Modular Arithmetic Using Low Order Redundant Bases

M.G. Parker and M. Benaissa, *Member*, *IEEE*

Abstract—N-digit, radix-α bases are proposed for VLSI implementation of redundant arithmetic, mod m, where $\langle \alpha^N \rangle_m = \pm 1$, $\langle \alpha^j \rangle_m \neq \pm 1$, for $0 < j < N$ and m is prime. These bases simplify arithmetic overflow and are well suited to redundant arithmetic. The representations provide competitive, multiplierless T-point Number Theoretic Transforms, mod m, where $T \mid N$ or $T \mid 2N$.

Index Terms—Modular arithmetic, redundant number systems, number theoretic transforms, residue and polynomial number systems.

1 Introduction

MODULAR integer arithmetic can realize VLSI signal processing, fault-tolerant and error-correction systems, using Number Theoretic Transforms (NTTs) [5] and/or Residue and Polynomial Residue Number Systems (RNS/PRNS) [3], [8]. This correspondence proposes low order redundant α-bases, defined over suitable integer moduli, m, with m prime and $\langle \alpha^N \rangle_m = \pm 1$, $\langle \alpha^j \rangle_m \neq \pm 1$ for $0 < j < N$ (i.e., α has order N, mod m, where N is low), where α is also the integer radix of the basis, mod m, and $\langle * \rangle_m$ means, "take the residue of *, mod m." These "low order" α-bases simplify arithmetic overflow, mod m, and utilize Redundant Number Representations (RNR) to limit carry-propagation and allow parallel computation of digit sums/products [6]. An integer, e, mod m, is represented using N radix-α digits, d_i,

$$e = \left\langle \sum_{i=0}^{N-1} d_i \alpha^i \right\rangle_m \quad d_i \in \mathbf{D}_\alpha, \quad \mathbf{D}_\alpha = \{0, 1, \ldots, |\alpha| - 1 + q\} + \gamma, \quad (1)$$

where q specifies redundancy and γ is chosen to centralize the digit-set, $\mathbf{D}_\alpha$, about zero, (i.e., $\gamma = -\lfloor (|\alpha| - 1 + q)/2 \rfloor$), α, q, γ, and N are all integers and α can be negative. The ensuing discussion refers to "RNR(q_c) α-radix bases," where q indicates redundancy and c that the digit-set, $\mathbf{D}_\alpha$, is centralized about zero. Highly-efficient T-point NTTs, mod m, are possible using these bases, where $T \mid 2N$, and the ideas generalize arithmetic and transforms over Fermat and Mersenne moduli [5] to nonbinary radices. In [9], only radix-3 is considered and a generalized form of Leibowitz arithmetic is proposed, not explicitly using redundant representations. In [2], a generalization of the form $M = (q^{pn} - 1)/(q^n - 1)$ is proposed, but only developed for radix-2 and, again, RNR is not explicitly used. It is argued here that RNR is well matched to nonbinary radices for VLSI implementation of modular arithmetic using conventional binary hardware.

- M.G. Parker is with the Telecommunications Research Group, Department of Electronic and Electrical Engineering, University of Bradford, Bradford, BD7 1DP, UK. E-mail: mgparker@bradford.ac.uk.
- M. Benaissa is with the School of Engineering, University of Huddersfield, Huddersfield, HD1 3DH, UK. E-mail: m.benaissa@hud.ac.uk.

Manuscript received 20 Jan. 1995; revised 11 Jan. 1996
For information on obtaining reprints of this article, please send e-mail to: transcom@computer.org, and reference IEEECS Log Number C96310.

