Corollary 8: For $S_{p_i}^n$, $1 \leq i \leq k$, let $\gamma_i$ be the greatest common root of $\delta$ and $\beta_i$ when they are commensurable, and let $\gamma_i$ be unity otherwise. Let

$$m \leq \min_i \{(n_i - 1) \log_2 \beta_i + \log_2 \gamma_i\}$$

and let $Q_i = R_i^n R_i^\delta$ for $1 \leq i \leq k$. If $Q = Q_i Q_j$ for any $1 \leq j \leq k$ where $Q$ is composed from the mappings $Q_i$, $1 \leq i \leq k$, then $Q = Q_j$.

The fact that $m$ must be bounded from above in (25) in order to guarantee control of accumulated error and avoid situations such as that exhibited in Fig. 10 demonstrates that the phrase "carry more digits" does not always mean that greater overall accuracy will follow, and such cliches should not be used as a substitute for a true understanding of the formal structure of floating-point number systems and base conversion.

**References**


The Correspondence Between Methods of Digital Division and Multiplier Recoding Procedures

JAMES E. ROBERTSON, MEMBER, IEEE

Abstract — This paper relates previous analyses of the binary SRT division to the theory of multiplier recoding. Since each binary quotient digit has three possible values, the quotient resulting from the SRT division is in recoded form; in this paper it is shown that the recoding is a function of the divisor, and the method for determining the characteristic Boolean function of the recoding is presented. The relationship between the division and the recoding is established by scaling the division in such a way that the scaled 'divisor' becomes a constant. Higher radix results are also discussed.

Index Terms — Binary arithmetic, division, minimal representation, multiplication, multiplier recoding, redundancy.

**Introduction**

Statistical analyses of the so-called SRT method of binary division have been conducted by Freeman [1] and Shively [8]. At each recursive step of this division procedure, three alternatives are possible: shift left, add and shift left, or subtract and shift left. One may therefore take the point of view that each quotient digit correspondingly has one of the three values 0, -1, or +1. Thus the division procedure results in a quotient in redundant recoded form. It is the first purpose of this paper to establish the correspondence between the quotient recordings and a class of multiplier recordings.

It is first necessary to review some aspects of the theory of multiplier recoding. For the binary case with recoded digital values of +1, 0, or -1, a recoding can in general be characterized by the choice of two Boolean functions. For the important case of arithmetically symmetric recordings (defined below), the two Boolean functions are duals of one another; hence this class of recordings is characterized by the choice of one Boolean function. A third class of recordings is next defined by restricting the Boolean functions in such a way that each function can be determined by the choice of a single binary numerical parameter in the interval 0 to 1. It is this third class of recordings that corresponds to the quotient recordings of the SRT division.

The next step in establishing the correspondence is that of scaling. It is obvious that the value of the quotient remains unchanged if both the divisor and dividend are multiplied...
ROBERTSON: DIGITAL DIVISION AND MULTIPLIER RECODING PROCEDURES

by the same parameter; similarly the detailed rules for each recursive step can be scaled in such a way that both the value and the particular recoded form of the quotient remain unchanged [3]. One further observation is that, although in practice the divisor for the SRT division is restricted to the range from 1/2 to 1, the method is valid for all divisors greater than 1.

Although the primary emphasis here is on the binary case, with redundancy limited to the use of three digital values -1, 0, and +1, procedures and examples are described for establishing a correspondence with known division methods involving greater redundancy with radix 2, and with higher radix division methods.

REVIEW OF THE THEORY OF MULTIPLIER RECODING

A multiplier y is represented in radix r by a sign digit y_0 and m nonsign digits y_1, y_2, ..., y_m. The algebraic value of a fraction y in radix complement representation is then

\[ y = -y_0 + \sum_{i=1}^{m} r^{-i}y_i. \]

The multiplier y is said to be in conventional form if the sign digit y_0 is either 0 or 1 and if each nonsign digit y_i has one of the r values 0, 1, 2, ..., r-1.

The effect of recoding is to transform y into m digits y_i (i=1, 2, ..., m), with each y_i selected from the extended set of values -(r-1), ..., -1, 0, 1, ..., r-1 or from a subset thereof. If the extended set of values over which y_i may range has more than r values, the recoded representation is redundant.

The basic equation for multiplier recoding follows from the observation that the addition of a mode digit m_i in digital position i is compensated for by the subtraction of rm_i in digital position i+1. The net effect at the i-th digital position is then

\[ y_i = y_{i+1} + m_i - rm_i. \]  

The requirement that the algebraic value of y shall remain unchanged by recoding is expressed as

\[ -y_0 + \sum_{i=1}^{m} r^{-i}y_i = \sum_{i=1}^{m} r^{-i}y_i = y, \]

and by substitution of (1), leads to the boundary mode conditions m_0 = y_0 and m_m = 0. Restriction of the range of y_i to -(r-1), ..., -1, 0, 1, ..., r-1 restricts each m_i to one of the two values 0, 1.

The more usual practice in multiplication is to inspect first the least significant digit y_m and direct attention thereafter in a serial fashion to digits of greater significance to the left of y_m. The recodings considered here are right-directed; that is, the most significant digits are inspected first, and attention is directed thereafter serially to digits of lesser significance to the right.

