Digit Selection for SRT Division and Square Root

Peter Kornerup, *Member, IEEE*

***Abstract*— The quotient digit selection in the SRT division algorithm is based on a few most significant bits of the remainder and divisor, where the remainder is usually represented in a redundant representation. The number of leading bits needed depends on the quotient radix and digit set, and is usually found by an extensive search, to assure that the next quotient digit can be chosen as valid for all points *(remainder,divisor)* in a set defined by the truncated remainder and divisor, i.e., an "uncertainty rectangle".**

This paper presents expressions for the number of bits needed from the truncated remainder and divisor (the truncation parameters), thus eliminating the need for a search through the truncation parameter space for validation.

The analysis is then extended to the digit selection in SRT square root algorithms, where it is shown that, in general, it may be necessary to increase the number of leading bits needed for digit determination in a combined divide and square root algorithm. An easy condition to check the number of bits needed is established, also checking the number of initial digits of the root may have to be found by other means, e.g., by table look-up.

The minimally redundant, radix-4 combined divide and square root algorithm is finally analyzed, and it is shown that in this case it can be implemented without such a special table to determine initial digits for the square root.

***Index Terms*— Digit selection, division, square root**

I. Introduction

THE SRT class of division algorithms is characterized by the use of redundant representations for the quotient, and most often as well for the remainder. Since the invention in the late fifties simultaneously by D. Sweeney, J.E. Robertson [2] and K.D. Tocher [3], and the introduction of the use of redundant representations for the remainders by D.E. Atkins [4], these methods have been extensively studied and implemented in processors. The famous "Pentium™ bug", where certain anomalies in the behavior of the floating point divide instruction were discovered, turned out to be caused by a few incorrect entries in the table employed by the quotient digit determination algorithm used for the radix 4 SRT implementation [5].

Due to the redundancy in the quotient digit set, there are overlaps between digit selection regions in the Robertson diagram, allowing a choice between two digit values. Hence even if the information on the remainder and divisor is incomplete (representing an uncertainty interval), it may be possible to choose one of the alternative quotient digit values. By allowing such a relaxed quotient digit determination, it is possible to base the quotient digit selection on leading digits of the divisor and of the remainder in a redundant representation. The determination of the truncation parameters, i.e., how many digits of the remainder and divisor will be needed, has been extensively studied since the paper by Atkins, e.g., in [6], [7], [8], [9], [10] to list a few. These all use extensive searches to check the validity of a given set of parameters, recently [11] reduced the search to four pairs of truncation parameters. To cite [10] "*It is not possible to determine the optimal choices of δ and f analytically, as several factors are involved in making these choices.*" (δ and f here being the number of fractional digits needed in the truncated divisor resp. remainder.) It is, however, well-known that there is a simple lower bound on δ, [12] has a lower bound on f, and [11] has upper bounds on δ and f.

It is shown here, that given a value of δ satisfying the bound mentioned above, it is indeed possible to determine analytically the other parameter f, such that a valid quotient digit selection function can be specified, eliminating the need for checking by search.

The analysis is then repeated for the SRT square root algorithm, where it is shown that in general it may be necessary to increase the number of bits needed for digit selection, compared to the equivalent parameters for division. It is also shown how to determine the number of initial digits that may have to be determined by other means. A combined divide and square root algorithm is developed, sharing digit selection intervals. For the radix 4 case it is shown that there is a simple way to determine the single needed initial digit, thus avoiding a special table for this purpose, along the lines also found in [13], however realized differently.

Section II introduces the fundamentals of SRT division and the notation, together with certain bounds used in the quotient digit selection. Section III then develops the theory leading to the determination of the truncation parameters, and thus the specification of the quotient digit selection function. A few examples then illustrate the results. The analysis is then in Section IV repeated for square root, concluding how to determine the truncation parameters for a combined divide and square root algorithm, and the number of initial digits that may have to be determined by other means for the square root determination. To complete the picture, the special case of minimally redundant radix 4 is analyzed, determining selection constants for the combined algorithm, including the possibility in this case also to determine the leading digit for the square root, avoiding a special table to initialize the algorithmin this case. Finally Section V concludes with some comments on

Dept. of Mathematics and Computer Science, University of Southern Denmark, Odense, Denmark, *E-mail: kornerup@imada.sdu.dk*

Work supported by the Danish Natural Science Research Council, grant no. 21-00-0679. This is a revised version of [1], presented at ARITH16, now extended with the analysis of square root.

the extensibility of the results to more complicated truncation procedures.

II. Fundamentals of SRT Division

SRT is not really a specific kind of division, rather it is a class of division methods, characterized by the following:

- The divisor is normalized.
- A redundant symmetric quotient digit set is used.
- Quotient digits selected by a few leading digits of remainder and divisor.
- The remainders may be in a redundant representation.

Let β be the quotient radix and $D = \{-a, \cdots, 0, \cdots, a\}$ the quotient digit set with $\beta/2 \le a \le \beta - 1$, and define the *redundancy factor*

$$\rho = \frac{a}{\beta - 1} \quad \text{with} \quad \tfrac{1}{2} < \rho \le 1 \tag{1}$$

where $\rho = 1$ corresponds to a maximally redundant digit set.

Let x be the dividend and y the positive divisor, r_i the remainder (with $r_0 = x$) and d_i the digit selected in the ith step. The purpose of the *digit selection function*, $\sigma(r_i, y)$, is to select the next quotient digit d_{i+1}, while keeping the new remainder $r_{i+1} = \beta r_i - d_{i+1} y$ bounded, say with bounds $\underline{B}$ and $\overline{B}$,

$$\underline{B} \le r_{i+1} \le \overline{B}. \tag{2}$$

Let the selection interval $[L_d, U_d]$ be the interval for βr_i, $\beta\underline{B} \le \beta r_i \le \beta\overline{B}$, for which it is possible to chose $d_{i+1} = d$ while keeping the updated remainder $r_{i+1} = \beta r_i - dy$ bounded by (2), i.e., for $L_d \le \beta r_i \le U_d$

$$\underline{B} \le r_{i+1} = \beta r_i - dy \le \overline{B}$$

must hold, corresponding to the Robertson diagram in Fig. 1.

From the diagram it is seen that

$$L_d = dy + \underline{B} \quad \text{and} \quad U_d = dy + \overline{B}, \tag{3}$$

in particular (3) must hold for $d_{i+1} = \pm a$, the extremal digit values, hence

$$L_{-a} = -ay + \underline{B} \quad \text{and} \quad U_a = ay + \overline{B}.$$

But as seen from Fig. 1, $\beta\underline{B} = L_{-a}$ and $\beta\overline{B} = U_a$, hence it follows that

$$\underline{B} = -\rho y \quad \text{and} \quad \overline{B} = \rho y$$

and from (3) we find

$$L_d = (d - \rho)y \quad \text{and} \quad U_d = (d + \rho)y. \tag{4}$$

To assure that at least one digit value can be chosen for any r_i, every value of βr_i must fall in at least one digit selection interval, i.e., it is necessary that

$$U_{d-1} \ge L_d,$$

hence by (4) we must require that $(d - 1 + \rho)y \ge (d - \rho)y$, or $\rho \ge \frac{1}{2}$ which is always satisfied since $a \ge \beta/2$. Actually by (1) recalling that we require $y > 0$

$$U_{d-1} - L_d = (2\rho - 1)y > 0, \tag{5}$$

thus consecutive selection intervals overlap, such that there are values of r_i for which in general there is a choice between two digit values (and possibly three for $\rho = 1$).

In summary, provided that $-\rho y \le r_i \le \rho y$, then the selection function can deliver at least one digit value d_{i+1} such that the new remainder satisfies $-\rho y \le r_{i+1} \le \rho y$. However, for $i = 0$, the dividend is used as the first remainder, $r_0 = x$, thus we must require that x and/or y are normalized such that $-\rho y \le x \le \rho y$, or $-\rho \le \frac{x}{y} \le \rho$. Any scaling applied for this normalization must then be used to correct the final quotient remainder pair.

Observation 1: With quotient radix $\beta \ge 2$, quotient digit set $\{-a, \cdots, 0, \cdots, a\}$, and $\rho = \frac{a}{\beta - 1}$, there exists a digit selection function $\sigma(r_i, y)$ that delivers a next quotient digit d_{i+1} such that the next remainder

$$r_{i+1} = \beta r_i - d_{i+1} y \quad \text{for} \quad i = 0, 1, \cdots$$

satisfies

$$-\rho y \le r_{i+1} \le \rho y,$$

provided that the dividend x, equal to the initial remainder r_0, and divisor $y > 0$ are normalized such that $-\rho \le \frac{x}{y} \le \rho$.