2 Choice of Modulus, m, Supporting a Low Order Basis

Restricting ourselves to m prime, m is chosen so that,

$$m \,\Big|\, \left(\alpha^N \pm 1\right). \quad (2)$$

In general the wordlength of (1), (N digits), is overlarge unless m is a large prime factor of $\alpha^N \pm 1$. Cyclotomic factorization [1] specifies $\alpha^N - 1 = \prod_{k|N} \Phi_k(\alpha)$ where,

$$\alpha^N - 1 = \Phi_N(\alpha)\Phi_1(\alpha) \quad \alpha^N + 1 = \Phi_{2N}(\alpha)\Phi_2(\alpha) \text{ for } N \text{ prime}$$
$$\alpha^N + 1 = \Phi_{2N}(\alpha) \text{ for } 2N \text{ a power of } 2, \quad (3)$$

where $\Phi_1(\alpha) = \alpha - 1$, $\Phi_2(\alpha) = \alpha + 1$, and $\Phi_k(\alpha)$ is the kth cyclotomic polynomial in α. Primality is not guaranteed for $\Phi_N(\alpha)$ or $\Phi_{2N}(\alpha)$, as specified in (3). For the cases of (3), an element, mod m, is represented by a radix-α low order basis of minimum or marginally overlarge N-digit wordlength and

$$m = \Phi_N(\alpha)/f \text{ or } \Phi_{2N}(\alpha)/f, \ m \text{ prime}, \ f \text{ a small integer, ideally } 1 \quad (4)$$

A few examples for positive α are given in Table 1 (there are many more). For a given m and N prime, an equivalent $-\alpha$ basis always exists, as $\Phi_N(\alpha) = \Phi_{2N}(-\alpha)$, so negative α is implicit in Table 1.

3 Redundant Number Representations (RNR)

Consider arithmetic, mod m, using centralized redundant digit-sets, $\mathbf{D}_\alpha$, with redundant input and output. Using binary logic to represent each digit-set, redundancy, ($q > 0$), must at least double the number of bits per digit-set if $\alpha = \pm 2$. For $|\alpha|$ not a power of 2, redundancy is achieved without any bit increase at all. For instance, with $\alpha = \pm 7$, RNR(0_c) and RNR(1_c) both require three bits per digit-set, i.e., redundancy demands no extra bits. In general, $q \leq 2^{\lceil \log_2(\alpha) \rceil} - \alpha$ is possible without any bit increase per digit-set. RNR(1_c) limits carry-propagation to three stages for addition and to three stages for multiplication, which can be reduced to one stage when $\alpha = \pm 2$. With RNR(2_c), addition is reduced to two stages, but still three stages for $\alpha = \pm 2$ and, for multiplication, three stages are required, but only two for $\alpha = \pm 3$. Higher redundancies achieve no more reduction in stages for addition or multiplication. (The definition of an arithmetic "stage" is clarified by the ensuing example). Values of $\alpha = \pm(2^w - 1)$ or $\pm(2^w - 2)$, for some positive integer, w, are well-suited to RNR radix-α bases using binary logic. More generally, using V-state multivalued logic, (MVL), α is ideally chosen so that,

$$\alpha = \pm\left(V^w - 1\right) \text{ or } \pm\left(V^w - 2\right), \quad w \text{ a positive integer}. \quad (5)$$

It is not essential that α satisfies (5), especially for large α. Equation (5) only identifies where RNR optimally matches V-state hardware.

4 Advantages of Low Order RNR Bases

Prime moduli, m, given by (4), are chosen over which to specify N-digit, radix-α, RNR(q_c) bases. The combination of low order basis and RNR ensures limited carry-propagation and only one long end-around carry for each arithmetic stage, mod m, as $\langle \alpha^N \rangle_m = \pm 1$. Wrap-arounds can be localized using two-planar VLSI layout, or by "folding" each stage back on itself in a single plane, to realize fully-systolic solutions. Fully-centralized digit-sets (i.e., $|\alpha| - 1 + q$ even) ensure trivial implementation of wrap-around inversion

