INT. J. MATH. EDUC. SCI. TECHNOL., 2003
VOL. 34, NO. 4, 561–574

Division in a binary representation for complex numbers

DAVID C. BLEST*
School of Mathematics & Physics, University of Tasmania, Private Bag 1-360,
Launceston, Tasmania 7250, Australia
e-mail: D.Blest@utas.edu.au

and TARIQ JAMIL
Department of Electrical and Electronics Engineering, Sultan Qaboos University,
Al-Khod 123, Muscat, Sultanate of Oman
e-mail: tjamil@squ.edu.om

(*Received 6 June 2002*)

Computer operations involving complex numbers, essential in such applications as Fourier transforms or image processing, are normally performed in a 'divide-and-conquer' approach dealing separately with real and imaginary parts. A number of proposals have treated complex numbers as a single unit but all have foundered on the problem of the division process without which it is impossible to carry out all but the most basic arithmetic. This paper resurrects an early proposal to express complex numbers in a single 'binary' representation, reviews basic complex arithmetic and is able to provide a fail-safe procedure for obtaining the quotient of two complex numbers expressed in the representation. Thus, while an outstanding problem is solved, recourse is made only to readily accessible methods. A variety of extensions to the work requiring similar basic techniques are also identified. An interesting side-line is the occurrence of fractal structures, and the power of the 'binary' representation in analysing the structure is briefly discussed.

1. Introduction

Complex numbers have a crucial role in modern science and engineering. Their application in most digital signal processing algorithms such as fast Fourier transforms and geometric analysis in graphics or image processing demands an efficient approach to handling complex arithmetic [1]. Such are the demands for frequent and speedy arithmetic of complex numbers that resort is being made to dedicated hardware [2]. However, even with recent vast advances in VLSI technology and the availability of over 100 million transistors on a single IC chip [3], virtually all complex arithmetic is merely add-on software to the basic operations of real arithmetic. Computer operations involving complex numbers are most commonly performed by dealing with the real and imaginary parts separately and then accumulating the individual results to get the final result of the complex operation. This 'divide-and-conquer' technique thus forsakes the advantages of using complex numbers in computer arithmetic. The possibility of representing

*Author to whom correspondence should be addressed.

International Journal of Mathematical Education in Science and Technology
ISSN 0020–739X print/ISSN 1464–5211 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0020739031000108592

complex numbers as single entities immediately reduces the number of arithmetic operations required when compared with current techniques, for example: addition of two complex numbers requires two separate additions, one for real and one for imaginary parts while efficient multiplication of two complex numbers can be reduced to three individual multiplications, three additions and two subtractions as indicated in the scheme $(a+ib)(c+id) = [(c-d)b + (a-b)c] + i[(c-d)b + (a+b)d]$. Representation as a single entity reduces this to one addition and one multiplication respectively. Moreover, as will be seen, there is no need to worry about arithmetic signs. However, as also will be seen, a single entity representation for a complex number requires twice the number of bits as for a real number. If increased speed and efficiency can be demonstrated for such a complex arithmetic then there is significant potential for its implementation in large-scale problems in signal processing and graphics manipulation.

It should be noted also that computations with complex entities as pairs of ordinary IEEE doubles have never been entirely satisfactory—for example, it is not reasonable to write machine code that reliably computes the sum, difference, product and quotient of complex numbers guaranteed accurate to the last bit, as it is with real data.

Various authors, for examples [4–6] have proposed methods and algorithms to make the standard calculations with complex numbers as efficient as possible. However, the possibility of dealing with a complex number manipulated as a single entity would appear to offer an opportunity to increase efficiency dramatically. Knuth [7] offers a recent survey of a number of proposals for single entity treatments of complex numbers and, as early as 1960, Knuth himself [8] proposed a 'quater-imaginary' number system in which every complex number can be expressed as linear combinations of the powers of the imaginary number $2i$ as base or radix and where the coefficients are drawn from the set $\{0, 1, 2, 3\}$. In the same note Knuth reported but did not elaborate the more general proposition of using powers of $\sqrt{-n}$ where n is a positive integer and the coefficients may be drawn from the set $\{0, 1, 2, \ldots, n-1\}$. A special case of obvious interest in computer implementation is what Knuth termed the bi-imaginary system with $n = 2$ using only the digits $0, 1$. It will be noted that in the above proposals even powers of the radix give real numbers and odd powers imaginary ones so that it remains possible to work separately with real and imaginary parts. Benedek and Panzone [9] have described the full range of options for choice of base when employing only the binary digits and these offer the possibility for further investigation of the division problem along the lines of the solution in this work.

