



VLSI Research Group

University of Windsor

A Residue Number System
Implementation of Real Orthogonal
Transforms

IEEE Transactions on Signal Processing

V. S. Dimitrov, G. A. Jullien, and W. C. Miller
Corresponding author G.A. Jullien

A Residue Number System Implementation of Real Orthogonal Transforms

V. S. Dimitrov, G. A. Jullien, and W. C. Miller Corresponding author G.A. Jullien

Abstract

Recent work has focused on performing residue computations that are quantized within a dense ring of integers in the real domain. The aims of this paper are to provide an efficient algorithm for the approximation of real input signals, with arbitrarily small error, as elements of a quadratic number ring, and to prove RNS moduli restrictions for simplified multiplication within the ring. The new approximation scheme can be used for implementation of real-valued transforms and their multidimensional generalizations.

Keywords: residue number systems, computer arithmetic, number theory, number systems, discrete transforms, finite rings, computational complexity

1.0 Introduction

Many digital signal processing (*DSP*) algorithms use input data that is real. There exists a wide spectrum of real-valued orthogonal transforms dedicated to overcome the need of using the complex *DFT*, among the most popular are the discrete Hartley transform (*DHT*), discrete cosine transform (*DCT*), discrete sine transform (*DST*), Haar transform, and slant transform [1]. One of the well-known algebraic ‘tricks’ to improve the computational efficiency of the *DFT* is the use of quadratic residue number systems (*QRNS*). This allows the multiplication of two Gaussian complex numbers to be performed using only 2 integer multiplications. A comprehensive survey of the theory and applications of the *QRNS*, can be found in reference [2].

In [3] Cozzens and Finkelstein proposed a parallel algorithm for computation of the *DFT* via computation of 1) algebraic-integer quantization of the input signal (an outer level of parallelism) and 2) an *RNS* implementation of the arithmetic operations over algebraic integers (an inner level of parallelism). Bequilard and O’Neil used a similar idea in deriving an efficient *RNS* implementation of 2-D *DCT* [4]. Because the structure of the algebraic-integer quantization causes severe limitations over the *RNS* moduli, the main problem is in finding a sufficiently good compromise between these conditions.

The aim of our paper is twofold. First, we propose an approximation scheme which can be used for all real-valued transforms. The technique we employ allows a representation of the real input samples as pairs of integers, where the error of the approximation can be made arbitrarily small; the remaining processing steps are error-free. The problem of finding such an approximation scheme has previously been posed by G.Ray [5]. For example,

working with the quadratic number ring $Z(\sqrt{2})$, one needs to approximate the input samples using the form $a + b\sqrt{2}$, where a and b are integers. Our algorithm is based on a new irrational number system, which can be viewed as a generalization of that proposed by Bergman [6].

Second, to simplify the multiplication over the field, we can apply a technique similar to the *QRNS* approach; in this case, however, the restrictions on the moduli are much weaker, compared to the Cozzens-Finkelstein approach. (The Cozzens-Finkelstein restrictions, are the main limitations to the wider applicability of their algorithm [15].) With a proper choice of the quadratic ring, we can reach a sufficiently large dynamic range using fairly small moduli. If the dynamic range of the computation is fixed, it is important to find those rings of quadratic integers, such that the cardinality of their direct product is greater than the dynamic range.

Our investigations are mainly based on approximation over quadratic number rings $Z(\sqrt{2})$ used in [4], and $Z(\sqrt{5})$ used in [7]. We show that when 6-bit moduli are allowed, the computation over $Z(\sqrt{5})$ corresponds to the 34 bits dynamic range, whereas the computation over $Z(\sqrt{2})$ supports 22 bits of dynamic range. However, for 7-bit moduli the situation is reversed - we have the dynamic range equal to 74 bits in $Z(\sqrt{2})$ versus a dynamic range of 67 bits in $Z(\sqrt{5})$. The proof of the results are mainly based on the properties of the Legendre symbol and some simple number-theoretic considerations.

Our new technique can be extended in a straightforward manner to other quadratic rings. Some asymptotic estimations can be obtained if one exploits Chebotarev density theorems from analytic number theory [8].

2.0 Computations in Quadratic Extensions of $\mathbf{Z}/M\mathbf{Z}$

The following notation is used in the rest of the paper. Let:

\mathbf{R} = the real numbers;

\mathbf{C} = the complex numbers;

\mathbf{Q} = the rationals;

\mathbf{Z} = the integers;

$\{m_1, m_2, \dots, m_L\}$ = a set of relatively prime moduli;

M = the product of the moduli $\prod_i m_i$, $(i=1,2,\dots,L)$;

$\mathbf{Z}(\sqrt{s})$ = a real quadratic number ring, whose elements are of the form $x + y \cdot \sqrt{s}$, if

$s \equiv 2, 3 \pmod{4}$ or $x + y \cdot \frac{1 + \sqrt{s}}{2}$ with $s \equiv 1 \pmod{4}$, s is squarefree.

$\mathbf{Z}_M(\sqrt{s})$ = a real quadratic number ring, whose elements are of the form above with M

a sufficiently large integer and $|x, y| \leq \frac{M}{2}$, $x, y \in \mathbf{Z}$. s is called a discriminant of the

real quadratic number ring. Addition and multiplication are performed modulo M .

The usual way to exploit the advantages of *RNS* computations, when one deals with real numbers, consists of two steps: scale the real numbers by a factor q , where q is a sufficiently large integer, and approximate them to the nearest integer. Although well known and widely used, this approach has at least two disadvantages. Firstly, in order to ensure a proper accuracy of the computation, q should be sufficiently large. Secondly, to avoid overflow one has to use either 1) many small prime moduli, which complicates the final reconstruction via the Chinese Remainder Theorem (*CRT*), or 2) a small number of large moduli. In this latter case, the possibility of using look-up tables to implement the basic modular arithmetic operations is limited. In fact, this scale and approximate strategy can be viewed as a rational approximation of real numbers of the form (p,q) , where q is a constant. Reed and Truong [16] were the first to suggest the use of quadratic integers to compute convolutions, showing that number-theoretic transforms (*NTTs*) over the rings of quadratic integers possess the cyclic convolution property. The problem of how to transform the input real sequence into quadratic integer form, however, remains open. For the quadratic number ring $Z(\sqrt{5})$ a proper algorithm has been proposed in [7].

