Computation and Formal Verification of SRT Quotient and Square Root Digit Selection Tables

David M. Russinoff
Department of Central Verification
Advanced Micro Devices, Inc.
Austin, Texas, U.S.A.
david@russinoff.com

***Abstract*—We present a comprehensive, self-contained, and mechanically verified proof of correctness of a maximally redundant SRT design for floating-point division and square root extraction, supported by verified procedures that (a) test the admissibility of a proposed digit selection table, (b) determine the minimal dimensions of an admissible table for a given arbitrary radix, and (c) generate these tables. For square root extraction, we also provide a verified procedure for generating an initial approximation that meets the accuracy requirement of the algorithm and ensures that the digit selection index derived from successive partial roots remains static throughout the computation. A radix-8 instantiation of these algorithms has been implemented in the floating-point unit of the AMD processor code-named *Steamroller*. To ensure their correctness, all of our results and procedures have been formalized and mechanically checked by the ACL2 prover. We present evidence of the value of this approach by comparing it to that of a more conventional published paper that reports similar results, which are shown to be fatally flawed.**

***Index Terms*—Interactive theorem proving, formal verification, SRT division.**

I. Introduction

THE Sweeney-Robertson-Tocher (SRT) algorithm for division and square root extraction is ubiquitous in contemporary microprocessor design [5], [7], [15] and notoriously prone to implementation error [13]. Nonetheless, most research on this topic has been limited to microarchitectural and performance concerns, ignoring the problem of correctness [6], [8], [11]. Investigation of the algorithm itself has mainly focused on establishing bounds on the dimensions of digit selection tables without providing any explicit procedures for generating these tables or verifying their correctness properties [2], [3], [12]. The rare inquiry that does purport to prove correctness [10] is typically lacking in mathematical rigor and consequently, as we shall demonstrate below, susceptible to error.

The revelation of the 1994 Pentium FDIV bug sparked some interest in the application of interactive theorem proving (ITP) to the verification of SRT designs [4], [9], [14], but this was limited to the special case of radix-4 division (two quotient bits per iteration) and was based on a simple high-level circuit design [16] that failed to account for various features that are common to commercial implementations, such as the redundant represention of partial remainders, which affects the requirements of the digit table by effectively doubling the approximation error. Recently, we described a formal proof of correctness of a real SRT RTL design that has been implemented in an AMD processor [15], but this was also a radix-4 divider and, like earlier efforts, ours did not address the more complicated problem of square root extraction.

The AMD processor code-named *Steamroller*, currently under development, includes a radix-8 (three bits per iteration) SRT floating-point module. This paper is a component of the formal verification of this module: a comprehensive analysis of the SRT algorithm for both division and square root extraction with arbitrary radix 2^ρ. All results and procedures presented here have been formalized in the ACL2 logic [1] and their correctness has been mechanically checked with the ACL2 theorem prover. The proof script, consisting of more than 800 lemmas, is provided as a supplement to this paper (mainly for the purpose of demonstrating its existence), which also includes an appendix containing more readable pseudo-code definitions of the underlying procedures for generating and verifying the required tables.

Since our main concern is the reliability of our results, we have ignored various well known opportunities for optimization in order to simplify the proof. In particular, our analysis is limited to the case of "maximal redundancy," which allows all quotient digit values in the set $\{1-2^\rho, \ldots, 2^\rho - 1\}$.

In Section III, we generalize the results of [15] by defining a criterion for quotient digit selection tables of arbitrarily high radix, which is proved necessary and sufficient to produce correct quotients and remainders. We also present a simple procedure that determines whether there exists a table of size $2^M \times 2^N$ that meets this criterion, for given ρ, M, and N, and another that generates such a table if possible.

One difference between the SRT algorithms for division and square root extraction is that the latter requires an initial approximation of the root to be used as input to the table. That is, the first several iterations must be performed by some other means before the table may be invoked on subsequent iterations. In Section IV, we establish a criterion for a table that may be used for square root extraction after K iterations and show that any table that satisfies this criterion for some K also satisfies the criterion for division. We also define computable procedures that determine the existence of a table of given dimensions that meets this criterion, and another that generates it if possible. In Section V, we derive a formula for computing the entries of a "seed table" through which an approximation of the square root, accurate to $K\rho$ bits, is derived from an approximation of the radicand.

Applying our results to the case $\rho = 3$, we find that the smallest admissible table for division is given by $M = 7$ and $N = 3$, and that the optimal parameters for square root are $M = 7$, $N = 4$, and $K = 2$. Thus, the digit selection table implemented in AMD's Steamroller floating-point unit consists of 128×16 3-bit entries and the square root seed table consists of 48 6-bit entries.

Another complication of the square root algorithm is that the digit selection index is derived from the partial root, which changes on every iteration, instead of a fixed divisor. The consequences of treating the leading bits of the partial root as constant, which tends to simplify the digit selection logic at the expense of increasing the table size, are investigated in [3]. Here we pursue an alternative approach, as described in Section V, involving an adjustment of the initial approximation to ensure that the leading bits of the partial quotient actually remain constant. This requires an initial full-width comparison to determine whether the approximation derived from the seed table is an underestimate of the square root, which adds a cycle of latency but simplifies the table access logic without increasing the table size. This scheme is also implemented in the Steamroller processor.

The primary contribution of this paper is the level of mathematical rigor that it brings to the subject of inquiry, which is a prerequisite for mechanical proof-checking. Published results in this area are typically not amenable to formalization. Instead, like most mathematical claims, these results (excluding those of the ITP efforts cited above) have generally relied on the process of social review along with simulation-based testing for the detection of possible errors. Experience has demonstrated convincingly, however, that more reliable verification methods are required to ensure the correctness of a commercial floating-point design, and, in particular, that of an SRT divider. Rigorous analysis of a high-radix digit selection table involves extensive computation that cannot be carried out by hand. Some of the results that we shall require (e.g., Lemmas 3.2 and 4.4) are most naturally proved by appeal to geometric intuition. Such proofs, however instinctively satisfying, cannot provide an appropriate level of confidence for our purpose.

In spite of the evident value of mechanical proof-checking in this area, there remains a perpetual need to justify this approach. To that end, in Section II we present a compelling illustration of the inadequacy of both informal proof and the standard review process.

II. Pitfalls of Informal Analysis

In a well known paper of 2005, Kornerup [10] aims to provide an analytical solution to the problem of determining the minimal dimensions of a valid SRT digit selection table for a given radix and redundancy factor, citing a published claim [11] that no such solution is possible. Since the motivation for this endeavor was a desire to eliminate the need for "extensive searches to check the validity of a given set of parameters" [10], it is worth noting that a C++ implementation of our procedure `ExistsSrtTable` (see supplementary appendix) performs this check for radices as large as 32 in less than one second. Nonetheless, the problem is of some theoretical interest.

The most obvious difference between [10] and the present paper is that the former exposition omits a number of critical details, especially pertaining to the analysis of the more complicated square root operation and its combination with division, some of which, as seen in the definitions and proofs of Section IV below, involve subtle analysis. A careful examination of the proofs exposes serious deficiencies in the statements of the theorems themselves, which are concealed by these omissions.

The analysis in [10] is based on the following parameters:

- β is the underlying radix of the computation (a power of 2);
- a is the maximum element of the digit set, $\{-a, \ldots, a\}$, which satisfies $\frac{\beta}{2} < a < \beta$;
- $\rho = \frac{a}{\beta-1}$ is the redundancy factor (note that we use the same symbol for a different purpose);
- t is the number of fractional bits of the shifted and truncated partial remainder (corresponding to $M - \rho - 1$ in our notation);
- u is the number of bits of the truncated divisor (corresponding to $N + 1$ in our notation).

In the prelude to the main result of [10] pertaining to division (Theorem 3), it is observed that for given β, a, and u such that

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

the smallest value of t for which there exists a "valid quotient digit selection function" is either t_0 or $t_0 + 1$, where t_0 is the smallest t satisfying

$$2^{-t} \leq \left(\rho - \frac{1}{2}\right) - (a - \rho)2^{-u}.$$

The following is established as a necessary and sufficient condition for the existence of such a function for any given β, a, u, and t: For all d and k satisfying $1 \leq d \leq a$ and $2^{u-1} \leq k < 2^u$,

$$\begin{aligned} &\Delta(d, u, t, k) \\ &= \lfloor 2^{t-u}(d - 1 + \rho)k - 1 \rfloor - \lceil 2^{t-u}(d - \rho)(k + 1) \rceil \\ &\geq 0. \end{aligned}$$

It is also noted that for fixed d, this inequality holds for all k provided that the following condition is satisfied for some k_1: $\Delta(d, u, t, k) = 0$ for $2^u \leq k < k_1$ and $\Delta(d, u, t, k_1) > 0$. The conclusion drawn from these observations (Theorem 3) is that for given β, a, u, and $t = t_0$, if this condition is satisfied in the single case $d = a$ for some k_1, then a valid digit selection function exists. The details of this step of the argument are among those that are omitted from the paper.

The first deficiency of this result is that it fails to deliver the promised "analytical" solution.[1] In fact, the number of values of k for which the above inequality must be tested is unknown and potentially as large as 2^{u-1}. But a more serious complaint against the theorem is that it is false. One counterexample is the maximally redundant case $\beta = 16$, $a = 15$, and $u = 9$. Here we have $t_0 = 2$, $\Delta(15, 9, 2, k) = 0$ for $2^{u-1} = 256 \leq k < 282$, and $\Delta(15, 9, 2, 282) = 1$ (and indeed, consequently, $\Delta(15, 9, 2, k) \geq 0$ for all $k < 2^u = 512$). However, $\Delta(14, 9, 2, 265) = -1$, and therefore, the digit selection function is invalid. Similarly, in the maximally redundant radix-32 case with $u = 11$, we again have $t_0 = 2$, and since $\Delta(31, 11, 2, k) = 0$ for $1024 \leq k < 1074$ and $\Delta(31, 11, 2, 1074) = 1$, the stated criterion is claimed to provide a valid selection function for $t = t_0$, but since $\Delta(30, 11, 2, 1041) = \Delta(28, 11, 2, 1042) = -1$, it does not. The other central result of the paper (Theorem 4), pertaining to square root extraction, is similarly flawed. It might also be noted that the results produced in the earlier paper [11], upon which these results were explicitly intended to improve, may be shown to be correct.

