Iterations of the R5 Dragon Curve

Kevin Ryde

Draft 11, November 2021

Abstract

Various results on the R5 dragon curve, including coordinates, area, boundary, enclosure sequence, convex hull, centroid, moment of inertia, area trees, and some fractionals. Also some results on the quartet curve and tree which are a similar base.

Contents

Notation

Some formulas have terms going in a repeating pattern of say 4 values according as an index $k \equiv 0$ to $3 \bmod 4$. It's convenient to write them like

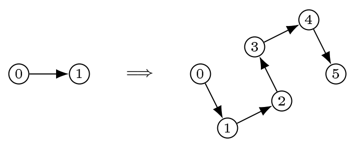
$$[5,\ 8,\ -5,\ 9] \qquad \text{values according as } k \bmod 4$$

meaning 5 when $k \equiv 0 \bmod 4$, or 8 when $k \equiv 1 \bmod 4$, etc. Likewise periodic patterns of other lengths.

Periodic patterns like this can also be expressed by powers of -1 or i (being roots of unity), but except in simple cases that tends to be less clear than the values.

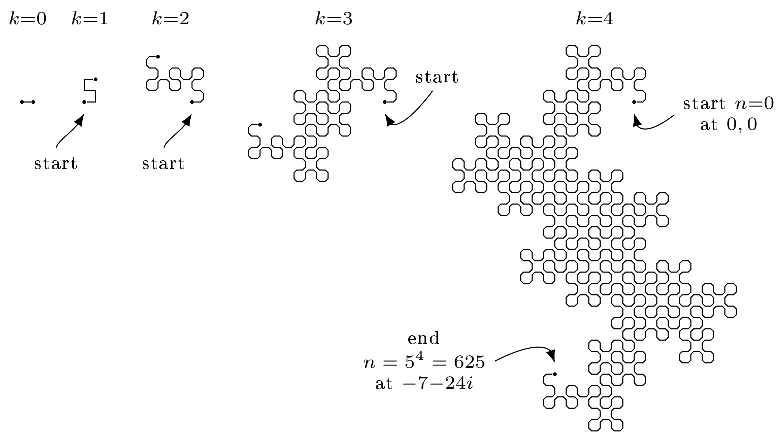
1 R5 Dragon Curve

The R5 dragon curve (as named by Jörg Arndt [1,2]) is defined recursively as a repeated replacement of each line segment by 5 segments in an "S" shape



Knuth [8] gives this as $(DDUU)^*$ in terms of Dekking's folding products. Those products fold by reverse and flip in odd sub-parts. A reverse and flip of $DDUU$ is no change and so the same as a plain segment replacement.

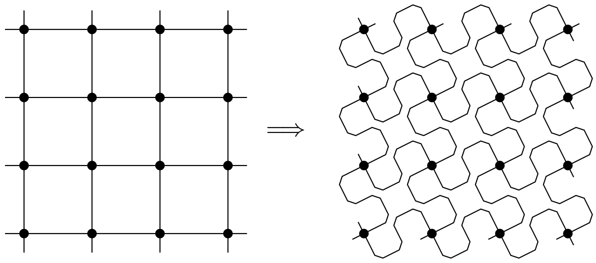
The curve touches at vertices. The following samples have vertices chamfered off to better see the turns and joins.



1.1 Plane Filling

Theorem 1 (Special case of Dekking)**.** *The R5 dragon curve touches at vertices but does not cross itself.*

Proof. Consider an infinite square grid with unit line segments connecting the points. Each line segment expands to an R5 "S" shape as follows. The corners of the new line segments are chamfered off to show how they meet the expansions from other lines.

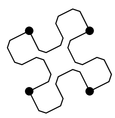


segment
expansions

The expanded grid is the same grid pattern rotated by arctan $\frac{2}{1}$.

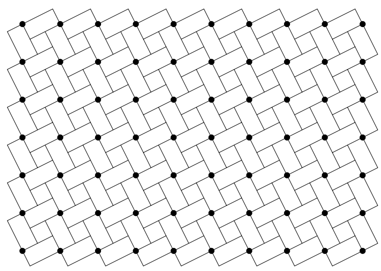
Any subset of the full grid expands to a new bigger set with the same number of crossings. The R5 dragon curve begins with a single line segment which is such a subset with no crossings. $\square$

R5 self-avoidance is a special case of the result by Dekking[6] that a folding sequence is self-avoiding if and only if its "θ-loop" is a simple curve (non-overlapping, non-crossing, possibly touching at vertices). The R5 base pattern is the same forward and reverse so its folding expansion is the same as its plain expansion.



R5 dragon θ-loop
of base patterns

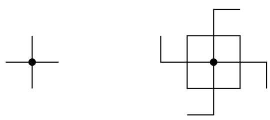
The base pattern is like a 2×1 brick. The grid expansion above puts a brick on each line segment, with an empty square in the middle. This is a classical tiling pattern[11].



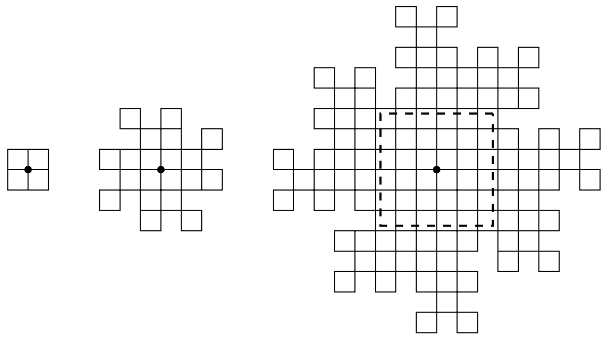
tiling of the plane
by 2×1 bricks
and unit squares

Theorem 2 (Special case of Dekking)**.** *Four copies of the R5 dragon curve arranged at right angles fill the plane.*

Proof. The initial cross expands

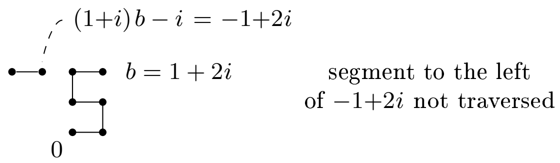


Take the central 2×2 block. With two expansions it grows



The dashed square is a 6×6 block at the origin. So each 2×2 (possibly overlapping) grows to at least 6×6. By repeated expansion they grow to an arbitrarily large square at the origin. □

Four-arm plane filling follows from the carousel theorem of Dekking [6]. A self-avoiding folding sequence is plane-filling if and only if its θ-loop is maximally simple (every grid point touched twice). Four arms of a plane-filling self-avoiding folding sequence perfectly fill if and only if the base pattern does not traverse the segment to the left of $(1+i)z - i$, where z is the endpoint of the base figure. For the R5 dragon, $z = b$ and the segment is to the left of $-1+2i$.