TABLE 1
EXAMPLE LOW ORDER REDUNDANT BASES, MOD m

α	N	m	γ	q	R	t	c	p_i:q_p
2	8	$\Phi_{16}(2) = 257$	–1	1	16	9	9	–
2	16	$\Phi_{32}(2) = 65,537$	–1	1	32	17	17	–
2	7	$\Phi_{7}(2) = 127$	–1	1	14	7	7	–
2	7	$\Phi_{14}(2) = 43$	–1	1	14	6	6	–
3	7	$\Phi_{7}(3) = 1,093$	–1	1	14	12	11	–
3	7	$\Phi_{14}(3) = 547$	–1	1	14	12	10	–
3	13	$\Phi_{13}(3) = 797,161$	–1	1	26	24	20	–
3	13	$\Phi_{26}(3) = 398,581$	–1	1	26	24	19	–
3	16	$\Phi_{32}(3)/2$	–1	1	32	30	25	–
3	32	$\Phi_{64}(3)/2$	–1	1	64	62	50	–
3	64	$\Phi_{128}(3)/2$	–1	1	128	126	101	–
5	5	$\Phi_{10}(5) = 521$	–3	2	15	12	10	2,3:1
5	13	$\Phi_{13}(5)$	–3	2	39	36	29	2,3:1
6	4	$\Phi_{8}(6) = 1,297$	–3	2	12	12	11	–
6	7	$\Phi_{7}(6) = 55,987$	–3	2	21	18	16	–
6	11	$\Phi_{22}(6)$	–3	2	33	30	26	–
7	4	$\Phi_{8}(7)/2 = 1,201$	–3	1	12	12	11	–
7	5	$\Phi_{5}(7) = 2,801$	–3	1	15	12	12	–
7	13	$\Phi_{13}(7)$	–3	1	39	36	34	–
7	17	$\Phi_{34}(7)$	–3	1	51	48	46	–
11	4	$\Phi_{8}(11)/2 = 7,321$	–6	2	16	16	13	3,4:1
15	3	$\Phi_{3}(15) = 241$	–7	1	12	8	8	–
15	3	$\Phi_{6}(15) = 211$	–7	1	12	8	8	–
23	5	$\Phi_{5}(23) = 292,561$	–12	1	25	20	19	2,3,4:1
24	4	$\Phi_{8}(24) = 331,777$	–12	2	20	20	19	4,7:4
27	3	$\Phi_{3}(27) = 757$	–14	2	15	10	10	4,7:1
29	4	$\Phi_{8}(29)/2 = 353,641$	–15	2	20	20	19	2,3,5:1
59	3	$\Phi_{3}(59) = 3,541$	–30	2	18	12	12	3,4,5:1
120	2	$\Phi_{4}(120) = 14,401$	–60	1	14	14	14	2,7,9:6
124	2	$\Phi_{4}(124) = 15,377$	–62	1	14	14	14	2,7,9:2
126	2	$\Phi_{4}(126) = 15,877$	–63	1	14	14	14	2,5,13:4

R = red′ bit-w′len, t = trans′ bit-w′len, $c = \lceil \log_2(m) \rceil$ = conv′ bit-w′len

when $\left\langle \alpha^N \right\rangle_m = -1$, by simply inverting digit-set states, and is an alternative to the Leibowitz technique [9], both for $\alpha = 2$ or otherwise. If $|\alpha| - 1 + q$ is odd then full digit-set centralization is impossible. However it is still possible, using digit-set offsets, to trivially absorb wrap-around digit inversion, and the use of offsets will be demonstrated in the ensuing example. It is expected that corresponding VLSI implementations will be advantageous in terms of area, speed and throughput, even if wordlengths are marginally overlarge. Wordlengths can sometimes be reduced from N digits to $N - 1$ digits, prior to transmission, as follows:

If $m = \Phi_N(\alpha)$, N prime, then $m = \sum_{i=0}^{N-1} \alpha^i$. Using RNR($1_c$) the range, r, covered by the first $N - 1$ digits is $r = \sum_{i=0}^{N-2} \alpha \times \alpha^i = m - 1$. If $m = \Phi_{2N}(\alpha)$, N prime, then $m = \sum_{i=0}^{N-1} (-\alpha)^i$. Using RNR($1_c$) the range covered by the first $N - 1$ digits is $r = \sum_{i=0}^{N-2} \alpha \times (-\alpha)^i = -m + 1$. In both cases r is sufficient to represent every element, mod m, using $N - 1$ digits.