Thus, the alleged theorems of [10] falsely guarantee the existence of valid quotient digit selection functions for certain sets of parameters. This circumstance, which has apparently gone unnoticed since the publication of the paper in 2005, will come as a surprise to those who believe that informal quasi-mathematical argument, when presented by a distinguished scientist and ratified by a process of expert review, can ensure floating-point design correctness as reliably as formal machine-checked proof. Moreover, any radix-16 or -32 SRT hardware divider based on these results is likely to have a bug very similar to that of the original Pentium FDIV instruction.

III. SRT Division and Quotient Digit Selection

Let x and d be rational numbers, pre-scaled so that $1 \leq d < 2$ and $|x| < d$. Our objective is to compute a sequence of approximations that converges to the quotient $\frac{x}{d}$. This is achieved by an iterative process that generates a sequence of *partial remainders*, $p_0 = x, p_1, \ldots, p_n$, and *partial quotients*, $q_0 = 0, q_1 \ldots, q_n$. On each iteration, the current partial remainder p_{k-1} is shifted by ρ bits, where 2^ρ is the underlying radix of the computation, and a multiple $m_k d$ of the divisor is subtracted to form the next partial remainder, while the *quotient digit* m_k contributes to the partial quotient:

Lemma 3.1: Given an integer ρ and rational numbers d and x, let $p_0 = x$, $q_0 = 0$, and for $k > 0$,

$$p_k = 2^\rho p_{k-1} - m_k d$$

and

$$q_k = q_{k-1} + 2^{-k\rho} m_k,$$

where m_k is an integer. Then for all $k \geq 0$,

$$p_k = 2^{k\rho}(x - q_k d).$$

Thus, if $d > 0$ and $-d \leq p_k < d$, then

$$-2^{-k\rho} \leq \frac{x}{d} - q_k < 2^{-k\rho}.$$

Proof: The claim is trivial for $k = 0$, and for $k > 0$,

$$\begin{aligned} 2^{k\rho}(x - q_k d) &= 2^{k\rho}(x - (q_{k-1} + 2^{-k\rho} m_k)d) \\ &= 2^{k\rho}(x - q_{k-1} d) - m_k d \\ &= 2^\rho p_{k-1} - m_k d \\ &= p_k. \end{aligned}$$

■

The quotient digit m_k is selected from the range $1 - 2^\rho \leq m_k \leq 2^\rho - 1$ and is required to preserve the invariant $-d \leq p_k < d$. Thus, our objective may be formulated as follows:

> Given a positive integer ρ and rationals d and p with $1 \leq d < 2$ and $-d \leq p < d$, find an integer m such that $|m| < 2^\rho$ and $-d \leq 2^\rho p - md < d$.

The crux of the SRT algorithm is that the value of m is read from a fixed table, using indices derived from truncated approximations of p and d. Let M and N denote the widths of the indices corresponding to p and d, respectively. We have 2^N approximations of d, occurring at equal sub-intervals (of length 2^{-N}) of the interval $1 \leq d < 2$, and 2^M approximations of p occurring at equal sub-intervals (of length 2^{2-M}) of the interval $-2 \leq p < 2$.

As illustrated in Figure 1 for the case $\rho = 2$, $M = 5$, and $N = 2$, the sub-intervals of $1 \leq d < 2$ are numbered from left to right. For given N, and for $j = 0, \ldots, 2^N - 1$, we shall denote the lower endpoint of sub-interval j as δ_j. Thus, j represents the fractional part of δ_j, i.e.,

$$\delta_j = 1 + 2^{-N} j.$$

The sub-intervals of $-2 < p < 2$ are numbered so that each i is the M-bit two's complement representation of the lower endpoint π_i of sub-interval i. Thus, for $i = 0, \ldots, 2^M - 1$,

$$\pi_i = \begin{cases} 2^{2-M} i & \text{if } i < 2^{M-1} \\ 2^{2-M} i - 4 & \text{if } i \geq 2^{M-1}. \end{cases}$$

These partitions produce a $2^M \times 2^N$ matrix of rectangles in the dp-plane, each of width 2^{-N} and height 2^{2-M}. Let R_{ij} denote the rectangle with lower left vertex (δ_j, π_i), and let S_{ij} denote the rectangle with the same lower left vertex and width and twice the height, i.e., for $0 \leq i < 2^M$ and $0 \leq j < 2^N$,

$$R_{ij} = \left\{(d, p) \mid \delta_j \leq d < \delta_j + 2^{-N}, \pi_i \leq p < \pi_i + 2^{2-M}\right\}$$

and

$$S_{ij} = \left\{(d, p) \mid \delta_j \leq d < \delta_j + 2^{-N}, \pi_i \leq p < \pi_i + 2^{3-M}\right\}.$$

The divisor d is approximated by some δ_j and at each iteration, the partial remainder p is approximated by some π_i. The index j is simply extracted from the leading fractional bits of d, and hence the error is bounded by

$$0 \leq d - \delta_j < 2^{-N}.$$

[1] Moreover, it is evident that the claim put forth in [11] is misinterpreted here, as it clearly addresses optimality with respect to implementation metrics (delay and area) rather than merely table dimensions. (This has been confirmed through private communication with the first author.)

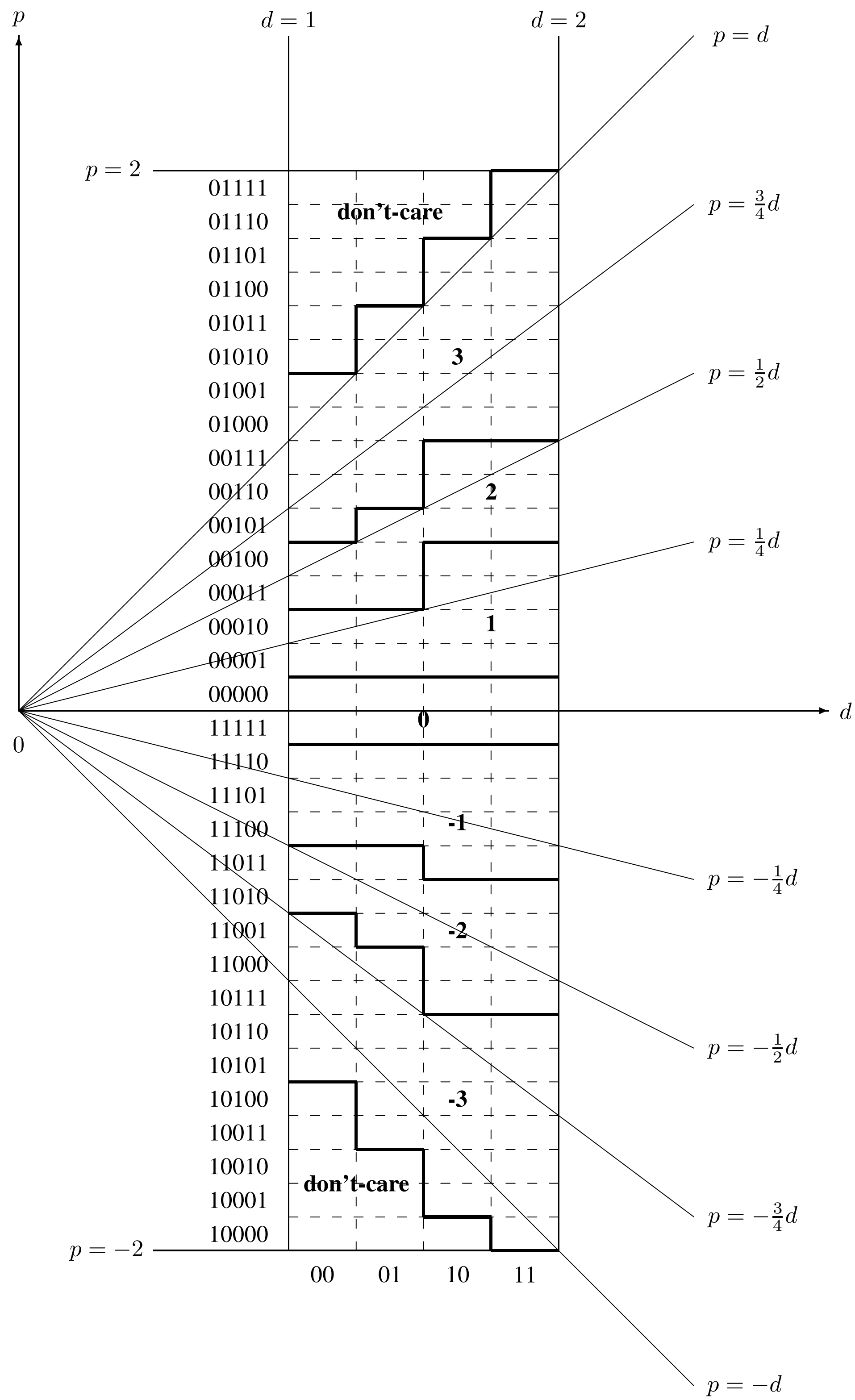


Fig. 1. $\rho = 2$, $M = 5$, $N = 2$

The approximation of p is more subtle because our implementation does not compute p explicitly. As a practical matter, a full carry-propagate addition cannot be executed in the same cycle as the table access, and consequently p is represented in a carry-save form, i.e., as a sum of two terms. These terms are both truncated to M bits and the results are added to produce the approximation of π_i. Thus, the resulting error may approach twice the distance between successive approximations:

$$0 \leq p - \pi_i < 2^{3-M},$$

i.e., (d, p) is confined to the *uncertainty rectangle* S_{ij}.

We shall develop a procedure for generating a table of minimal dimensions that provides a quotient digit $m = \phi(i,j)$ satisfying $-d \leq 2^\rho p - md < d$ for all $(d,p) \in S_{ij}$. Note that this constraint is equivalent to

$$\frac{m-1}{2^\rho} \leq \frac{p}{d} < \frac{m+1}{2^\rho},$$

and therefore the sign of each table entry $\phi(i,j)$ is determined by that of π_i and need not be stored explicitly by an implementation. Thus, such a table consists of at most 2^{M+N} ρ-bit entries.