1.2 Turn

Number each point starting n=0 at the origin. Per Arndt [2], the replications give a turn sequence which is 90° turns of

$$turn(n) = \begin{cases} +1 \ \ \text{(left)} & \text{if } \mathit{LowestNon0}(n) = 1 \text{ or } 2 \\ -1 \ \ \text{(right)} & \text{if } \mathit{LowestNon0}(n) = 3 \text{ or } 4 \end{cases} \qquad n \ge 1 \tag{1}$$

$$= {+}{+}{-}{-}{+}{+}{+}{-}{-}{+}{+}{+}{-}{-}{-}{+}{+}{-}{-}{-}{+}{+}{-}{-}\ldots \qquad \text{A337004}$$

$$\mathit{LowestNon0}(n) = \text{base 5 lowest non-0 digit of } n$$

$$= 1, 2, 3, 4, 1, 1, 2, 3, 4, 2, 1, 2, 3, 4, 3, 1, 2, \ldots \qquad n \ge 1 \qquad \text{A277543}$$

Or the next turn after point n,

$$turn(n{+}1) = \begin{cases} +1 & \text{if } \mathit{LowestNon4}(n) = 0 \text{ or } 1 \\ -1 & \text{if } \mathit{LowestNon4}(n) = 2 \text{ or } 3 \end{cases} \qquad n \ge 0$$

$$\mathit{LowestNon4}(n) = \text{base 5 lowest non-4 digit of } n$$

$$= \mathit{LowestNon0}(n{+}1) - 1$$

$$= 0, 1, 2, 3, 0, 0, 1, 2, 3, 1, 0, 1, 2, 3, 2, 0, 1, \ldots \qquad n \ge 0$$

$turn(n)$ and $turn(n{+}1)$ are related simply by $n{+}1$ changing low 4s into low 0s and carry to increment the digit above.

n	$\cdots$	t	$4\cdots4$	base-5 digits, $t = 0$ to 3
$n{+}1$	$\cdots$	$t{+}1$	$0\cdots0$	

Figure 1: *turn*

Predicates for left and right turns are

$$
\begin{aligned}
\mathit{TurnLpred}(n) &= \begin{cases} 1 & \text{if } n \ge 1 \text{ and } \mathit{LowestNon0}(n) = 1 \text{ or } 2 \\ 0 & \text{otherwise} \end{cases} \\
&= 1,1,0,0,1,1,1,0,0,1,1,1,0,0,0,1,1,\ldots \quad n \ge 1 \\
\mathit{TurnRpred}(n) &= \begin{cases} 1 & \text{if } n \ge 1 \text{ and } \mathit{LowestNon0}(n) = 3 \text{ or } 4 \\ 0 & \text{otherwise} \end{cases} \\
&= 0,0,1,1,0,0,0,1,1,0,0,0,1,1,1,0,0,\ldots \quad n \ge 1 \qquad \text{A175337}
\end{aligned}
$$

Generating functions for these sequences follow by considering the base-5 digits of those n which are left or right turn. A left turn is k many low zeros then digit 1 or 2 so $n = (1 \text{ or } 2).5^k + m.5^{k+1}$ for integer m. In a generating function $1/(1 - x^{5^{k+1}})$ is 1 at multiples of 5^{k+1} then multiply x^{5^k} or $x^{2.5^k}$ to add $(1 \text{ or } 2).5^k$. Similarly a right turn is k low zeros then digit 3 or 4 so $n = (3 \text{ or } 4).5^k + m.5^{k+1}$ so multiply $x^{3.5^k}$ or $x^{4.5^k}$.

$$
\begin{aligned}
g\mathit{TurnLpred}(x) &= \sum_{k=0}^{\infty} \frac{x^{5^k} + x^{2.5^k}}{1 - x^{5^{k+1}}} &&= \sum_{k=0}^{\infty} \frac{x^{5^k}(1 + x^{5^k})}{1 - x^{5^{k+1}}} \\
g\mathit{TurnRpred}(x) &= \sum_{k=0}^{\infty} \frac{x^{3.5^k} + x^{4.5^k}}{1 - x^{5^{k+1}}} &&= \sum_{k=0}^{\infty} \frac{x^{3.5^k}(1 + x^{5^k})}{1 - x^{5^{k+1}}}
\end{aligned}
$$

A generating function for *turn* follows similarly, and is also difference

$$
\mathit{turn}(n) = \mathit{TurnLpred}(n) - \mathit{TurnRpred}(n)
$$

At (2), factor $1 - x^{5^k}$ cancels from numerator and denominator, though it's then less clear that the denominator is a replication.

$$
\begin{aligned}
g\mathit{turn}(x) &= \sum_{k=0}^{\infty} \frac{x^{5^k} + x^{2.5^k} - x^{3.5^k} - x^{4.5^k}}{1 - x^{5^{k+1}}} \\
&= \sum_{k=0}^{\infty} \frac{x^{5^k}(1 + x^{5^k})^2}{1 + x^{5^k} + x^{2.5^k} + x^{3.5^k} + x^{4.5^k}} \qquad (2)
\end{aligned}
$$

If a generating function for just an initial part of the sequence is required then stopping the sum at k suffices for $n < 5^{k+1}$ where the next term would begin (a left turn with $k{+}1$ low zeros and digit 1 above).

For computer calculation of the turn sequence on a binary machine, Arndt [1] gives code in C with base 5 digits in 3 bits each 000 through 100 and a loop to increment in that form. This has the attraction of not requiring divisions or moduli to locate the lowest non-0.