Here we propose an approximation algorithm, applicable for all quadratic number rings.

2.1 Generalized irrational number systems

In 1957 Bergmann published the following representation for positive real numbers:

$$A = \sum_i a_i \omega^i, \quad A \text{ - real, } a_i \in \{0, 1\}, \quad \omega = \frac{1 + \sqrt{5}}{2}.$$

This number representation is a key point in the algorithm for number-theoretic transforms over the ring $Z(\sqrt{5})$ [7]. Here we extend this algorithm to other quadratic rings including higher-order algebraic rings and transcendental rings.

Let γ be an arbitrary irrational number, z is positive real, and $0 < \gamma < 1$. Consider the following form:

$$z = \sum_{i=1}^{\infty} d_i \gamma^{-i} \quad (1)$$

where $0 \leq d_i \leq \lfloor \gamma \rfloor$ are integers. We shall call this form a *generalized irrational number system* (GINS). When $1 < \gamma < 2$ then z is represented as a sum of different negative powers of γ . The existence of such a representation follows from a theorem proved by Erdos, Horvath and Joo [14]. Let d_i be referred to as *digits* of z in the GINS. The procedure of computing the digits d_i of z in *GINS* is as follows:

Step 1: $z_0 = z; i = 0;$

Step 2: $z_1 = z_0 \cdot \gamma; z_1 = \text{int}(z_1) + \text{frac}(z_1); d_i = \text{int}(z_1); z_0 = \text{frac}(z_1); i = i + 1;$

Step 3: Goto *Step 2*;

The negative powers of γ can be approximated in advance using the form:

$$\gamma^{-i} \approx \alpha_i + \beta_i \cdot \gamma \quad (2)$$

Let us suppose that the representation of z in $GINS$ has K digits. Then z can be represented as shown in eqn. (3):

$$z = \sum_{i=1}^K d_i \cdot (\alpha_i + \beta_i \gamma) = \left(\sum_{i=1}^K d_i \alpha_i \right) + \left(\sum_{i=1}^K d_i \beta_i \right) \gamma \quad (3)$$

The approximation of z into the form $\alpha + \beta \cdot \gamma$ with error smaller than ε comprises: 1) finding a proper value for K ; 2) precomputing and storing sufficiently good approximations for γ^{-i} .

Let us substitute $\alpha = \sum_{i=1}^K d_i \alpha_i$ and $\beta = \sum_{i=1}^K d_i \beta_i$ and consider the inequality:

$$|z - (\alpha + \beta \cdot \gamma)| = \left| \sum_{i=1}^{\infty} d_i \gamma^{-i} - \sum_{i=1}^K d_i (\alpha_i + \beta_i \cdot \gamma) \right| < \varepsilon \quad (4)$$

From eqn. (4) it follows that K should be selected such that:

$$\left| \sum_{i=K+1}^{\infty} d_i \gamma^{-i} + \delta \right| < \varepsilon \quad (5)$$

where δ is the sum of the errors of precomputed approximations:

$$\delta = \sum_{i=1}^K (\gamma^{-i} - (\alpha_i + \beta_i \cdot \gamma)).$$

Because δ can be made arbitrarily small, we focus our

attention on the first term in eqn. (5). We have:

$$\left| \sum_{i=K+1}^{\infty} d_i \gamma^{-i} \right| \leq \gamma \cdot \sum_{i=K+1}^{\infty} \gamma^{-i} = \frac{1}{\gamma^K} \cdot \sum_{i=0}^{\infty} \gamma^{-i} = \frac{1}{(1-\gamma) \cdot \gamma^{K+1}} < \varepsilon \quad (6)$$

Therefore K should be selected such that:

$$K > \frac{\log \frac{1}{\varepsilon \cdot (\gamma - 1)}}{\log \gamma} - 1 \quad (7)$$

2.2 Some examples

Let us consider the algorithm for approximating real numbers into the form $a + b \frac{1 + \sqrt{5}}{2}$,

that is, as elements of the ring $Z[\sqrt{5}]$. Let $\omega = \frac{1 + \sqrt{5}}{2}$.

First of all, we mention that the negative powers of ω can be presented *exactly* using the form $a + b\omega$, due to the identity [10]:

$$\omega^{-i} = (-1)^i F_{i+1} + (-1)^{i+1} F_i \omega \quad (8)$$

where F_i is the i -th Fibonacci number, defined as the following recurrence relation:

$$F_0=0, F_1 = 1 \text{ and } F_i = F_{i-1} + F_{i-2} \text{ if } i > 1.$$

This means that the error of the approximations of the precomputed powers, in this case, is equal to zero. As an example, the representation of the first twenty negative powers is:

$$\omega^{-1} = -1 + \omega; \omega^{-2} = 2 - \omega; \omega^{-3} = -3 + 2\omega; \omega^{-4} = 5 - 3\omega; \omega^{-5} = -8 + 5\omega;$$

$$\begin{aligned}
 \omega^{-6} &= 13 - 8\omega; \omega^{-7} = -21 + 13\omega; \omega^{-8} = 34 - 21\omega; \omega^{-9} = -55 + 34\omega; \\
 \omega^{-10} &= 89 - 55\omega; \omega^{-11} = -144 + 89\omega; \omega^{-12} = 233 - 144\omega; \\
 \omega^{-13} &= -377 + 233\omega; \omega^{-14} = 610 - 377\omega; \omega^{-15} = -987 + 610\omega; \\
 \omega^{-16} &= 1597 - 987\omega; \omega^{-17} = -2584 + 1597\omega; \omega^{-18} = 4181 - 2584\omega; \\
 \omega^{-19} &= -6765 + 4181\omega; \omega^{-20} = 10946 - 6765\omega.
 \end{aligned}$$

Now let us apply the above algorithm for $x = 0.6723$.