To simplify the analysis of the digit selection, we are assuming that $y > 0$. We will also assume that the digit selection and remainder updating takes place in binary arithmetic, and implicitly also that the quotient radix β is of the form $\beta = 2^m$ for some $m \ge 1$.

III. Quotient Digit Selection

Due to the overlap $U_{d-1} - L_d = (2\rho - 1)y > 0$ it is not necessary to know the exact value of the remainder r_i to be able to select a correct next digit d_{i+1}. The digit selection intervals are conveniently illustrated in a *P-D diagram* or *Taylor diagram* as in Fig. 2, showing the intervals as functions of the divisor y, assumed normalized $\frac{1}{2} \le y < 1$, and the shifted remainder βr_i.

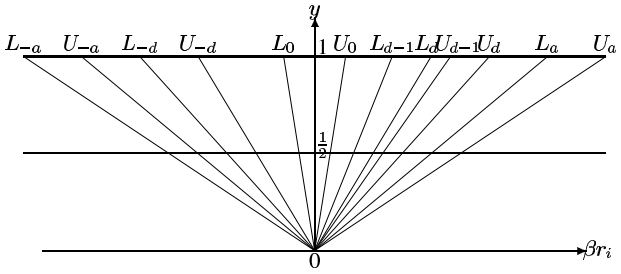


Fig. 2. P-D diagram for digit selection with normalized divisor.

For any given fixed value of the divisor y, it is now possible to choose partition points, $S_d(y)$, in the selection intervals $[L_d(y), U_d(y)]$, or rather in the stricter overlap intervals $S_d(y) \in [L_d(y), U_{d-1}(y)]$, such that the selection function $\sigma(r_i, y)$ returning d_{i+1} is defined by

$$\begin{aligned} \beta r_i \ge 0: &\quad S_d(y) \le \beta r_i < S_{d+1}(y) \;\Rightarrow\; d_{i+1} = d. \\ \beta r_i < 0: &\quad -S_{d+1}(y) \le \beta r_i < -S_d(y) \;\Rightarrow\; d_{i+1} = -d, \end{aligned}$$

using the symmetry around the y-axis, allowing us to restrict the analysis to $d \ge 0$, assuming that the remainder $r_i \ge$

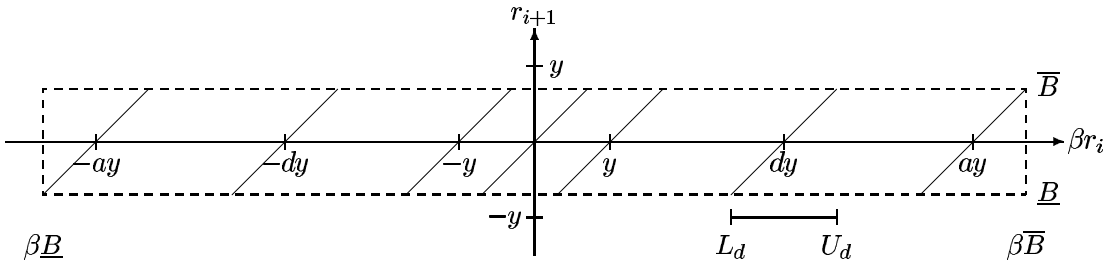


Fig. 1. Robertson diagram for SRT division

0 without loss of generality, as there is a simple way of mapping negative remainders (and divisors) into their positive equivalents [1], adjusting the sign of the digit appropriately.

To simplify the following discussion, we shall often assume that y is fixed and drop the argument y in the notation of selection intervals $[L_d(y), U_d(y)]$ and partition points $S_d(y)$. However, these can be pictured as functions of y as in Fig. 3.

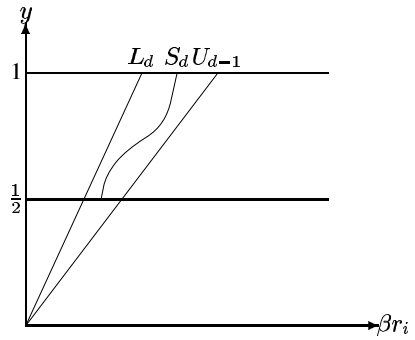


Fig. 3. $S_d(y)$ as a function of y

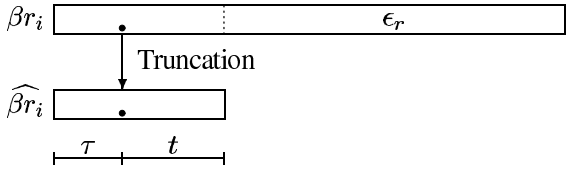
Due to the overlap between selection intervals, the partition points can be chosen such that it is sufficient to check a few of the leading digits of the (possibly redundant) value of βr_i. Let $\widehat{\beta r_i}$ denote a truncated value of βr_i, and $\mathrm{ulp}(\widehat{\beta r_i})$ denote the unit in the last place of the truncated value, say $\mathrm{ulp}(\widehat{\beta r_i}) = 2^{-t}$. Define the truncation error, ε_r, by

$$\beta r_i = \widehat{\beta r_i} + \varepsilon_r$$

then for various binary representations of r_i we have:

2's complement:	$0 \le \varepsilon_r < \mathrm{ulp}(\widehat{\beta r_i})$
2's compl. carry-save:	$0 \le \varepsilon_r < 2\mathrm{ulp}(\widehat{\beta r_i})$
borrow-save[2]:	$-\mathrm{ulp}(\widehat{\beta r_i}) < \varepsilon_r < \mathrm{ulp}(\widehat{\beta r_i})$

as illustrated below, where τ is the number of integer bits, which we shall not be concerned with here.



It is essential to note that truncation of the remainder for digit selection is assumed to take place on the redundant representation, before conversion into non-redundant representation. Thus it is not necessary to convert the full-length remainder.

Similarly, considering only a few leading digits of the divisor, let $\widehat{y}$ denote the truncated value of y with $\mathrm{ulp}(\widehat{y}) = 2^{-u}$ for $u \ge 1$, and truncation error δ defined by $y = \widehat{y} + \delta$, where $0 \le \delta < \mathrm{ulp}(\widehat{y})$ for y in 2's complement. The function

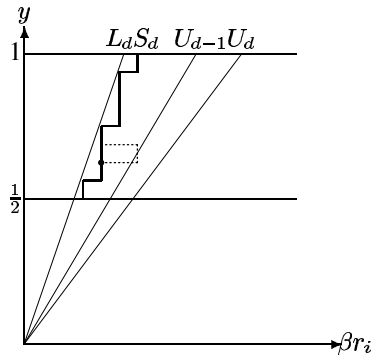


Fig. 4. $S_d(y)$ as a function of truncated divisor $\widehat{y}$

$S_d(y)$ now becomes a step function, delimiting rectangles within which a particular quotient digit can be chosen, as pictured in Fig. 4. We shall assume that $S_d(y)$ is chosen as far to the left as possible.

The step function $S_d(y)$ is now determined by a set of constants $\widehat{S}_d(\widehat{y})$, corresponding to the various truncated values $\widehat{y}$. These constants are assumed specified to the same accuracy as $\widehat{\beta r_i}$ (i.e., $\mathrm{ulp}(\widehat{\beta r_i}) = 2^{-t}$). For fixed $\widehat{y} = k2^{-u}$, where $\mathrm{ulp}(\widehat{y}) = 2^{-u}$, $\widehat{S}_d(\widehat{y})$ can then be written as an integer multiple $s_{d,k}2^{-t}$ of $\mathrm{ulp}(\widehat{\beta r_i}) = 2^{-t}$.

The dotted rectangle in Fig. 4 located with lower left-hand corner at $(\widehat{S}_d(\widehat{y}), \widehat{y})$, for y in non-redundant 2's complement and βr_i redundant carry-save 2's complement, shows the set of points $(\beta r_i, y)$ satisfying

$$\widehat{S}_d(\widehat{y}) \le \beta r_i < \widehat{S}_d(\widehat{y}) + 2\mathrm{ulp}(\widehat{\beta r_i}) \text{ and } \widehat{y} \le y < \widehat{y} + \mathrm{ulp}(\widehat{y}) \tag{6}$$

where the point $(\widehat{S}_d(\widehat{y}), \widehat{y})$ and the truncations have to be chosen such that the next quotient digit $d_{i+1} = d$ can be selected for any point in the rectangle (6).