A variation can be made using bit values 011 through 111. Doing so allows an increment with the CPU adder then some bit twiddling.

for $u = n$ with base-5 digits as bits 011 through 111,

$$turn(u) = \begin{cases} 1 & \text{if } \mathit{LowOctal4Middle}(u) = 0 \\ -1 & \text{if } \mathit{LowOctal4Middle}(u) = 1 \end{cases}$$

$$\mathit{LowOctal4Middle}(u) = \mathrm{BITAND}\big(u,\ (\mathit{MaskToLowOctal4}(u) + 1)/4\big)$$

$$\mathit{MaskToLowOctal4}(u) = \mathit{MaskToLowBit1}\big(\mathrm{BITAND}(u,\ 100100\ldots100)\big)$$

$$\mathit{MaskToLowBit1}(u) = \mathrm{BITXOR}(u, u{-}1) \qquad \text{A038712}$$

$$\mathit{increment}(u) = u{+}1 + \mathit{LowThrees}(u{+}1) \tag{3}$$

$$\mathit{LowThrees}(u) = \mathrm{BITAND}\big(\lfloor \mathit{MaskToLowOctal4}(u)/8 \rfloor,\ 011011\ldots011\big)$$

MaskToLowBit1 is a bit mask of the lowest 1-bit and all bits below it. *MaskToOctal4* is a bit mask of the lowest octal digit ≥ 4 and all bits below. *LowOctal4Middle* is the middle bit of the lowest octal digit ≥ 4. This determines the turn since base-5 digits 1,2 are $100, 101$ whereas digits 3,4 are $110, 111$.

LowThrees locates low octal 0-digits with a ≥ 4 above them and gives those 0s as 3s. In *increment*, the $+1$ propagates carry through low 111 bits but leaves them as 000 which is not the desired representation. Adding *LowThrees* (or BITOR if preferred) changes them to 011.

The first turn at $n{=}1$ is $u = \ldots 33334$ octal. *increment* requires all high 0 digits to be represented by 3s.

On expanding the curve, 4 turns are inserted in each segment. A segment is between each existing turn, and at start and end of the curve.

Figure 2

	$\boxed{\text{L}}$		$\boxed{\text{L}}$		$\boxed{\text{L}}$		$\boxed{\text{R}}$		$\boxed{\text{R}}$		existing turns
LLRR		LLRR		LLRR		LLRR		LLRR		LLRR	new turns

Arndt [1, figure 1.31-R] gives this in a morphism 0, 1 for *TurnRpred*, and similar in OEIS A175337 with "F" for each segment. Each term gets a new 0,0,1,1 before it.

$$\mathit{TurnRpred} = \quad 0 \to 0,0,1,1,0 \quad 1 \to 0,0,1,1,1 \quad \text{starting from } 0$$

The pairs RR and LL each side of an existing turn become runs either RRR or LLL according as the existing turn is R or L. So run lengths in the turn sequence are an initial 2 then pairs either 2,3 or 3,2 according as $turn = +1$ or -1 respectively. Counting the first run as $m{=}0$, run lengths are

$$\begin{aligned} \mathit{TurnRun}(m) &= \begin{cases} 2 & \text{if } m{=}0 \ \text{(lefts)} \\ \frac{5}{2} + \frac{1}{2}\, turn(\frac{m}{2}) & \text{if } m \text{ even } \geq 2 \ \text{(lefts)} \\ \frac{5}{2} - \frac{1}{2}\, turn(\frac{m+1}{2}) & \text{if } m \text{ odd} \ \text{(rights)} \end{cases} \\ &= \begin{cases} 2 & \text{if } m{=}0 \\ \frac{5}{2} + \frac{1}{2}(-1)^m\, turn(\lceil \frac{m}{2} \rceil) & \text{if } m \geq 1 \end{cases} \\ &= 2,\ 2,3,\ 2,3,\ 3,2,\ 3,2,\ 2,3,\ 2,3,\ 2,3,\ldots \\ turn &= \quad +1,\ +1,\ -1,\ -1,\ +1,\ +1,\ +1, \\ g\mathit{TurnRun}(x) &= -\tfrac{1}{2} + \tfrac{5}{2}\frac{1}{1-x} + \tfrac{1}{2}\Big(1 - \frac{1}{x}\Big)\, gturn(x^2) \end{aligned}$$

For a curve of finite k, the run lengths end with a final 2. This is like the initial 2. By symmetry, the run length sequence for a finite k is equal to its own reversal.

The n which is the start of a run follows from the new and existing turns too. In each LLRR, the left $n \equiv 1 \bmod 5$ is the start of a run unless preceded by an existing L. The right $n \equiv 3 \bmod 5$ is always the start of a run. Expressing this with an index $m \geq 0$,

$$\begin{aligned} \mathit{TurnRunStart}(m) &= 1 + \sum_{j=0}^{m-1} \mathit{TurnRun}(j) \\ &= \tfrac{5}{2}m + \begin{cases} 1 - \mathit{TurnLpred}(\tfrac{5}{2}m) & \text{if } m \text{ even} \\ \tfrac{1}{2} & \text{if } m \text{ odd} \end{cases} \\ &= 1, 3, 5, 8, 10, 13, 16, 18, 21, 23, 25, 28, 30, \ldots \end{aligned} \tag{4}$$

(4) can use $1 - \mathit{TurnLpred} = \mathit{TurnRpred}$ for $m \geq 1$, and 5 is not required since that is simply a low 0 which does not change the result.

$$\mathit{TurnRunStart}(m) = \left\lfloor \tfrac{5}{2}m \right\rfloor + \begin{cases} \mathit{TurnRpred}(\tfrac{1}{2}m) & \text{if } m \text{ even} \geq 2 \\ 1 & \text{otherwise} \end{cases}$$