For a right-directed recoding, (1) is applied recursively with i assuming the values 1, 2, ..., m, in ascending order. For each value of i, the values of m_{i-1} and y_i are known, and m_i and y_i are to be determined.

For the binary case (r=2), each y_i and m_i have one of the values 0, 1 and may be treated as Boolean variables, and y_i has one of the three values -1, 0, 1. The sign of y_i is m_{i-1}, the magnitude of y_i is y_i \oplus m_i (where \oplus is the symbol for EXCLUSIVE OR); therefore

\[ y_i = (-1)^{m_{i-1}}(y_i \oplus m_i). \]

The unknown mode digit m_i can be determined by the Boolean equation

\[ m_i = m_{i-1} \lor \overline{y_i} \lor f_i \lor m_{i-1} \lor g_i \]

in which m_{i-1} and g_i, for example, means m_{i-1} AND g_i, the symbol \lor represents the INCLUSIVE OR, and \overline{y_i} means NOT y_i. The Boolean functions f_i and g_i are in theory arbitrary functions of conventional multiplier digits other than y_i. For the right-directed recodings discussed by Penhollow [5], the functional dependence of f_i and g_i was restricted to digits y_{i+j} for j \geq 1, that is, on multiplier digits to the right of y_i.

It is sometimes convenient to impose the restriction of arithmetic symmetry on the recoding. A recoding is arithmetically symmetric if the recoded representation of -y can be found from the recoded representation of y by replacing in each digital position 0's by 0's, 1's by -1's, and -1's by +1's. The effect of arithmetic symmetry on (3) is that f_i and g_i are dual Boolean functions, i.e.,

\[ g_i = f_i \overline{y_i}, \]

THE SIMPLEST MINIMAL RIGHT-DIRECTED BINARY RECODING

A binary recoding is minimal if the probability that y_i is nonzero is minimal. The simplest such right-directed recoding [5] is obtained if, in (3),

\[ f_i = y_{i+1} y_{i+2}, \quad g_i = f_i \overline{y_{i+1} y_{i+2}}. \]

Since g_i = f_i \overline{y_i}, the recoding is arithmetically symmetric.

The tabular equivalent of (2) and (3), with f_i and g_i defined by (4), is given in Table I. Table II presents the numerical example of the recoding of the binary fraction 45/256 = 0.00101101, using the rules embodied in Table I.

DIVISION AS A METHOD OF RECODING

The recursion relationship applicable to many varieties of division is

\[ X_i = rX_{i-1} - qd, \]

with r the radix, i is an index ranging from 1 to m, the number of quotient digits, d is the divisor, X_{i-1} and X_{i} are successive partial remainders, and q_i is the ith quotient digit. The initial partial remainder X_0 is the dividend. The basic problem in division is that of selection of q_i; this selection is based on the values of the shifted partial remainder rX_{i-1} and the divisor d.

The problem of selection of q_i is eased if q_i is redundantly represented, that is, if each q_i may have more than r values. The rules for selection of q_i, coupled with (5), characterize
a method of division, aside from special procedures preliminary to and following the recursive steps.

Division may be regarded as a recoding procedure if the resultant quotient is represented in other than conventional form. Let

\[ Q' = \frac{Qd}{d}, \]  

(6)

where \( Q \) is the number to be recoded. Both the divisor \( d \) and the product \( Qd \) are in conventional form; division by \( d \) then produces \( Q' \), an algebraically equivalent but recoded version of \( Q \). The simplest division with redundancy in the representation of the quotient is the binary SRT method. With \( r = 2 \), each quotient digit \( q_i \) may have one of the three values \(-1, 0, \) or \(+1\). The selection rules are particularly simple; they are:

1) if \(-1/2 \leq 2X_{i-1} < 1/2, q_i = 0;\)
2) if \(d \geq 0\) and \(1/2 \leq X_{i-1} < 1/2\), then \(q_i = 1\), and if \(d \geq 0\) and \(2X_{i-1} < -1/2\), then \(q_i = -1;\)
3) if \(d < 0\) and \(2X_{i-1} < -1/2\), then \(q_i = 1\), and if \(d < 0\) and \(1/2 \leq 2X_{i-1} < 1/2\), then \(q_i = -1.\)

Although the method is normally used with the divisor normalized (i.e., either \(1/2 \leq d < 1\) or \(-1 \leq d < -1/2\), the method remains valid if the upper limit on the magnitude of the divisor is removed (i.e., either \(d \geq 1\) or \(d < -1\)).

The recoding equivalent to that of Table II is obtained if, in (6), \(d = 2/3\), \(Q = 45/256\), and \(Qd = 15/128\). The division \(Qd/d\) illustrated in Table III, using the rules of the SRT division, yields a quotient \(Q'\), each digit \(q_i\) of which is identical with the corresponding digit \(y_i\) resulting from the right-directed recoding of Tables I and II. The division example of Table III is artificial in the sense that the divisor \(d = 2/3\) cannot be represented in a binary computer of finite precision. In the table, \(2/3\) is represented as 0.1010 or a variant thereof, with the length of the stroke indicating the period of the repetitive pattern.

### Table I

<table>
<thead>
<tr>
<th>(m_{i-1})</th>
<th>(y_i)</th>
<th>(y_{i+1})</th>
<th>(y_{i+2})</th>
<th>(m_i)</th>
<th>(y_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0/1</td>
<td>0/1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* 0/1 indicates that the digit may be either 0 or 1.
† 1 = -1.