For the shifted remainder βr_i in borrow-save, $\widehat{S}_d(\widehat{y})$ would just have to be chosen as the midpoint of the lower edge, but for the following analysis we will assume that the representation is carry-save, with the rectangle shown fully drawn to more detail in Fig. 5, where the modifications for borrow-save are trivial.

[2]Often also denoted signed-digit

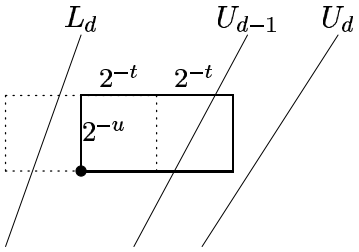


Fig. 5. Overlapping uncertainty rectangles

Using (4) for $d > 0$ the rectangle has to be to the right of the line $L_d(y) = (d-\rho)y$, yielding the following condition on the upper left-hand corner:

$$L_d(\widehat{y} + \text{ulp}(\widehat{y})) = (d-\rho)(\widehat{y} + \text{ulp}(\widehat{y})) \leq \widehat{S}_d(\widehat{y}). \quad (7)$$

These rectangles are overlapping, since they are of width $2\text{ulp}(\widehat{\beta r}_i)$, but are positioned at a horizontal spacing of $\text{ulp}(\widehat{\beta r}_i)$. Since $\widehat{S}_d(\widehat{y})$ is chosen as small as possible, for any point in the (dotted) rectangle overlapping from the left, the digit value $d-1$ must be chosen. Thus the midpoint of the bottom edge must be to the left of the line $U_{d-1}(y)$, yielding this additional condition:

$$\widehat{S}_d(\widehat{y}) + \text{ulp}(\widehat{\beta r}_i) \leq U_{d-1}(\widehat{y}) = (d-1+\rho)\widehat{y}. \quad (8)$$

But the lower right-most corner must also be to the left of the line $U_d(y)$, hence we must also require

$$\widehat{S}_d(\widehat{y}) + 2\text{ulp}(\widehat{\beta r}_i) \leq U_d(\widehat{y}) = (d+\rho)\widehat{y}. \quad (9)$$

It is easy to see that for $t \geq 1$ (which we shall see later is always the case), the upper bound on $\widehat{S}_d(\widehat{y})$ obtained from (8) is smaller than or equal to the bound found from (9). Thus combining conditions (6) and (8) to determine the size and position of the rectangles we must require

$$(d-\rho)(\widehat{y}+\text{ulp}(\widehat{y})) \leq \widehat{S}_d(\widehat{y}) \leq (d-1+\rho)\widehat{y} - \text{ulp}(\widehat{\beta r}_i). \quad (10)$$

But $\widehat{S}_d(\widehat{y})$ has to be an integer multiple of $\text{ulp}(\widehat{\beta r}_i) = 2^{-t}$, hence defining $\widehat{S}_d(\widehat{y}) = s_{d,k}2^{-t}$ we must require:

$$\left\lceil 2^{t-u}(d-\rho)(k+1) \right\rceil = s_{d,k} \leq \left\lfloor 2^{t-u}(d-1+\rho)k - 1 \right\rfloor \quad (11)$$

for $d > 0$, using $\text{ulp}(\widehat{y}) = 2^{-u}$ and defining $\widehat{y} = k2^{-u}$, for integer k, $2^{u-1} \leq k < 2^u$.

Recall that we required $\widehat{\beta r}_i \geq 0$ and thus it should also be possible to choose $d = 0$, but obviously then $\widehat{S}_0(\widehat{y}) = 0$ for all $\widehat{y}$, or $s_{0,k} = 0$ for all k, $2^{u-1} \leq k < 2^u$. Note that the right-most bounds on the uncertainty rectangles for $d = 0$ are implicitly chosen by the choice of $\widehat{S}_d(\widehat{y})$ for $d = 1$.

Without restrictions on t, u and ρ, there is only the integer term -1 which can be moved in and out of the floor and ceiling functions, but reorganizing terms then condition (11) for $d > 0$ can be written as:

$$\begin{aligned} &\left\lceil 2^{t-u}(d-\rho)k + 2^{t-u}(d-\rho) + 1 \right\rceil \\ &\quad \leq \left\lfloor 2^{t-u}(d-\rho)k + 2^{t-u}(2\rho-1)k \right\rfloor, \end{aligned} \quad (12)$$

where the ceiling and floor expressions are linear functions of k:

$$\lceil Ak + B \rceil \leq \lfloor (A+C)k \rfloor, \quad (13)$$

with $A \geq 0$, $B \geq 1$ and $C > 0$ for $d \geq 1$. Clearly it is necessary that $Ck \geq B$ for this condition to be satisfied, but it is easily seen that $Ck - B \geq 1$ is a sufficient condition, since then there is at least one integer between the ceiling and floor expressions. Hence if the condition

$$2^{t-u}\left((2\rho-1)k - (d-\rho)\right) \geq 2,$$

holds for the minimal value $k = 2^{u-1}$ and the maximal value $d = a$, then this is sufficient for (12) to hold.

Thus the stronger condition derived from $Ck - B \geq 1$

$$2^{-t} \leq \left((\rho - \tfrac{1}{2}) - (a-\rho)2^{-u}\right)/2 \quad (14)$$

may be used to find values of t, $2^{-t} = \text{ulp}(\widehat{\beta r}_i)$ and u, $2^{-u} = \text{ulp}(\widehat{y})$, for which (12) is satisfied for all $d \in \{1, \cdots, a\}$ and all k such that $2^{u-1} \leq k \leq 2^u - 1$, i.e., $\frac{1}{2} \leq \widehat{y} < 1$. Note, however, that the solutions to (14) need not be optimal in the sense that t is minimal, since it might be sufficient to require $Ck \geq B$, i.e., only require

$$2^{-t} \leq \left((\rho - \tfrac{1}{2}) - (a-\rho)2^{-u}\right), \quad (15)$$

which would allow a solution for t which is one smaller than the solution to (14). We shall below return to the choice between the two conditions.

Obviously, the right-hand side of both (14) and (15) must be strictly positive for solutions to exist for t, hence we want a u satisfying

$$2^{-u} < \frac{\rho - \frac{1}{2}}{a - \rho}, \quad (16)$$

provided that $a > \rho$, or $\beta > 2$, since $\beta = 2$ is the only case where $\rho = a(= 1)$. As seen in Example 3 below, the case $\beta = 2$ can be handled separately. Hence (16) is a sufficient condition on u for all $\beta > 2$, $d \in \{1, \cdots, a\}$ and all k such that $2^{u-1} \leq k \leq 2^u - 1$.

Returning to (13) and the choice between (14) and (15) to determine a minimal value of t, we need the following lemma:

Lemma 2: Given constants $A \geq 0$, $C > 0$, then there exists integers $k_0 \leq k_1$ such that the inequality

$$\lceil Ak + B \rceil \leq \lfloor (A+C)k \rfloor \quad (17)$$

holds for all $k \geq k_0$, provided that:

i) $\lceil Ak + B \rceil = \lfloor (A+C)k \rfloor$ for $k_0 \leq k \leq k_1 - 1$,
and ii) $\lceil Ak_1 + B \rceil < \lfloor (A+C)k_1 \rfloor$.

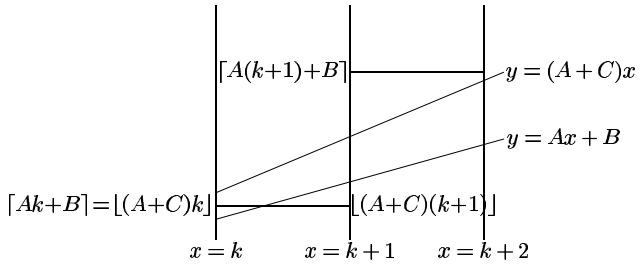
Proof: Define $\Delta(x) = \lfloor (A+C)x \rfloor - \lceil Ax + B \rceil$. Since $\Delta(x) \leq (A+C)x - (Ax+B) = Cx - B$ and $\Delta(x)$ is integral, it follows that $\Delta(x) \leq \lfloor Cx - B \rfloor$. If on the other hand $\lfloor Cx - B \rfloor = n$ then $\Delta(x) \geq n - 1$ (there are at least $n-1$ integers between $Ax + B$ and $(A+C)x$), thus

$$\lfloor Cx - B \rfloor - 1 \leq \Delta(x) \leq \lfloor Cx - B \rfloor. \quad (18)$$

Hence $\Delta(x) - 1 \leq \lfloor Cx - B \rfloor - 1 \leq \lfloor Cy - B \rfloor - 1 \leq \Delta(y)$ for $y \geq x$, thus:

$$y \geq x \;\Rightarrow\; \Delta(y) \geq \Delta(x) - 1. \quad (19)$$

$\Delta(x)$ may not always increase with x as shown in the following figure



where $\Delta(k) = 0$, but $\Delta(k+1) = \Delta(k+2) = -1$.