Knuth was not able to find an effective algorithm for division, citing the dependence of division on the order properties of the real numbers. Without a secure division algorithm, calculation of functions such as logarithms, exponentials, roots, etc., of course is not available. In addition, in the bi-imaginary case, i is represented by a non-terminating "*bi*-mal" [compare *deci*-mal] immediately introducing the possible problems of truncation and rounding errors. A second proposal by Penney [10] offers the basic detail [that is the detail of representation, addition and multiplication of complex integers] for the radix $-1 + i = \sqrt{-2i}$ and uses only the binary digits — an alternative base is $-1 - i$ but, interestingly, replacing -1 by $+1$ in either case does not allow representation of all complex integers. Penney does not refer to the division process which has the same problems identified by Knuth. The purpose of this paper is to develop the full detail of representation and arithmetic using the radix $-1+i$ and to indicate how the division process can be given a computer implementation.

Without appearing to have knowledge of Knuth's work, Stepanenko [1] addresses the same general proposal, concentrating on the radix $\sqrt{-2} = \pm i\sqrt{2}$ and using the binary digits. Stepanenko proposes to solve the problem of the quotient of two complex numbers by multiplication after first finding the reciprocal of the denominator using a well-known algorithm derived from the Newton iterative method for finding roots of nonlinear equations. However Stepanenko admits that the process '*...reduces to a good choice of the initial approximation...*' and does not indicate how this, in fact, very limited range of possibilities may be identified. Indeed Stepanenko moves on to propose '*implementation of the method of succesive subtraction of the divisor*'. For efficient calculation, definiteness of process is a *sine qua non*. The current paper adopts the Newton approach to division but is able to secure convergence of the iterative process by direct calculation of an appropriate first approximation.

2. The base $-1+i$

In the following it is shown that any real or complex number z may be represented in terms of sums of powers of the radix as

$$z_{|(-1+i)} = \sum_{j=-\infty}^{j=K} a_j(-1+i)^j \tag{1}$$

where $a_j \in \{0,1\}$, $a_K = 1$ and $a_i = 0\ \forall i > K$. In parallel with the normal decimal representation this will be written $1a_{K-1}a_{K-2}\ldots a_2a_1a_0.a_{-1}a_{-2}\ldots$ where the 'point' separates, on its right, the coefficients of negative powers. It will be noted that plus/minus signs are not required in this representation. It should also be noted that an infinite number of fractional digits is not required for completeness in this number system [11] supporting the practical requirement for finite representations in any computer implementation.

Table 1 shows the zero power and first four positive and negative powers of $-1+i$ and useful results for the expression of powers of reals.

Thus $1_{|(-1+i)} = 1.0\ldots.$ Similarly $i \rightarrow 11.0$, and it is a matter of simple arithmetic to check the results shown in table 2.

It will be noted that at least some decimal fractions have representations which include positive powers of the radix. In addition, using these representations, we have the trivial but significant results that $1.0 + 1.0 = 1100.0$ and $11.0 + 111.0 = 0.0$, *the zero rule*, and it will be noted that such results are independent of the actual bit, that is if two numbers with a 1 in position m are added then this results, in the sum, with 0s in positions m and $m+1$ and 1s added in positions $m+2$ and $m+3$. These two rules plus the obvious cases where there are $0+0=0$, $1+0=0+1=1$ are sufficient to determine addition of numbers with the given base. However, for

$n =$	$-4s$	-4	-3	-2	-1	0	1	2	3	4	$4s$
$(-1+i)^n =$	$(-1/4)^s$	$(-1/4)$	$(1-i)/4$	$(i/2)$	$(-1-i)/2$	1	$-1+i$	$-2i$	$2+2i$	-4	$(-4)^s$

Table 1. Powers of $(-1+i)$ for small n, s any positive integer.

Decimal	Base $(-1+i)$	Decimal	Base $(-1+i)$
$-1 \rightarrow$	11101.0	$4 \rightarrow$	111010000.
$-i \rightarrow$	111.0	$1/2 \rightarrow$	1.11
$2 \rightarrow$	1100.0	$-1/4 \rightarrow$	0.0001
$3 \rightarrow$	1101.0	$1/16 \rightarrow$	0.00000001
$16 \rightarrow$	100000000.0	$1/4 \rightarrow$	1.1101
$-4 \rightarrow$	10000.0	$1/8 \rightarrow$	0.000011
$(-4)^s \rightarrow$	$100.0 \cdots 4s$ zeros$\cdots 00.0$	$(-1/4)^s \rightarrow$	$0.\cdots 4s\ -1$ zeros$\cdots 1$

Table 2. Representations of useful real numbers in base $(-1+i)$.

computer implementation, addition will require a satisfactory implementation of the zero rule in order to ensure that calculations are not swamped by carry propagation in the case of the first rule, especially when these are part of a multiplication. One of us [TJ] continues to address this problem.