### Table II

Recoding of 45/256. The Result of the Recoding is 0.0011001. \(T = 45/256\).

<table>
<thead>
<tr>
<th>i</th>
<th>Digits Inspected</th>
<th>Corresponding Values</th>
<th>Result of Recoding</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(m_0, y_{13} y_{12} y_{11})</td>
<td>0001</td>
<td>(m_0 = 0, y_{13} = 0)</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>(m_1, y_{12} y_{11} y_{10})</td>
<td>0010</td>
<td>(m_1 = 0, y_{12} = 0)</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>(m_2, y_{11} y_{10} y_{0})</td>
<td>0101</td>
<td>(m_1 = 1, y_{11} = 0)</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>(m_3, y_{10} y_{9} y_{8})</td>
<td>0011</td>
<td>(m_4 = 0, y_{10} = 1)</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>(m_4, y_{9} y_{8} y_{7})</td>
<td>1110</td>
<td>(m_5 = 1, y_{9} = 0)</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>(m_5, y_{8} y_{7} y_{6})</td>
<td>1101</td>
<td>(m_6 = 1, y_{8} = 0)</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>(m_6, y_{7} y_{6} y_{5})</td>
<td>1010</td>
<td>(m_7 = 1, y_{7} = 0)</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>(m_7, y_{6} y_{5} y_{4})</td>
<td>1100</td>
<td>(m_8 = 0, y_{6} = 1)</td>
<td>2, 3</td>
</tr>
</tbody>
</table>

Comments: 1) \(m_0 = y_0\) is the left boundary condition.
2) \(y_9\) and \(y_{10}\) are assumed to be zero.
3) The right boundary condition \(m_8 = 0\) is satisfied.

### Table III

SRT Division of 15/128 by 2/3. The Result of the Division is 45/256 = 0.0011001.

| \(2 \times x_0\) | 0.001111 | 15/64 |
| \(x_1 = 2 \times x_0\) | 0.001111 | \(q_1 = 0\) |
| \(2 \times x_1\) | 0.011111 | 15/32 |
| \(x_2 = 2 \times x_1\) | 0.001111 | \(q_2 = 0\) |
| \(2 \times x_2\) | 0.1111 | 15/16 |
| \(-d\) | 1.0110101 | \(q_3 = 1\) |
| \(x_3 = 2 \times x_2 - d\) | 0.010001 | 13/48 |
| \(2 \times x_3\) | 0.100001 | 13/24 |
| \(-d\) | 1.011010 | \(q_4 = 1\) |
| \(x_4 = 2 \times x_3 - d\) | 1.111000 | \(-3/24 = -1/8\) |
| \(2 \times x_4\) | 1.110000 | \(-1/4\) |
| \(x_5 = 2 \times x_4\) | 1.110000 | \(q_5 = 0\) |
| \(2 \times x_5\) | 1.100000 | \(-1/2\) |
| \(x_6 = 2 \times x_5\) | 1.100000 | \(q_6 = 0\) |
| \(2 \times x_6\) | 1.000000 | \(-1\) |
| \(+d\) | 0.101000 | \(q_7 = 1\) |
| \(x_7 = 2 \times x_6 + d\) | 1.101010 | \(-1/3\) |
| \(2 \times x_7\) | 1.010101 | \(-2/3\) |
| \(+d\) | 0.101100 | \(q_8 = 1\) |
| \(x_8 = 2 \times x_7 + d\) | 0.000000 | 0 |

Fig. 1 illustrates the mapping of \(|rX_{i-1}|\) onto \(|X_i|\) for the SRT division with \(d = 2/3\). The analysis of the division of Fig. 1 is similar to that of Freeman, but differs in two important respects:

1) The interval of shifted partial remainder magnitudes (i.e., of \(|2X_{i-1}|\)) is (0, 1) rather than (1/2, 1).
2) One step corresponds to the generation of one quotient digit, rather than an addition or subtraction followed by normalization.

The quantity 1/2, the abscissa of the discontinuity in the mapping function in Fig. 1, is called the comparison constant since the comparison of \(|2X_{i-1}|\) with this constant...
determines whether \( q_i \) is zero or \( \pm 1 \). The comparison constant plays the following important role in the analysis of the SRT division. If \((1/2)^-\) and \((1/2)^+\) are defined by

\[
\left( \frac{1}{2} \right)^- = \lim_{\Delta \to 0} \left( \frac{1}{2} - \Delta \right), \quad \left( \frac{1}{2} \right)^+ = \lim_{\Delta \to 0} \left( \frac{1}{2} + \Delta \right), \quad \Delta \geq 0,
\]

then endpoints of intervals are \( 1/2 \) and the images of \((1/2)^-\) and \((1/2)^+\), namely \( 0, 1/3, 1/2, 2/3, \) and \( 1 \). The intervals thus found are \( \alpha, \beta, \gamma, \) and \( \delta \) and the steady-state probability that \( |2x_{i-1}| \) lies within each interval is \( 1/2 \) for \( \alpha \) and \( 1/6 \) for \( \beta, \gamma, \) and \( \delta \). The probability that \( |q_i| \) is \( 1 \) is then the probability that \( |2x_{i-1}| \) is in the interval \( (1/2, 1) \) and is \( 1/3 \). Since the shift average \( \langle s \rangle \) is the reciprocal of this probability, \( \langle s \rangle = 3 \).