Since $C > 0$, by (18) there exists a minimal k_0 such that $\Delta(k) \geq 0$ for $k \geq k_0$ until eventually there is a minimal $k_1 \geq k_0$ such that $\Delta(k_1) \geq 1$. By (19) the lemma then has been proven. □

We can now combine the previous discussion with the lemma, into the following result:

Theorem 3: (SRT digit selection constants)
For radix β SRT division for $\beta > 2$ with digit set $D = \{-a, \cdots, a\}$ and $\rho = \frac{a}{\beta-1}$, the selection constants $\widehat{S_d}(\widehat{y}) = s_{d,k}2^{-u}$ can be determined for $1 \leq d \leq a$ and $\widehat{y} = k \cdot \mathrm{ulp}(\widehat{y})$ as

$$s_{d,k} = \left\lceil 2^{t-u}(d-\rho)(k+1) \right\rceil$$

for $k = 2^{u-1}, \cdots, 2^u - 1$, using truncation parameters $\mathrm{ulp}(\widehat{S_d}(\widehat{y})) = \mathrm{ulp}(\widehat{\beta r_i}) = 2^{-t}$ and $\mathrm{ulp}(\widehat{y}) = 2^{-u}$, where u has to satisfy

$$2^{-u} < \frac{\rho - \frac{1}{2}}{a - \rho}.$$

To determine t for given u, let t_0 be the smallest t satisfying

$$2^{-t} \leq (\rho - \tfrac{1}{2}) - (a-\rho)2^{-u},$$

and let

$$\Delta(u,t,k) = \lfloor 2^{t-u}(a-1+\rho)k - 1 \rfloor - \lceil 2^{t-u}(a-\rho)(k+1) \rceil,$$

then

$$t = \begin{cases} t_0 & \textbf{if } \Delta(u,t_0,2^{u-1}) \geq 1 \\ t_0 & \textbf{if } \Delta(u,t_0,k) = 0, \textbf{ for } k = 2^{u-1}, \cdots, k_1 - 1, \\ & \qquad \textbf{and } \Delta(u,t_0,k_1) \geq 1 \\ t_0 + 1 & \textbf{otherwise}. \end{cases}$$

Proof: The expression for $s_{d,k}$ is from (11), and the condition on u from (16) was shown to be sufficient for $d > 0$, by using $\max d = a$, and $\min k = 2^{u-1}$, for all values of t.

Rewriting (11) we found that u and t must satisfy the condition

$$\lceil Ak + B \rceil \leq \lfloor (A+C)k \rfloor, \tag{20}$$

or equivalently $\Delta(u,t,k) \geq 0$, for $d = a$ and for all k, $2^{u-1} \leq k < 2^u$, where

$$\begin{aligned} A &= 2^{t-u}(d-\rho) \geq 0 \\ B &= 2^{t-u}(d-\rho) + 1 \\ C &= 2^{t-u}(2\rho - 1) > 0. \end{aligned}$$

Using Lemma 2 the two first choices for t imply that (11) holds for all $k \geq k_0 = 2^{u-1}$. As we saw before, $Ck - B \geq 0$ for all d, $1 \leq d \leq a$ translates into the condition

$$2^{-t_0} \leq (\rho - \tfrac{1}{2}) - (a-\rho)2^{-u}.$$

If the conditions for the first two choices for t fail, then the stronger condition

$$2^{-t} \leq \left((\rho - \tfrac{1}{2}) - (a-\rho)2^{-u} \right)/2,$$

corresponding to $t = t_0 + 1$, implies that $Ck - B \geq 1$ for $k = k_0 = 2^{u-1}$ and $d = a$, and again by the lemma this implies that condition (11) holds for all $k \geq k_0$, and then also for all d, $1 \leq d \leq a$. □

Example 1 *(Minimally redundant radix 4 SRT)*
Let $\beta = 4$ with minimally redundant digit set $D = \{-2, -1, 0, 1, 2\}$, hence $a = 2$ and $\rho = \frac{2}{3}$, and from (16) we derive $u \geq 4$. Choosing $u = 4$ by Theorem 3, $t = t_0$ has to be the smallest solution to

$$\begin{aligned} 2^{-t} &\leq (\rho - \tfrac{1}{2}) - (a-\rho)2^{-u} \\ &= (\tfrac{2}{3} - \tfrac{1}{2}) - (2 - \tfrac{2}{3})2^{-4} \\ &= \tfrac{1}{6} - \tfrac{1}{12} = \tfrac{1}{12}, \end{aligned}$$

hence $t_0 = 4$. For $k = 2^{u-1} = 8, 9, 10$, $\Delta(4,4,k) = 0$ but $\Delta(4,4,11) = 1$. Thus by the second choice for t in Theorem 3, $t = t_0 = 4$. Hence we can compute the values of the constants $\widehat{S}(\widehat{y}) = s_{d,k}2^{-t}$ for $d > 0$ as

$$s_{d,k} = \lceil 2^{t-u}(d-\rho)(k+1) \rceil = \left\lceil (d - \tfrac{2}{3})(k+1) \right\rceil,$$

resulting in the following table of comparison constants

$k = 16\widehat{y}$	8	9	10	11	12	13	14	15
$16\widehat{S}_1$	3	4	4	4	5	5	5	6
$16\widehat{S}_2$	12	14	15	16	18	19	20	22

Utilizing the definitions (4) of the functions $L_d(y)$ and $U_d(y)$ we have

$d =$	-2	-1	0	1	2
$L_d(y)$	$\frac{-8}{3}y$	$\frac{-5}{3}y$	$\frac{-2}{3}y$	$\frac{1}{3}y$	$\frac{4}{3}y$
$U_d(y)$	$\frac{-4}{3}y$	$\frac{-1}{3}y$	$\frac{2}{3}y$	$\frac{5}{3}y$	$\frac{8}{3}y$

which together with the table of $\widehat{S}_d$ yields the P-D diagram for the first quadrant in Fig. 6. In practice the left part of the diagram (for $\widehat{\beta r_i} < 0$) is not needed, when utilizing symmetries in the uncertainty rectangles.

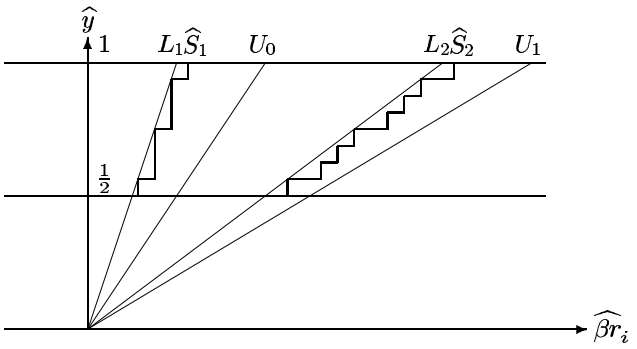


Fig. 6. P-D diagram for minimally redundant radix 4 SRT

□

Example 2 Continuing the previous example, if instead of choosing the minimal $u = 4$ the value $u = 5$ is chosen, we find

$$\begin{aligned} 2^{-t} &\leq (\rho - \tfrac{1}{2}) - (a - \rho)2^{-u} \\ &= (\tfrac{2}{3} - \tfrac{1}{2}) - (2 - \tfrac{2}{3})2^{-5} \\ &= \tfrac{1}{6} - \tfrac{1}{24} = \tfrac{1}{8}, \end{aligned}$$

having the minimal solution $t = t_0 = 3$. However, for $k = 2^{u-1} = 16$ we find for $\Delta(u, t_0, k)$

$$\begin{aligned} \Delta(5, 3, 16) &= \left\lfloor 2^{-2} \cdot \tfrac{5}{3} \cdot 16 - 1 \right\rfloor - \left\lceil 2^{-2} \cdot \tfrac{4}{3} \cdot 17 \right\rceil \\ &= \left\lfloor \tfrac{17}{3} \right\rfloor - \left\lceil \tfrac{17}{3} \right\rceil \\ &= 5 - 6 = -1, \end{aligned}$$

thus by Theorem 3 it is necessary to increase $t = t_0 + 1 = 4$, hence $(u, t) = (5, 4)$ is also a valid pair of truncation parameters, but obviously not as good as $(4, 4)$ found in the previous example. As a check we find $\Delta(5, 4, 16) = \left\lfloor \frac{37}{3} \right\rfloor - \left\lceil \frac{34}{3} \right\rceil = 12 - 12 = 0$ and $\Delta(5, 4, 17) = \left\lfloor \frac{79}{6} \right\rfloor - \left\lceil \frac{36}{3} \right\rceil = 13 - 12 = 1$. □

Example 3 *(Radix 2 SRT)*
For $\beta = 2$ and digit set $\{-1, 0, 1\}$ the condition on u in Theorem 3 cannot be used since $a = \rho = 1$. However, for $u = t = 1$ and $k = 1$, corresponding to $2^{u-1} = 1 = k < 2^u = 2$, it is easily seen that (11) is satisfied for $d = 1$.