Bit-wordlengths for redundant ("R"), transmission ("t") and conventional ("c" $= \lceil \log_2(m) \rceil$) representations are given in Table 1 for each value of m, along with suggested redundancy parameters, γ and q. Table 1 shows that it is possible that $t = c$ and, for large α, marginally less than a power of two, $R = c$ is possible, so wordlengths are not necessarily increased. If α is not small digit-set arithmetic can be decomposed over RNS. RNS moduli, p_i, will satisfy,

$$|\alpha| - 1 + q_p = \prod_i p_i \tag{6}$$

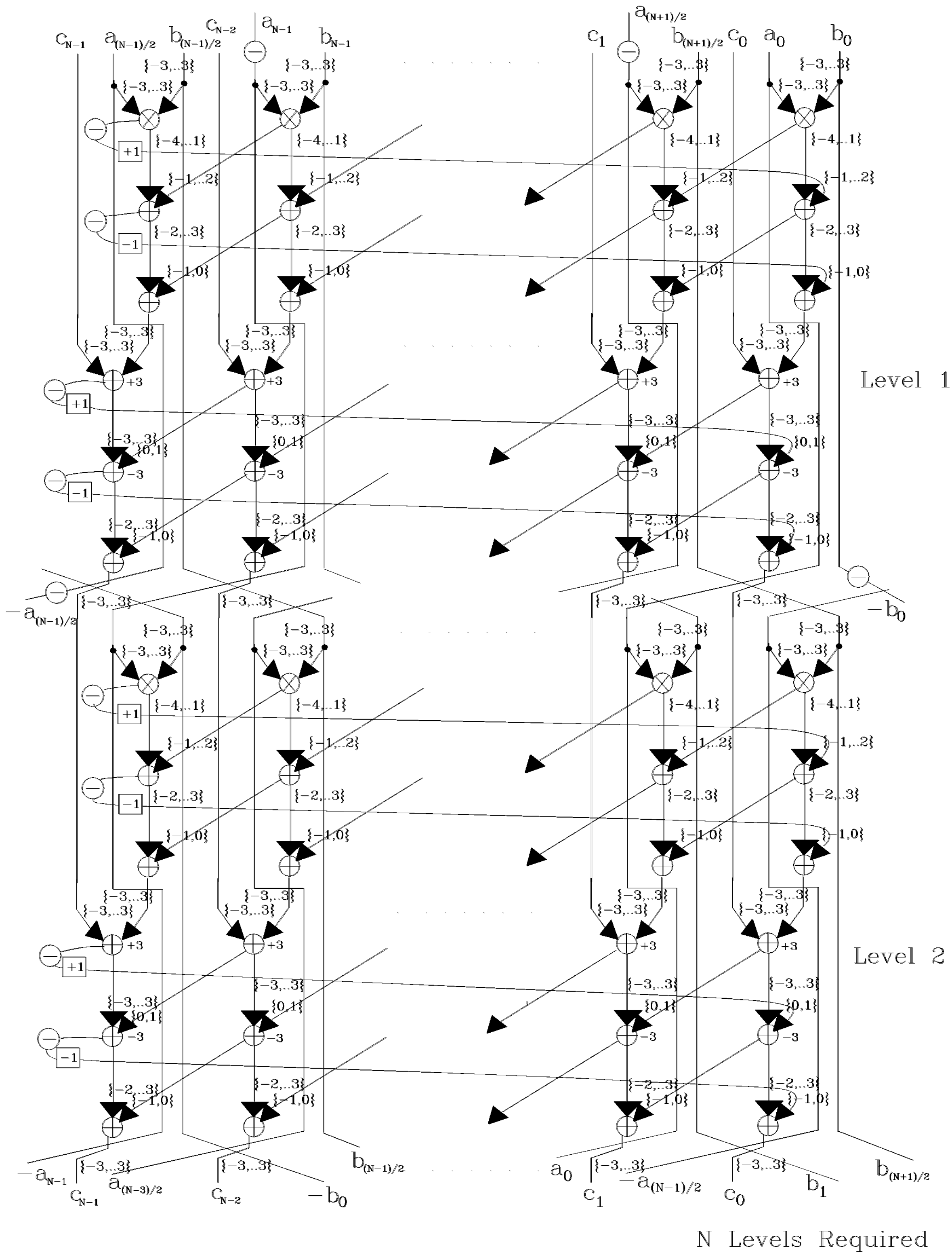

Fig. 1. Six-basis RNR(1_c) multiplier, mod $m|(6^N + 1)$, N odd.

where q_p is the redundancy using digit-set RNS. Suggested p_i and q_p are given in Table 1. This method is developed in [4].