The following definition presents a formulation of the table requirements that allows straightforward computational verification:

Definition 3.1: Let ρ, M, and N be positive integers and let ϕ be an integer-valued function of two integer variables. Then ϕ is an admissible radix-2^ρ $M \times N$ SRT division table if for all i and j, if $0 \leq i < 2^M$, $0 \leq j < 2^N$, and

$$-\delta_j - 2^{-N} - 2^{3-M} < \pi_i < \delta_j + 2^{-N},$$

then

$$\max(1 - 2^\rho, L_{ij}) \leq \phi(i,j) \leq \min(2^\rho - 1, U_{ij}),$$

where

$$L_{ij} = \begin{cases} \min\left(2^\rho - 1, \lceil \frac{2^\rho(\pi_i + 2^{3-M})}{\delta_j} \rceil - 1\right) & \text{if } i < 2^{M-1} \\ & \text{or } i = 2^M - 1 \\ \min\left(2^\rho - 1, \lceil \frac{2^\rho(\pi_i + 2^{3-M})}{\delta_j + 2^{-N}} \rceil - 1\right) & \text{otherwise} \end{cases}$$

and

$$U_{ij} = \begin{cases} \max\left(1 - 2^\rho, \lfloor \frac{2^\rho \pi_i}{\delta_j + 2^{-N}} \rfloor + 1\right) & \text{if } i < 2^{M-1} \\ \max\left(1 - 2^\rho, \lfloor \frac{2^\rho \pi_i}{\delta_j} \rfloor + 1\right) & \text{if } i \geq 2^{M-1}. \end{cases}$$

Lemma 3.2: Let ρ, M, and N be positive integers and let ϕ be an integer-valued function of two integer variables. Then ϕ is an admissible radix-2^ρ $M \times N$ SRT division table if and only if for all i, j, p, and d, if $0 \leq i < 2^M$, $0 \leq j < 2^N$, $(d,p) \in S_{ij}$, and $-d \leq p < d$, then $m = \phi(i,j)$ satisfies $-2^\rho < m < 2^\rho$ and $-d \leq 2^\rho p - md < d$.

Proof: First note that if S_{ij} lies either entirely above the line $p = d$ or entirely below the line $p = -d$, then no constraint is imposed on $\phi(i,j)$. In the first case, the lower right vertex of S_{ij}, $(\delta_j + 2^{-N}, \pi_i)$, must lie on or above $p = d$, a condition expressed by the inequality

$$\pi_i \geq \delta_j + 2^{-N}.$$

The second case similarly depends on the location of the upper right vertex, $(\delta_j + 2^{-N}, \pi_i + 2^{3-M})$, and is characterized by

$$\pi_i + 2^{3-M} \leq -(\delta_j + 2^{-N}).$$

Thus, the constraint on $\phi(i,j)$ is in force only if neither of these inequalities holds, i.e.,

$$-\delta_j - 2^{-N} - 2^{3-M} < \pi_i < \delta_j + 2^{-N}.$$

Suppose that this condition holds for indices i and j and let $m = \phi(i,j)$. Then all $(d,p) \in S_{ij}$ with $-d \leq p < d$ must satisfy

$$\frac{m-1}{2^\rho} \leq \frac{p}{d} < \frac{m+1}{2^\rho}.$$

Since $p < d$, the upper bound is satisfied trivially if $m = 2^\rho - 1$. Therefore, a necessary and sufficient condition to ensure that this bound holds generally is that if $m \neq 2^\rho - 1$, then S_{ij} does not intersect the region between the lines $p = d$ and $p = \frac{m+1}{2^\rho}d$, or equivalently, that S_{ij} lies entirely below the latter of the two. The maximum value of the quotient $\frac{p}{d}$ in S_{ij} occurs at either the upper left or the upper right vertex, depending on the sign of their common p-coordinate, $\pi_i + 2^{3-M}$. If $i < 2^{M-1}$ or $i = 2^M - 1$, then $\pi_i + 2^{3-M} > 0$ and the critical vertex is the upper left, $(\delta_j, \pi_i + 2^{3-M})$, so the requirement is

$$\frac{\pi_i + 2^{3-M}}{\delta_j} \leq \frac{m+1}{2^\rho}.$$

If $2^{M-1} \leq i < 2^M - 1$, then $\pi_i \leq 0$ and consideration of the upper right vertex yields

$$\frac{\pi_i + 2^{3-M}}{\delta_j + 2^{-N}} \leq \frac{m+1}{2^\rho}.$$

In all cases, the required upper bound is satisfied if and only if $m \geq L_{ij}$.

Similarly, since $p \geq -d$, the lower bound

$$\frac{p}{d} \geq \frac{m-1}{2^\rho}$$

is satisfied trivially if $m = 1 - 2^\rho$. To guarantee that this bound holds generally, it must be shown that if $m \neq 1 - 2^\rho$, then each point in S_{ij} lies on or above the line $p = \frac{m-1}{2^\rho}d$. The minimum value of $\frac{p}{d}$ in S_{ij} occurs at either the lower left or the lower right vertex, depending on the sign of π_i.

If $\pi_i \geq 0$, then the critical vertex is the lower right, $(\delta_j + 2^{-N}, \pi_i)$ and the requirement is

$$\frac{\pi_i}{\delta_j + 2^{-N}} \geq \frac{m-1}{2^\rho}.$$

If $\pi_i < 0$, then consideration of the lower left vertex yields

$$\frac{\pi_i}{\delta_j} \geq \frac{m-1}{2^\rho}.$$

In all cases, the required lower bound is satisfied if and only if $m \leq U_{ij}$. ∎

The following is an immediate consequence of Lemmas 3.1 and 3.2:

Theorem 1: Let ρ, M, and N be positive integers and let ϕ be an admissible radix-2^ρ $M \times N$ SRT division table. Let

x and d be rational numbers such that $1 \leq d < 2$ and $|x| < d$. Let $p_0 = x$, $q_0 = 0$, and for $k > 0$,

$$p_k = 2^\rho p_{k-1} - m_k d$$

and

$$q_k = q_{k-1} + 2^{-k\rho} m_k,$$

where m_k is an integer. Assume that for all $k > 0$, if $|p_{k-1}| < 2$, then $m_k = \phi(i,j)$, where $(d, p_{k-1}) \in S_{ij}$. Then for all $k \geq 0$, $|p_k| < 2$ and

$$2^{-k\rho} \leq \frac{x}{d} - q_k < 2^{-k\rho}.$$

Definition 3.1 provides simple procedures that (a) determine the existence of an admissible SRT table for given radix and dimensions and (b) construct one if possible:

Lemma 3.3: Let ρ, M, and N be positive integers. There exists an admissible radix-2^ρ $M \times N$ SRT division table if and only if for all i and j with $0 \leq i < 2^M$ and $0 \leq j < 2^N$, if

$$-\delta_j - 2^{-N} - 2^{3-M} < \pi_i < \delta_j + 2^{-N},$$

then $L_{ij} \leq U_{ij}$. In this case, one such table is defined by

$$\phi(i,j) = \max(1 - 2^\rho, L_{ij}).$$

These procedures are implemented by the functions `ExistsDivTable` and `SRTTableEntry`, as displayed in the appendix. By direct computation of the former, it is readily shown that for $\rho = 2$, the smallest admissible division table has dimensions $M = 5$ and $N = 2$, and that for $\rho = 3$, the smallest table has $M = 7$ and $N = 3$. These two tables, as generated by `SRTTableEntry`, are displayed in Figures 1 and 2, in which each value $\phi(i,j)$ is indicated by a label associated with R_{ij}.

For each of these tables, the following conditions may be verified by inspection of the graph for each entry $\phi(i,j) = m$:

(1) If S_{ij} intersects the region $-d \leq p < d$, then m is defined and $-2^\rho < m < 2^\rho$.
(2) If m is defined and $m \neq 2^\rho - 1$, then each point in S_{ij} lies below the line $p = \frac{m+1}{2^\rho} d$.
(3) If m is defined and $m \neq 1 - 2^\rho$, then each point in S_{ij} lies on or above the line $p = \frac{m-1}{2^\rho} d$.

As argued in the proof of Theorem 1, it follows that ϕ is an admissible radix 2^ρ division SRT table.

Note that in some cases, there is a choice between two acceptable values of m. If S_{ij} lies within the region bounded by $p = \frac{m}{2^\rho} d$ and $p = \frac{m+1}{2^\rho} d$, where $-2^\rho < m < 2^\rho$, then the required inequalities are satisfied by both m and $m+1$. For example, in the radix-4 table of Figure 1, although we have assigned 2 as the value of $\phi(00100, 10)$, since $S_{00100,10}$ lies between $p = \frac{1}{2} d$ and $p = \frac{1}{4} d$, we could have chosen 1 instead.

It is clear that a necessary and sufficient condition for the existence of an admissible $M \times N$ radix-2^ρ table is that each S_{ij} straddles at most one of the lines $p = \frac{m}{2^\rho} d$, $m = 1 - 2^\rho, \ldots, 2^\rho - 1$. For example, if we attempt to construct a 6×3 radix-8 table, thereby doubling the height of the rectangular elements shown in Figure 2, we find that the uncertainty rectangle $S_{001101,000}$ intersects both $p = \frac{3}{4} d$ and $p = \frac{7}{8} d$, requiring that $m = \phi(001101, 000)$ satisfy both $m \leq 6$ and $m \geq 7$. In fact, these indices are identified by executing the function `ExistsDivTable`.

The admissibility of a division table may be checked visually by examining the bold "staircases" that bound the regions of constant value. Suppose that R_{ij} and $R_{(i-1)j}$ are separated by a segment of such a staircase, i.e., $\phi(i,j) = m$ and $\phi(i-1,j) = m-1$. Since R_{ij} is contained in both S_{ij} and $S_{(i-1)j}$, it must lie above the line $p = \frac{m-1}{2^\rho} d$ and below $p = \frac{m}{2^\rho} d$. That is, a staircase that separates the regions on which $\phi = m$ and $\phi = m-1$ must lie entirely above $p = \frac{m-1}{2^\rho} d$, and when shifted up through one sub-interval, it must still lie below $p = \frac{m}{2^\rho} d$.

IV. SRT Square Root Extraction and Digit Selection

Given a rational number x in the range $\frac{1}{4} < x < 1$, our objective is to construct a sequence of *partial roots*, $q_0 = 0, q_1, \ldots$, that converge to $\sqrt{x}$. For $k > 0$,

$$q_k = q_{k-1} + 2^{-k\rho} m_k,$$

where 2^ρ is the underlying radix and the *root digit* m_k is again an integer in the interval $-2^\rho < m_k < 2^\rho$, selected to maintain a bound on the *partial remainders*, which may be defined as

$$p_k = 2^{k\rho}(x - q_k^2),$$

or alternatively by the recurrence formula

$$p_k = 2^\rho p_{k-1} - m_k(2q_{k-1} + 2^{-k\rho} m_k).$$

The equivalence of these two expressions is established by the following:

Lemma 4.1: Let ρ be an integer and let x be a rational number. Let $q_0 = 0$, $p_0 = x$, and for $k > 0$,

$$q_k = q_{k-1} + 2^{-k\rho} m_k$$

and

$$p_k = 2^\rho p_{k-1} - m_k(2q_{k-1} + 2^{-k\rho} m_k),$$

for some integer m_k. Then for $k \geq 0$, $p_k = 2^{k\rho}(x - q_k^2)$.