Thus only one bit of the fraction part of y is needed for $\widehat{y}$, and this bit is always 1 since $\frac{1}{2} \leq y < 1$, implying that the quotient selection is independent of y. Note that $\mathrm{ulp}(\widehat{\beta r_i}) = 2^{-1} = \frac{1}{2}$. From (4) the lower and upper bounds are then found to be

$d =$	-1	0	1
$L_d(y)$	$-2y$	$-y$	0
$U_d(y)$	0	y	$2y$

and it now easy to see that $\widehat{S}_0 = -\frac{1}{2}$ and $\widehat{S}_1 = 0$, since the uncertainty rectangle has height $\frac{1}{2}$ and width 1. The resulting full P-D diagram is shown in Fig. 7. □

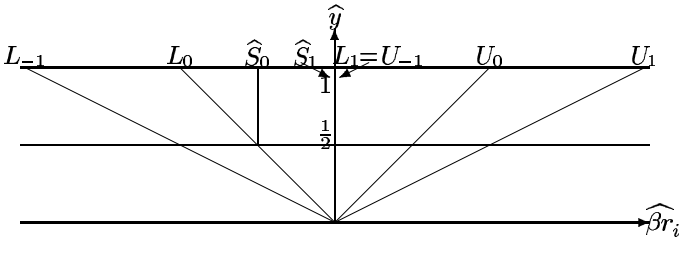


Fig. 7. P-D diagram for radix 2 SRT with redundant remainder

IV. SRT Square Root

The SRT square root algorithm assumes the radicand is normalized in a "relaxed" way, since preferably the scaling factor applied for normalization should be an even power of the radix, used in the representation of the radicand x, and in the arithmetic. Normally this will mean that a binary representation and arithmetic is used, and that the radix used for representing the quotient is of the form $\beta = 2^m$. We will initially assume that normalization is to the interval $\frac{1}{4} \leq x < 1$.

Let x be the normalized radicand and (q_i, r_i) the root,remainder pair such that $x = q_i^2 + r_i$. The purpose of the digit selection function is to select the next root digit $d = d_{i+1}$, while keeping the new remainder $r_{i+1} = \beta r_i - (2q_i + d\beta^{-i-1})d$ bounded. Let the remainder bounds as before be $\underline{B}_i$ and $\overline{B}_i$, such that

$$\underline{B}_i \leq r_i \leq \overline{B}_i, \tag{21}$$

where we shall see that the bounds here turn out to depend on the iteration index i.

Let $[L_d(i), U_d(i)]$ be the selection interval of βr_i for which we can select the digit $d_{i+1} = d$, while keeping the scaled remainder r_{i+1} bounded:

$$\begin{aligned} \beta r_i &\in [L_d(i), U_d(i)] \\ \Downarrow & \\ \underline{B}_{i+1} &\leq r_{i+1} = \beta r_i - (2q_i + d\beta^{-i-1})d \leq \overline{B}_{i+1}, \end{aligned}$$

and update the root by $q_{i+1} = q_i + d_{i+1}\beta^{-i-1}$. Hence we have the bounds

$$\begin{aligned} L_d(i) &= \underline{B}_{i+1} + 2q_i d + d^2\beta^{-i-1} \\ &\leq \beta r_i \leq \overline{B}_{i+1} + 2q_i d + d^2\beta^{-i-1} = U_d(i). \end{aligned} \tag{22}$$

With $d = a$, the maximal digit value, we have for βr_i maximal, $\beta r_i = \beta \overline{B}_i$

$$\overline{B}_{i+1} = \beta \overline{B}_i - (2q_i + a\beta^{-i-1})a \tag{23}$$

whose solution is $\overline{B}_i = 2\rho q_i + \rho^2\beta^{-i}$, which can be checked by insertion. Similarly the following expression for $\underline{B}_i$ can be found, such that the remainder r_i for all i must satisfy

$$\underline{B}_i = -2\rho q_i + \rho^2\beta^{-i} \leq r_i \leq 2\rho q_i + \rho^2\beta^{-i} = \overline{B}_i \tag{24}$$

To find the overlap between selection intervals we have using (22) that

$$\begin{aligned} L_d(i) &= \underline{B}_{i+1} + 2q_i d + d^2\beta^{-i-1} \\ &= 2q_i(d - \rho) + (d - \rho)^2\beta^{-i-1} \\ U_d(i) &= \overline{B}_{i+1} + 2q_i d + d^2\beta^{-i-1} \\ &= 2q_i(d + \rho) + (d + \rho)^2\beta^{-i-1}, \end{aligned}$$

where it is necessary for the digit selection that each value of βr_i falls in at least one selection interval, i.e., $U_{d-1}(i) \geq L_d(i)$, or

$$\begin{aligned} &U_{d-1}(i) - L_d(i) \\ &\quad = (2\rho - 1)\left(2q_i + (2d - 1)\beta^{-i-1}\right) \geq 0. \end{aligned} \tag{25}$$

Note that this inequality not only depends on d as in division, but also on the iteration index $i = 1, 2, \cdots$. Since $2\rho - 1 > 0$, for (25) to hold for all $d \in \{-a, \cdots, a\}$, i.e., for all $d \geq -a = -\rho(\beta - 1)$, and since

$$1 - 2d \leq 1 + 2\rho(\beta - 1) = 2\rho\beta - (2\rho - 1) \leq 2\rho\beta,$$

in order to assure an overlap of selection intervals it is required that

$$q_i \geq \rho\beta^{-i}.$$

Since $|\sqrt{x} - q_i| < \rho\,\mathrm{ulp}(q_i) = \rho\,\beta^{-i}$ and $\frac{1}{2} \le \sqrt{x} < 1$ this condition will be satisfied for all $i \ge i'$ for some $i' \ge 1$. Since $\rho \le 1$ it is sufficient to require $q_i \ge \beta^{-i}$, in particular it is possible to use $q_0 = d_0 = 1$ as the initial value. This is necessary whenever $\rho < 1$, since the maximally obtainable value of $q_n = \sum_1^n d_i\beta^{-i}$ without a β^0 term is

$$q_n \le \sum_1^n a\beta^{-i} < \frac{a}{\beta - 1} = \rho.$$

For maximally redundant digit sets where $\rho = 1$, it is feasible to start with $d_0 = 0$ and $d_1 = \left\lceil \frac{\beta}{2} \right\rceil$ yielding $1 > q_1 \ge \frac{1}{2} \ge \beta^{-1}$.

A. Combining SRT Square Root with Division

Looking for a square root algorithm as close as possible to division, we want the selection intervals $[L_d(i), U_{d-1}(i)] = [L_d^i(2q_i), U_{d-1}^i(2q_i)]$ to be independent of i, and as functions of $2q_i$ to coincide with the bounds for division as functions of the divisor y, at least for $i \ge i'$ for some $i' \ge 1$. Since the bounds for root extraction are

$$\begin{aligned} U_{d-1}^i(2q_i) &= 2q_i(d-1+\rho) + (d-1+\rho)^2\beta^{-(i+1)} \\ L_d^i(2q_i) &= 2q_i(d-\rho) + (d-\rho)^2\beta^{-(i+1)}, \end{aligned} \tag{26}$$

where the second term in $U_{d-1}^i(2q_i)$ is non-negative and small for large i, this term may be discarded, thus yielding the smaller upper bound

$$U_{d-1}^*(2q_i) = 2q_i(d-1+\rho), \tag{27}$$

which is now identical to the upper bound $U_{d-1}(y) = y(d - 1 + \rho)$ for division, when y is substituted by $2q_i$.

However, where the range of the divisor y is the interval $[\frac{1}{2}; 1)$, this is unfortunately not identical to the range of $2q_i \approx 2\sqrt{x} \in [1; 2)$, using the previously assumed normalization of the radicand $x \in [\frac{1}{4}; 1)$. But changing the normalization of the radicand x such that $x \in [\frac{1}{16}; \frac{1}{4})$, implies $2q_i \approx 2\sqrt{x} \in [\frac{1}{2}; 1)$, now coinciding with the interval of the divisor y.