The *GINS* representation of x , using 20 digits, is:

$$\begin{aligned}
 0.6723 &= (0.10000010101001001010)_\omega \\
 &= \omega^{-1} + \omega^{-7} + \omega^{-9} + \omega^{-11} + \omega^{-14} + \omega^{-17} + \omega^{-19}
 \end{aligned}$$

Replacing the negative powers of ω we find:

$$\begin{aligned}
 0.6723 &\approx (-1 + \omega) + (-21 + 13\omega) + (-55 + 34\omega) + (-144 + 89\omega) \\
 &\quad + (610 - 377\omega) + (-2584 + 1597\omega) + (-6765 + 4181\omega) \\
 &= -8960 + 5538\omega
 \end{aligned}$$

Note that the nonexistence of consecutive ones in the ω -representation of x is not merely a ‘lucky event’, but can be justified by the fact that ω is a root of the equation

$x^2 - x - 1 = 0$. The overall proportion of ones in the representation of a given real num-

ber in the ω -representation is, on average, $\frac{5 - \sqrt{5}}{10} = 0.276393\dots$ [12], which is consid-

erably better than the standard binary number system average of 0.5.

The second example is based on the approximation of the real numbers of the form

$a + b\sqrt{2}$, where a and b are integers. The negative powers of $\sqrt{2}$ can not be represented

exactly in this form, and therefore we have to compute their approximations in advance.

The corresponding pairs for the first 14 negative powers of $\sqrt{2}$ are shown in Table 1:

Table 1. The first 14 negative powers of $\sqrt{2}$ and their approximations

i	$\sqrt{2}^{(-i)}$	$a + b\sqrt{2}$	i	$\sqrt{2}^{(-i)}$	$a + b\sqrt{2}$
1	0.70710678	(-408,289)	2	0.5	(-288,204)
3	0.35355339	(204,-144)	4	0.25	(-144,102)
5	0.17677669	(102,-72)	6	0.125	(-72,51)
7	0.08838834	(51,-36)	8	0.0625	(-444,314)
9	0.04419417	(314,-222)	10	0.03125	(-222,157)
11	0.02209702	(157,-111)	12	0.015625	(297,-210)
13	0.01104853	(775,-548)	14	0.0078125	(437,-309)

Let us consider the approximation of the number $x = 0.8036$. The representation of x in the $\sqrt{2}$ -irrational number system with 20 digits precision is:

$$x \approx (\sqrt{2})^{-1} + (\sqrt{2})^{-7} + (\sqrt{2})^{-14} \quad (9)$$

After substituting the corresponding powers we have:

$$x \approx (-408 + 289\sqrt{2}) + (51 - 36\sqrt{2}) + (437 - 309\sqrt{2}) = 80 - 56\sqrt{2} \quad (10)$$

We should mention that in the $\sqrt{2}$ -irrational number system, the minimal distance between ones is 3, that is, combinations of digits...011... or...101... are impossible. This follows from the fact that $\sqrt{2}$ is smaller than the positive root of the cubic equation $x^3 - x^2 - 1 = 0$. The complexity of this algorithm depends strongly upon the average number of ones in the corresponding irrational number system. This is encapsulated in the following theorem:

Theorem 1: Let γ be an arbitrary irrational number between 1 and 2. Then the average number of ones in the γ -based GINS is:

$$F(\gamma) = \begin{cases} \frac{\gamma - 1}{(2\gamma - 1) \cdot \log_2 \gamma} & \text{if } \left(\gamma \leq \frac{1 + \sqrt{5}}{2} \right) \\ \frac{(\gamma - 1)^2}{(1 + (\gamma - 1)^2) \cdot \log_2 \gamma} & \text{if } \left(\gamma > \frac{1 + \sqrt{5}}{2} \right) \end{cases} \quad (11)$$

Proof: (see Appendix I)

□

3.0 Mapping into a quadratic-like residue number system

In this section we shall consider the case of quadratic number rings in which the discriminant, s , is of the form $4k + 2$ or $4k + 3$ and $\gamma = \sqrt{s}$. All of the proofs below can be extended to the case of discriminants of the form $4k + 1$. Let M be an odd integer, which is pairwise prime to the discriminant s .

Let $Z_M[\gamma] = \{ a + \gamma b \mid a, b \in Z_M \}$ be a set of M^2 elements, defined in such a manner that addition is given by $(a + \gamma \cdot b) + (c + \gamma \cdot d) = (a + c) + \gamma \cdot (b + d)$ and multiplication is given by $(a + \gamma \cdot b) \cdot (c + \gamma \cdot d) = (a \cdot c + s \cdot b \cdot d)_M + \gamma \cdot (a \cdot d + b \cdot c)_M$. The set $Z_M[\gamma]$ is a commutative ring. To show this it is only necessary to show that any arbitrary three elements of $Z_M[\gamma]$ satisfies the postulates of the ring.

Let h be a solution of the congruence:

$$x^2 \equiv d \pmod{M} \quad (12)$$

(In the next section we shall give a necessary and sufficient condition for the existence of such a solution.)

Now we map an element $a + \gamma \cdot b$ in $Z_M[\gamma]$ into $(a + h \cdot b)_M$. It is easy to show that such a mapping is a homomorphic mapping. If one uses both solutions of the congruence

$x^2 \equiv s \pmod{M}$, $\pm h$ for mapping an element into $(\Delta, \bar{\Delta})$, where $\Delta = (a + h \cdot b)_M$ and

$\bar{\Delta} = (a - (h \cdot b))_M$, we show, in Theorem 2 below, that such a mapping is isomorphic and

that the set $\{\Delta, \bar{\Delta}\}$, $\Delta, \bar{\Delta} \in Z_M$ is the direct sum of two copies of Z_M of M^2 elements.

This theorem is a special case of lemma 2.3, proved by Cozzens and Finkelstein in [3]. A similar result for the case of Gaussian integers is proposed in [9].