Proof: The claim is trivial for $k = 0$, and for $k > 0$,

$$\begin{aligned}
2^{k\rho}(x - q_k^2) &= 2^{k\rho}(x - (q_{k-1} + 2^{-k\rho} m_k)^2) \\
&= 2^{k\rho}(x - (q_{k-1}^2 + 2^{1-k\rho} q_{k-1} m_k + 2^{-2k\rho} m_k^2)) \\
&= 2^{k\rho}(x - q_{k-1}^2) - m_k(2q_{k-1} + 2^{-k\rho} m_k) \\
&= 2^\rho p_{k-1} - m_k(2q_{k-1} + 2^{-k\rho} m_k) \\
&= p_k.
\end{aligned}$$

■

Our objective is to select root digits that preserve the invariants $\frac{1}{2} \leq q_k < 1$ and $2^{-k\rho} \leq \sqrt{x} - q_k < 2^{-k\rho}$. We derive two equivalent formulations of the latter inequality:

Lemma 4.2: Let ρ be an integer and let x be a positive rational number. Let $q_0 = 0$, $p_0 = x$, and for $k > 0$,

$$q_k = q_{k-1} + 2^{-k\rho} m_k$$

and

$$p_k = 2^\rho p_{k-1} - m_k(2q_{k-1} + 2^{-k\rho} m_k),$$

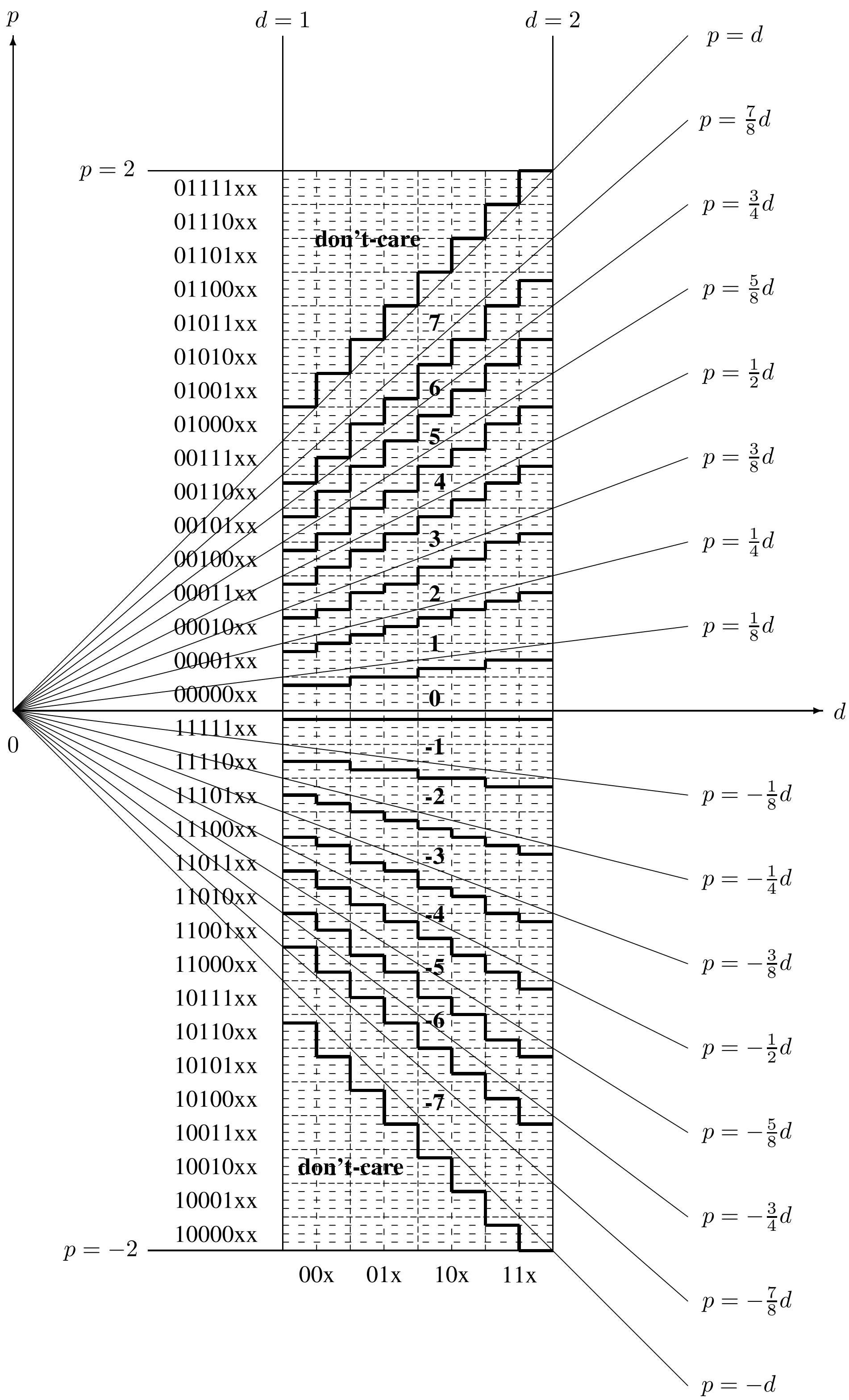


Fig. 2. $\rho = 3$, $M = 7$, $N = 3$

for some integer m_k. Then for $k > 0$, if $q_k \geq \frac{1}{2}$, then the following are equivalent:

(a) $q_k - 2^{-k\rho} \leq \sqrt{x} < q_k + 2^{-k\rho}$;
(b) $-2q_k \leq p_k - 2^{-k\rho} < 2q_k$;
(c) $\frac{m_k-1}{2^\rho}\left(2q_{k-1} + (m_k - 1)2^{-k\rho}\right) \leq p_{k-1} < \frac{m_k+1}{2^\rho}\left(2q_{k-1} + (m_k + 1)2^{-k\rho}\right)$.

Proof: The equivalence of (a) and (b) follows from Lemma 4.1: since $q_k \geq 2^{-k\rho}$,

$$\begin{aligned} & q_k - 2^{-k\rho} \leq \sqrt{x} < q_k + 2^{-k\rho} \\ \Leftrightarrow \quad & (q_k - 2^{-k\rho})^2 \leq x < (q_k + 2^{-k\rho})^2 \\ \Leftrightarrow \quad & q_k^2 - 2^{1-k\rho}q_k + 2^{-2k\rho} \leq x < q_k^2 + 2^{1-k\rho}q_k + 2^{-2k\rho} \\ \Leftrightarrow \quad & -2q_k + 2^{-k\rho} \leq 2^{k\rho}(x - q_k^2) < 2q_k + 2^{-k\rho} \\ \Leftrightarrow \quad & -2q_k + 2^{-k\rho} \leq p_k < 2q_k + 2^{-k\rho}. \end{aligned}$$

To show that (b) is equivalent to (c), note that since

$$\begin{aligned} 2q_k + 2^{-k\rho} &= 2(q_{k-1} + 2^{-k\rho}m_k) + 2^{-k\rho} \\ &= 2q_{k-1} + (2m_k + 1)2^{-k\rho} \\ &= (m_k + 1)\left(2q_{k-1} + (m_k + 1)2^{-k\rho}\right) \\ &\quad -m_k(2q_{k-1} + 2^{-k\rho}m_k), \end{aligned}$$

the upper bound $p_k < 2q_k + 2^{-k\rho}$ is equivalent to

$$\begin{aligned} & 2^\rho p_{k-1} - m_k(2q_{k-1} + 2^{-k\rho}m_k) \\ & \quad < \quad (m_k + 1)\left(2q_{k-1} + (m_k + 1)2^{-k\rho}\right) \\ & \qquad\quad -m_k(2q_{k-1} + 2^{-k\rho}m_k) \end{aligned}$$

or

$$p_{k-1} < \frac{m_k + 1}{2^\rho}\left(2q_{k-1} + (m_k + 1)2^{-k\rho}\right).$$

Similarly, the lower bound $p_k \geq -2q_k + 2^{-k\rho}$ may be replaced by

$$p_{k-1} \geq \frac{m_k - 1}{2^\rho}\left(2q_{k-1} + (m_k - 1)2^{-k\rho}\right).$$

■

We shall once again pursue a table-based approach to the selection of m_k. As suggested by the similarity between the partial remainder recurrence formulas for division and square root, and between the bounds $-d \leq p_k < d$ and Condition (b) of Lemma 4.3, we shall find that in various cases of interest, the same table may be used for both, with the variable $2q_{k-1}$ used for the table index in the square root computation in place of the constant d. This imposes a bound, however, on $m_k^2 2^{-k\rho}$, the term that distinguishes the two formulas. Consequently, the table is used to derive m_k for $k > K$, for some K, after the first K iterations are performed by some other method.

The following lemma guarantees that if $\frac{1}{2} \leq q_K < 1$, then the same bounds are satisfied by all subsequent q_k and that for all $k \geq K$, $|p_k| < 2$:

Lemma 4.3: Let ρ be a positive integer and let x be a rational number, $\frac{1}{4} < x < 1$. Let $q_0 = 0$, $p_0 = x$, and for $k > 0$,

$$q_k = q_{k-1} + 2^{-k\rho}m_k$$

and

$$p_k = 2^\rho p_{k-1} - m_k(2q_{k-1} + 2^{-k\rho}m_k),$$

for some integer m_k. Assume that $\frac{1}{2} \leq q_{k-1} < 1$ for some $k > 1$.

(a) If $q_{k-1} - 2^{(1-k)\rho} \leq \sqrt{x} < q_{k-1} + 2^{(1-k)\rho}$, then $|p_{k-1}| < 2$.
(a) If $q_k - 2^{-k\rho} \leq \sqrt{x} < q_k + 2^{-k\rho}$ and $|m_k| < 2^\rho$, then $\frac{1}{2} \leq q_k < 1$.