Below we shall see how to determine an index i' such that an overlap of selection intervals is assured for all $i \ge i'$. Recalling that for SRT division it is assumed that dividend x and divisor y are normalized such that $-\rho \le \frac{x}{y} \le \rho$, both an approximate quotient q_i and an approximate root q_i can be written as $q_i = \sum_1^i d_i\beta^i$, i.e., as a proper fraction.

Rewriting the right-hand expression of (26) into

$$L_d^i(2q_i) = \left(2q_i + (d-\rho)\beta^{-(i+1)}\right)(d-\rho),$$

we notice that this can be considered a value of the bound for division $L_d(z) = z(d-\rho)$, where the argument $z = 2q_i + \delta_i^d$ is a perturbed version of the argument y used for division, the perturbation being $\delta_i^d = (d-\rho)\beta^{-(i+1)}$. As for division we shall again assume that $\beta r_i > 0$, handling negative remainders by symmetry. Observing then that

$$0 \le d - \rho \le a - \rho \quad \text{for} \quad 1 \le d \le a,$$

the perturbation δ_i^d can be made arbitrarily small by requiring $i \ge i'$ for some sufficiently large i'. Since the line $L_d(z)$ is to be the left boundary for the uncertainty rectangles, the (non-negative) perturbation can be compensated for by increasing the height of these rectangles by the maximal perturbation.

We shall now turn our attention to the selection function, in the form of determining the step-wise functions $S_d(z)$, where $z = 2q_i$ here takes the place of the divisor y in division. In particular we will first attempt to determine a value i', such that the perturbations $\max_d(\delta_i^d)$ are sufficiently small for $i \ge i'$, using the same truncation parameters u and t as for division. Recalling (10) for choosing the step function $\widehat{S}_d(\widehat{z})$ for division, but now adding the perturbation in the left bound we get the condition:

$$(d-\rho)(\widehat{z} + \mathrm{ulp}(\widehat{z}) + \delta_i^d) \le \widehat{S}_d(\widehat{z}) \le (d-1+\rho)\widehat{z} - \mathrm{ulp}(\widehat{\beta r_i}), \tag{28}$$

where $\widehat{z}$ now is either $\widehat{y}$ or $\widehat{2q_i}$.

As for division, $\widehat{S}_d(\widehat{z})$ has to be an integer multiple of $\mathrm{ulp}(\widehat{\beta r_i}) = 2^{-t}$, hence defining $\widehat{S}_d(\widehat{z}) = s_{d,k}2^{-t}$ we must require:

$$\begin{aligned} s_{d,k} &= \left\lceil 2^{t-u}(d-\rho)(k+1+2^u\delta_i^d) \right\rceil \\ &\le \left\lfloor 2^{t-u}(d-1+\rho)k - 1 \right\rfloor \end{aligned} \tag{29}$$

using $\mathrm{ulp}(\widehat{z}) = 2^{-u}$, and defining $\widehat{z} = k2^{-u}$, for integer k, $2^{u-1} \le k < 2^u$, assuming that the range of z is the half-open interval $\frac{1}{2} \le z < 1$.

Now we want to determine a minimal bound $\varepsilon > 0$ as a function of i', such that

$$\max_d(\delta_i^d) = (a-\rho)\beta^{-(i+1)} \le \varepsilon = (a-\rho)\beta^{-(i'+1)} \tag{30}$$

for $i \ge i'$, preferably without changing the values of the discretisation parameters u and t.

However, this will not be possible in general. Let $\Delta(u, t, k, \varepsilon)$ be defined as in Theorem 3, but modified with the inclusion of a perturbation ε

$$\begin{aligned} &\Delta(u,t,k,\varepsilon) = \\ &\left\lfloor 2^{t-u}(a-1+\rho)k - 1 \right\rfloor - \left\lceil 2^{t-u}(a-\rho)(k+1+2^u\varepsilon) \right\rceil. \end{aligned} \tag{31}$$

Consider the case of radix 4 SRT as in Example 1 where $\Delta(4,4,k,0) \ge 0$ for $8 \le k \le 15$, but $\Delta(4,4,8,\varepsilon) \le -1$ for any $\varepsilon > 0$, since $2^{t-u}(a-\rho)(k+1)$ happens to be integral. Thus (29) cannot be satisfied and it will be necessary to increase u and/or t.

Inserting the bound ε from (30) in (31), changing the function Δ now to be a function of i' instead of ε we obtain

$$\begin{aligned} \Delta'(u,t,k,i') = &\left\lfloor 2^{t-u}(a-1+\rho)k - 1 \right\rfloor \\ &- \left\lceil 2^{t-u}(a-\rho)\left(k+1+2^u(a-\rho)\beta^{-(i'+1)}\right) \right\rceil, \end{aligned} \tag{32}$$

whose values must be non-negative for all values of k, $2^{u-1} \le k \le 2^u - 1$ for a chosen set of parameters u, t and i'.

Theorem 4: The digit selection of radix $\beta \ge 2$ SRT square root can for $i \ge i'$ be implemented using the same algorithms as SRT division, when i' is chosen such that

$$\Delta'(u,t,2^{u-1},i') \ge 1$$

or

$$((\Delta'(u,t,k,i') = 0 \ \text{ for } \ k = 2^{u-1}, \cdots, k_1 - 1) \ \textbf{and} \ \Delta'(u,t,k_1,i') \ge 1)$$

for suitable values of u and t. These values must satisfy $u \geq u'$ and $t \geq t'$, where u' and t' are the values determined for division by Theorem 3. It is assumed that the divisor y is normalized to the interval $[\frac{1}{2}; 1)$ and the radicand x to the interval $[\frac{1}{16}; \frac{1}{4})$.

Proof: By Theorem 3, the choices of u and t assure that digit selection for division is possible for $\beta > 2$, what remains is to show that they are also valid for digit selection for square root when $i \geq i'$ in these cases. The condition of the theorem on u, t and i' implies by Lemma 2 that $\Delta'(u, t, k, i') \geq 0$ for all k, $2^{u-1} \leq k \leq 2^u - 1$, such that condition (29), or equivalently (28) holds for all digits $d > 0$. But the latter condition on the step function $S(z)$ defined by the points $\widehat{S}_d(\widehat{z})$ is equivalent to requiring that the "height extended" uncertainty rectangles (see Fig. 8) for all k, d, δ_i^d and $i \geq i'$ lie between the lines $L_d(z)$ and $U^*_{d-1}(z)$, both as defined for division.

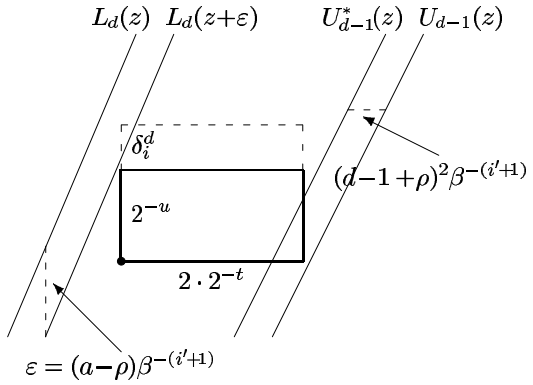


Fig. 8. Extended uncertainty rectangle

This follows from the equivalent of (7), specifying that the upper left-hand corner must be to the right of the line $L_d(z)$, as specified by

$$L_d(\widehat{z} + \mathrm{ulp}(\widehat{z}) + \varepsilon) = (d - \rho)(\widehat{z} + \mathrm{ulp}(\widehat{z}) + \varepsilon) \leq \widehat{S}_d(\widehat{z}),$$

and similarly that the midpoint of the lower edge of the rectangle must be to the left of the line $U^*_{d-1}(z)$

$$\widehat{S}(\widehat{z}) + \mathrm{ulp}(\widehat{\beta r_i}) \leq U^*_{d-1}(\widehat{z}) = (d - 1 + \rho))\widehat{z}.$$

These bounds are equivalent to (28), thus the uncertainty rectangles, for division as well as for square root, for all digits $d > 0$ and $2^{u-1} \leq k < 2^u$, for $i \geq i'$ are properly located.