It will be noted that the result $1.0 + 11101.0 = 0.0$, i.e $1 + (-1) = 0$ in base 10, enables subtraction of two numbers in the representation. Thus if the minuend is

$$a_n a_{n-1} a_{n-2} \ldots a_{j+4} a_{j+3} a_{j+2} a_{j+1} \mathbf{0}\, a_{j-1} \ldots$$

and subtrahend is

$$b_n b_{n-1} b_{n-2} \ldots b_{j+4} b_{j+3} b_{j+2} b_{j+1} \mathbf{1}\, b_{j-1} \ldots$$

then the result of the subtraction $a_j - b_j$ is that $a_j = 0 \rightarrow 1$, $a_{j+1} \rightarrow a_{j+1}$, $a_{j+2} \rightarrow a_{j+2} + 1$, $a_{j+3} \rightarrow a_{j+3} + 1$, $a_{j+4} \rightarrow a_{j+4} + 1$, $b_j = 1 \rightarrow 0$ where the usual rules for addition apply to the required sums of the as but otherwise no other coefficient is changed. The other three possible subtractions of coefficients are immediate.

The multiplication process for numbers in the base $-1 + i$ representation is the same as for ordinary multiplication except that the special rules (above) for addition must be employed. The zero rule plays an important role in reducing the summands resulting from intermediate additions. It will be noted that for the product of complex numbers only one such multiplication is required compared with the three multiplications if real and imaginary parts are treated separately.

The ability to express $0, 1, 2, 3$ and powers of -4 as simple 'binary' expressions suggests a process to express any positive real f in the required representation. This may be divided into three steps: use of the division algorithm to convert ordinary decimal expressions to base 4; a systematic process converting from base 4 to base (-4); use of tables 1 and 2 to convert from base (-4) to base $(-1+i)$. First, to express f in base 4:

$$f_{|4} = \sum_{t=-M}^{T} q_t 4^t \text{where } T \text{ is a positive integer such that } 4^{T+1} > N \geq 4^T$$

and M is fixed by machine limitations in any implementation. This 'normalized' representation is unique when $0 \leq q_t < 4$. If the constraint on q_t is removed the expression is not unique and we say that we have an unnormalized base 4 representation. In the usual notation we write the base 4 representation as

$$(1 q_{T-1} \ldots q_5 q_4 q_3 q_2 q_1 q_0 . q_{-1} q_{-2} \ldots q_{-M})_{|4}$$

and this may be converted to a base -4 representation by replacing each digit in odd location $\dots q_{-3} q_{-1}, q_1, q_3, q_5, \dots$ with its negative to obtain

$$(\dots(-q_5)q_4(-q_3)q_2(-q_1)q_0.(-q_{-1})q_{-2}(-q_{-3})\dots)_{|-4}$$

which is unnormalized. Normalization follows from the identity

$$-\alpha \times [-4]^s = (-\alpha + 4) \times [-4]^s + 1 \times [-4]^{s+1}$$

for integer α and s. Thus, any negative q_t is replaced by an integer in the set $\{1, 2, 3\}$ by adding 4 to the digit and adding 1 to the digit q_{t+1}. This operation removes all negative numbers but can create digits with a value of 4 after the addition of a 1, say at q_τ. Normalization is completed on replacement of the 4 by 0 and subtraction of 1 from $q_{\tau+1}$. It may be noted that with all negative bases, all integers, whether positive or negative, have a unique positive representation.

Having obtained the representation in base -4 each of the q_t can be substituted in favour of the binary representations of tables 1 and 2.

As an example of the process we find the base $(-1+i)$ representation of 60.351:

$$\begin{aligned}
60.351_{|10} &= (3, 3, 0.1, 1, 2, 1, 3, 1, 2, 3, \dots)_{|4} \\
&= (3, -3, 0.-1, 1, -2, 1, -3, 1, -2, 3, \dots)_{|-4} \\
&= (4, 1, 1.3, 2, 2, 2, 1, 2, 2, 3, \dots)_{|-4} \\
&= (-1, 0, 1, 1.3, 2, 2, 2, 1, 2, 2, 3, \dots)_{|-4} \\
&= (1, 3, 0, 1, 1.3, 2, 2, 2, 1, 2, 2, 3, \dots)_{|-4}\text{(normalized)} \\
&= 1, 1101, 0000, 0001, 0001.1101, 1100, 1100, \\
&\quad 1100, 0001, 1100, 1100, 1101, \dots_{|(-1+i)}
\end{aligned}$$

where the commas are introduced only to identify the grouping of 0, 1, 2 or 3 multiples of the appropriate power of -4.

Representations of negative reals are obtained by multiplying the representation of the positive real by $-1 \equiv 11101.0$; positive or negative imaginary numbers are obtained by multiplication of the corresponding real representation by $i \equiv 11.0$ (or alternatively, for negative imaginaries, the positive real by $-i \equiv 111.0$); complex numbers by adding the representations of real and imaginary parts. All representations are then in the form of equation 1.