Theorem 3. *The m'th left or right turn point n is given by the following recurrences, for turns indexed by m starting first turn as m=0,*

$$\mathit{TurnLeft}(m) \tag{5}$$

$$= \begin{cases} 1 & \text{if } m=0 \\ 5^k \ + \mathit{TurnLeft}\big(m - \tfrac{1}{2}(5^k+1)\big) & \text{if } m \leq 5^k,\ m \neq 0 \\ 2.5^k + \mathit{TurnLeft}\big(m - (5^k+1)\big) & \text{if } 5^k < m < \tfrac{1}{2}(3.5^k+1) \\ 3.5^k + \mathit{TurnLeft}\big(m - \tfrac{1}{2}(3.5^k+1)\big) & \text{if } \tfrac{1}{2}(3.5^k+1) \leq m < 2.5^k \\ 4.5^k + \mathit{TurnLeft}\big(m - 2.5^k\big) & \text{if } m \geq 2.5^k \end{cases}$$

where, for $m \geq 1$, have k biggest with $\tfrac{1}{2}(5^k+1) \leq m$

$$= 1, 2, 5, 6, 7, 10, 11, 12, 16, 17, 21, 22, \ldots$$

$$\mathit{TurnRight}(m)$$

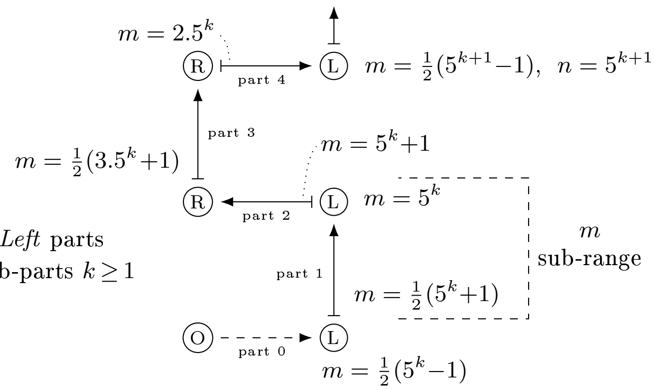
$$= \begin{cases} 5^k \ + \mathit{TurnRight}\big(m - \tfrac{1}{2}(5^k-1)\big) & \text{if } m < 5^k-1 \\ 2.5^k + \mathit{TurnRight}\big(m - (5^k-1)\big) & \text{if } 5^k-1 \leq m < \tfrac{1}{2}(3.5^k-3) \\ 3.5^k & \text{if } m = \tfrac{1}{2}(3.5^k-3) \\ 3.5^k + \mathit{TurnRight}\big(m - \tfrac{1}{2}(3.5^k-1)\big) & \text{if } \tfrac{1}{2}(3.5^k-1) \leq m < 2.5^k-1 \\ 4.5^k & \text{if } m = 2.5^k-1 \\ 4.5^k + \mathit{TurnRight}\big(m - 2.5^k\big) & \text{if } m \geq 2.5^k \end{cases}$$

where k biggest with $\tfrac{1}{2}(5^k-1) \leq m$

$$= 3, 4, 8, 9, 13, 14, 15, 18, 19, 20, 23, 24, \ldots$$

Proof. In an expansion level k, there are 5^k segments and $5^k - 1$ turns between them. Since the curve is symmetric in 180° rotation, there are half lefts and half rights $\frac{1}{2}(5^k - 1)$ each.

The recurrences follow from how many turns of each direction in each sub-part and between sub-parts. Expansion level $k+1$ comprises the following sub-parts for level $k \geq 1$ (so that there is at least one turn within each such k).



TurnLeft parts
$k+1$, sub-parts $k \geq 1$

Part 0 has $\frac{1}{2}(5^k-1)$ left turns so that the L after it is $m = \frac{1}{2}(5^k-1)$ and the first m within part 1 is $m = \frac{1}{2}(5^k+1)$. Taking k as the biggest with $\frac{1}{2}(5^k+1) \leq m$ is then m ranging from the first turn in part 1 to the L at $m = \frac{1}{2}(5^{k+1}+1)$ inclusive.

The turns within part 1, and the L after it, are the m sub-range shown dashed. Each of those is an n with a high 1-digit. So the recurrence is 5^k and reduce to part 0 and the L following it by subtracting $\frac{1}{2}(5^k-1)$.

The recurrence is each of the other parts similarly. There is no L after parts 3 or 4 and when they reduce to part 0 the new m is small enough not to reach the L following part 0.

k=0 occurs only for $m = 1, 2$ (the two left turns in a $k+1 = 1$ curve) and the cases reduce to m=0 which is the special case in (5).

Similarly *TurnRight*,

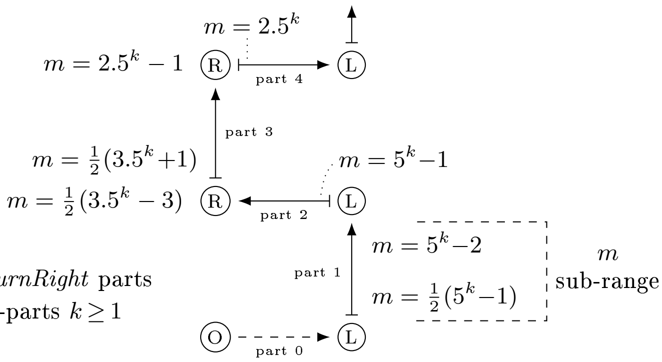


Figure 3: *TurnRight* parts
$k+1$, sub-parts $k \geq 1$

For R, the sub-ranges sought, and reduced to part 0, do not include the following turn, since parts 2 and 3 would want an R after, but part 0 has an L after. Hence the special cases for those m which are the R after parts 2 and 3. $\square$

Theorem 4. *$n = TurnLeft(m)$ can be calculated by the following digit procedure*

$n \leftarrow 2m$
for each base-5 digit position high to low in n
  if digit $= 1$ or 3 then $n \leftarrow n - 1$
  else if digit $= 2$ then
    $n \leftarrow n - 2$
    if now digit $\neq 2$ then result $n + 2$
result $n + 1$

And $n = TurnRight(m)$ can be calculated by the following digit procedure

$n \leftarrow 2m + 2$
for each ternary digit position high to low in n
  if digit $= 1$ or 3 then $n \leftarrow n + 1$
  else if digit $= 2$ then
    $n \leftarrow n + 2$
    if now digit $\neq 2$ then result $n - 1$
result n

The digit tested at each digit position is in the successively modified n, not just the original $2m$ or $2m+2$.