Theorem 2: The direct sum of 2 copies of Z_M is:

$$S_M = Z_M + Z_M = \{ \{\Delta, \bar{\Delta}\}, \Delta, \bar{\Delta} \in Z_M \}$$

where $\pm h$ are the solutions of $x^2 \equiv s \pmod{M}$, $\Delta = (a + h \cdot b)_M$, $\bar{\Delta} = (a - (h \cdot b))_M$,

$(\Delta, \bar{\Delta}) + (\zeta, \bar{\zeta}) = (\Delta + \zeta, \bar{\Delta} + \bar{\zeta})$ and $(\Delta, \bar{\Delta}) \cdot (\zeta, \bar{\zeta}) = (\Delta \cdot \zeta, \bar{\Delta} \cdot \bar{\zeta})$ is ring of M^2 ele-

ments isomorphic to the ring $Z_M[\gamma]$

Proof:

If $a + \gamma \cdot b \in Z_M[\gamma]$, then let Φ be the mapping:

$\Phi: (a + \gamma \cdot b) \rightarrow ((a + h \cdot b)_M, (a - (h \cdot b))_M) = (\Delta, \bar{\Delta})$, where $\Delta = (a + h \cdot b)_M$ and $\bar{\Delta} = (a - (h \cdot b))_M$.

To show that Φ is an isomorphism one must show that Φ is both one-to-one and onto and that Φ preserves addition and multiplication. To show that Φ is onto, one needs to demonstrate that given an arbitrary element $(\Delta, \bar{\Delta}) \in S_M$, there exists an element $a + \gamma \cdot b \in Z_M[\gamma]$ such that $\Phi(a + \gamma \cdot b) = (\Delta, \bar{\Delta})$ is an element of S_M . Following the definition of the mapping we have:

$$a + h \cdot b \equiv \Delta \pmod{M} \quad (13)$$

$$a - (h \cdot b) \equiv \bar{\Delta} \pmod{M} \quad (14)$$

Summing eqn. (13) and (14) yields $2a \equiv \Delta + \bar{\Delta} \pmod{M}$, and subtracting eqn. (13) and (14) yields $2hb \equiv \Delta - \bar{\Delta} \pmod{M}$. Because M is odd and $GCD(M, 2) = 1$, then $GCD(M, 2) = 1$ and $GCD(M, h) = 1$, and so there exists inverse elements 2^{-1} and h^{-1} (modulo M). We now solve eqn. (13) and (14) for a and b , namely:

$$a \equiv 2^{-1}(\Delta + \bar{\Delta}) \pmod{M} \quad (15)$$

$$b \equiv 2^{-1}h^{-1}(\Delta - \bar{\Delta}) \pmod{M} \quad (16)$$

From eqn. (15) and (16) we see that $(\Delta, \bar{\Delta}) \in S_M$ is an image of $a + \gamma \cdot b \in Z_M[\gamma]$ under the mapping Φ . This proves that Φ is an onto mapping.

To show that Φ is one-to-one, assume $\Phi(a + \gamma \cdot b) = \Phi(c + \gamma \cdot d)$. It follows that

$$((a + h \cdot b)_M, (a - (h \cdot b))_M) = ((c + h \cdot d)_M, (c - (h \cdot d))_M) \quad (17)$$

This implies

$$a + h \cdot b \equiv c + h \cdot d \pmod{M} \quad (18)$$

$$a - (h \cdot b) \equiv c - (h \cdot d) \pmod{M} \quad (19)$$

The congruencies of eqn. (18) and (19) are equivalent to $2a \equiv 2c \pmod{M}$, that is $a \equiv c \pmod{M}$ and $2 \cdot h \cdot b \equiv 2 \cdot h \cdot d \pmod{M}$, that is $b \equiv d \pmod{M}$.

Hence $a + \gamma \cdot b = c + \gamma \cdot d$ and Φ is a one-to-one mapping.

To show that Φ preserves addition and multiplication let $a + \gamma \cdot b$ and $c + \gamma \cdot d$ be arbitrary elements in $Z_M[\gamma]$. Then:

$$\begin{aligned} \Phi((a + \gamma \cdot b) + (c + \gamma \cdot d)) &= \Phi((a + c) + \gamma \cdot (b + d)) \\ &= ((a + c) + h \cdot (b + d), (a + c) - (h \cdot (b + d))) \\ &= (a + h \cdot b, a - (h \cdot b)) + (c + h \cdot d, c - (h \cdot d)) \\ &= \Phi(a + \gamma \cdot b) + \Phi(c + \gamma \cdot d) \end{aligned}$$

and

$$\begin{aligned} \Phi((a + \gamma \cdot b) \cdot (c + \gamma \cdot d)) &= \Phi((a \cdot c + s \cdot b \cdot d) + \gamma \cdot (a \cdot d + b \cdot c)) \\ &= (a \cdot c + s \cdot b \cdot d + h \cdot (a \cdot d + b \cdot c)), \\ &= (a \cdot c + s \cdot b \cdot d) - (h \cdot (a \cdot d + b \cdot c)) \\ &= ((a + h \cdot b) \cdot (c + h \cdot d), (a - (h \cdot b)) \cdot c - (h \cdot d)) \\ &= (a + h \cdot b, a - (h \cdot b)) \cdot (c + h \cdot d, c - (h \cdot d)) \\ &= \Phi(a + \gamma \cdot b) \cdot \Phi(c + \gamma \cdot d) \end{aligned}$$

Hence $Z_M[\gamma]$ is isomorphic to the ring S_M and the theorem is proved, where eqn. (15) and (16) define the inverse mapping.

□

3.1 An example

Let us consider the multiplication of the real numbers, $x = 0.7623$ and $y = 0.6592$, using this new technique.