EXAMPLE: Let $\alpha = 6$, $q = 1$, and $\gamma = -3$. Then a radix-6, RNR(1_c) multiplier can be proposed, as shown in Fig. 1, with the data flow shown in Fig. 2, where,

$$a = \left\langle \sum_{i=0}^{N-1} a_i 6^i \right\rangle_m,\ b = \left\langle \sum_{i=0}^{N-1} b_i 6^i \right\rangle_m,\ c = \left\langle \sum_{i=0}^{N-1} c_i 6^i \right\rangle_m$$

$$c = \langle a \times b \rangle_m,\ m \Big| \left(6^N + 1\right), \text{ and } a_i, b_i, c_i \in \{-3, \dots, 3\}$$

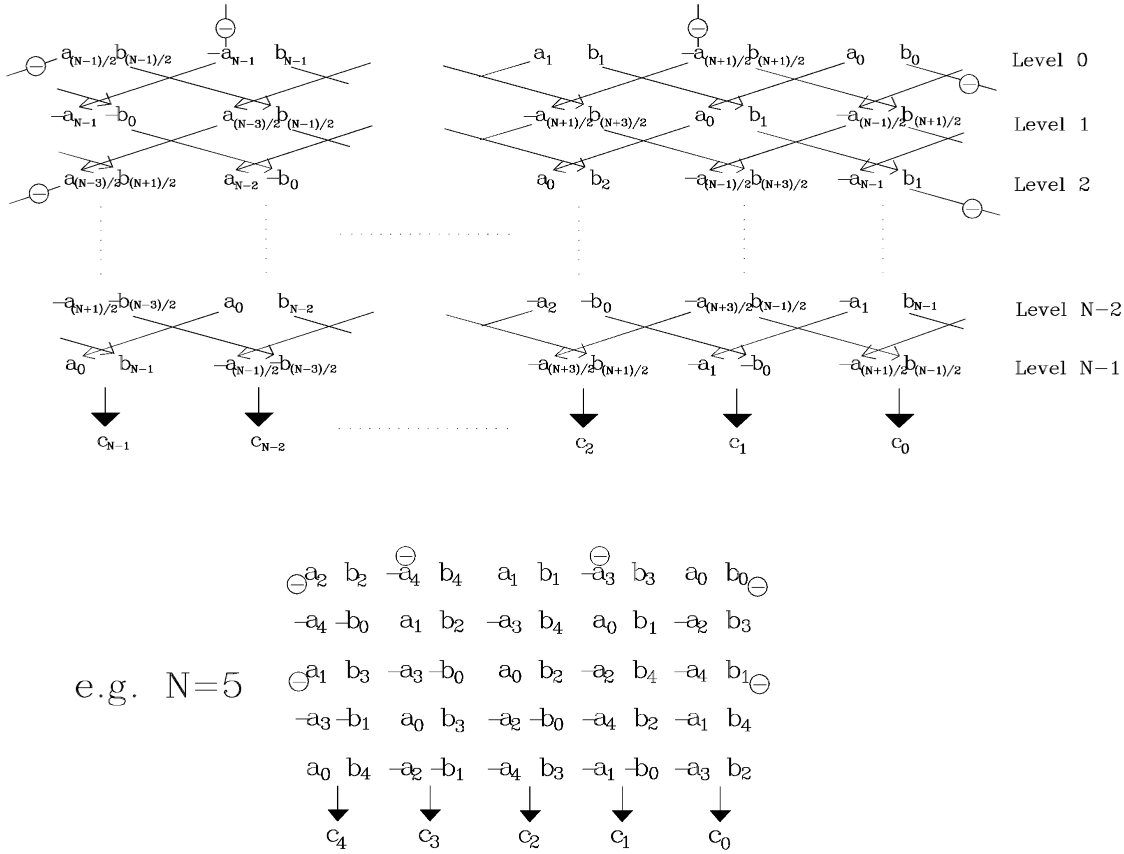

Fig. 2. Data flow for α-basis multiplier, mod $m|(\alpha^N + 1)$, N odd.