Fig. 2 shows the effect of scaling Fig. 1 by the factor \( z = 1/2d = 3/4 \).

If the scaled partial remainder is denoted by \( u_i \), then \( u_i = zX_i \) for each value of \( i \). Thus each partial remainder, including the dividend \( x_0 \), is scaled. The divisor scales to \( 1/2 \) and the comparison constant scales from \( 1/2 \) in Fig. 1 to \( 1/2z = 3/8 \) in Fig. 2. The images of the comparison constant, i.e., the endpoints of the intervals, are also scaled by the factor \( z \). The probabilities associated with the intervals remain invariant; the probability densities, due to the scaling on interval lengths, are scaled by the factor \( z^{-1} \). Equation (6), with \( z = 1/2d \) scales to \( Q' = (1/2)Q/1/2 \), and the quotient \( Q' \) resulting from the scaled division are identical not only in algebraic value but also in digit pattern to the quotient resulting from the conventional SRT division.

In effect, the conventional SRT division is characterized by a comparison constant which is always \( 1/2 \); the divisor value is the parameter that determines the properties of the recoding. The choice of the scaling factor \( z = 1/2d \) results in a procedure in which the "divisor" is always \( 1/2 \) and the comparison "constant" \((1/2)z\) becomes the parameter which determines the properties of the recoding. More complicated methods of division involve addition or subtraction of more than one divisor multiple of the form \( nd \), and perhaps values of the radix \( r \) greater than 2. (The radix \( r \) is an integer, and \( n \) is usually an integer, but is sometimes a simple rational
If \( ru_{i-1} \) is defined as

\[
ru_{i-1} = - m_{i-1} + \sum_{j=0}^{m-i-1} r^{-j-1} y_{i+j},
\]

then

\[
ru_i = - m_i + \sum_{j=0}^{m-i} r^{-j-1} y_{i+j+1} = - m_i + \sum_{j=1}^{m-i} r^{-j} y_{i+j}.
\]

Substitution in (8) yields

\[
ru_{i-1} = r^{-1}(y_i + ru_i),
\]

or

\[
u_i = ru_{i-1} - r^{-1} y_i.
\]  

Equation (9) can be interpreted as the recursion relationship for division, (5), scaled so that the divisor becomes 1/r; that is, \( z = 1/rd \). The quantity \( ru_{i-1} \) is identified with the scaled shifted partial remainder, which is known, \( u_i \) is the partial remainder to be determined, and \( y_i \) is identified as the quotient digit to be determined by the selection rules. The mode digit \( m_i \) is identified as the sign digit of the partial remainder \( u_i \).

The above analysis indicates that (5) for division and (1) for recoding are of the same order of generality. Just as it is necessary to augment (5) with a specific set of selection rules in order to completely characterize a division, it is also necessary to augment (1) with specific rules for determining the values of recorded digits. One of the advantages of the correspondence established here is now obvious. Sets of selection rules augmenting (5) are known for a wide variety of division procedures, including use of a higher radix than the binary, and including use of quotient digit values which are not integers. On the other hand, rules augmenting (1) have been limited to radix 2 with recorded digit values of -1, 0, and 1. Thus, the translation of the quotient digit selection rules into rules for right-directed recodings, to be discussed in connection with specific methods of division in examples to follow, can be expected to provide additional insight into the theory of right-directed recodings.

THE CLASS OF RIGHT-DIRECTED RECODINGS CORRESPONDING TO THE SCALED SRT DIVISION

The analysis of the previous section indicates that for a binary radix the scaled partial remainder \( 2u_{i-1} \) is related to the Boolean variables employed during recoding by the equation

\[
2u_{i-1} = - m_i + \sum_{j=0}^{m-i} 2^{-j-1} y_{i+j}.
\]

Equation (3), for an arithmetically symmetric right-directed recoding, is

\[
m_i = m_{i-1} \bar{y}_i \lor \bar{y}_i f_i \lor m_{i-1} f_i^P.
\]  

Thus, \( 2u_{i-1} \) contains all the information necessary for the determination of \( m_i \) since the functional dependence of \( f_i \) and \( f_i^P \) is on the \( y_{i+j} (j \geq 1) \). Furthermore, the quotient
digit resultant from the division is equivalent to the recoded multiplier digit \( y_i \), which is easily determined once \( m_i \) is known, from the equation

\[
y_i = (-1)^{m_i-1} (y_i \oplus m_i).
\]

It remains to be shown how, given the quotient digit selection rules, the function \( f_i \) can be determined.

The quotient digit selection rules for the scaled SRT division with the scaling factor \( z = 1/2d \) becomes

1) \[
\frac{z}{2} \leq \frac{z}{2} \leq 2u_{i-1}, y_i = 0
\]

2) \[
\frac{z}{2} \leq 2u_{i-1}, y_i = 1
\]

3) \[
2u_{i-1} < -\frac{z}{2}, y_i = -1.
\]

These rules, with the recursion relationship [(9) with \( r = 2 \)], namely

\[
u_i = 2u_{i-1} - \frac{1}{2} y_i,
\]

completely describe the scaled division. In these equations, the scaling factor \( z \) is a parameter in the range \( 0 \leq z \leq 1 \), corresponding to \( 0 \leq d \leq 1 \). The choice of \( z \) determines \( f_i \) (and its dual) for substitution in (10). Note that (11) implies arithmetic symmetry; hence the use of \( f_i \) in (10) rather than the arbitrary function \( g_i \) of (3). The proof of arithmetic symmetry follows from the fact that two divisions are performed with scaled dividends \( u_0 \) and \( -u_0 \), (11) guarantees that in the two resultant quotients, corresponding digits will either both be zero or negatives of one another.