Proof. The effect of the *TurnLeft* procedure is to hold the *TurnLeft* result so far in the high base-5 digits of n, and $2m$ in the low digits.

	high	k	low	
$n =$	result		$2m$	base-5

The effect of $2m$ is to have the cases in (5) distinguished by the base-5 digit at position k. The digit there is the new desired digit of n (the $4.5^k + \ldots$ etc) in the recurrences, with the exception of $m = 5^k$.

$n \leftarrow n{-}1$ or $n \leftarrow n{-}2$ in the procedure are adjustments for the reduced m of the recurrences. These are where the Ls between sub-parts are extra Ls over Rs. In parts 0 or 4 there are no net extra Ls,

Case $m{=}5^k$ is $2m{=}200\ldots00$ in base-5, so its digit seen is 2 but its recurrence case is only 1.5^k. The procedure notices this when subtraction $n \leftarrow n{-}2$ borrows up through all the low 0s so the digit changes from 2 (to 0 if it's the least significant, or to 1 otherwise).

Other $2m$ with a digit 2 do not change by borrow this way since they are at least 2 bigger, so at least 2 to subtract from.

As a remark, the $n \leftarrow n{-}1$ subtractions for digits 1 or 3 never borrow from digit k. Since base 5 is odd, the even $2m$ must have an even number of odd digits. When the high digit is odd 1 or 3, there must be another odd below it and any borrow from decrementing will stop there. This allows a computer calculation to have a subtraction of either -1 or -2 in the same loop, and watch for borrow reducing digit k to indicate the $n{+}2$ early exit.

Corresponding considerations give the *TurnRight* procedure. It has $2m{+}2$ in the low digits of n so that a range low like $m = 5^k{-}1$ becomes $2m{+}2 = 2.5^k$.

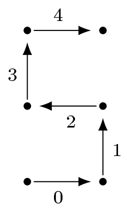
Its carry test is for $m = \frac{1}{2}(3.5^k{-}3)$ which is the R between parts 1 and 2 in figure 3. This is $2m{+}2 = 244\ldots44$ so digit 2, but digit 3 is desired. $n \leftarrow n{+}2$ carries through the 4s to change that 2 (to 4 when least significant, or to 3 otherwise) and indicate the exception.

Other $2m{+}2$ with digit 2 are at least 2 smaller so do not carry this far. And similar to the remark above, digit 1 or 3 increment cases have at least one more odd digit so their carry never reaches digit k either. □

If integers are represented in base-5 then testing for the respective digits is simple. On a binary computer, it may be desirable to convert to a vector of base-5 digits and apply increments or decrements there, either by an explicit loop or by bit twiddling. The *increment* bit twiddling at (3) suits the *TurnRight* procedure. A similar decrement and bit twiddling could suit the *TurnLeft* procedure.

1.3 Direction

The total turn is found from base 5 digits of n. Reckoning the first segment as n=0, each part 1 and 3 are rotation +90°. Each part 2 is a rotation 180°.



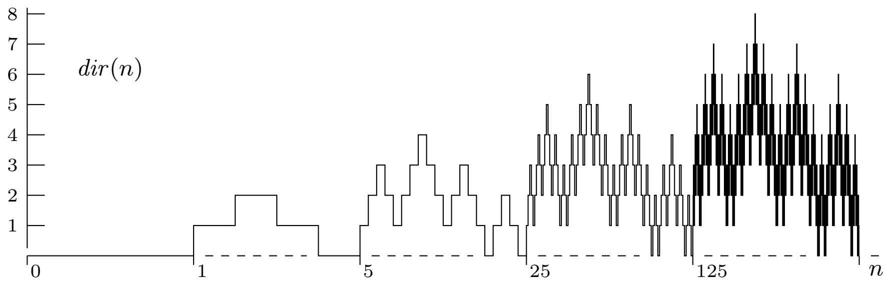
$$
\begin{aligned}
dir(n) &= \sum_{i=0}^{n-1} turn(i) \qquad\qquad \text{direction of segment } n \\
&= \text{count(1-digits)} \;+\; 2 \times \text{count(2-digits)} \;+\; \text{count(3-digits)} \\
&= 0, 1, 2, 1, 0, 1, 2, 3, 2, 1, 2, 3, 4, 3, 2, 1, 2, 3, 2, 1, \ldots
\end{aligned}
$$

$$
\begin{aligned}
gdir(x) &= \frac{1}{1-x}\; gturn(x) \\
&= \sum_{k=0}^{\infty} \frac{x^{5^k}\,(1+x^{5^k})^2}{(1-x)\,(1+x^{5^k}+x^{2.5^k}+x^{3.5^k}+x^{4.5^k})}
\end{aligned} \tag{6}
$$

The generating function is a usual factor $1/(1{-}x)$ for cumulative turns. The direct interpretation of each term is to put 1 where the digit at position k in n is 1 or 3, and to put 2 where digit 2. The numerator at (7) factorizes to (6).

$$
\begin{aligned}
&\frac{x^{5^k}{+}\cdots{+}x^{2.5^k-1} \;+\; 2x^{2.5^k}{+}\cdots{+}2x^{3.5^k-1} \;+\; x^{3.5^k}{+}\cdots{+}x^{4.5^k-1}}{1-x^{5^{k+1}}} \\
&\qquad = \frac{x^{5^k}{-}x^{4.5^k} \;+\; x^{2.5^k}{-}x^{3.5^k}}{(1-x)\,(1-x^{5^{k+1}})}
\end{aligned} \tag{7}
$$

Some of the structure of *dir* can be illustrated in a plot.



Blocks of $n=5^k$ to $5^{k+1}-1$ are scaled to the same width, and linear within those blocks. In successive blocks, the overall shape is preserved, just some additional excursions up added. They are where new low digit $1,2,3$ adds $1,2,1$ respectively.