What remains is to handle the case $\beta = 2$, where it is easily seen since $\rho = 1$ that the perturbations $\delta_i^d \leq 0$. For $u = t = 1$, checking the cases by insertion, it is found that (29) is satisfied for all values of $d \in \{-1, 0, 1\}$ and $k = 1$ for all $i \geq 0$, since (29) in this case is simply a rewriting of the condition that $\Delta'(1, 1, 1, 0) \geq 0$. □

Values of Δ' can thus be used to verify parameter choices, e.g., for $\beta = 4$, minimally redundant, where we know that $u = t = 4$ is sufficient for division, but that u and/or t must be increased. Trying $u = 4$, $t = 5$ and $i' = 3$ we find

$$\{\Delta'(4, 5, k, 3) \mid k = 8 \,..\, 15\} = \{0, 2, 2, 2, 4, 4, 4, 6\},$$

but also for $i' = 2$

$$\{\Delta'(4, 5, k, 2) \mid k = 8 \,..\, 15\} = \{0, 1, 1, 2, 3, 3, 4, 5\},$$

both satisfying Theorem 4. Choosing instead $u = 5$ and $t = 4$ turns out also to be possible:

$$\{\Delta'(5, 4, k, 2) \mid k = 16 \,..\, 31\} = \{0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2\}.$$

As an illustration some examples are shown in Table I, including the ones found above. The smallest possible values of u, t and i' have been chosen, allowing for non-negative values of $\Delta'(u, t, k, i')$ for all k such that $2^{u-1} \leq k \leq 2^u - 1$.

TABLE I
PARAMETERS FOR SOME COMBINED DIVIDE AND SQRT ALGORITHMS.

β	a	ρ	u	t	i'
2	1	1	1	1	0
4	2	$\frac{2}{3}$	4	5	2
4	2	$\frac{2}{3}$	5	4	2
4	3	1	4	4	1
8	4	$\frac{4}{7}$	7	6	3
8	7	1	5	4	2

However, recall that it is not necessary to check all values of k, it is sufficient to check a few initial values, e.g., for $\beta = 8$, minimally redundant

$$\{\Delta'(7, 6, k, 3) \mid k = 64 \,..\, 127\} = \{1, 1, 0, 1, 1, 1, 2, 1, \cdots, 5, 5, 4, 5, 5, 6, 5\}$$

the first value $\Delta'(7, 6, 64, 3) = 1$ is sufficient to assure that the remaining values are non-negative.

Finally note that while y is a constant in division, normally given in non-redundant representation, for square root q_i is computed for each step. Hence q_i is likely to be in a redundant representation, but can be converted to non-redundant representation "on-the-fly" [14], and then truncated to accuracy $\mathrm{ulp}(\widehat{2q_i}) = 2^{-u}$ for digit selection.

[12] investigated whether the truncation parameters t and u can be chosen such that some constant value, $z = 2q$, found by table look-up in non-redundant form, can be used as an approximation of $\widehat{2q_i}$ during the iterative phase of the square root algorithm, i.e., for $i > i'$, i' suitably chosen.

Assuming that the quotient radix is of the form $\beta = 2^m$ where $m \geq 1$, having determined $2q_{i'}$ (and implicitly digits $d_1, \cdots, d_{i'}$) by table look-up, then $\mathrm{ulp}(2q_{i'}) = 2^{-mi'+1}$. From the proof of Theorem 4 it is easy to see that the vertical location and size of a uncertainty rectangle can be determined by a modified condition (32) on Δ', to remain fixed for $i > i'$ at $\widehat{z} = \widehat{2q} = \widehat{2q_{i'}}$ with $\mathrm{ulp}(\widehat{2q}) = 2^{-mi'+1}$. Since

$$|2q_i - \widehat{2q}| \leq 2(|q_i - \sqrt{x}| + |\sqrt{x} - \widehat{q}|) \leq 2\mathrm{ulp}(\widehat{2q}) = 2^{-mi'+2}$$

it is possible to choose u, t and i' such that all subsequent points $(2q_i + \delta_i^d, \beta r_i)$ for $i > i'$ are inside such suitably located uncertainty rectangles. We shall, however, not pursue

this possibility further, but instead see that it is possible to avoid the initial table look-up phase in the frequently used case of $\beta = 4$.

B. Radix 4 Divide and Square Root without Initial PLA

Let $\beta = 4$ with minimally redundant digit set $D = \{-2, -1, 0, 1, 2\}$. From the table above we choose $u = 4$, $t = 5$ and $i' = 2$. We can now compute the values of the constants $\widehat{S}(\widehat{y}) = s_{d,k}2^{-t}$ for $d > 0$ as

$$s_{d,k} = \lceil 2^{t-u}(d-\rho)(k+1+2^u\varepsilon)\rceil = \lceil 2(d-\tfrac{2}{3})(k+\tfrac{4}{3})\rceil\ ,$$

resulting in the following table of comparison constants valid for square root digit selection for $i \geq 2$ with $z = 2q_i$, and for all $i \geq 1$ for division with $z = y$, assuming $z \in [\frac{1}{2}; 1)$:

$k = 16\widehat{z}$	8	9	10	11	12	13	14	15
$32\widehat{S}_1$	7	7	8	9	9	10	11	11
$32\widehat{S}_2$	25	28	31	33	36	39	41	44

Utilizing that the definitions of the functions $L_d(z)$ and $U_d^*(z)$ are the same as for division, combining with the table of $\widehat{S}_d$, yields the P-D diagram for the first quadrant as shown in Fig. 9.

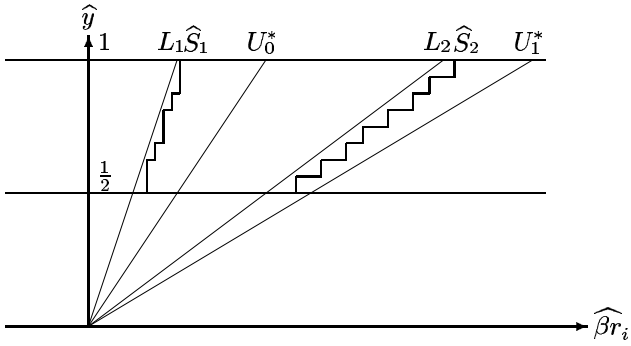


Fig. 9. P-D diagram for a combined divide and square root, radix 4 SRT, $i > 1$.

Since $q_i = \sum_1^i d_i 2^{-i} \approx \sqrt{x} \in [\frac{1}{4}; \frac{1}{2})$ and $i' = 2$, it is sufficient initially to determine d_1 for the square root algorithm, before the general SRT algorithm can be applied. Following [13] for completeness we now want to see if it is possible to find common selection intervals for d_1, valid for division as well as square root, thus avoiding a special look-up table to determine the initial digit.

For the square root of the radicand x, there are two possibilities for the choice of the first digit, $d_1 \in \{1, 2\}$, as seen by the following based on choosing the initial remainder $r_0 = x$:

$\underline{d_1 = 1}$**:** Here $q_1 = \frac{1}{4}$, thus $r_1 = 4x - \frac{1}{4}$, and the bounds $\underline{B_1} \leq 4x - \frac{1}{4} \leq \overline{B_1}$ are satisfied for $\frac{1}{144} \leq x \leq \frac{25}{144}$.

$\underline{d_1 = 2}$**:** Here $q_1 = \frac{1}{2}$, thus $r_1 = 4x - 1$, and the bounds $\underline{B_1} \leq 4x - 1 \leq \overline{B_1}$ are satisfied for $\frac{16}{144} \leq x \leq \frac{56}{144}$.

using the remainder bounds $\underline{B_i} \leq r_i \leq \overline{B_i}$ from (24). Note that the whole range $\frac{1}{16} \leq x < \frac{1}{4}$ is covered, with an overlap interval $[\frac{16}{144}; \frac{25}{144}]$ of width $\frac{1}{16}$ where both values of d_1 are possible. Translating these bounds on x into bounds on $\beta r_0 = 4x$ we find

$$\begin{aligned} d_1 = 1 \quad &\text{can be chosen for} \quad \tfrac{1}{4} \leq \beta r_0 \leq \tfrac{25}{36} \\ d_1 = 2 \quad &\text{can be chosen for} \quad \tfrac{16}{36} \leq \beta r_0 \leq 1, \end{aligned}$$

with an overlap of width $\frac{1}{4}$ between these intervals.

Note that $q_0 = 0$, hence the value of $\widehat{2q_0} = 0$ cannot be used for digit selection together with $\widehat{\beta r_0} = \widehat{4x}$, since $\widehat{z} = \widehat{2q_0}$ is supposed to be in the interval $[\frac{1}{2}; 1)$. But we are free to choose any suitable value for $\widehat{2q_0}$ (or k) in the comparisons or table look-up to determine $d_1 \in \{1, 2\}$, we may even translate or scale the initial remainder βr_0, as long as we assure that the standard digit selection chooses the proper value of d_1 according to the selection intervals just found.