Proof: Since $q_{k-1} = \sum_{i=1}^{k-1} 2^{-i\rho}m_i$ is an integral multiple of $2^{(1-k)\rho}$, $q_{k-1} < 1$ implies $q_{k-1} \leq 1 - 2^{(1-k)\rho}$.

Suppose $q_{k-1} - 2^{(1-k)\rho} \leq \sqrt{x} < q_{k-1} + 2^{(1-k)\rho}$. By Lemma 4.2,

$$p_{k-1} \geq -2q_{k-1} + 2^{(1-k)\rho} > -2$$

and

$$p_{k-1} < 2q_{k-1} + 2^{(1-k)\rho} \leq 2(1 - 2^{(1-k)\rho}) + 2^{(1-k)\rho} < 2.$$

Now suppose $q_k - 2^{-k\rho} \leq \sqrt{x} < q_k + 2^{-k\rho}$ and $|m_k| < 2^\rho$. Then

$$\begin{aligned} q_k &= q_{k-1} + 2^{-k\rho}m_k \leq 1 - 2^{(1-k)\rho} + 2^{-k\rho}m_k \\ &< 1 - 2^{(1-k)\rho} + 2^{-k\rho}2^\rho \\ &= 1. \end{aligned}$$

If $q_{k-1} < \frac{1}{2}$, then $q_k \leq \frac{1}{2} - 2^{-k\rho} < \sqrt{x} - 2^{-k\rho}$, contradicting our assumption. ■

With $2q_{k-1}$ replaced by d in Lemma 4.2 (c), our objective may be formulated as follows:

> Given positive integers ρ and K and rational numbers d and p such that $1 \leq d < 2$ and $|p| < 2$, find an integer m such that $-2^\rho < m < 2^\rho$ and for all $k > K$, if $-d + 2^{(1-k)\rho} \leq p < d + 2^{(1-k)\rho}$, then
> $$\begin{aligned} & \frac{m-1}{2^\rho}\left(d + (m-1)2^{-k\rho}\right) \\ & \leq \quad p \\ & < \quad \frac{m+1}{2^\rho}\left(d + (m+1)2^{-k\rho}\right). \end{aligned}$$

We have the following formulation of the requirements of a square root digit selection table for a given iteration k, analogous to Definition 3.1:

Definition 4.1: Let ρ, M, N, and k be positive integers and let ϕ be an integer-valued function of two integer variables. Then ϕ is an admissible radix-2^ρ $M \times N$ SRT square root table for iteration k if for all i and j, if $0 \leq i < 2^M$, $0 \leq j < 2^N$, $k > K$, and

$$-\delta_j - 2^{-N} - 2^{3-M} + 2^{(1-k)\rho} < \pi_i < \delta_j + 2^{-N} + 2^{(1-k)\rho},$$

then the following conditions hold for $m = \phi(i, j)$:

(a) $-2^\rho < m < 2^\rho$.
(b) If $m \neq 2^\rho - 1$, then

$$\pi_i + 2^{3-M} \leq \begin{cases} \frac{m+1}{2^\rho}\left(\delta_j + (m+1)2^{-k\rho}\right) \\ \quad \text{if } i < 2^{M-1} \text{ or } i = 2^M - 1 \\ \frac{m+1}{2^\rho}\left(\delta_j + 2^{-N} + (m+1)2^{-k\rho}\right) \\ \quad \text{if } 2^{M-1} \leq i < 2^M - 1. \end{cases}$$

(c) If $m \neq 1 - 2^\rho$, then

$$\pi_i \geq \begin{cases} \frac{m-1}{2^\rho}\left(\delta_j + 2^{-N} + (m-1)2^{-k\rho}\right) & \text{if } i < 2^{M-1} \\ \frac{m-1}{2^\rho}\left(\delta_j + (m-1)2^{-k\rho}\right) & \text{if } i \geq 2^{M-1}. \end{cases}$$

Lemma 4.4: Let ρ, M, N, and k be positive integers, $k > 1$, and let ϕ be an integer-valued function of two integer variables. Then ϕ is an admissible radix-2^ρ $M \times N$ SRT square root table for iteration k if and only if or all i, j, p, and d, if $0 \leq i < 2^M$, $0 \leq j < 2^N$, $k > K$, $(d,p) \in S_{ij}$, and $-d \leq p - 2^{(1-k)\rho} < d$, then $m = \phi(i,j)$ satisfies $-2^\rho < m < 2^\rho$ and

$$\begin{aligned}\frac{m-1}{2^\rho}\left(d + (m-1)2^{-k\rho}\right) &\leq p \\ &< \frac{m+1}{2^\rho}\left(d + (m+1)2^{-k\rho}\right).\end{aligned}$$

Proof: Consider the following four lines in the dp-plane:
ℓ_1: $p = d + 2^{(1-k)\rho}$
ℓ_2: $p = -d + 2^{(1-k)\rho}$
ℓ_3: $p = \frac{m+1}{2^\rho}\left(d + (m+1)2^{-k\rho}\right)$
ℓ_4: $p = \frac{m-1}{2^\rho}\left(d + (m-1)2^{-k\rho}\right)$.

For given i and j, the constraints of the lemma hold vacuously if S_{ij} lies either entirely above the line ℓ_1 or entirely below ℓ_2, as determined by the lower right vertex, $(\delta_j + 2^{-N}, \pi_i)$, or the upper right vertex, $(\delta_j + 2^{-N}, \pi_i + 2^{3-M})$, respectively. Thus, the constraints are in force only if

$$-\delta_j - 2^{-N} - 2^{3-M} + 2^{(1-k)\rho} < \pi_i < \delta_j + 2^{-N} + 2^{(1-k)\rho}.$$

We may assume that this condition holds.

We must show that the upper bound

$$p < \frac{m+1}{2^\rho}\left(d + (m+1)2^{-k\rho}\right)$$

is satisfied by every $(d,p) \in S_{ij}$ with $-d + 2^{(1-k)\rho} \leq p < d + 2^{(1-k)\rho}$, i.e., below ℓ_1 and on or above ℓ_2, if and only if Condition (b) of Definition 4.1 holds. Since the bound is satisfied trivially if $m = 2^\rho - 1$, we may assume that $m < 2^\rho - 1$. Of the two lines ℓ_1 and ℓ_3, ℓ_1 has the greater slope and p-intercept and therefore lies above ℓ_3 for $d > 0$. But for $d \geq 1$ and $k \geq 2$,

$$\begin{aligned}\frac{m+1}{2^\rho}\left(d + (m+1)2^{-k\rho}\right) + d &\geq \frac{m+1}{2^\rho} + 1 \\ &> \frac{1-2^\rho}{2^\rho} + 1 \\ &= 2^{-\rho} \\ &\geq 2^{(1-k)\rho},\end{aligned}$$

and hence ℓ_3 lies above ℓ_2 in the region of interest. It follows that the required upper bound holds for all $(d,p) \in S_{ij}$ with $-d + 2^{(1-k)\rho} \leq p < d + 2^{(1-k)\rho}$ if and only if S_{ij} lies entirely below ℓ_3, or equivalently, both upper vertices lie on or below ℓ_3. Suppose first that $i < 2^{M-1}$ or $i = 2^M - 1$, so that $\pi_i + 2^{3-M} > 0$. If the slope $m+1$ is negative, then since

$$\delta_j + (m+1)2^{-k\rho} > 1 - 2^\rho 2^{-k\rho} = 1 - 2^{(1-k)\rho} \geq 0,$$

we have

$$\begin{aligned}\pi_i + 2^{3-M} &> 0 \\ &\geq \frac{m+1}{2^\rho}\left(\delta_j + (m+1)2^{-k\rho}\right) \\ &> \frac{m+1}{2^\rho}\left(\delta_j + 2^{-N} + (m+1)2^{-k\rho}\right)\end{aligned}$$

and both vertices lie above the line. If the slope is non-negative, then the critical vertex is the upper left. In either case, a necessary and sufficient condition is that

$$\pi_i + 2^{3-M} \leq \frac{m+1}{2^\rho}\left(\delta_j + (m+1)2^{-k\rho}\right).$$

On the other hand, if $2^{M-1} \leq i < 2^M - 1$, then $\pi_i + 2^{3-M} \leq 0$. If the slope $m+1$ is positive, then every point $(d,p) \in S_{ij}$ lies below ℓ_3, since

$$\begin{aligned}p &< \pi_i + 2^{3-M} \\ &\leq 0 \\ &\leq \frac{m+1}{2^\rho}\left(\delta_j + (m+1)2^{-k\rho}\right) \\ &\leq \frac{m+1}{2^\rho}\left(d + (m+1)2^{-k\rho}\right),\end{aligned}$$

and if $m + 1 \leq 0$, then the critical vertex is the upper right. Thus, a necessary and sufficient condition is

$$\pi_i + 2^{3-M} \leq \frac{m+1}{2^\rho}\left(\delta_j + 2^{-N} + (m+1)2^{-k\rho}\right).$$

The analysis of the lower bound,

$$p \geq \frac{m-1}{2^\rho}\left(d + (m-1)2^{-k\rho}\right),$$

is similar. Since the bound is satisfied trivially if $m = 1 - 2^\rho$, we may assume $m > 1 - 2^\rho$. For $d \geq 1$, ℓ_4 lies below ℓ_1, since

$$\begin{aligned}&d + 2^{(1-k)\rho} - \frac{m-1}{2^\rho}(d + (m-1)2^{-k\rho}) \\ &= \left(1 - \frac{m-1}{2^\rho}\right)d + 2^{(1-k)\rho}\left(1 - \left(\frac{m-1}{2^\rho}\right)^2\right) \\ &\geq 1 - \frac{m-1}{2^\rho} \\ &> 0,\end{aligned}$$

and ℓ_4 lies above ℓ_2, since

$$\begin{aligned}&\frac{m-1}{2^\rho}(d + (m-1)2^{-k\rho}) - (-d + 2^{(1-k)\rho}) \\ &\geq \left(1 + \frac{m-1}{2^\rho}\right)d - 2^{(1-k)\rho} \\ &\geq 1 + \frac{2-2^\rho}{2^\rho} - 2^{-\rho} \\ &= 2^{-\rho} \\ &> 0.\end{aligned}$$