Successive new highs are at n=0 and then in n digit 2s add 2 each, and single 1 adds 1,

$$\text{new high } dir(n) \text{ at } n = \begin{cases} \frac{1}{2}(3.5^k-1) = 1222\ldots \\ \frac{1}{2}(5.5^k-1) = 2222\ldots \end{cases} \qquad \text{A238366}$$

All 2s is the maximum *dir* in curve k, and is the middle segment.

$$\max_{n=0}^{5^k-1} dir(n) = 2k, \quad \text{at } n = \tfrac{1}{2}(5^k-1) = 222\ldots22 \text{ base-5} \qquad (8)$$
$$= 0, 2, 12, 62, 312, 1562, \ldots \qquad \text{A125831}$$

Each return to $dir=0$ is where n comprises only base-5 digits $0,4$.

$$dir(n)\text{=0 at } n = \text{index } j \text{ in binary, change to digits 0,4 in base-5}$$
$$= 0, 4, 20, 24, 100, 104, 120, 124, \ldots \qquad \text{A097251}$$

v–existing
The number of left and right turns from 1 to n inclusive are

$$TurnsL(n) = \sum_{j=1}^{n} TurnLpred(n)$$
$$= 1, 2, 2, 2, 3, 4, 5, 5, 5, 6, 7, 8, 8, 8, 8, \ldots$$
$$TurnsR(n) = \sum_{j=1}^{n} TurnRpred(n)$$
$$= 0, 0, 1, 2, 2, 2, 2, 3, 4, 4, 4, 4, 5, 6, 7, \ldots$$

$$gTurnsL(x) = \frac{1}{1-x}\, gTurnLpred(x) = \sum_{k=0}^{\infty} \frac{x^{5^k} + x^{2.5^k}}{(1-x)\,(1-x^{5^{k+1}})}$$
$$gTurnsR(x) = \frac{1}{1-x}\, gTurnRpred(x) = \sum_{k=0}^{\infty} \frac{x^{3.5^k} + x^{4.5^k}}{(1-x)\,(1-x^{5^{k+1}})}$$

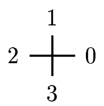
Each turn is left or right so the sum lefts plus rights is simply n. The difference lefts minus rights is net dir. In the generating functions, this difference is $gdir$ form (7).

$$\mathit{TurnsL}(n) + \mathit{TurnsR}(n) = n \tag{9}$$
$$\mathit{TurnsL}(n) - \mathit{TurnsR}(n) = \mathit{dir}(n) \tag{10}$$

Sum and difference of (9),(10) give

$$\mathit{TurnsL}(n) = \tfrac{1}{2}\big(n + \mathit{dir}(n)\big)$$
$$\mathit{TurnsR}(n) = \tfrac{1}{2}\big(n - \mathit{dir}(n)\big)$$

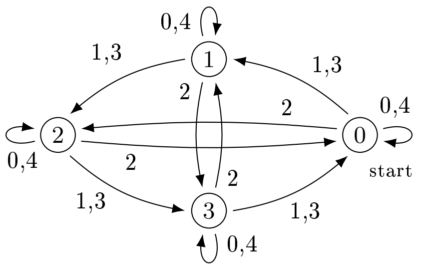
$dir(n)$ mod 4 is a net segment direction East, North, West, or South.



dir mod 4
$= 0, 1, 2, 1, 0, 1, 2, 3, 2, 1, 2, 3, 0, 3, 2, 1, \ldots$

$$\begin{aligned} \mathit{dir}(n) &\equiv 0 \text{ at } n = 0,\ \ 4, 12, 20, 24, 32, 36, 38, 42, \ldots \\ &\equiv 1 \text{ at } n = 1,\ \ 3,\ \ 5,\ \ 9, 15, 19, 21, 23, 25, \ldots \\ &\equiv 2 \text{ at } n = 2,\ \ 6,\ \ 8, 10, 14, 16, 18, 22, 26, \ldots \\ &\equiv 3 \text{ at } n = 7, 11, 13, 17, 27, 31, 33, 35, 39, \ldots \end{aligned}$$

A state machine for dir mod 4 steps forward $0, 1$ or 2 according to the digit of n. This is the same for digits taken high to low or low to high. (Similar would apply for dir and some other modulus.)



$dir(n)$ mod 4
by digits of n

morphism $\quad 0 \to 01210 \quad 1 \to 12321 \quad 2 \to 23032 \quad 3 \to 30103 \quad$ starting 0

1.4 Coordinates

The location of a vertex n is given by the sub-part expansions. It's convenient to write this is complex numbers for the offsets and rotations. The locations in the base figure are

$$\mathit{digit}(a) = 0,\ 1,\ 1{+}i,\ i,\ 2i \qquad \text{according as } a = 0 \text{ to } 4 \tag{11}$$
$$b = 1{+}2i \qquad \text{end of base figure}$$

For a new digit a at the high end of n, the location is the sub-part shifted and rotated. This corresponds to expansion by unfolding.

$$\mathit{point}(a.5^k + n) \;=\; \mathit{digit}(a).b^k \;+\; \mathit{point}(n).i^{\mathit{dir}(a)} \qquad n < 5^k \tag{12}$$

For a new digit a at the low end of n, the location is scaled up by b and the base figure position suitably rotated. This corresponds to expansion by segment replacement (with rotation to keep the initial segment fixed East).

$$point(5n+a) \;=\; b.point(n) + digit(a).i^{dir(n)} \tag{13}$$

Both (12) and (13) are a sum of b powers for base-5 digits.

$$\begin{aligned}
&n = a_{k-1}\ldots a_1 a_0 \;\; \text{base-5 digits} \\
&point(n) = b^{k-1}\, digit(a_{k-1}) && \text{high digit} \quad (14)\\
&\qquad + b^{k-2}\, digit(a_{k-2})\; i^{\,dir(a_{k-1})} \\
&\qquad + b^{k-3}\, digit(a_{k-3})\; i^{\,dir(a_{k-1}a_{k-2})} \\
&\qquad + \cdots \\
&\qquad + b^{1}\quad digit(a_1)\quad i^{\,dir(a_{k-1}a_{k-2}\cdots a_2)} \\
&\qquad + b^{0}\quad digit(a_0)\quad i^{\,dir(a_{k-1}a_{k-2}\cdots a_2 a_1)} && \text{low digit} \\
&\qquad = 0,\, 1,\, 1{+}i,\, i,\, 2i,\, 1{+}2i,\, 1{+}3i,\, 3i, \ldots
\end{aligned}$$

Each direction $dir(a_{k-1}\ldots a_{j+1})$ is of the digits above a_j, not including a_j itself.