The quadratic ring selected in this example is the ring $Z_M(\sqrt{2})$. For this ring an efficient choice for the modulus, M , is a Mersenne prime; that is, a prime number of the form

$2^p - 1$. The reasons why this choice is particularly efficient are: 1) the congruence

$x^2 \equiv 2 \pmod{(2^p - 1)}$ is solvable for all odd primes p and the solutions $\pm h$ are $\pm 2^{\frac{p+1}{2}}$,

which replaces multiplications by h with single shifts; 2) the inverse elements of $\pm h$ are

$\pm 2^{\frac{p-1}{2}}$, since the inverse element of 2 is 2^{p-1} this makes the inverse mapping also very

easy.

Let us chose the modulus of the ring $Z_M(\sqrt{2})$ to be $2^{19} - 1 = 524287$. The approximations in this example, using the algorithm introduced in the previous section, are:

$$0.7623 \approx 104 - 73\sqrt{2}$$

$$0.5692 \approx -56 + 40\sqrt{2}.$$

The solutions of the congruence $x^2 \equiv 2 \pmod{524287}$ are ± 1024 . Therefore, the forward map of x and y yields:

$$\begin{aligned} (104 - 73 \cdot 1024, 104 + 73 \cdot 1024)_M &= (-74648, 74856) \\ (-56 + 40 \cdot 1024, -56 - 40 \cdot 1024)_M &= (40904, -41016) \end{aligned}$$

The element-by-element multiplication of the conjugate pairs, modulo M , yields

$$(\Delta, \bar{\Delta}) = (45696, -69024).$$

The final step is the inverse map, based on equation eqn. (15) and (16):

$$\begin{aligned} z &= x \cdot y = \alpha + \beta\sqrt{2} = (2^{-1}(\Delta + \bar{\Delta}), 2^{-1}h^{-1}(\Delta - \bar{\Delta})) \\ &= (2^{18} \cdot (45696 + (-69024)), 2^{18} \cdot 2^9 \cdot (4596 - (-69024)))_M \\ &= (262144 \cdot (45696 + (-69024)), 262144 \cdot 512 \cdot (45696 - (-69024)))_M \\ &= (-11664, 8248) \end{aligned}$$

which is, as we expect, exactly equal to $(104 - 73\sqrt{2}, (-56 + 40\sqrt{2}))$.

Performing the final recovery we have $z = -11664 + 8248\sqrt{2}$ which is approximately equal to 0.4335.

Although the solution of eqn. (12) is precomputed, it would still be useful to have an efficient algorithm for solving this congruence, especially for large values of M .

The problem of finding square roots in finite fields has been extensively studied in computational number theory due to its importance in many cryptographic systems. There are no known polynomial-time algorithms for finding square roots in finite fields, but there exist sufficiently efficient randomized algorithms [18][19][20][21]. We have the following esti-

mation for the computational complexity of this problem: if M is an odd prime, then the solution of eqn. (12) needs $O(\log^2 M)$ steps. The solution can be found using Peralta's probabilistic algorithm [20]. The last estimation shows us that even for large M , the solution of eqn. (12) causes no problems.

3.2 Residue number system implementation

Residue number systems over Z_M are based on the Chinese Remainder Theorem (CRT). The theorem states that if $M = p_1 p_2 \dots p_L$ is the factorization of the positive integer M into relative prime factors, then: $Z_M \cong Z_{p_1} \otimes Z_{p_2} \otimes \dots \otimes Z_{p_L}$, that is, Z_M is isomorphic to the direct product of the smaller rings Z_{p_i} . A similar result holds for more general number rings, based on the general definition of a prime number. In $Z[\gamma]$, the primes are the elements that cannot be expressed as a product of two numbers neither of which is 1 or -1. With this definition in hand, the Chinese remainder theorem for $Z[\gamma]$ can be stated as:

$$Z_M[\gamma] = Z_{p_1}[\gamma] \otimes Z_{p_2}[\gamma] \otimes \dots \otimes Z_{p_L}[\gamma] \quad (20)$$

where $M \in Z[\gamma]$ has a relatively prime decomposition $M = p_1 p_2 \dots p_L$. We refer to this version of the residue number system as a *quadratic-like residue number system*. To ensure that the further decomposition, shown in the previous section, is possible, the numbers p_i should be selected such that they are primes in Z , but not in $Z[\gamma]$. This means, that the congruences:

$$x^2 \equiv \gamma \pmod{p_i} \quad (21)$$

must be solvable for all i , that is γ must be a quadratic residue modulo p_i for all i .

The Legendre symbol is useful here: if p is a prime number, then $\left(\frac{a}{p}\right)$ is defined as:

$\left(\frac{a}{p}\right) = 1$ if a is a quadratic residue modulo p , and $\left(\frac{a}{p}\right) = -1$ if a is a non-quadratic residue modulo p ; otherwise it is equal to zero.

We can use Legendre's symbol to identify the quadratic reciprocity law. If p and q are odd primes, then:

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} \quad (22)$$

If $p = 2$ and q is an odd prime, then:

$$\left(\frac{2}{q}\right) = (-1)^{\frac{q^2-1}{8}} \quad (23)$$

Now let us consider three cases for γ :

3.2.1 $\gamma = \sqrt{2}$

For the quadratic ring $Z_p[\gamma] = Z_p[\sqrt{2}]$ the quadratic reciprocity law implies that p must have the form $8k \pm 1$.

3.2.2 $\gamma = \sqrt{3}$

In this case the use of the quadratic reciprocity law yields:

$$\left(\frac{3}{p}\right)\left(\frac{p}{3}\right) = (-1)^{\frac{p-1}{2}} \quad (24)$$

Consider the two cases for p . First let p be a prime of the form $4k + 1$; here we have

$\left(\frac{3}{p}\right)\left(\frac{p}{3}\right) = 1$ and because we require $\left(\frac{3}{p}\right) = 1$, then p should be such that the congru-

ence $x^2 \equiv p \pmod{3}$ is solvable. This restricts p to be of the form $3k + 1$; so, in this case, p should be of the form $12k + 1$. A similar proof for the case $p = 4k - 1$ shows that the other possibility is for p is to be of the form $12k - 1$.

3.2.3 $\gamma = \frac{1 + \sqrt{5}}{2}$

In this case the condition for p is $5k \pm 1$.