In Fig. 1, circled "+" (+ x), "×," "−," imply addition (+ offset x), multiplication, and negation, respectively, and squared "+x" implies offset by x. The three stages of digit-set multiplication are as follows,

Stage	Carry	Sum	Computation
1		$\{-3, \ldots, 3\} \times \{-3, \ldots, 3\}$	$\{-3, \ldots, 3\} \times \{-3, \ldots, 3\} \in \{-1, \ldots, 2\}6 + \{-4, \ldots, 1\}$
2	$\{-1, \ldots, 2\}$	$\{-4, \ldots, 1\}$	$\{-1, \ldots, 2\} + \{-4, \ldots, 1\} \in \{-1, 0\}6 + \{-2, \ldots, 3\}$
3	$\{-1, 0\}$	$\{-2, \ldots, 3\}$	$\{-2, \ldots, 3\} + \{-1, 0\} \in \{-3, \ldots, 3\}$
	–	$\{-3, \ldots, 3\}$	

But $\left\langle 6^N \right\rangle_m = -1$, so the most-significant-digit, (msd), carries are inverted on wrap-around. For stage 1, the msd output carry digit-set, $\{-1, \ldots, 2\}$, is inverted to become $\{-2, \ldots, 1\}$. An offset of 1 is then added to $\{-2, \ldots, 1\}$ to give $\{-1, \ldots, 2\} = \{-2, \ldots, 1\} + 1$, before passing the digit to the input carry of the least-significant-digit, (lsd), of stage 2. Similarly the msd output carry digit-set, $\{-1, 0\}$, of stage 2 is inverted and offset by -1 to give $\{-1, 0\} = -1\{-1, 0\} - 1$. The two offsets, 1 and -1, cancel to give a partial product without offset.

The three stages of digit-set addition of partial products are as follows,

Stage	Carry	Sum	Computation
4		$\{-3, \ldots, 3\} + \{-3, \ldots, 3\} + 3$	$\{-3, \ldots, 3\} + \{-3, \ldots, 3\} + 3 \in \{0, 1\}6 + \{-3, \ldots, 3\}$
5	$\{0, 1\}$	$\{-3, \ldots, 3\} - 3$	$\{0, 1\} + \{-3, \ldots, 3\} - 3 \in \{-1, 0\}6 + \{-2, \ldots, 3\}$
6	$\{-1, 0\}$	$\{-2, \ldots, 3\}$	$\{-1, 0\} + \{-2, \ldots, 3\} \in \{-3, \ldots, 3\}$
	–	$\{-3, \ldots, 3\}$	

Stages 4 and 5 offset the sum by $\sum_{i=0}^{N-1} 3(6^i)$ and $\sum_{i=0}^{N-1} -3(6^i)$, mod m, respectively. These offsets cancel. Once again the msd carries undergo inversion and offset, and these offsets also cancel. The final product output, c, therefore undergoes no offset. All offsets and inversion are implicitly implemented in the associated cells and are without cost. Figs. 2 and 3 show the data flow through a multiplier where $m|(\alpha^N + 1)$, N odd and even, respectively, where the circled "−" imply digit-set negation. The negation is implicitly implemented and costs nothing. Fig. 1 is shown for N odd.

Examples for $m|(\alpha^N - 1)$ are implemented in a similar fashion but without negation on wrap-around.