Consequently, the bound is satisfied for all $(d,p) \in S_{ij}$ with $-d \leq p - 2^{(1-k)\rho} < d$ if and only if each point in S_{ij} lies on or above ℓ_4, as determined by its lower vertices. If $i < 2^{M-1}$, i.e., $\pi_i \geq 0$, then since

$$\delta_j + (m-1)2^{-k\rho} \geq 1 - 2^\rho 2^{-k\rho} = 1 - 2^{(1-k)\rho} \geq 0,$$

if $m - 1 \leq 0$, then for all $(d,p) \in S_{ij}$,

$$\begin{aligned}p &\geq \pi_i \geq 0 \geq \frac{m-1}{2^\rho}\left(\delta_j + (m+1)2^{-k\rho}\right) \\ &\geq \frac{m-1}{2^\rho}\left(d + (m+1)2^{-k\rho}\right),\end{aligned}$$

and if $m-1>0$, then the critical vertex is the lower right. Therefore, the requirement is

$$\pi_i \geq \frac{m-1}{2^\rho}\left(\delta_j + 2^{-N} + (m-1)2^{-k\rho}\right).$$

If $\pi_i < 0$, then a similar argument yields the condition

$$\pi_i \geq \frac{m-1}{2^\rho}\left(\delta_j + (m-1)2^{-k\rho}\right).$$

■

The preceding results of this section may be summarized:

Theorem 2: Let ρ, M, N, and K be positive integers and let ϕ be an admissible radix-2^ρ $M \times N$ SRT square root table for every iteration $k > K$. Let x be a rational number, $\frac{1}{4} < x < 1$. Let $q_0 = 0$, $p_0 = x$, and for $k = 1, \ldots, n$,

$$q_k = q_{k-1} + 2^{-k\rho} m_k$$

and

$$p_k = 2^\rho p_{k-1} - m_k(2q_{k-1} + 2^{-k\rho} m_k),$$

where m_k is an integer. Assume that $\frac{1}{2} \leq q_K < 1$, $q_K - 2^{-K\rho} \leq \sqrt{x} < q_K + 2^{-K\rho}$, and for $k > K$, if $\frac{1}{2} \leq q_{k-1} < 1$ and $|p_{k-1}| < 2$, then $m_k = \phi(i,j)$, where $(2q_{k-1}, p_{k-1}) \in S_{ij}$. Then for all $k \geq K$, $|p_k| < 2$ and $-2^{-k\rho} \leq \sqrt{x} - q_k < 2^{-k\rho}$.

Proof: We shall prove by induction that for $k \geq K$, $\frac{1}{2} \leq q_k < 1$ and $q_k - 2^{-k\rho} \leq \sqrt{x} < q_k + 2^{-k\rho}$. Suppose that these conditions hold for $k-1$. Then $-2q_{k-1} + 2^{(1-k)\rho} \leq p_{k-1} < 2q_{k-1} + 2^{(1-k)\rho}$ by Lemma 4.2, $|p_{k-1}| < 2$ by Lemma 4.3, and consequently, for some i and j, $m_k = \phi(i,j)$ and $(2q_{k-1}, p_{k-1}) \in S_{ij}$. Therefore, by hypothesis, $|m_k| < 2^\rho$ and

$$\begin{aligned} &\frac{m_k - 1}{2^\rho}\left(2q_{k-1} + (m_k - 1)2^{-k\rho}\right) | \\ &\leq \quad p_{k-1} \\ &< \quad \frac{m_k + 1}{2^\rho}\left(2q_{k-1} + (m_k + 1)2^{-k\rho}\right). \end{aligned}$$

By Lemma 4.2, $q_k - 2^{-k\rho} \leq \sqrt{x} < q_k + 2^{-k\rho}$, and by Lemma 4.3, $\frac{1}{2} \leq q_k < 1$. ■

We shall develop a procedure for generating a table for a given radix that meets the requirements of both division and square root, of minimal dimensions and with minimal K. In Section V, we turn to the problem of generating the initial partial quotient and remainder, q_K and p_K.

First we note that any table that meets the requirements for square root extraction may be used for division as well:

Lemma 4.5: If ϕ is an admissible radix-2^ρ $M \times N$ square root table for all iterations $k > K$, then ϕ is an admissible radix-2^ρ $M \times N$ division table.

Proof: Suppose that

$$-\delta_j - 2^{-N} - 2^{3-M} < \pi_i < \delta_j + 2^{-N},$$

where $0 \leq i < 2^M$, $0 \leq j < 2^N$. Then for some K',

$$-\delta_j - 2^{-N} - 2^{3-M} + 2^{(1-k)\rho} < \pi_i < \delta_j + 2^{-N} + 2^{(1-k)\rho}$$

for all $k > K'$. Let $m = \phi(i,j)$. For all $k > \max(K, K')$, Conditions (a), (b), and (c) of Definition 4.1 hold. It follows that if $m \neq 2^\rho - 1$, then

$$\pi_i + 2^{3-M} \leq \begin{cases} \frac{m+1}{2^\rho}\delta_j & \text{if } i < 2^{M-1} \text{ or } i = 2^M - 1 \\ \frac{m+1}{2^\rho}\left(\delta_j + 2^{-N}\right) & \text{if } 2^{M-1} \leq i < 2^M - 1, \end{cases}$$

which implies $m \geq L_{ij}$. Similarly, if $m \neq 1 - 2^\rho$, then

$$\pi_i \geq \begin{cases} \frac{m-1}{2^\rho}\left(\delta_j + 2^{-N}\right) & \text{if } i < 2^{M-1} \\ \frac{m-1}{2^\rho}\delta_j & \text{if } i \geq 2^{M-1}, \end{cases}$$

which implies $m \leq U_{ij}$. ■

While Definition 4.1 provides a procedure to determine whether a square root table is admissible for a given iteration k, we would like a procedure for determining admissibility for all sufficiently large k. This is provided by the following:

Definition 4.2: Let ρ, M, N, and K be positive integers and let ϕ be an integer-valued function of two integer variables. Then ϕ is a K-admissible radix-2^ρ $M \times N$ SRT table if for all i and j, if $0 \leq i < 2^M$, $0 \leq j < 2^N$, and

$$-\delta_j - 2^{-N} - 2^{3-M} < \pi_i < \delta_j + 2^{-N} + 2^{-K\rho},$$

then the following conditions hold for $m = \phi(i,j)$:

(a) $-2^\rho < m < 2^\rho$.
(b) $m \geq L_{ij}$.
(c) If $m \neq 1 - 2^\rho$, then

$$\pi_i \geq \begin{cases} \frac{m-1}{2^\rho}\left(\delta_j + 2^{-N} + (m-1)2^{-(K+1)\rho}\right) \\ \quad \text{if } i < 2^{M-1} \\ \frac{m-1}{2^\rho}\left(\delta_j + (m-1)2^{-(K+1)\rho}\right) \\ \quad \text{if } i \geq 2^{M-1}. \end{cases}$$

A K-admissible table is essentially one that is admissible for every iteration $k > K$:

Lemma 4.6: Let ρ, M, N, and K be positive integers and let ϕ be an integer-valued function of two integer variables.

(a) If ϕ is a K-admissible radix-2^ρ $M \times N$ SRT table, then for all $k > K$, ϕ is an admissible SRT square root table for iteration k.
(b) Let ϕ be an admissible radix-2^ρ $M \times N$ SRT square root table for iteration k for all $k > K$ and let

$$\phi'(i,j) = \begin{cases} 1 - 2^\rho & \text{if } -\delta_j - 2^{-N} - 2^{3-M} < \pi_i \\ & \quad \leq -\delta_j - 2^{-N} - 2^{3-M} + 2^{-K\rho} \\ \phi(i,j) & \text{otherwise.} \end{cases}$$

Then ϕ' is a K-admissible radix-2^ρ $M \times N$ SRT table.

Proof: Suppose that ϕ satisfies Definition 4.2. Let $k > K$ and

$$-\delta_j - 2^{-N} - 2^{3-M} + 2^{(1-k)\rho} < \pi_i < \delta_j + 2^{-N} + 2^{(1-k)\rho}.$$

Then

$$-\delta_j - 2^{-N} - 2^{3-M} < \pi_i < \delta_j + 2^{-N} + 2^{-K\rho},$$

Of the conditions imposed by Definition 4.1, (a) and (c) follow from the corresponding conditions of Definition 4.2. To establish (b), note that if $m \neq 2^\rho - 1$, then we have $m \geq L_{ij}$, where

$$L_{ij} = \begin{cases} \left\lceil \frac{2^\rho(\pi_i + 2^{3-M})}{\delta_j} \right\rceil - 1 & \text{if } i < 2^{M-1} \text{ or } i = 2^M - 1 \\ \left\lceil \frac{2^\rho(\pi_i + 2^{3-M})}{\delta_j + 2^N} \right\rceil - 1 & \text{if } 2^{M-1} \leq i < 2^M - 1, \end{cases}$$

which implies

$$\pi_i + 2^{3-M} \leq \begin{cases} \frac{m+1}{2^\rho}\delta_j & \text{if } i < 2^{M-1} \text{ or } i = 2^M - 1 \\ \frac{m+1}{2^\rho}\left(\delta_j + 2^{-N}\right) & \text{if } 2^{M-1} \leq i < 2^M - 1. \end{cases}$$