Per (27), the real and imaginary parts of each b^k can be calculated by Lucas sequences if x and y are wanted separately. The i direction factor selects $\pm$ either real or imaginary in each term.

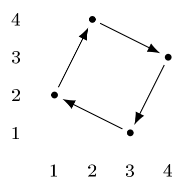
The direction $\arg b^k$ is at a multiple of

$$\begin{aligned}
\arg b = \arctan 2 &= 63.434948^\circ \\
&= 1.107148\ldots \quad \text{radians} && \text{A105199}
\end{aligned}$$

This is never a multiple of a full circle, meaning b^k is never on the x axis, since its imaginary part is never 0,

$$\operatorname{Im} b^k \bmod 5 \;\equiv\; [1,2,4,3] \qquad k{\ge}1 \;\text{A070402} \;(15)$$

The pattern of Re and Im mod 5 is



$\mathrm{Re}, \mathrm{Im}\; b^k \bmod 5$

The reverse coordinate calculation is to take a segment at z and direction offset d and find its n. This can be done following *point* formula (14) from low to high.

$$\text{unpoint}(z,\, d) \qquad z = \text{Gaussian integer},\;\; d = \begin{cases} \pm 1 & \text{if } z \text{ even} \\ \pm i & \text{if } z \text{ odd} \end{cases}$$

loop until $(z{=}0)$ or $(z{=}\pm i$ and $d{=}{-}z)$

$$a = \begin{cases} 0 & \text{if } z/d \equiv 0 \mod b \\ 1 & \text{if } z/d \equiv 1{+}i \mod b \\ 2 & \text{if } z/d \equiv i \mod b \\ 3 & \text{if } z/d \equiv 1 \mod b \\ 4 & \text{if } z/d \equiv 2i \mod b \end{cases} \qquad \text{base-5 digit } a$$

$d \leftarrow d/i^{dir(a)}$ undo low dir

$z \leftarrow z - d.digit(a)$ step to multiple of b

$z \leftarrow z/b$ divide out b

end loop

if $z{=}0$ and $d{=}1$ **then** n in arm 0
if $z{=}i$ and $d{=}{-i}$ **then** $5^k{-}n$ in arm 1
if $z{=}0$ and $d{=}{-1}$ **then** n in arm 2
if $z{=}{-i}$ and $d{=}i$ **then** $5^k{-}n$ in arm 3
where k is the number of digits of n generated

Terms in (14) are multiples of b except for the lowest, so the low $z \bmod b$ determines the low digit a. At the start of the loop

$$d = i^{dir(a_{k-1}\dots a_1 a)}$$

so that dividing it out (rotation) leaves

$$z/d \equiv digit(a)/i^{dir(a)}$$

which uniquely determines a from z and d. The effect of $/i^{dir(a)}$ is to swap the odd digits 1 and 3 of what would otherwise be the digit points (11).

For computer calculation, everything can be done in Cartesian x, y rather than full complex arithmetic. The $\bmod b$ cases are

$$x{+}2y \bmod 5 \equiv 0,3,2,1,4 \quad \text{is digit } a = 0 \,\text{to}\, 4 \text{ respectively} \tag{16}$$

For arm 0, the d parameter is the segment direction as a complex number direction so

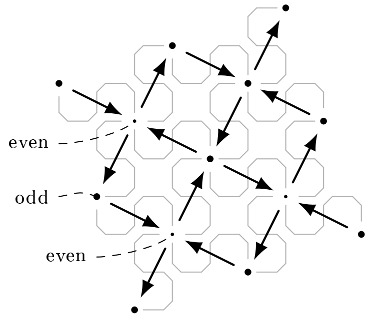
$$unpoint(point(n),\, i^{dir(n)}) = n \quad \text{in arm 0}$$

Similarly for arm 2 with $-point(n)$ and $-i^{dir(n)}$. But arms 1 and 3 are effectively calculated in reverse. Their original d is the direction from point n backwards

$$unpoint(i.point(n),\, -i\,.\,i^{-dir(n-1)}) \;= n \quad \text{in arm 1} \tag{17}$$

In both cases there are two d values giving the two visits to each Gaussian integer point z. They could be expressed by a sign $s = \pm 1$ if preferred. Such an s can be used through the calculation, but then the parity of z must be accounted for at each step (it is included in d). So digit a is then determined $\bmod 2b$, or Cartesians at (16) $\bmod 10$.

Geometrically, subtracting $digit(a)$ and dividing out b has the effect of undoing the lowest level of segment expansion. The segment direction d ensures the subtraction goes back to the start of the higher segment.



In arms 1 and 3, the reversal noted in (17) means this returning to segment start ends at $z = \pm i$ and $d = -z$. The loop gives repeated high digit 4s there, being reversals back from a 5^k.

From theorem 2, every point except the origin is visited exactly twice by arms of the curve. The $other(n)$ point number of the other visit can be calculated from n without the location z as such.