For example:

$$\begin{aligned}
&5.5_{10} \to 111011100.11_{|(-1+i)} \\
&-5.5_{10} \to 111011100.11 \times 11101.0 = 11001101.11_{|(-1+i)} \\
&12.25_{|10} \to 100010001.1101_{|(-1+i)} \\
&-12.25i_{|10} \to 100010001.1101 \times 111.0 = 11101110000.0011_{|(-1+i)} \\
&(-5.5 - 12.25i)_{|10} \to 11001101.11 + 11101110000.0011 = 10111101.1111_{|(-1+i)}
\end{aligned}$$

Conversion from binary to decimal representation is straightforward, using table 2. Given a 'binary' representation for a complex number $z = x + iy$ the usual requirements of finding the complex conjugate $\bar{z}$, real part $x = (z + \bar{z})/2$, imaginary part $y = i(x - z)$, magnitude squared $|z|^2 = z\bar{z}$ are easily accomplished within the representation. In particular, it is possible to develop a systematic—and hence computer feasible—digit by digit replacement rule for direct calculation of the complex conjugate. Such possible projects are left to the interested reader.

3. Evaluation of quotients in the base (−1+*i*)

An attempt to evaluate, by long division and showing remainders, the simple quotient 8/2 in the base $(-1+i)$ gives an immediate indication of the problems that arise even with real numbers:

$$\frac{111000000}{1100} = 100000 + \frac{1000000}{1100}$$

but

$$\frac{1000000}{1100} = 1000 + \frac{1110100000}{1100}$$

and, once more

$$\frac{1110100000}{1100} = 1010000 + \frac{11101100000}{1100}$$

It is seen that the numerators grow and are propagated to the left with no indication that the process can terminate. It is thus necessary to identify an alternative approach to the division process. One possibility seeks an independent evaluation of the reciprocal of the denominator in a quotient

$$\frac{s}{z} = \frac{p+iq}{x+iy}$$

so that if, say, $z^{-1} = w = u + iv$ then $s/z = sz^{-1} = sw = (p+iq)(u+iv)$ which may be evaluated using the rules for multiplication derived in the previous section.

A standard algorithm for producing reciprocals, z^{-1}, without division, obtained using the Newton–Raphson method and which is applicable to both real and complex functions, is

$$w_{s+1} = w_s(2 - zw_s) \text{ for } s = 0, 1, 2, \ldots \text{ and } w = \lim_{s\to\infty} w_s \qquad (2)$$

If such a process can be implemented then there is an immediate reduction in effort compared with the more complicated Newton approach to simultaneous roots of two functions of two variables. As is pointed out by Stepanenko [1], the iteration process '*reduces to a good choice of the initial approximation*', w_0. For the base $(-1+\mathrm{i})$ there is a simple choice of such an approximation which guarantees convergence as is indicated in the following. Moreover, once such an initial approximation is found, convergence of the Newton method is rapid. Thus, if the error in the $(s+1)$st approximation w_s for $s = 0, 1, 2, \ldots$ is $\epsilon_{s+1} = w_{s+1} - 1/z$ where $w_{s+1} = w_s(2 - zw_s)$ then

$$\begin{aligned}\epsilon_{s+1} &= 2w_s - zw_s^2 - 1/z \\ &= -z(w_s - 1/z)^2 = -z\epsilon_s^2\end{aligned}$$

indicating the quadratic convergence offered by Newton's method. Moreover

$$\left|\frac{\epsilon_{s+1}}{\epsilon_s}\right| = |z\epsilon_s| < 1$$

if

$$\left|z(w_s - 1/z)\right| = |zw_s - 1| < 1$$

again, for all $s = 0, 1, 2, \ldots$.

Thus an initial approximation, w_0, must also satisfy

$$|zw_0 - 1| < 1 \tag{3}$$

Given that z has the expression given in equation (1)

$$(-1+i)^{-K}z = 1 + a_{K-1}(-1+i)^{-1} + a_{K-2}(-1+i)^{-2} + \ldots$$

so that, in comparison with equation (3)

$$\left|(-1+i)^{-K}z - 1\right| = \left|\sum_{m=-\infty}^{-1} a_{K+m}(-1+i)^m\right|$$

This last expression is simply a sum of the negative powers of the radix. It is shown in the next section that, except in a very small proportion of cases, the general expression

$$\alpha + i\beta = \sum_{j=-\infty}^{-1} a_j(-1+i)^j \tag{4}$$

has magnitude less than unity. Thus in all such cases the initial approximation

$$w_0 = (-1+i)^{-K} \equiv 0.0\ldots(K-1) \text{ zeros followed by } 1 \tag{5}$$

guarantees that equation (3) is satisfied and thus the iteration process of equations (2) converges. Computer assisted enumeration of the possibilities indicates that less than 1% of such initial approximations fail to ensure convergence. Again in the next section it is shown that in these cases the alternative

$$w_0' = iw_0 = i \times (-1+i)^{-K} \equiv (0.0\ldots(K-2) \text{ zeros followed by } 11) \tag{6}$$

that is to multiply the basic initial approximation by $i \equiv 11.0$, is then sufficient.