Now suppose ϕ is an admissible SRT square root table for every iteration $k > K$ and let ϕ' be defined as above. Let

$$-\delta_j - 2^{-N} - 2^{3-M} < \pi_i < \delta_j + 2^{-N} + 2^{-K\rho},$$

for some i and j and let $m = \phi'(i, j)$. If $\pi_i \le -\delta_j - 2^{-N} - 2^{3-M} + 2^{-K\rho}$, then

$$\begin{aligned} L_{ij} &< \frac{2^\rho(\pi_i + 2^{3-M})}{\delta_j + 2^{-N}} \\ &\le \frac{2^\rho(-\delta_j - 2^{-N} + 2^{-K\rho})}{\delta_j + 2^{-N}} \\ &= \frac{2^{(1-K)\rho}}{\delta_j + 2^{-N}} - 2^\rho \\ &< 1 - 2^\rho \\ &= m \end{aligned}$$

and Definition 4.2 is satisfied. In the remaining case,

$$-\delta_j - 2^{-N} - 2^{3-M} + 2^{-K\rho} < \pi_i < \delta_j + 2^{-N} + 2^{-K\rho}$$

and the three conditions of Definition 4.1 must hold for $k = K + 1$. Since (a) and (c) coincide with the corresponding conditions of Definition 4.2, we need only show that $m \ge L_{ij}$. Since this is clearly true if $m = 2^\rho - 1$, we may assume that $m \le 2^\rho - 2$. Suppose $\pi_i \ge \delta_j$. Then $i < 2^{M-1}$ and it follows from (b) that

$$\begin{aligned} \pi_i + 2^{3-M} &\le \frac{m+1}{2^\rho}\delta_j + \frac{(m+1)^2}{2^\rho}2^{-(K+1)\rho} \\ &< (1 - 2^{-\rho})\delta_j + 2^{-\rho}2^{(1-K)\rho} \\ &= \delta_j + 2^{-\rho}(2^{(1-K)\rho} - \delta_j) \\ &\le \delta_j + 2^{-\rho}(1 - 1) \\ &= \delta_j, \end{aligned}$$

a contradiction. Therefore, we may also assume $\pi_i < \delta_j$. But then for all $k > K$,

$$-\delta_j - 2^{-N} - 2^{3-M} + 2^{(1-k)\rho} < \pi_i < \delta_j + 2^{-N} + 2^{(1-k)\rho},$$

and hence

$$\pi_i + 2^{3-M} \le \begin{cases} \frac{m+1}{2^\rho}\left(\delta_j + (m+1)2^{-k\rho}\right) \\ \quad \text{if } i < 2^{M-1} \text{ or } i = 2^M - 1 \\ \frac{m+1}{2^\rho}\left(\delta_j + 2^{-N} + (m+1)2^{-k\rho}\right) \\ \quad \text{if } 2^{M-1} \le i < 2^M - 1. \end{cases}$$

Consequently,

$$\pi_i + 2^{3-M} \le \begin{cases} \frac{m+1}{2^\rho}\delta_j & \text{if } i < 2^{M-1} \text{ or } i = 2^M - 1 \\ \frac{m+1}{2^\rho}\left(\delta_j + 2^{-N}\right) & \text{if } 2^{M-1} \le i < 2^M - 1, \end{cases}$$

which implies $m \ge L_{ij}$. ■

Thus, for given ρ, M, N, and K, Definition 4.2 may be used to determine whether there exists a table that is admissible all square root iterations $k > K$, and consequently for division as well. If such a table does exist, then it may be generated by the same procedure that was developed for division tables:

Lemma 4.7: Let ρ, M, N, and K be positive integers. There exists a K-admissible radix-2^ρ $M \times N$ SRT table if and only if for all i and j with $0 \le i < 2^M$ and $0 \le j < 2^N$, if

$$-\delta_j - 2^{-N} - 2^{3-M} < \pi_i < \delta_j + 2^{-N} + 2^{-K\rho}$$

and

$$L_{ij} > 1 - 2^\rho$$

(where L_{ij} is defined as in Definition 3.1), then

$$\pi_i \ge \begin{cases} \frac{L_{ij}-1}{2^\rho}\left(\delta_j + 2^{-N} + (L_{ij}-1)2^{-(K+1)\rho}\right) & \text{if } i < 2^{M-1} \\ \frac{L_{ij}-1}{2^\rho}\left(\delta_j + (L_{ij}-1)2^{-(K+1)\rho}\right) & \text{if } i \ge 2^{M-1}. \end{cases}$$

In this case, one such table is defined by

$$\phi(i, j) = \max(1 - 2^\rho, L_{ij}).$$

Proof: Let $0 \le i < 2^M$, $0 \le j < 2^N$, and

$$-\delta_j - 2^{-N} - 2^{3-M} < \pi_i < \delta_j + 2^{-N} + 2^{-K\rho}.$$

Suppose that if $L_{ij} > 1 - 2^\rho$, then the conclusion of the lemma holds. Then clearly, the requirements of Definition 4.2 are satisfied by $m = \max(1 - 2^\rho, L_{ij})$.

Conversely, suppose that some m satisfies Definition 4.2 and that $L_{ij} > 1 - 2^\rho$. Then $L_{ij} \le m < 2^\rho$ and $\pi_i \ge f(m)$, where

$$f(m) = \begin{cases} \frac{m-1}{2^\rho}\left(\delta_j + 2^{-N} + (m-1)2^{-(K+1)\rho}\right) & \text{if } i < 2^{M-1} \\ \frac{m-1}{2^\rho}\left(\delta_j + (m-1)2^{-(K+1)\rho}\right) & \text{if } i \ge 2^{M-1}, \end{cases}$$

and we need only show that $\pi_i \ge f(L_{ij})$. But note that f is an increasing function of m for $m > -2^\rho$, since

$$\begin{aligned} 2^\rho(f(m+1) - f(m)) &\ge \delta_j + (2m-1)2^{-(K+1)\rho} \\ &\ge 1 + (1 - 2^{\rho+1})2^{-(K+1)\rho} \\ &> 1 - 2^{1-K\rho} \\ &\ge 0. \end{aligned}$$

Therefore, $\pi_i \ge f(m) \ge f(L_{ij})$. ■

If $\phi(i, j) = m$ for a K-admissible table ϕ, then as noted in the proof of Lemma 4.4, S_{ij} must lie above the line $p = \frac{m-1}{2^\rho}\left(d + (m-1)2^{-K\rho}\right)$. Consequently, in addition to the criterion given in Section III, the staircase that separates the regions $\phi = m$ and $\phi = m - 1$ must lie above that line.

Consider the 5×2 division table of Figure 1. Since the lower left vertex of $R_{11010,00}$ lies on $p = -\frac{1}{2}d$, this point does not lie above $p = -\frac{1}{2}\left(d - 4^{1-K}\right)$; therefore, this is not a K-admissible square root table for any positive K. Moreover, the same is true of every $5 \times N$ radix-4 table for every N. There does, however, exist a 6×2 2-admissible table. This is confirmed by execution of the function `ExistsDivSqrtTable`, displayed in the appendix, which implements the procedure specified by Lemma 4.7. Such a table may be generated by executing `SRTTableEntry`.

Now consider the 7×3 radix-8 table of Figure 2. Since the lower right vertex of $R_{0011110,001}$ lies on the line $p = \frac{3}{4}d$, this table cannot be used for the square root. As confirmed by `ExistsDivSqrtTable`, however, there exist both 8×3 and 7×4 2-admissible radix-8 tables.

V. Square Root Seed Tables

In order to employ a K-admissible SRT table to compute square roots as described in Theorem 2, we shall require an efficient method of deriving, for a given radicand x, the initial root digits $m_1, \ldots, m_K$ to be used in the iterative computation of p_K and q_K, which must satisfy the constraints of the theorem. Our strategy is to read the $(K\rho)$-bit integer $S = 2^{K\rho}q_K$ from a table using the $(K\rho)$-bit integer $\ell = \lfloor 2^{K\rho}x \rfloor$ as an index. Lemma 5.1 (a) below provides a set of conditions on the table entry S at index ℓ that ensures that q_K meets its requirements.

As noted in Section I, we would like to arrange for the column of the digit selection table that is determined by the partial root q_k to be independent of k for $k \geq K$. Thus, we would like to ensure that the most significant $N+1$ bits of q_k, consisting of the leading 1 and the N-bit table index, coincide with the corresponding bits of q_K. If we assume $K\rho > N+1$, as in the case of interest $K = 2$, $\rho = 3$, $N = 4$, then a sufficient condition is that the leading $K\rho - 1$ bits match, i.e., for all $k > K$,

$$\lfloor 2^{K\rho-1}q_k \rfloor = \lfloor 2^{K\rho-1}q_K \rfloor.$$

Lemma 5.1 (b) provides a formula for deriving an adjusted value S' from the seed table entry S that retains the properties of q_K required by Theorem 2 and, as established by (c), satisfies this additional condition as well. Note that this derivation requires a full-width comparison of $\sqrt{x}$ and q_K, which may be implemented by reading the value of S^2 from a parallel table and comparing it with x during the pre-processing phase.

As a simplifying assumption, we ignore the case $K = \rho = 1$, which is of no practical interest:

Lemma 5.1: Let ρ and K be positive integers with $K\rho > 1$. Let x be rational, $\frac{1}{4} < x < 1$, and $\ell = \lfloor 2^{K\rho}x \rfloor$.

(a) Let S be an integer satisfying

$$2^{K\rho-1} \leq S < 2^{K\rho}$$

and

$$2^{-K\rho}(S-1)^2 \leq \ell < \ell + 1 \leq 2^{-K\rho}(S+1)^2,$$

and let $Q = 2^{-K\rho}S$. Then $\frac{1}{2} \leq Q < 1$ and

$$Q - 2^{-K\rho} \leq \sqrt{x} < Q + 2^{-K\rho}.$$

(b) Let

$$S' = \begin{cases} S & \text{if } S \text{ is odd or } \sqrt{x} > Q \\ S-1 & \text{if } S \text{ is even and } \sqrt{x} < Q \end{cases}$$

and $Q' = 2^{-K\rho}S'$. Then $\frac{1}{2} \leq Q' < 1$ and

$$Q' - 2^{-K\rho} \leq \sqrt{x} < Q' + 2^{-K\rho}.$$

(c) Let $q_K = Q'$ and for all $k > K$, let $q_k = q_{k-1} + 2^{-k\rho}m_k$, where m_k is an integer and $|m_k| < 2^\rho$. Then for all $k \geq K$, if $q_k - 2^{-k\rho} \leq \sqrt{x} < q_k + 2^{-k\rho}$, then

$$\lfloor 2^{K\rho-1}q_k \rfloor = \lfloor 2^{K\rho-1}Q' \rfloor.$$

Proof: (a) The bounds on Q hold trivially. To derive the bounds on $\sqrt{x} - Q$, note that substituting $2^{K\rho}Q$ for S in the second hypothesis yields

$$(Q - 2^{-K\rho})^2 \leq 2^{-K\rho}\ell < 2^{-K\rho}(\ell+1) \leq (Q + 2^{-K\rho})^2.$$

Since $2^{-K\rho}\ell \leq x < 2^{-K\rho}(\ell+1)$, this implies

$$(Q - 2^{-K\rho})^2 \leq x < (Q + 2^{-K\rho})^2,$$

and the claim follows.

(b) We may assume $S' = S - 1$, for otherwise $S' = S$, $Q' = Q$, and the claims follow immediately. Since $Q > \sqrt{x} > \frac{1}{2}$, $S = 2^{K\rho}Q > 2^{K\rho-1}$, and hence $2^{K\rho-1} \leq S' < 2^{K\rho}$, which implies $\frac{1}{2} \leq Q' < 1$. Moreover,

$$\sqrt{x} - 2^{-K\rho} < Q - 2^{-K\rho} = Q' < Q \leq \sqrt{x} + 2^{-K\rho},$$

i.e., $Q' - 2^{-K\rho} \leq \sqrt{x} < Q' + 2^{-K\rho}$.