Given the value of \( z \), the function \( f_i \) can be determined in the following way. Express \( z \) as a binary fraction of the form \( Z = k \), where \( k \) is an integer \( 0 \leq k \leq \infty \) and \( Z \) is an odd integer such that \( 0 \leq Z \leq 2^k \). \( f_i \) is then a function of \( k \) digits \( y_{i+j} \), for \( i = 1, 2, \ldots \), \( k \), and has \( 2^k - Z \) terms in its canonical expansion. Each minterm has \( k \) digits; the \( j \)th digit may be either \( y_{i+j} \) or \( \bar{y}_{i+j} \), and is assigned a weight \( w_j \), equaling 1 and 0, respectively. The \( i \)th minterm itself can be associated with a weight \( W_i \),

\[W_i = \prod_{j=1}^k 2^{-w_j},\]

which never exceeds \( 1 - 2^{-k} \). The canonical expansion of \( f_i \) is then the disjunction of all minterms with weights in the range \( z \leq W_i \leq 1 - 2^{-k} \). As an example of the determination of \( f_i \), consider the simplest minimal right-directed recoding, which corresponds to a scaled SRT division of Table IV, with \( z = 3/4 \). For \( z = 3/4 \), \( Z = 3, 2^k = 4 \), and \( k = 2 \). \( f_i \) is then a function of \( y_{i+1} \) and \( y_{i+2} \), and has \( 2^2 - Z = 1 \) minterm. The four minterms involving \( y_{i+1} \) and \( y_{i+2} \) are, respectively, \( y_{i+1} y_{i+2}, y_{i+1} \bar{y}_{i+2}, \bar{y}_{i+1} y_{i+2}, \) and \( \bar{y}_{i+1} \bar{y}_{i+2} \). The only acceptable minterm with weight \( z \geq 3/4 \) is the last one, therefore \( f_i = y_{i+1} \bar{y}_{i+2} \) and \( f_i^p = y_{i+1} + y_{i+2} \).

**Analysis of the SRT Family of Right-Directed Recodings**

The SRT division has been intensively analyzed by Freiman and Shively. The crucial step of identifying the Markov chain as the mathematical model of the SRT division was made by Freiman. In Shively's model, one time step corresponds to the generation of one quotient digit, including zero values with partial remainder magnitudes in the range \([0, 1]\). In contrast, one time step in Freiman's model corresponds to the generation of one nonzero quotient digit with as many partial remainder normalizations as are required, and partial remainder magnitudes are in the interval \([1/2, 1]\). Although Shively's results were originally given for the SRT division with the divisor magnitude \( |d| \) in the range \([1/2, 1]\), and with \( |d| \) as the independent variable, his results are restated here for the scaled SRT division with the scaling factor \( z = 1/2d \) as the independent variable, with \( z \) in the range \([0, 1]\). Thus, the trivial extension of Shively's results to the divisor range \( 1 < |d| \leq \infty \) (or \( 0 \leq z < 1/2 \)) is presented here.

Shively's results are summarized with reference to Fig. 3, which shows the locations of the lower order discontinuities (see 4a, 4b below) in the steady-state probability density of the scaled partial remainder magnitude \( u \) as functions of the scaling factor \( z \).

1) The range of \( u \) is \([0, 1 - z]\) for \( 0 \leq z \leq 1/2 \) and is \([0, z]\) for \( 1/2 < z \leq 1 \). Therefore, the probability density is zero outside these ranges, namely in the triangular region bounded by \( u = 1, u = z, \) and \( u = 1 - z \).

2) The probability density \( p(u, z) \) is symmetric about \( u = 1/2 \); for a fixed value \( z_0 \) of \( z \)

\[p\left(\frac{1}{2} + \theta, z_0\right) + p\left(\frac{1}{2} - \theta, z_0\right) = 2 \quad \text{for} \quad 0 < \theta < \frac{1}{2}.
\]

Symmetry of this form implies that a plot of the locations of discontinuities in the probability density in the \( u, z \) plane for \( u = 1/2 \) is the mirror image of the plot of discontinuities for \( u < 1/2 \). Furthermore, for each value \( z_0 \) of \( z \), if \( p(1/2 + \theta, z_0) \) is known, \( p(1/2 - \theta, z_0) \) can be determined from (13). In particular, symmetry and result 1, above, imply that the probability density of \( u \) is uniformly 2 for the triangular region, bounded by \( u = 0, u = z, \) and \( u = 1 - z \).

3) For fixed \( z = z_0 \), \( p(u, z_0) \) is a monotonic function of \( u \) for \( 0 < u < 1 \). This result guarantees that \( p(u, z_0) \) is uniform between negative jumps in the value of \( p(u, z_0) \) as \( u \) increases. The two next results are concerned, respectively, with deterministic procedures for finding the locations and magnitudes of the discontinuities in \( p(u, z) \). Some indication of the complexity of the pattern of discontinuities is given by Fig. 4.