It is thus possible just using $2\widehat{\beta r_0} = 2 * \widehat{4x}$ instead of $\widehat{4x}$ for the initial root digit selection. Recall that the entries in the table of selection constants are lower bounds for the digit selections, since $2 * \frac{16}{36} = \frac{32}{36} < \frac{31}{32}$, then $\frac{31}{32}$ can be used as a lower bound for selecting $d_1 = 2$, choosing $k = 10$ for $i = 1$. The corresponding lower bound for selecting $d_1 = 1$ is then $2 * \frac{1}{4} = \frac{16}{32}$, with an upper bound way beyond the lower bound for choosing $d_1 = 2$. Note that multiplication by the factor 2 is permissible, since the uncertainty rectangle of βr_0 here only has half the width because $\widehat{\beta r_0} = \widehat{4x}$ is in non-redundant representation, thus it has only half the error as values of $\widehat{\beta r_i}$ used in later iterations.

Finally, there is another minor problem with the digit selection for square root determination. If the initial digit $d_1 = 2$ is chosen, then $q_1 = \frac{1}{2}$, which is not in the interval $[\frac{1}{4}; \frac{1}{2})$, and there are no selection constants corresponding to the value $k = 16$ in the table above. Of course entries for $k = 16$ (and other larger values of k) could easily be added, but for a table look-up implementation this would imply that one additional bit would be needed in the address k for the look-up. Since the entries for $k = 16$ would be 12 and 47, it is easily seen from the P-D diagram in Fig. 6 that there is plenty of "room" to change the values for $k = 15$ from 11 resp. 44 to the values 12 resp. 47, and just use the selection constants for $k = 15$ instead of the missing ones for $k = 16$. Thus the following table of selection constants is valid for radix 4 division and square root for all $i \geq 1$:

$k = 16\widehat{z}$	8	9	10	11	12	13	14	15, 16
$32\widehat{S}_1$	7	7	8	9	9	10	11	12
$32\widehat{S}_2$	25	28	31	33	36	39	41	47

using $k = 10$ and $2\widehat{\beta r_0} = 2 * \widehat{4x}$ instead of $\widehat{4x}$ for the initial square root digit selection.

V. Conclusions

It has been shown that for SRT division, it is indeed possible to analytically define the truncation parameter t for the shifted remainder, $2^{-t} = \mathrm{ulp}(\widehat{\beta r_i})$, given a value of the divisor truncation parameter u satisfying a certain bound, e.g., the minimal such value of u. Thus the quotient digit selection

function can be defined without the need to extensively check the validity of some chosen parameters.

The analysis was then modified for the SRT square root algorithm, where it is well known that initially some digits may have to be determined by other means, but that it is possible to use the same algorithm to determine digits d_i for $i \geq i'$, with i' suitably chosen. It has been shown that by possibly increasing the values of the truncation parameters determined for division, such minimal values of i' may easily be found from a condition very similar to the one used to determine the division parameters.

A combined divide and square root algorithm has been developed, sharing digit selection intervals. To complete the picture, the particularly important case of minimally redundant radix 4 is analyzed in detail. Along the line of [13], but differently, it is shown that the initial digit of the square root here can be found by the otherwise unchanged shared divide and square root algorithm, by a slightly different initialization of the square root case.

We have here used standard truncation, just discarding digits below a certain position. In some of the previous papers, e.g., [7], [10], [8], also different truncations of the "save" and "carry"-parts in carry-save representations have been analyzed, and similarly for different truncations of the positive and negative parts in borrow-save represented remainders. It is also possible to reduce the truncation error by including a carry-bit from a few extra positions beyond the truncation point, as suggested in [7]. Another (equivalent) possibility is to apply the digit-parallel PN-recoding from [15]. Such extra (user specified) truncation parameters could also be included in the analysis presented here, but will of course complicate it.

REFERENCES

[1] P. Kornerup, "Revisiting SRT Quotient Digit Selection," in *Proc. 16th IEEE Symposium on Computer Arithmetic*. IEEE Computer Society, June 2003, pp. 38–45.

[2] J. Robertson, "A New Class of Digital Division Methods," *IRE Transactions on Electronic Computers*, vol. EC-7, pp. 218–222, 1958, reprinted in [16].

[3] K. Tocher, "Techniques of Multiplication and Division for Automatic Binary Computers," *Quarterly Journal of Mechanics and Applied Mathematics*, vol. 11, pp. 364–384, 1958.

[4] D. Atkins, "Higher-Radix Division Using Estimates of the Divisor and Partial Remainders," *IEEE Transactions on Computers*, vol. C-17, pp. 925–934, 1968, reprinted in [16].

[5] T. Coe and P. Tang, "It takes Six Ones to Reach a Flaw," in *Proc. 12th IEEE Symposium on Computer Arithmetic*, IEEE Computer Society, 1995.

[6] G. Taylor, "Radix 16 SRT Dividers with Overlapped Quotient Selection Stages," in *Proc. 7th IEEE Symposium on Computer Arithmetic*. IEEE Computer Society, 1985, pp. 64–71.

[7] T. Williams and M. Horowitz, "SRT division diagrams and their usage in designing custom integrated circuits for division," Stanford University, Tech. Rep. CSL-TR-87-326, 1986.

[8] N. Burgess and T. Williams, "Choices of Operand Truncation in the SRT Division Algorithm," *IEEE Transactions on Computers*, vol. 44, no. 7, pp. 933–938, July 1995.

[9] M. Ercegovac and T. Lang, *Division and Square Root: Digit-Recurrence Algorithms and Implementations*. Kluwer Academic Publishers, 1994.

[10] S. Oberman and M. Flynn, "Minimizing the Complexity of SRT Tables," *IEEE Transactions on VLSI systems*, vol. 6, no. 1, pp. 141–149, March 1998.

[11] B. Parhami, "Precision Requirements for Quotient Digit Selection in High-Radix Division," in *Proc. 35-th Asilomar Conference on Circuits, Systems and Computers*. IEEE Press, 2001, pp. 1670–1673.

[12] L. Ciminiera and P. Montushi, "Higher Radix Square Rooting," *IEEE Transactions on Computers*, vol. 39, no. 10, pp. 1220–1231, October 1990.

[13] M. Ercegovac and T. Lang, "Radix-4 Square Root Without Initial PLA," *IEEE Transactions on Computers*, vol. C-39, no. 9, pp. 1016–1024, August 1990.

[14] ——, "On-the-Fly Conversion of Redundant into Conventional Representations," *IEEE Transactions on Computers*, vol. C-36, no. 7, pp. 895–897, July 1987, reprinted in [17].

[15] M. Daumas and D. Matula, "Further Reducing the Redundancy of Notation Over a Minimally Redundant Digit Set," *Journal of VLSI Signal Processing*, vol. 33, no. 1/2, pp. 7–18, 2003.

[16] E. E. Swartzlander, Ed., *Computer Arithmetic, Vol I*. Dowden, Hutchinson and Ross, Inc., 1980, reprinted by IEEE Computer Society Press, 1990.

[17] ——, *Computer Arithmetic, Vol II*. IEEE Computer Society Press, 1990.

Peter Kornerup was born in Aarhus, Denmark, 1939. He received the mag.scient. degree in mathematics from Aarhus University, Denmark, in 1967. After a period with the University Computing Center, from 1969 involved in establishing the computer science curriculum at Aarhus University, where he helped found the Computer Science Department in 1971. Through most of the 70's and 80's he served as Chairman of that department. Since 1988 he has been Professor of Computer Science at Odense University, now University of Southern Denmark, where he has also served a period as the Chairman of the department. He spent a leave during 1975/76 with the University of Southwestern Louisiana, Lafayette, LA; four months in 1979 and shorter stays in many years with Southern Methodist University, Dallas, TX; one month with Université de Provence i Marseille in 1996 and two months with Ecole Normale Supérieure de Lyon i 2001. His interests include compiler construction, microprogramming, computer networks and computer architecture, but in particular his research has been in computer arithmetic and number representations, with applications in cryptology and digital signal processing.

Prof. Kornerup has served on the program committees for numerous IEEE, ACM and other meetings, in particular he has been on the Program Committees for the 4th through the 16th IEEE Symposia on Computer Arithmetic, and served as Program Co-Chair for these symposia in 1983, 1991 and 1999. He has been guest editor for a number of journal special issues, and served as an associate editor of the IEEE Transactions on Computers during 1991–95.