4. Sums of negative powers of the base (−1+*i*)

In order to establish convergence of the iterative process it is necessary to be able to decide when terms of the form of equation (4) satisfy $\left|\alpha + i\beta\right| = \sqrt{\alpha^2 + \beta^2} < 1$. It is possible to limit the possibilities by first establishing bounds for the real and imaginary parts α, β. For example, by examining the first eight negative powers of $(-1+i)$ and, noting that subsequent powers are simply $(1/16)^s$ multiples for integers $s = 1, 2, 3, \ldots$, it is possible to accumulate all terms which include a positive real component; no term representing a step in a contrary direction can be recouped because the denominators are doubled after each pair of consecutive powers:

$$
\begin{aligned}
&\frac{1-i}{4}+\frac{1+i}{8}+\frac{1}{16}+\frac{1-i}{64}+\frac{1+i}{128}+\frac{1}{256}+\dots \\
&\quad=(-1+i)^{-3}+(-1+i)^{-5}+(-1+i)^{-8}+(-1+i)^{-11}+(-1+i)^{-13}+(-1+i)^{-16}+\dots \\
&\quad=\left((-1+i)^{-3}+(-1+i)^{-5}+(-1+i)^{-8}\right)\times\left(1+\frac{1}{16}+\left(\frac{1}{16}\right)^2+\left(\frac{1}{16}\right)^3+\dots\right) \\
&\quad=\left(\frac{7}{16}-i\frac{1}{8}\right)\times\frac{16}{15}=\frac{7}{15}-i\frac{2}{15}
\end{aligned}
$$

Thus $\alpha \le 7/15$. Similar calculations show that $-13/15 \le \alpha \le 7/15$ and $-14/15 \le \beta \le 11/15$. Figure 1 displays these bounds along with the circle of convergence $\alpha^2+\beta^2<1$.

It will be seen that there are three regions R_j, for $j=1,2,3$, in which the choice $w_0=(-1+i)^{-K}$ would not lead to convergence. However, regions R_1 and R_3 are not reached in the summation process for $\alpha+i\beta$ as is now shown.

As may be seen from table 2, if successive negative powers of $-1+i$ are taken in slices of eight consecutive terms then there is a factor of 1/16 between each slice. Consideration of the 256 possible numbers $0.p_1p_2p_3p_4p_5p_6p_7p_8$ shows that only 1 point does not lie within the unit circle. This is $0.10110100_{|(-1+i)} \equiv -1/2-i(7/8)$ with magnitude $\sqrt{65}/8$. Thus the maximum magnitude which may be accumulated from taking consecutive slices of eight terms is

$$\frac{\sqrt{65}}{8}\times\left(\frac{1}{16}+\left(\frac{1}{16}\right)^2+\left(\frac{1}{16}\right)^3+\dots\right)=\frac{\sqrt{65}}{120}$$

If we now take a circle, again centred at the origin, with radius $1-\sqrt{65}/120$ we find that six additional points lie in its exterior. The seven points are listed in table 3.

Thus any complex number which has any of these numbers represented in the first eight terms in its $(-1+i)$ expansion is potentially in region R_2 although some will remain inside the unit circle. However, a suitable initial approximation for the iteration process can be found for those cases where

$$\sum_{j=-\infty}^{-1} a_j(-1+i)^j=\alpha+i\beta\in R_2$$

Decimal	Base $(-1+i)$
$-3/4-i(5/8)$	0.10010100
$-13/16-i(9/16)$	0.10010110
$-3/4-i(9/16)$	0.10010111
$-1/2-i(7/8)$	0.10110100
$-7/16-i(7/8)$	0.10110101
$-9/16-i(13/16)$	0.10110110
$-1/2-i(13/16)$	0.10110111

Table 3. Points exterior to the circle $x^2+y^2=(1-\sqrt{65}/120)^2$.

Putting $w'_0 = i(-1+i)^{-K}$ gives

$$w'_0 z = i(-1+i)^K z = i(1+\alpha+i\beta) = -\beta + i(\alpha+1)$$

thus

$$w'_0 z - 1 = -(\beta+1) + i(\alpha+1)$$

and

$$\left|w'_0 z - 1\right|^2 = (\beta+1)^2 + (\alpha+1)^2$$

Inspection of figure 1 is sufficient to show that R_2 is well inside the circle $(\alpha+1)^2 + (\beta+1)^2 = 1$. Thus, if $\left|\alpha + i\beta\right| \geq 1$ then taking $w'_0 = i(-1+i)^K$ ensures convergence.

Clearly some numbers $\alpha + i\beta$ lie within both circles and this allows a decision rule for the first approximation in the Newton process. By analysis similar to that above, the set of all numbers for which the radix expansion commences $0.1001 = -(3/4) - i(1/2)$ or $0.101101 = -(1/2) - i(3/4)$ are inside the circle $(\alpha+1)^2 + (\beta+1)^2 = 1$. This set clearly includes the points listed in table 3.