4.0 Some asymptotic estimates

In order to compare the applicability of different quadratic number rings, we use the following theorem:

Theorem 3: (Chebotarev Density Theorem):

Let D be the degree of the splitting ring of the polynomial $q(x)$. Define $\pi(x, q)$ to be

$\#\{p \leq x, p - \text{prime} \mid q(x) \text{ splits completely mod } p\}$. Then

$$\pi(x, q) = O\left(\frac{x}{D \log x}\right) \quad (25)$$

Proof: See [8].

□

A very important corollary is the prime number theorem which states that if $\pi(x) =$

$$\#\{primes \leq x\}, \text{ then } \pi(x) \approx \frac{x}{\log x}.$$

For quadratic number rings the Chebotarev density theorem implies that 1/2 of the primes cause $x^2 - \gamma$ to split completely. Therefore, from an asymptotical point of view, the Chebotarev density theorem ‘equalizes’ the applicability of the different quadratic number rings.

For small rings, however, we can observe considerable differences. The applicable primes that can be used for three different number rings are provided below:

$$Z_p(\sqrt{2}) \Rightarrow \{7, 17, 23, 31, 47, 71, 73, 79, 89, 97, 103, 113, 127\}$$

$$Z_p(\sqrt{3}) \Rightarrow \{11, 13, 23, 37, 47, 59, 61, 71, 73, 83, 97, 109\}$$

$$Z_p(\sqrt{5}) \Rightarrow \{11, 19, 29, 31, 41, 59, 61, 71, 79, 89, 101, 109\}$$

Let us consider the cases for 5, 6 and 7-bit moduli. We have the following estimations for the dynamic range of the computations performed over those rings (see Table 2, 3 and 4):

Table 2. The applicable 5-bit primes and dynamic range for the three rings

	<i>applicable primes, p</i>	<i>product of the primes</i>	<i>dynamic range</i>
$Z_p(\sqrt{2})$	7,17,23,31	84847	16.37 bits
$Z_p(\sqrt{3})$	11,13,23	3289	11.68 bits
$Z_p(\sqrt{5})$	11,19,29,31	187891	17.52 bits

Table 3. The applicable 6-bit primes and dynamic range for the three rings

	<i>applicable primes, p</i>	<i>product of the primes</i>	<i>dynamic range</i>
$Z_p(\sqrt{2})$	7,17,23,31,47	3987809	21.93 bits
$Z_p(\sqrt{3})$	11,13,23,37,47,59,61	20584736029	34.26 bits
$Z_p(\sqrt{5})$	11,19,29,31,41,59,61	27725009069	34.69 bits

Table 4. The applicable 7-bit primes and dynamic range for the three rings

	<i>applicable primes, p</i>	<i>product of the primes</i>	<i>dynamic range</i>
$Z_p(\sqrt{2})$	7,17,23,31,47,71,73, 79,89,97,103,113,127	20836452453876790368337	74.14 bits
$Z_p(\sqrt{3})$	11,13,23,37,47,59,61, 71, 73, 83, 97,109	93627372451137852613	74.14 bits
$Z_p(\sqrt{5})$	11,19,29,31,41,59,61, 71, 79, 89,101,109	152368432449359034821	67.04 bits

As we observe, different sizes of the primes used lead to different optimal choices of the ring over which to perform the residue computations.

5.0 Implementing real discrete-valued transforms

Let us write our orthogonal transform in the following form:

$$y(k) = \sum_{n=0}^{N-1} x(n)h(k, n), \quad k = 0, \dots, N-1 \quad (26)$$

where $h(k,n)$ is usually called a kernel of the transform and $x(n)$ is a real input data sequence of length N . As examples, for the *DCT* [11] we have $h(k,n) = 2 \cos \frac{\pi(2n+1)k}{2N}$, and for the discrete Hartley transform, we have $h(k,n) = \cos \frac{2\pi nk}{N} + \sin \frac{2\pi nk}{N}$, where, in both cases, the computation of $F(k)$ consists of N inner product computations. The elements of the kernel are real numbers and they can be precomputed and approximated as elements of a ring $Z[\gamma]$.

5.1 The real-valued algorithm

Now we propose the following algorithm for computing eqn. (26).

Step 1: Transform the kernel of the transform into conjugate pairs in the quadratic-like residue number system representation. Because the kernel of the transform consists of constant numbers, we can precompute this step in advance.

Step 2: Perform the transformation of the real input sequence, $x(n)$, into the quadratic integer form: $x(n) \approx x_1(n) + x_2(n)\gamma$; this step can be accomplished using the approximation algorithm proposed in Section 2.1.

Step 3: Transform the pairs $(x_1(n), x_2(n))$ into conjugate pairs in the quadratic-like residue number system representation:

$$(\Delta_1(n), \bar{\Delta}_1(n))_{p_1}, (\Delta_2(n), \bar{\Delta}_2(n))_{p_2}, \dots, (\Delta_k(n), \bar{\Delta}_k(n))_{p_k}$$

using, as moduli, prime numbers p_i ($i=1,2,\dots,k$) such that γ is a quadratic residue modulo p_i for

each i . This step can be accomplished exploiting the technique proposed in Section 3.0.

Step 4: Perform the transform using one of the large variety of fast algorithms. Perform inversion of the conjugate pairs obtained using equations eqn. (15) and (16). The resulting pairs are:

$$(y_{11}(n), y_{21}(n))_{p_1}, (y_{12}(n), y_{22}(n))_{p_2}, \dots, (y_{1k}(n), y_{2k}(n))_{p_k}$$

Step 5: Reconstruct the final result using the Chinese remainder theorem to solve the system of congruences:

$$y_1(n) \equiv y_{1i}(n) \pmod{p_i} \quad \text{for } i = 1, 2, \dots, k \quad (27)$$

$$y_2(n) \equiv y_{2i}(n) \pmod{p_i} \quad \text{for } i = 1, 2, \dots, k \quad (28)$$

$$\text{Step 6: } y(n) = y_1(n) + y_2(n)\gamma$$

5.2 A complexity analysis for the real-valued algorithm

Let us analyze the complexity of this algorithm:

Step 1 requires $O(\log n)$ operations for each input sample, that is, its complexity is $O(n \cdot \log n)$.