4) The locations of the discontinuities in \( p(u, z) \) are found as follows.

a) The zeroth order pair of discontinuities is \( u_0 = 1 - z \) and \( u_0 = z \). In the triangular region of Fig. 3 bounded by \( u = 1, u = 1 - z, \) and \( u = z, p(u, z) = 0. \) Since \( p(u, z) \) is positive and nonzero elsewhere in the unit square, there is a negative jump in \( p(u, z) \) in the upper half \( u > 1/2 \) at \( u = 1 - z (z \leq 1/2) \)
Fig. 3. Discontinuities of order three or less.

Fig. 4. Discontinuities of order five or less.
and at \( u = z \geq 1/2 \). By symmetry, there is a negative jump at \( p(u, z) \) in the lower half \( (u < 1/2) \) at \( u = z \leq 1/2 \) and at \( u = 1 - z \geq 1/2 \).

b) Given the \( k \)th order pair of discontinuities, with the equations \( u_k = 2z - b_k \) and \( u^*_k = (b_k - 1) - 2^k z \), the \((k+1)\)st order pair is formed as follows [cf. (11), (12)].

i) If \( u_k < z/2 \), then \( u_{k+1} = 2u_k - 2^{k+1} z - 2b_k \). For \( u_k > 1 - z/2 \), \( u_{k+1} = (2b_k + 1) - 2^{k+1} z \).

ii) If \( u_k < z/2 \), then \( u_{k+1} = 2u_k - (b_k + 2) - 2^{k+1} z \). For \( u_k > 1 - z/2 \), \( u_{k+1} = 2(2u_k + 1) - 2^{k+1} z - (2b_k + 1) \).

Thus, \( b_k \) is either \( 2b_k \) or \( 2b_k + 1 \), depending on whether \( u_k \) or \( u^*_k \) is less than \( z/2 \). Note that the two lines forming each pair of lines of discontinuity are symmetric. The two lines are symmetric if \( u_k + u^*_k = 1 \); since the rules for generating pairs of lines result in \( u_k + u^*_k = 1 \), symmetry is preserved.

c) For an interval of \( z \), the process of generating pairs of lines of discontinuity terminates when \( u_k \) and \( u^*_k \) are in the interval \((z/2, 1 - z/2)\) of \( u \). Let the value of \( k \) for the terminating pair of lines of discontinuity be \( T - 1 \). The lines are \( u_{T-1} = 2^{-1} z - b_{T-1} \) and \( u^*_{T-1} = (b_{T-1} + 1) - 2^{-1} z \). These lines intersect at \( u = 1/2 \), \( z = (2b_{T-1} + 1)/(2^T) \), and the interval of \( z \) for which no additional lines of discontinuity can occur is

\[
\frac{2b_{T-1}}{2^T - 1} < z < \frac{2(b_{T-1} + 1)}{2^T + 1}.
\]

Examples are the following.

<table>
<thead>
<tr>
<th>( T )</th>
<th>( u_{T-1} )</th>
<th>( u^*_{T-1} )</th>
<th>Intersection with ( u = 1/2 )</th>
<th>Interval of ( z )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( z )</td>
<td>( 1 - z )</td>
<td>( z = 1/2 )</td>
<td>( 0 &lt; z &lt; 2/3 )</td>
</tr>
<tr>
<td>2</td>
<td>( 2 - z )</td>
<td>( 2 - 2z )</td>
<td>( z = 3/4 )</td>
<td>( 2/3 &lt; z &lt; 2/3 )</td>
</tr>
<tr>
<td>4</td>
<td>( 8 - z )</td>
<td>( 8 - 8z )</td>
<td>( z = 15/16 )</td>
<td>( 4/3 &lt; z &lt; 4/3 )</td>
</tr>
<tr>
<td>3</td>
<td>( 4 - 3 )</td>
<td>( 4 - 4 )</td>
<td>( z = 7/8 )</td>
<td>( 6/3 &lt; z &lt; 8/3 )</td>
</tr>
<tr>
<td>4</td>
<td>( 8 - 7 )</td>
<td>( 8 - 8 )</td>
<td>( z = 15/16 )</td>
<td>( 14/15 &lt; z &lt; 16/17 )</td>
</tr>
</tbody>
</table>

5) Symmetry requires that the magnitudes of the jumps at each of the two discontinuities of a symmetric pair be equal. If the magnitude of the jump at \( u_k \) is \( a_k \), then the sum for the symmetric pair \( u_k, u^*_k \) is \( 2a_k \). The magnitudes of the jumps form a geometric sequence; that is, \( a_{k+1} = (1/2) a_k \). For an interval of \( z \) having \( T \) pairs of discontinuities, the jump magnitudes are

\[
a_k = \frac{2^{-k}}{2(1 - 2^{-k})}, \quad k = 0, 1, \ldots, T - 1.
\]

Equation (14) follows from the requirement that

\[
\sum_{k=0}^{T-1} 2a_k = 2,
\]

since the total change in the probability density from \( u = 0 \) to \( u = 1 \) is 2.

6) The probability \( P_0(z) \) that the recorded digit \( y_i \) is zero is

\[
P_0(z) = \int_0^{z/2} p(u, z) \text{d}u
\]

\[
P_0(z) = z \quad \text{for} \quad 0 \leq z \leq 2/3
\]

\[
P_0(z) = 2/3 \quad \text{for} \quad 2/3 \leq z \leq 5/6.
\]

\( P_0(z) \) decreases in a nonmonotonic way from 2/3 to 1/2 in the interval 5/6 \leq z \leq 1.