For the first expansion the additional choices $p_5 = 1$, $p_6 = 0$, $p_7 = 0$, $p_8 = 1$ allow the maximum magnitude

$$|0.00001001_{|(-1+i)}| = \left|\frac{3}{16} + i\frac{1}{8}\right| = \frac{\sqrt{5}}{16} = 0.13975$$

As seen, an upper bound including all subsequent slices of eight negative powers of the radix is $(\sqrt{65}/120) = 0.13975$. Thus all points which may be constructed around the point $-(3/4) - i(1/2)$ as centre lie within a circle of radius $0.13975 + 0.06719 = 0.207$ (to three decimal places). It is a simple matter to check that this circle is entirely within the circle $(\alpha+1)^2 + (\beta+1)^2 = 1$. Similarly, chosing $p_7 = 1$ and

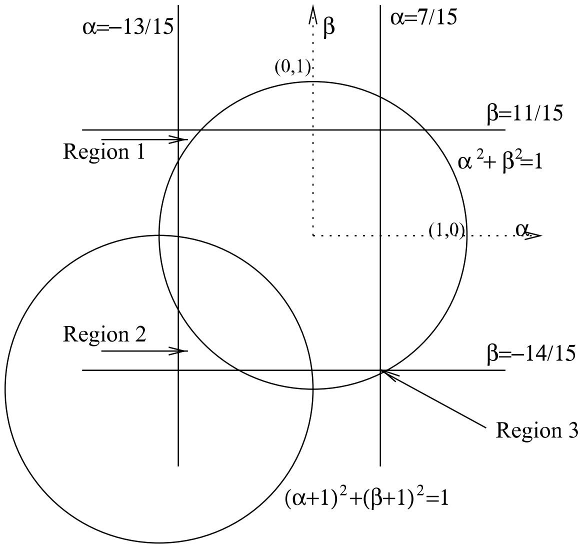

Figure 1. Convergence region: constraints on α and β.

Iteration no	Basic option	Using decision rule
0	$(-1+i)^{-3} = (1/4) - i(1/4)$	$(-1+i)^{-2} + (-1+i)^{-3} = (1/4) + i(1/4)$
1	$0.6015625000 \; - i0.1640625000$	$0.3984375000 \; + i0.1640625000$
2	$0.4632949826 + 0.04745101930$	$0.3363418577 + i0.08388519279$
3	$0.5976191750 - i0.2411555110$	$0.3417212911 + i0.1023169993$
4	$0.6258891057 + i0.5352610170$	$0.3409324421 + i0.1030755980$
5	$0.4245672037 - i0.6446767638$	$0.3409316153 + i0.1030723488$
6	$1.926416227 - i0.009392284842$	$0.3409316156 + i0.1030723489$
7	$\approx -6 + i3$	$0.3409316156 + i0.1030723489$
8	Diverging	Converged

Table 4. Reciprocal of $2.6875 - 0.8125i$.

$p_8 = 0$, allows an upper bound for the radius of a circle centred at $(1/2) - i(3/4)$ of $(\sqrt{2}/16) + (\sqrt{65}/120) = 0.156$, again well within the second circle of convergence. For these cases the second option for an initial condition should be chosen. Moreover taking the second option considerably improves convergence for those points which, while lying within the unit circle at the origin, are near to its boundary. Furthermore, the decision rule may be applied by inspection of the leading digits (base $-1+i$) of the number for which a reciprocal is required since if $\alpha + i\beta = 0.1001\ldots$ or $0.101101\ldots$ the leading significant digits of z are $11001\ldots$ or $1101101\ldots$ respectively. The following examples illustrate these cases:

(i) *To find the reciprocal w of $z = 2.6875 - 0.8125i$.*
$z_{|(-1+i)} = 1101.1010011$. Here the choice $w_0 = (-1+i)^{-3} \equiv (1/4) - i(1/4)$ gives $zw_0 - 1 = (17/32) - i(7/8)$ for which the magnitude is greater than unity. Thus in order to ensure convergence the initial approximation must be

$$iw_0 \equiv 11.0 \times (-1+i)^{-3} = (-1+i)^{-2} + (-1+i)^{-3} \equiv \frac{1}{4} + i\frac{1}{4}$$

when convergence is obtained, see table 4.

(ii) *To find the reciprocal of $z = -27 - 16i$*

$z_{|(-1+i)} = 11011010001.0$. According to the decision rule the initial approximation for the Newton process is $w_0 = (-1+i)^{-9} + (-1+i)^{-10} \equiv -(1/32)$. However in this case, without the benefit of the rule, the first option suggests $w_0 = (-1+i)^{-10} \equiv -i(1/32)$ when

$$|\alpha + i\beta| = |zw_0 - 1| = |0.1011010001| \equiv 0.980772\ldots$$

that is just inside the boundary of the unit circle at the origin. A comparison of the convergence in each case is offered in table 5.