5 Applications

Consider the NTT,

$$X[k] = \left\langle \sum_{n=0}^{T-1} x[n] \left(\alpha^{L/T}\right)^{nk} \right\rangle_m \quad \text{where } \alpha \text{ has order } L, \text{mod } m, \text{ and } T|L \qquad (7)$$

Using an N-digit, low order, redundant basis, with $L = N$ or $2N$, (7) will require multiplication implemented as (skew) cyclic rotations and $T(T - 1)$ additions. These NTTs are highly efficient and can, in turn, realize efficient T-point (skew) cyclic convolutions or PRNS, mod $x^T \pm 1$, using the same arithmetic. Furthermore, longer blocklength NTTs can be decomposed into smaller length T-point NTTs using a combination of the Prime Factor Algorithm and repeated applications of Rader's algorithm, as detailed in [7]. These NTTs compete with Fermat and Mersenne Transforms, provide a wide choice of moduli and many more NTT blocklengths. Identical blocklength NTTs over different moduli, m_i, can be combined using RNS, to increase dynamic range.

6 Assessment and Conclusion

Low order redundant bases have been defined over suitable prime moduli, m, to simplify arithmetic overflow and limit carry-

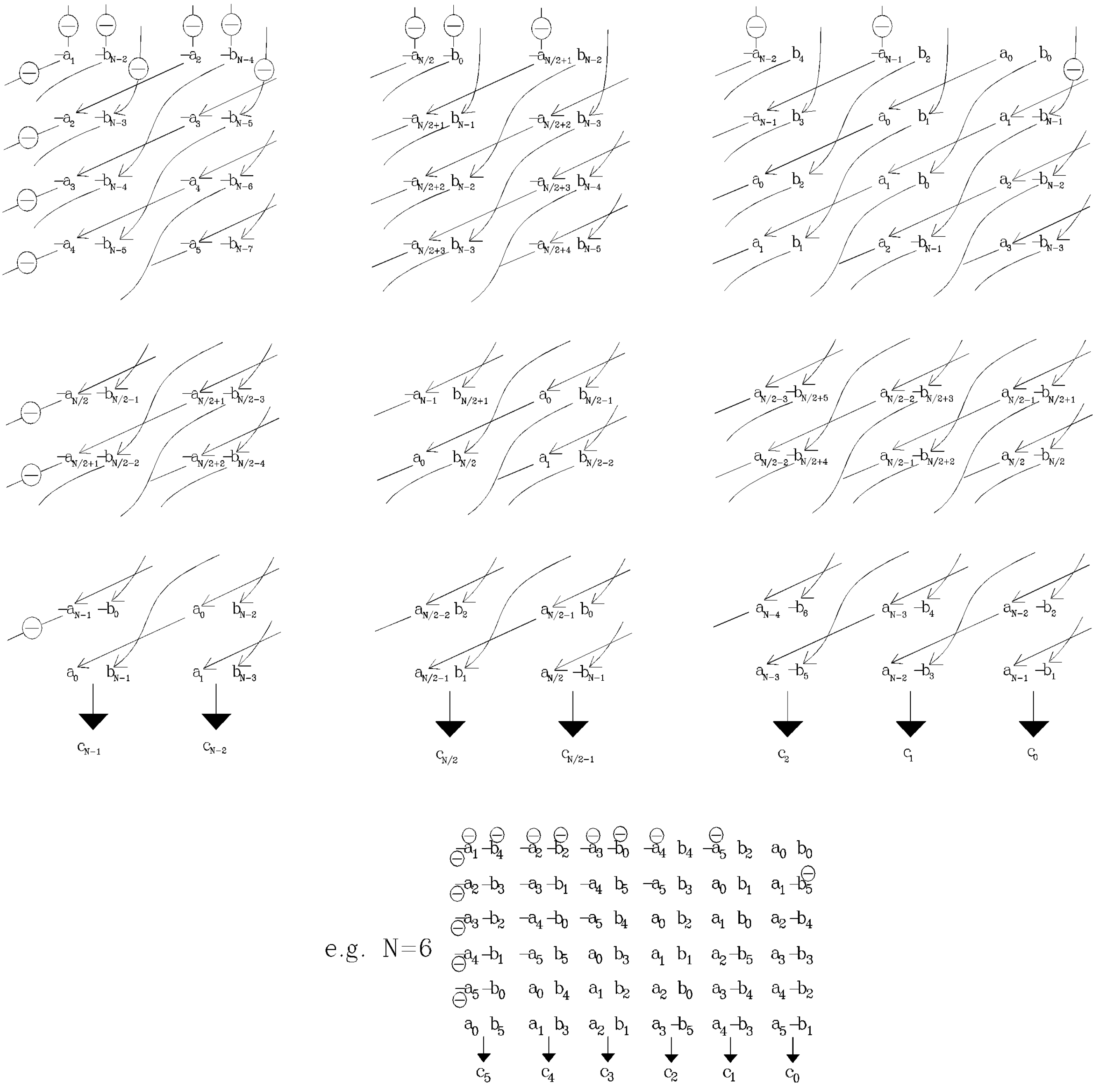

Fig. 3. Data flow for α-basis multiplier, mod $m|(\alpha^N + 1)$, N even.