The probability \( P_i(z) \) that the magnitude of the recorded digit \( |y_i| \) is 1 is

\[
P_i(z) = 1 - P_0(z)
\]

and the shift average \( \langle s \rangle \) is

\[
\langle s \rangle = \frac{1}{P_i(z)}
\]

Shively's results, as restated above, form the basis for the analysis of the family of right-directed encodings corresponding to the scaled SRT division. The encodings which have attracted the greatest attention in the past are the minimal ones, for which \( P_i(z) \) has its minimum value of 1/3. The minimal recordings of the SRT family correspond to a choice of \( z \) in the range 2/3 \leq z \leq 5/6. The simplest minimal recording of Table I is obtained by choosing \( z \) to be the simplest binary fraction in the minimal range; namely, \( z = 3/4 \), which yields \( f_i = y_{i+1} y_{i+2} \). Other relatively simple minimal recordings correspond to the choice of \( z = 11/16 \), or \( f_i = y_{i+1} y_{i+2} y_{i+3} y_{i+4} \) and to the choice of \( z = 13/16 \) or \( f_i = y_{i+1} y_{i+2} y_{i+3} y_{i+4} \). (These expressions for the \( f_i \) have been simplified.)

The minimal recordings of the SRT family are a subset of the minimal right-directed recordings found by Penhollow. The Penhollow minimal recordings differ in two ways.

1) For the SRT family, if a minterm of weight \( W_k \) is included in the canonical expansion of \( f_i \), all minterms having weight greater than \( W_k \) are also included. This restriction does not apply to the Penhollow recordings, for example,

\[
f_i = y_{i+1} (y_{i+2} y_{i+3} y_{i+4} + y_{i+3} y_{i+4} + y_{i+4})
\]

yields a minimal right-directed recording which is not a member of the SRT family, since the minterms of the canonical expansion of \( f_i \) have weights 11/16, 13/16, 7/8, and 15/16. The minterm of weight 3/4, needed in the canonical expansion for SRT minimum recording with \( k = 4 \) and \( z = 13/16 \), is not included in the canonical expansion of \( f_i \).

2) The restriction of arithmetic symmetry was not imposed by Penhollow. Let \( F_i \) be the class of all minimal right-directed recording functions \( f_i \), and let \( G_i \) be the class of all dual functions \( f_i^D \). Then any function from the class \( F_i \) and any function from the class \( G_i \) can be used as \( f_i \) and \( g_i \), respectively, in (3) to yield a minimal Penhollow recording. Arithmetic symmetry, in contrast, requires that once \( f_i \) is chosen, the one member of class \( G_i \) which is \( f_i^D \) must be used as \( g_i \), rather than any member of class \( G_i \).
Other recordings of the SRT family of historic interest are the following.

1) \( z = 0, f_z = 1, f^p_z = 0 \). In this case, the recording equation reduces to \( m_1 = \bar{y}_i, \) and each \( y \) is either +1 or -1. The corresponding division method is nonrestoring division.

2) \( z = 1/2, f_z = f^p_z = y_j + 1 \). This recording corresponds to differentiation in the following physical sense. Assume that a magnetic surface is magnetized positively for each digit \( y \) which is one, and is magnetized negatively for each digit which is zero. The recording corresponds to the pulse pattern obtained as the magnetic surface is moved past a reading head if a positive pulse is associated with +1, a negative pulse with a -1, and the absence of a pulse with zero.

This recording is also related to the transformation from binary to the Gray code.

3) \( z = 1, f_z = 0, f^p_z = 1 \). The recording equation reduces to \( m_1 = m_{i-1}, \) so that \( y_i = y_i \) if \( m_0 = y_0 = 0, \) and \( y_i = \bar{y}_i \) if \( m_0 = y_0 = 1. \) Thus positive numbers are left unchanged by this recording, and a negative number is replaced by the digitwise negative of its diminished radix complement.

4) \( z = 2/3. \) The recording for \( z = 2/3 \) is the canonical recording discussed by Reitwiesner [6], and shown to be the simplest minimal left-directed recording by Penhallow. The fact that 2/3 cannot be represented by a finite number of binary digits indicates that, as a right-directed recording, it would be necessary to inspect all digits to the right of \( y \).

**The Right-Directed Recoding Corresponding to the Stretch Division With \( d = 2/3 \)**

In the stretch division [2], additional redundancy is introduced into the representation of the quotient in such a way that the probability that the quotient digit is zero, and hence the shift average, is increased. The divisor \( d \) and the multiples \( (3/4)d \) and \( (3/2)d \) are available for addition or subtraction at each step; hence the possible values for each quotient digit are \(-3/2, -1, -3/4, 0, 3/4, 1, \) and 3/2.