5. The set of all sums of negative powers of the radix

As has been noted, if the terms of equation (4) are taken in slices of eight consecutive terms, then there is a geometrical ratio of (1/16) between consecutive slices. As the a_j take values $\{0, 1\}$ the set of all points described in equation (4) describes the twin-dragon fractal (boundary dimension ≈ 1.5236), see for example [12]. An impression of this fractal is given in figure 2, produced using

Iteration no	Basic option	Using decision rule
0	$(-1+i)^{-10} = i(1/32)$	$(-1+i)^{-9}i + (-1+i)^{-10}i = -(1/32)$
1	$-0.02636718750 + i0.04687500000$	$-0.03613281250 + i0.01562500000$
2	$-0.5373859404 + i0.002975463884$	$-0.02554035185 + i0.01774597167$
3	$-0.02462793245 + i0.04338020620$	$-0.02746754448 + i0.01641531287$
4	$-0.04950148699 + i0.008663798134$	$-0.02741156761 + i0.01624271162$
5	$-0.02114503844 + i0.03217394123$	$-0.02741156756 + i0.01624271184$
6	$-0.03639720748 + i0.01820184685$	$-0.02741156750 + i0.01624271182$
7	$-0.02477139420 + i0.01652408003$	$-0.02741156750 + i0.01624271182$
8	$-0.02724883215 + i0.01639386504$	Converged
9	$-0.02741184549 + i0.01624503221$	
10	$-0.02741116751 + i0.0162436574$	
11	$-0.02741116750 + i0.0162436582$	
12	$-0.02741116750 + i0.0162436582$	
	Converged	

Table 5. Reciprocal of $-27 - 16i$.

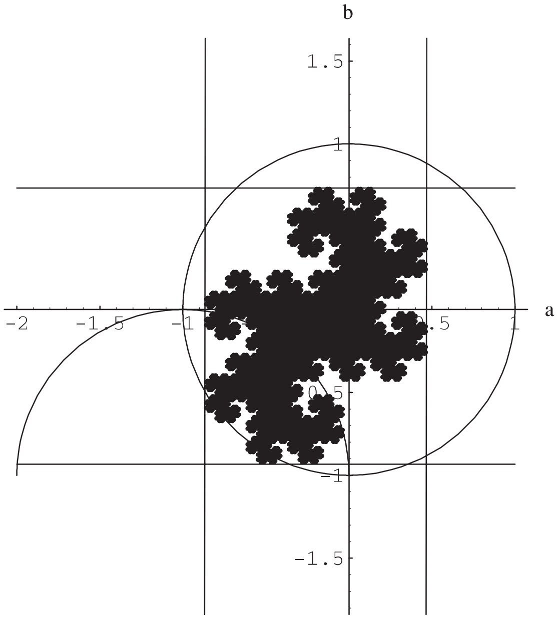

Figure 2. The twin-dragon fractal.

Mathematica$^{\text{TM}}$ and using 2^{13} points; also included, for reference, are the unit circles at the origin and at $(-1, -1)$ (upper semi-circle only) and the critical values for α and β derived in section 4. It will be noted from the figure that those points which are generated by equation (4) and which are near to the boundary of the circle at the origin are all well inside the circle centred at $(-1, -1)$ in further confirmation of the availability for improvement in convergence of the Newton process for reciprocals.

An important feature of the twin-dragon fractal is its ability to tile the plane. The radix notation discussed in this paper offers a convenient way to demonstrate this feature—since any complex number can be expressed in the form of equation (1) and complex integers can be expressed in terms of positive powers of the radix, this is obvious—and to manipulate the fractal image. For example, if the fractal is given by

$$F = \left\{ \sum_{j=-\infty}^{-1} a_j(-1+i)^j | a_j \in \{0,1\} \right\}$$

then multiplication by $(-1+i)$, which simply shifts the radix point one place to the right, gives

$$\begin{aligned} \left\{ (-1+i) \times \sum_{j=-\infty}^{-1} a_j(-1+i)^j \right\} &= \left\{ \sum_{j=-\infty}^{-1} a_j(-1+i)^{j+1} \right\} \\ &= (0+F) \cup (1+F) \end{aligned}$$

where,

$$z + F = \left\{ \sum_{j=-\infty}^{-1} \left[z + a_j(-1+i)^j \right] | a_j \in \{0,1\} \right\}.$$

Thus the result of the multiplication is the union of the original fractal and an identical fractal moved one unit in the x-direction. Similarly multiplication by $(-1+i)^2 = -2i$ [two radix point shifts to the right] gives $F \cup (1+F) \cup (i+F) \cup$ $(-1+i+F)$ offering additional shifts of one unit in the y-direction and a unit in the y-direction with -1 in the x-direction. The result of multiplying F by $(-i+1) = \sqrt{2} \times ((-1+i)/\sqrt{2})$, is a magnification by $\sqrt{2}$ and a clockwise rotation through 45°. The results of these first two transformations can be seen in figure 3.