(c) First note that for $k \geq K$, $q_k = Q' + \sum_{i=K+1}^{k} 2^{-i\rho}m_i$, where $|m_i| \leq 2^\rho - 1$, and hence

$$\begin{aligned} |q_k - Q'| &= \left| \sum_{i=K+1}^{k} 2^{-i\rho}m_i \right| \\ &< 2^{-(K+1)\rho}(2^\rho - 1)\sum_{i=0}^{\infty} 2^{-i\rho} \\ &= 2^{-K\rho}. \end{aligned}$$

Since

$$\lfloor 2^{K\rho}q_k \rfloor \leq 2^{K\rho}q_k < 2^{K\rho}(Q' + 2^{-K\rho}) = S' + 1,$$

we have $\lfloor 2^{K\rho}q_k \rfloor \leq S' = 2^{K\rho}Q'$ and

$$\lfloor 2^{K\rho-1}q_k \rfloor = \left\lfloor \frac{\lfloor 2^{K\rho}q_k \rfloor}{2} \right\rfloor \leq \left\lfloor \frac{2^{K\rho}Q'}{2} \right\rfloor = \lfloor 2^{K\rho-1}Q' \rfloor.$$

For the reverse inequality, we may assume $q_k < Q'$. We may also assume $\sqrt{x} < Q$; otherwise, $Q' = Q$ and since q_k and Q' are both integral multiples of $2^{-k\rho}$,

$$q_k \leq Q' - 2^{-k\rho} = Q - 2^{-k\rho} < \sqrt{x} - 2^{-k\rho},$$

contradicting $\sqrt{x} < q_k + 2^{-k\rho}$. It follows that S' is odd. Therefore, since $q_k > Q' - 2^{-K\rho}$,

$$2^{K\rho-1}q_k > 2^{K\rho-1}Q' - \frac{1}{2} = \frac{S'}{2} - \frac{1}{2} = \left\lfloor \frac{S'}{2} \right\rfloor = \lfloor 2^{K\rho-1}Q' \rfloor,$$

which implies $\lfloor 2^{K\rho-1}q_k \rfloor \geq \lfloor 2^{K\rho-1}Q' \rfloor$. ∎

The next lemma establishes the existence of a compliant seed table and gives a formula for computing its entries, implemented by the function `Seed`, specified in the appendix:

Lemma 5.2: Let ρ and K be positive integers and let $\psi(\ell) = \left\lceil \sqrt{2^{K\rho}(\ell+1)} \right\rceil - 1$, where $2^{K\rho-2} \leq \ell < 2^{K\rho}$. Then

$$2^{K\rho-1} \leq \psi(\ell) < 2^{K\rho}$$

and

$$2^{-K\rho}(\psi(\ell) - 1)^2 \leq \ell < \ell + 1 \leq 2^{-K\rho}(\psi(\ell)+1)^2.$$

Proof: Under the assumption that $\psi(\ell)$ is an integer, its definition is equivalent to

$$\sqrt{2^{K\rho}(\ell+1)} - 1 \leq \psi(\ell) < \sqrt{2^{K\rho}(\ell+1)}.$$

The lower bound yields

$$2^{K\rho}(\ell+1) \leq (\psi(\ell)+1)^2$$

and

$$\begin{aligned}\psi(\ell) &\geq \sqrt{2^{K\rho}(\ell+1)} - 1 \\ &\geq \sqrt{2^{K\rho}(2^{K\rho-2}+1)} - 1 \\ &> \sqrt{2^{K\rho}2^{K\rho-2}} - 1 \\ &= 2^{K\rho-1} - 1,\end{aligned}$$

which implies $\psi(\ell) \geq 2^{K\rho-1}$. From the upper bound, we have

$$\psi(\ell) < \sqrt{2^{K\rho}(\ell+1)} \leq \sqrt{2^{K\rho}2^{K\rho}} = 2^{K\rho}$$

and, since $4\ell \geq 2^{K\rho}$,

$$\begin{aligned}\sqrt{2^{K\rho}(\ell+1)} - \sqrt{2^{K\rho}\ell} &\leq \sqrt{4\ell(\ell+1)} - \sqrt{4\ell^2} \\ &< \sqrt{4\ell(\ell+1)+1} - \sqrt{4\ell^2} \\ &= 2\ell+1-2\ell \\ &= 1,\end{aligned}$$

which implies

$$\psi(\ell) < \sqrt{2^{K\rho}(\ell+1)} < \sqrt{2^{K\rho}\ell} + 1$$

and hence,

$$(\psi(\ell)-1)^2 < 2^{K\rho}\ell.$$

■

Along with the initial approximation q_K, the corresponding partial remainder p_K is required to initiate the iterative approximation of $\sqrt{x}$. While this may be computed directly as $p_K = 2^{K\rho}(x - q_K^2)$, an iterative computation is likely to be more efficient. For example, if $\rho = 3$ and $K = 2$, a two-cycle iteration using existing hardware is preferable to a computation involving a (6×6)-bit multiplication. In any case, to apply Theorem 4.4, it must be observed that q_K is actually generated by the recurrence formula from the root digits extracted from the table entry:

Lemma 5.3: Let ρ, K, and S be positive integers with $S < 2^{K\rho}$ and let $Q = 2^{-K\rho}S$. Let $q_0 = 0$, and for $k = 1, \ldots, K$,

$$q_k = q_{k-1} + 2^{-k\rho}m_k,$$

where $m_k = S[(K-k+1)\rho - 1 : (K-k)\rho]$. Then $q_K = Q$.

Proof: By induction, for $1 \leq k \leq K$,

$$q_k = 2^{-k\rho}S[K\rho - 1 : (K-k)\rho],$$

for if

$$q_{k-1} = 2^{(1-k)\rho}S[K\rho - 1 : (K-k+1)\rho],$$

then

$$\begin{aligned}q_k &= 2^{(1-k)\rho}S[K\rho-1 : (K-k+1)\rho] + 2^{-k\rho}m_k \\ &= 2^{-k\rho}2^{\rho}S[K\rho-1 : (K-k+1)\rho] \\ &\quad +2^{-k\rho}S[(K-k+1)\rho-1 : (K-k)\rho] \\ &= 2^{-k\rho}S[K\rho-1 : (K-k)\rho].\end{aligned}$$

In particular,

$$q_K = 2^{-K\rho}S[K\rho - 1 : 0] = 2^{-K\rho}S = Q.$$

■

References

[1] ACL2 Web site, `www.cs.utexas.edu/users/moore/acl2`.

[2] Atkins, Daniel E., "Higher-Radix Division Using Estimates of the Divisor and Partial Remainder," *IEEE Transactions on Computers*, Vol. C-17, No. 10, October 1968.

[3] Ciminiera, Luigi and Paolo Montuschi, "Higher Radix Square Rooting," *IEEE Transactions on Computers*, Vol. 39, No. 10, October 1990.

[4] Clarke, Edmund M., Steven M. German, and Xudong Zhou, "Verifying the SRT Division Algorithm Using Theorem Proving Techniques," *Formal Methods in System Design*, 14:1, January 1999.

[5] Coke, Jim et al., "Improvements in the Intel Core 2 Penryn Processor Family Architecture and Microarchitecture," *Intel Technology Journal*, Vol. 12, Issue 3, October 2008.

[6] Fandrianto, Jan, "Algorithm for High Speed Shared Radix 8 Division and Radix 8 Square Root," 9th IEEE Symposium on Computer Arithmetic, 1989.

[7] Gerwig, G., H. Wetter, E.M. Schwarz, J. Haess, C.A. Krygowski, B.M. Fleischer, and M. Kroener, "The IBM eServer z990 floating-point unit," *IBM Journal of Research and Development*, Volume 48, Number 3/4, 2004.

[8] Hobson, Richard F. and Michael W. Fraser, "An Efficient Maximum-Redundancy Radix-8 SRT Division and Square-Root Method," *IEEE Journal of Solid State Circuits*, Vol. 30, No. 1, January 1995.

[9] Kapur, Deepak and M. Subramaniam, "Mechanizing Verification of Arithmetic Circuits: SRT Division," Invited Talk, Proc. FSTTCS-17, Kharagpur, India, Springer LNCS 1346, pp. 103-122, Dec. 1997.

[10] Kornerup, Peter, "Digit Selection for SRT Division and Square Root," *IEEE Transactions on Computers*, Vol. 54, No. 3, March 2005.

[11] Oberman, Stuart and Michael Flynn, "Minimizing the Complexity of SRT Tables", *IEEE Transactions on VLSI Systems*, 6:1, March 1998.

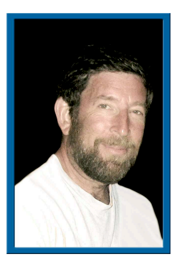
[12] Parhami, Behrooz, "Tight Upper Bounds on the Minimum Precision of the Divisor and the Partial Remainder in High-Radix Division," *IEEE Transactions on Computers*, Vol. 52, No. 11, November 2003.

[13] Pratt, V., "Anatomy of the Pentium Bug," *TAPSOFT '95: Theory and Practice of Software Development*, LNCS 915, Springer-Verlag, May 1995.

[14] Ruess, Harald and Natarajan Shankar, "Modular Verification of SRT Division," *Formal Methods in System Design*, 14:1, January 1999.

[15] Russinoff, David M., "Mechanical Verification of a Commercial SRT Divider," in *Design and Verification of Microprocessor Systems for High-Assurance Applications*, edited by Davis S. Hardin, Springer, 2010. `www.russinoff.com/papers/srt.html`.

[16] Taylor, G.S., "Compatible Hardware for Division and Square Root," *Proceeding of the 5th Symposiom on Computer Arithmetic*, IEEE Computer Society Press, 1981.



David M. Russinoff holds a BS in Mathematics from the Massachusetts Institute of Technology, a PhD in Mathematics from New York University, and an MS in Computer Sciences from the University of Texas at Austin. His research centers on the application of mathematical methods, with an emphasis on interactive theorem proving, to the verification of hardware designs, especially arithmetic circuits. Russinoff spent seventeen years at Advanced Micro Devices, Inc., during which he was responsible for the formal verification of the floating-point units of AMD's line of microprocessors.