Step 2 requires $O(n)$ operations.

Step 3 requires $O(k)$ operations

Step 4 depends to a great extent on the structure of the transform used; however, with one exception (the Haar transform, for which there exists a linear-time algorithm), the complexity of fast algorithms for real-valued transforms is $O(n \cdot \log n)$.

Step 5 depends on the number of primes used in the quadratic-like number system. The complexity of solving the systems of congruencies of eqn. (27) and eqn. (28) is

$(O(\log k)^2)$, where k is the number of prime moduli. This estimation is due to Collins [13]. Therefore, the overall complexity of this step is $O(n(\log k)^2)$.

Step 6 requires $O(n)$ operations.

In summary, the total complexity of the proposed algorithm for computing a real-discrete valued transform is $O(n \cdot \log n + n(\log k)^2)$. Usually the number of prime moduli is much smaller than the number of the input samples, or it is fixed based on practical considerations; in this case, then, the complexity is $O(n \cdot \log n)$.

6.0 Conclusions

In this paper a new technique for implementing real discrete-valued transforms is proposed. The approach presented is based on an approximation of the real input samples as elements of quadratic number rings prior to residue number system processing. The proper choice of the number ring leads to large dynamic range of the computations. The weaker restrictions on the prime moduli used in the residue computations allow us to break the

main bottleneck of the DFT implementation using RNS in a ring of algebraic integers.

After the first approximation step, all of the subsequent steps are error-free.

7.0 Appendix

In this appendix we shall present a proof of theorem 1 and some final comments.

Proof of theorem 1:

The procedure for obtaining the digits of an arbitrary real number (Algorithm 1) can be explained using Markov chains. Consider a system with two states, 0 and 1 (the digits) and

the transition matrix $T_{0,1} = \begin{bmatrix} t_{00} & t_{01} \\ t_{10} & t_{11} \end{bmatrix}$ describing the probability of transfer from state ‘ i ’

to state ‘ j ’ ($i,j=0,1$). The main feature of the Bergman arithmetic (where $\gamma = \frac{1 + \sqrt{5}}{2}$) is

the nonexistence of consecutive ones in the number representation. It follows from a simple observation that if in Step 2 of Algorithm 1 the result of the multiplication is greater

than 1 (which corresponds to state 1), then, after subtracting 1, the result will be smaller

than $\gamma - 1$ and the next digit definitely will be 0. Clearly, for γ -based *GINS*, where γ is

smaller than $\frac{1 + \sqrt{5}}{2}$, the same property, that is, the nonexistence of consecutive ones, is

valid, whereas for greater values of γ it is not. For this reason we separate our analysis in two parts.

First case: $1 < \gamma \leq \frac{1 + \sqrt{5}}{2}$

Let us denote as $p_0^{(n)}$ (resp. $p_1^{(n)}$) the probability that the n -th digit in the γ -based *GINS* of a randomly chosen real number, between 0 and 1, is 0 (resp. 1). For $p_0^{(n)}$ and $p_1^{(n)}$ we have the following recursive relationships:

$$p_0^{(n)} = t_{00}p_0^{(n-1)} + t_{10}p_1^{(n-1)} \quad (29)$$

$$p_1^{(n)} = t_{01}p_0^{(n-1)} + t_{11}p_1^{(n-1)} \quad (30)$$

In this case, the elements of the matrix $T_{0,1}$ are: $t_{00} = \frac{1}{\gamma}$, $t_{10} = 1$, $t_{01} = 1 - \frac{1}{\gamma}$ and $t_{11} = 0$. The solution of eqn. (29) and (30) for $p_1^{(n)}$ is:

$$p_1^{(n)} = \Theta \left[(-1)^n \Theta^n + \frac{1 - \Theta^n}{1 - \Theta} \right] \quad (31)$$

where $\Theta = 1 - \frac{1}{\gamma}$. Asymptotically:

$$p_1^{(\infty)} = \lim_{n \rightarrow \infty} p_1^{(n)} = \frac{\Theta}{1 - \Theta} = \frac{\gamma - 1}{2\gamma - 1} \quad (32)$$

Second case: $\frac{1 + \sqrt{5}}{2} < \gamma < 2$

Now the elements of the transition matrix are: $t_{00} = \frac{1}{\gamma}$, $t_{10} = \frac{1}{\gamma(\gamma - 1)}$, $t_{01} = 1 - \frac{1}{\gamma}$ and

$t_{11} = 1 - \frac{1}{\gamma(\gamma - 1)}$. Therefore:

$$p_1^{(n)} = \frac{\gamma-1}{\gamma} \left(\Theta^n + \frac{1-\Theta^n}{1-\Theta} \right) \quad (33)$$

where $\Theta = \frac{\gamma-1}{\gamma(\gamma-1)}$. Hence,

$$p_1^{(\infty)} = \lim_{n \rightarrow \infty} p_1^{(n)} = \frac{(\gamma-1)^2}{1+(\gamma-1)^2} \quad (34)$$

As was proved by Borel [17], almost all positive real numbers have an equal number of zeros and ones in their binary representation. More generally, we have that if g is an integer greater than one, then

$$z = \frac{w_1(z)}{g} + \frac{w_2(z)}{g^2} + \dots, \quad 0 \leq z \leq 1 \quad (35)$$

where every digit $w_i(z)$ is in $\{0, 1, \dots, g-1\}$. Borel's theorem states: For almost all z

($0 \leq z \leq 1$):

$$\lim_{n \rightarrow \infty} \frac{H_n^{(k)}(z)}{n} = \frac{1}{g} \quad (36)$$

where $H_n^{(k)}$ denotes the number of those w from the first n , which are equal to k ,

$0 \leq k \leq g-1$.