One obvious comment is that points at the edges of adjacent images will have two, or possibly three, distinct representations in terms of the radix—a situation akin to that of the equality of the decimals 1 and 0.99999…. For example,

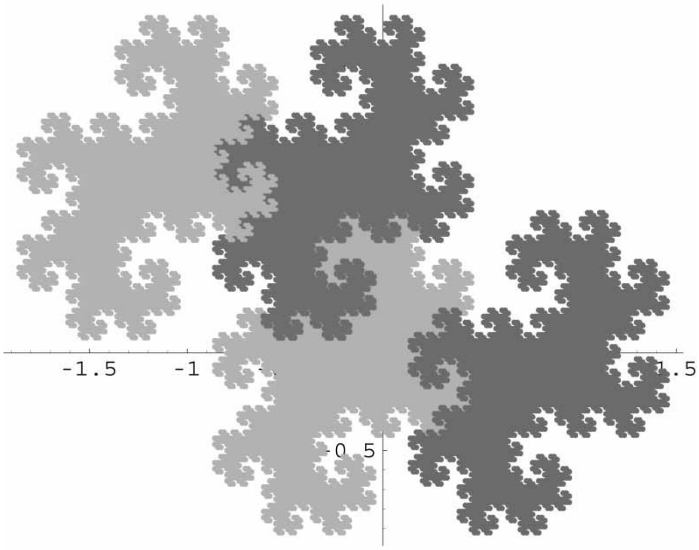

Figure 3. Elements of fractal structure.

the complex number $(7/15) - i(2/15) = 0.\dot{0}010100\dot{1}_{|(-1+i)}$ in F has image $-(8/15)-i(2/15) = 11101.\dot{0}010100\dot{1}_{|(-1+i)}$ in $(-1+F)$. But, in F,

$$-\frac{8}{15} - i\frac{2}{15} = \left(-\frac{1}{2} - i\frac{1}{8}\right)\left(\frac{16}{15}\right) = \left[(-1+i)^{-1} + (-1+i)^{-2} + (-i+1)^{-6}\right]\left(\frac{16}{15}\right)$$
$$= 0.\dot{1}100010\dot{0}_{|(-1+i)}$$

The relationship between such pairs of representations of points on the 'boundary' of the fractal and the calculation of triple points and their role in characterizing the fractal remain as possibilities for additional investigation. In addition, a similar investigation to that of this section might be made of fractals generated by the sum of negative powers of the radix for list of possible systems identified by Benedek and Panzone [9].

6. Conclusion

This paper has offered the mathematical detail of a 'binary' representation of complex numbers using $(-1+i)$ as basis and has elaborated a fail-safe procedure for the division process based on the Newton–Raphson iteration for solution of nonlinear equations. Particular attention was paid to the choice of the first approximation in the iteration and two cases requiring different initial approximations were identified. It was noted that a judicious choice of initial approximation can also be used to accelerate convergence of the Newton process.

While there are necessarily different rules for the arithmetic of complex numbers as single entities compared with the usual binary arithmetic, these should not be more costly in terms of speed of computer implementation and, while more bits are required for the representation, this should be compared with the need to find representations for separate real and imaginary parts in current computer arithmetic for complex numbers. The real advantage of a single representation of complex numbers was shown in the operations of multiplication and division where many individual multiplications, costly in terms of computer application, may be avoided to obtain the same result. However, subtraction, multiplication and division all depend in some way on addition and it was noted that further progress will depend on an efficient implementation of the zero rule discussed in section 2.

While this latter problem is still under investigation a number of possible extensions to this work requiring only the basic properties of complex numbers, but within the binary representation, have been identified. To these might be added the calculation of the square root of a complex number for which the Newton–Raphson method will be required.

References

[1] Stepanenko, V. N., 1996, *Cybernetics and Systems Analysis*, **32**, 585.

[2] Aoki, T., Amada, H., and Higuchi, T., 1997, *Proceedings 13th IEEE Symposium on Computer Arithmetic*, pp. 200–207.

[3] Geppert, L., 1999, *IEEE Spectrum*, **36**(7), 23.

[4] Lai, F., and Wu, C. E., 1991, *IEEE Trans. Comput.*, **40**, 952.

[5] Smith, D. M., 1998, *ACM Trans. Math. Software*, **24**, 359.

[6] Stewart, G. W., 1985, *ACM Trans. Math. Software*, **11**, 238.

[7] Knuth, D. E., 1998, *The Art of Computer Programming, vol. 2, Seminumerical Algorithms*, 3rd edn (Addison Wesley Longman), pp. 204–206.

[8] Knuth, D. E., 1960, *Commun. ACM*, **3**, 245.

[9] Benedek, A., and Panzone, R., 1997, *Proceedings , 4th "Dr. Antonio A.R. Monteiro" Congress on Mathematics*", Bahia Blanca, pp. 17–32.

[10] Penney, W. A., 1965, *J. ACM*, **12**, 247.

[11] Nielsen, A. M., and Kornerup, P., 1999, *IEEE Trans. Comput.*, **48**(11), pp. 1153–1165.

[12] Barnsley, M. F., 1993, *Fractals Everywhere*, 2nd edn. (Academic Press).