propagation, mod m. Redundancy is suited to non-power-of-two radices. VLSI implementations of adders and multipliers, mod m, will be symmetric and systolic, as demonstrated by the example, which also highlights the implicit inclusion of offset and negation. The area and latency of a multiplier will be $O(N^2)$ and $O(N)$, respectively, where N is the wordlength in digits, and the circuits can be pipelined down to an arithmetic "stage," where carries propagate over two or three arithmetic "stages." As a comparison, the Montgomery modular multiplier of [10] achieves approximately the same figures for complexity. However, the systolic cells of [10] will be larger and slower as each cell also has to realize a modular reduction. Moreover there is some unavoidable post-processing for [10]. On the other hand, [10] does not need to accommodate wrap-around interconnections, has less restrictions on choice of modulus, m (and is therefore more suited to cryptographic applications), and, unlike the circuits of this paper, is fully systolic in a single VLSI plane. Future work will apply the low order basis to Montgomery multipliers. Low order redundant bases are ideal for defining T-point NTTs, mod m, $T|N$ or $T|2N$, where the radix, α, has order N or $2N$, mod m, (N prime or $2N$, a power of 2). These arithmetic and transform modules can realize a wide range of signal processing, fault-tolerant, and error-correction systems.

ACKNOWLEDGMENT

This work was undertaken while M.G. Parker was at the University of Huddersfield.

REFERENCES

[1] R.E. Blahut, *Fast Algorithms for Digital Signal Processing*. Reading, Mass.: Addison-Wesley, 1985.
[2] V.S. Dimitrov, T.V. Cooklev, and B.D. Donevsky, "Generalized Fermat-Mersenne Number Theoretic Transform," *IEEE Trans. Circuits and Systems-II*, vol. 41, no. 2, pp. 133-138, Feb. 1994.
[3] H. Krishna, B. Krishna, K.-Y. Lin, and J.-D. Sun, *Computational Number Theory and Digital Signal Processing*. CRC Press, 1994.
[4] M-B. Lin and A.Y. Oruc, "A Fault-Tolerant Permutation Network Modulo Arithmetic Processor," *IEEE Trans. VLSI Systems*, vol. 2, no. 3, pp. 312-319, Sept. 1994.
[5] J.H. McClellan and C.M. Rader, *Number Theory in Digital Signal Processing*. Prentice Hall, 1979.
[6] B. Parhami, "Generalized Signed-Digit Number Systems: A Unifying Framework for Redundant Number Representations," *IEEE Trans. Computers*, vol. 39, no. 1, pp. 89-98, Jan. 1990.
[7] M.G. Parker and M. Benaissa, "Unusual-Length Number-Theoretic Transforms Using Recursive Extensions of Rader's Algorithm," *IEE Proc-Visualization Image Signal Processing*, vol. 142, no. 1, pp. 31-34, Feb. 1995.
[8] M.G. Parker, "VLSI Algorithms and Architectures for the Implementation of Number-Theoretic Transforms, Residue and Polynomial Residue Number Systems," PhD thesis, School of Eng., Univ. of Huddersfield, Mar. 1995.
[9] S. Sunder and A. Antoniou, "Arithmetic for Ternary Number-Theoretic Transforms," *IEEE Trans. Circuits and Systems-II*, vol. 40, no. 8, pp. 529-531, Aug. 1993.
[10] C.D. Walter, "Systolic Modular Multiplication," *IEEE Trans. Computers*, vol. 42, no. 3, pp. 376-378, Mar. 1993.