRC) techniques. The resulting new RNS-to-weighted converters are very fast and hardware-efficient due to the following reasons:

1. The new converters rely on a multioperand adder of size smaller than the size of the system's dynamic range.
2. The multioperand adder is a mod \(\left(2^a - 1\right)\) adder.

A carry save adder (CSA) based implementation of a mod \(\left(2^a - 1\right)\) multioperand adder is faster and more hardware-efficient than the implementation of a mod \(M\) adder with \(M \neq 2^a - 1\).

References


Figure 1: An efficient implementation of the new RNS-to-binary converter