Let us consider a real number z , $0 \leq z \leq 1$, presented in binary number system with precision 2^{-n} . Borel's theorem merely confirms our intuitive expectations that the average number of zeros and ones is equal to $\frac{n}{2}$.

However, for γ -based *GINS* ($1 < \gamma < 2$), we have a greater proportion of zeros. The function, estimating the average number of ones in representing a real number z with precision 2^{-n} in γ -based *GINS* is:

$$F(\gamma, n) = p_1^{(n)} \log_\gamma 2 \quad (37)$$

Letting n to go to infinity, we have:

$$F(\gamma) = \begin{cases} \frac{\gamma - 1}{(2\gamma - 1) \cdot \log_2 \gamma} & \text{if } \left(\gamma \leq \frac{1 + \sqrt{5}}{2} \right) \\ \frac{(\gamma - 1)^2}{(1 + (\gamma - 1)^2) \cdot \log_2 \gamma} & \text{if } \left(\gamma > \frac{1 + \sqrt{5}}{2} \right) \end{cases} \quad (38)$$

That completes the proof of the theorem.

□

Final comments

Simple calculations show that the function $F(\gamma)$ decreases in the interval $\left(1, \frac{1 + \sqrt{5}}{2}\right)$ and increases in the interval $\left(\frac{1 + \sqrt{5}}{2}, 2\right)$. Therefore, the average number of ones is minimal

when $\gamma = \frac{1 + \sqrt{5}}{2}$, that is, in Bergman arithmetic. In this case $F(\gamma) = 0.398122\dots$,

hence we have on average more than 20% fewer ones compared to the binary number system. When γ tends to 2, $F(\gamma)$ tends to 0.5, as expected. When γ tends to 1, $F(\gamma)$ tends to $\ln 2 = 0.693\dots$.

It is interesting to note, that if γ is a number greater than the larger root ψ of the equation

$x = 4^{\frac{x-1}{2x-1}}$, which is equal to 1.2890888..., the proportion of ones is smaller than that in the binary number system, whereas if $1 < \gamma < \psi$ it is greater than 0.5

8.0 References

- [1] N.Ahmed and K.R.Rao, Orthogonal transforms in digital signal processing, New-York: Springer-Verlag, 1975.
- [2] M.Soderstrand, W.K.Jenkins, G.A.Jullien and F.Taylor, Residue number systems: Modern applications in digital signal processing, IEEE Press, 1986.
- [3] J.H.Cozzens and L.A.Finkelstein, Computing the discrete Fourier transform using residue number systems in a ring of algebraic integers, IEEE Trans. on Information Theory, vol. 31, 1985, pp. 580-588.
- [4] A.L. Bequillard and S.D.O'Neil, Systolic RNS computation of the two-dimensional DCT in a ring of algebraic integers, Proceedings of the 20th Annual Conference on Information Sciences and Systems, Princeton University, Princeton, NJ, March 1986, pp. 783-789.

- [5] G.A.Ray, Residue arithmetics in rings of algebraic integers, IEEE Int. Conference ASSP, 1990, pp. 1527-1530.
- [6] G.Bergman, A number system with an irrational base, Mathematical Magazine, vol. 31, pp. 98-119, 1957.
- [7] V.S.Dimitrov, T.V.Cooklev and B.D.Donevsky, Number theoretic transforms over the golden section quadratic field, IEEE Trans. on Signal Processing, vol. 43, Aug. 1995, pp. 1790-1797
- [8] Computational methods in number theory, ed. H.W.Lenstra,Jr. and R.Tijdeman, third edition, Mathematical Centre Tracts, vol. 155, 1987.
- [9] T.K.Truong,J.J.Chang,I.S.Hsu,D.Y.Pei and I.S.Reed,Techniques for computing the DFT using the Quadratic Fermat Residue Number System, IEEE Trans. on Computers, vol.C-35, no.11,Nov.1986,pp.1008-1012
- [10] S.Vajda, Fibonacci and Lucas numbers and the golden section: theory and applications, New York, Wiley, 1989
- [11] N.Ahmed,T.Natarajan and K.R.Rao,Discrete Cosine Transform, IEEE Trans. on Computers, vol.C-23,Jan.1974,pp.90-93
- [12] V.S.Dimitrov and T.V.Cooklev,Two algorithms for modular exponentiation using nonstandard arithmetics, IEICE Trans.on Fundamentals, vol.E78-A,1995,pp.82-87
- [13] G.E.Collins, Computing time analysis for some arithmetic and algebraic algorithms, Proc. 1968 Summer Institute on Symbolic and Mathematical Computations, IBM Corporation, Cambridge, Mass.,1969, pp.197-231

- [14] P.Erdos, M.Horvath and I.Joo, On the uniqueness of the expansion $1 = \sum q^{-n_i}$,
Acta Mathematica Hungarica, vol.58, 1991, pp.333-342
- [15] R.A.Games,D.Moulin,S.D.O'Neil and J.Rushanan,Algebraic-integer quantization
and residue number system processing, IEEE Int.Conf. ASSP, vol.D3.15,
1989,pp.948-951
- [16] I.S.Reed and T.K.Truong,Convolutions over residue classes of quadratic integers,
IEEE Trans. on Information Theory, vol.IT-21,1976,pp.468-475
- [17] E.Borel, Les probabilités denumbrables et leurs applications arithmetiques, Rend.
Circ. Math, Palermo, vol.27, 1909, pp.247-271
- [18] J.C.Lagarias, Worst-case complexity bounds for algorithms in the theory of integral
quadratic forms, Journal of Algorithms, vol.1, 1980, pp.142-186
- [19] M.O.Rabin, Probabilistic algorithms in finite fields, SIAM J. on Computing, vol.9,
1980, pp.273-280
- [20] R.Peralta, A simple and fast probabilistic algorithm for computing square roots mod-
ulo a prime number, IEEE Trans. on Information Theory, vol.32, 1986, pp.846-847
- [21] E.Bach, Realistic analysis of some randomized algorithms, Journal of Computer and
System Science, 1991, vol. 18, pp.110-121