In addition to the comparison constant \( k = 1/2 \) to determine if another step of normalization can be performed, two more comparison constants are required. Ideally \( k = 7/8[d] \) would be used to separate the partial remainder range for which \( |q| = 3/4 \) from the partial remainder range for which \( |q| = 1. \) Similarly \( k = 5/4[d] \) would be used to separate the range for which \( |q| = 1 \) from the range for which \( |q| = 3/2. \)

With \( d = 2/3, \) and after scaling so that the scaled divisor is 1/2, the equation relating successive partial remainders is

\[
x_i = 2x_{i-1} - \frac{1}{2} y_i
\]

and the rules for selection of \( y_i \) are

\[
\text{if } \frac{5}{8} \leq 2x_{i-1} < \frac{7}{16}, y_i = -1
\]

\[
\text{if } \frac{7}{16} \leq 2x_{i-1} < \frac{3}{8}, y_i = -\frac{3}{4}
\]

The tabular equivalent of the rules of the recording are given in Table V. Note that the introduction of fractional values of \( y_i \) requires that the original digits \( y_{j+1} \) and \( y_{j+2} \) be replaced by the values \( y_{j+1}^* \) and \( y_{j+2}^* \) in subsequent steps of the recording.

A particular numerical example of this recording procedure is the following:

\[
0 \cdot 1 0 1 0 1 0 1 1 0 1 1 1 1 0 1 0 0 0 0 0 0
\]

The probability of a zero after recording is 3/4, corresponding to a shift average of 4.

**A Right-Directed Decimal Recoding**

As an example of a higher radix recoding, a decimal right-directed recoding can be derived from the decimal division example of [7]. In this example, two serial steps determine, first, \( q_{i+1} = 5, 0, 5 \) with comparison constants ideally equal to \( \pm 2.5d, \) and second, \( q_{i+2} = 3, 1, 0, 1, 2 \) with comparison constants ideally equal to \( \pm 0.5d \) and \( \pm 1.5d. \) The quotient digit \( q_{i+1} \) is then

\[
q_{i+1} = q_{i+2} + q_{i+1}^*.
\]
TABLE VI
A DECIMAL RECODING EXAMPLE

| 10u₀ | 0.1415926535 | 0.05 ≤ 0.10u₀ < 0.15, \( q_{i+1} = 1 \) |
| 10u₁ | 0.415926535 | 0.35 ≤ 0.10u₁ < 0.45, \( q_{i+1} = 4 \) |
| 10u₂ | 0.5926535 | 0.15 ≤ 0.10u₂ < 0.25 \( q_{i+1} = 2 \) |
| 10u₃ | -0.073465 | -0.45 ≤ 0.10u₃ < -0.35 \( q_{i+1} = 4 \) |
| 10u₄ | -0.073465 | -0.15 ≤ 0.10u₄ < -0.05 \( q_{i+1} = 1 \) |
| 10u₅ | 0.26535 | 0.25 ≤ 0.10u₅ < 0.35 \( q_{i+1} = 3 \) |
| 10u₆ | -0.3465 | -0.35 ≤ 0.10u₆ < -0.25 \( q_{i+1} = 3 \) |
| 10u₇ | -0.465 | -0.45 ≤ 0.10u₇ < 0.45 \( q_{i+1} = 3 \) |
| 10u₈ | +0.35 | 0.35 ≤ 0.10u₈ < 0.45 \( q_{i+1} = 4 \) |
| 10u₉ | -0.5 | -0.45 ≤ 0.10u₉ < 0.5 |}

After scaling by the factor 1/10d, and combining the two serial steps into one, the rules for the scaled division, for each scaled partial remainder \( u_i \), become

- If \( 10u_i < -0.45 \), \( q_{i+1} = 5 \)
- If \( -0.45 \leq 10u_i < -0.35 \), \( q_{i+1} = 4 \)
- If \( -0.35 \leq 10u_i < -0.25 \), \( q_{i+1} = 3 \)
- If \( -0.25 \leq 10u_i < -0.15 \), \( q_{i+1} = 2 \)
- If \( -0.15 \leq 10u_i < -0.05 \), \( q_{i+1} = 1 \)
- If \( -0.05 \leq 10u_i < +0.05 \), \( q_{i+1} = 0 \)
- If \( 0.05 \leq 10u_i < 0.15 \), \( q_{i+1} = 1 \)
- If \( 0.15 \leq 10u_i < 0.25 \), \( q_{i+1} = 2 \)
- If \( 0.25 \leq 10u_i < 0.35 \), \( q_{i+1} = 3 \)
- If \( 0.35 \leq 10u_i < 0.45 \), \( q_{i+1} = 4 \)
- If \( 0.45 \leq 10u_i \), \( q_{i+1} = 5 \)

The example of the recoding of 0.1415926535, treated here as a scaled division with a divisor value of 0.1, and with the recursion relationship \( u_{i+1} = 10u_i - 0.1q_{i+1} \), is shown in Table VI.

The result of the recoding is 0.1424733545. The recoding procedure requires inspection of the sign and first decimal digit, and knowledge of whether or not the second decimal digit is greater than or equal to five.

Summary

This report establishes a correspondence between results in binary recording theory obtained by Penhollow [5] and the analyses of binary division by Freiman [1] and Shively [8]. The correspondence is advantageous for the following reasons.

1) The results of the binary division analyses become applicable to recoding theory.

2) Other known methods of division, when subjected to similar correspondence relationships, yield new methods of recoding applicable to multiplication procedures.

3) For the author, the correspondences have emphasized the differences between multiplication and division. One essential difference is that a division requires a (theoretically infinite) family of recordings of the quotient, whereas in multiplication one member of the family usually suffices.

References