Theorem 5. *For a given $n \geq 1$ the $other(n)$ at the same location is given by the base-5 digits of n put though the following state machine low to high.*

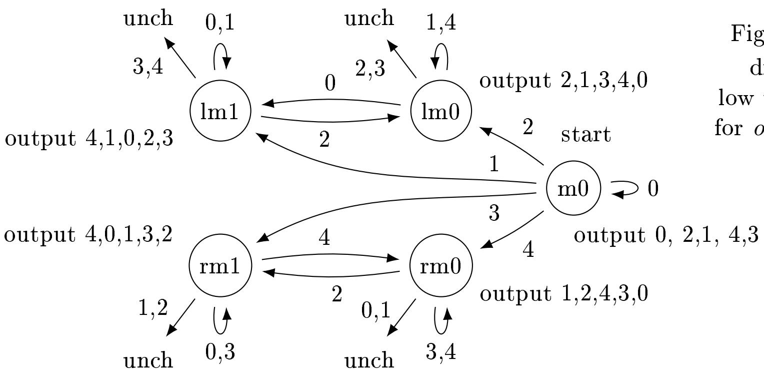


Figure 4:
digits
low to high
for $other(n)$

$$\begin{aligned} other(n) &= 0, \ \ 3, \ 14, \ 1, \ 8, \ 15, \ 18, \ 69, \ 4, 13, \ 70, \ 73, 16, \ 9, \ 2, \dots \\ arm \ &= 0, -1, -1, \ 1, \ 0, -1, -1, -1, \ 0, \ \ 0, -1, -1, \ \ 0, \ 0, \ 1, \dots \end{aligned}$$

In a given state, a digit of n gives the respective output digit and is a transition to a new state. If "unch" is reached then all further digits are unchanged for the output.

If n ends in rm0 then a high 0 digit is reckoned on n and goes to "unch". This is when $other(n)$ has more digits than n.

If n ends in lm0 then a high 0 digit is reckoned on n which goes to lm1. If n ends in lm1 then n is on the right boundary and $other(n)$ is in arm -1, at 5^k − output from the origin.

If n ends in rm1 then n is on the left boundary and $other(n)$ is in arm 1, at 5^k − output from the origin.

Proof. Suppose that m is at the same location as n but direction $+\delta$, and a certain dz offset away from n.

$$dir(m) \equiv dir(n) + \delta \quad \bmod 4 \tag{18}$$
$$point(m) = point(n) + i^{dir(n)}.dz$$

Factor $i^{dir(n)}$ on dz makes it like a low term of (14). This allows step (21) to require only the low digit of n.

Let a be the low digit of n and c the low digit of m so that

$$n = 5n' + a \qquad\qquad m = 5m' + c$$

From the low digit *point* formula (13), a and c are related by

$$i^{dir(m')}.\, digit(c) \;\equiv\; i^{dir(n')}.\, digit(a) + i^{dir(n)}.\, dz \quad \bmod b \tag{19}$$

and dividing through $i^{dir(m)}$ is, with (18) and *dir* as sum digit directions,

$$i^{-dir(c)}.\, digit(c) \;\equiv\; i^{-dir(a)-\delta}.\, digit(a) + i^{-\delta}.\, dz \quad \bmod b \tag{20}$$

$c = 0$ to 4 gives all complex integers mod b on the left side, so c is determined by δ, dz and a.

$$i^{-dir(c)}.\, digit(c) \;\equiv\; 0,\ 1{+}i,\ i,\ 1,\ 2i \qquad \text{for } c = 0 \text{ to } 4$$

New direction difference δ' is low digits dropped from (18)

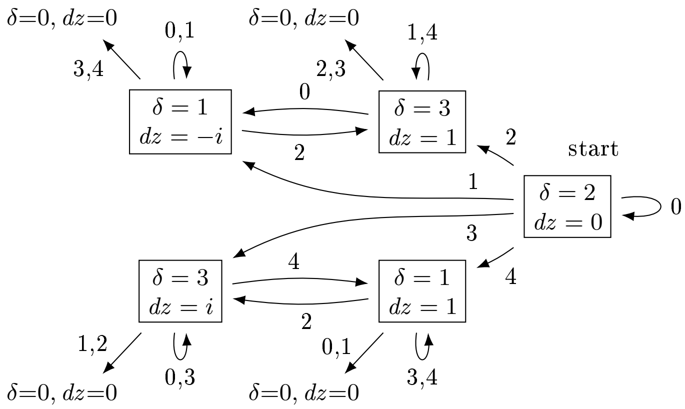
$$\delta' = \delta - dir(c) + dir(a) \;\bmod 4$$

New location offset dz' is the low digits taken off (13). The whole m' is not known yet, but $dir(m') = dir(n') + \delta'$ mod 4 is enough for the i power.

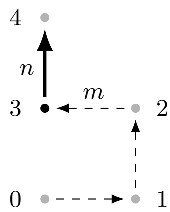
$$\begin{aligned}
dz'.i^{dir(n')} &= point(m') - point(n') \\
&= \quad \big(point(m) - i^{dir(m')}.digit(c)\big)/b \\
&\quad\; - \big(point(n) - i^{dir(n')}.digit(a)\big)/b \\
&= \big(dz.i^{dir(n)} - i^{dir(n')+\delta'}.digit(c) + i^{dir(n')}.digit(a)\big)/b
\end{aligned}$$
$$dz' = \big(dz.i^{dir(a)} - i^{\delta'}.digit(c) + digit(a)\big)/b \tag{21}$$

From (19), the bracketed part is a multiple of b.

These steps begin from $dz = 0$ so that m and n are the same location, and $\delta = 2$ for leaving there in the opposite direction. On reaching δ=0, dz=0, all further $c = a$ unchanged. Each possible digit a from n then gives the following transitions between combinations of δ, dz. These are per figure 4 and outputs there are c at (20).



High 0 digits on n loop in states lm1 or rm1. To see the rule for these as adjacent arms, first for lm1 suppose n has an extra high digit 3, so it goes to "unch" with high output c=2 on m.



So the other visit to $point(n)$ is at m along a curve directed from 2. Taking 3 as the origin means it is $5^k - m$ along a curve directed away from that origin in an arm at $-90°$.

For rm1, suppose n has an extra high digit 1, so it goes to "unch" with high 0 on m. Taking 1 as the origin means it is $5^k - m$ along a curve directed away from that origin in an arm at $+90°$. □

States lm0 and lm1 are reached by lowest non-0 digit 1 or 2 which is $turn(n)$ = 1 left from (1). The low digit is flipped 1↔2. Thereafter in lm0 and lm1, any digit 1s are unchanged. In lm0, the non-1 digits rotate +1 forwards, skipping 1. In lm1, the non-1 digits rotate −1 backwards, skipping 1. On reaching "unch" any further digits of n are unchanged in $other(n)$.

States rm0 and rm1 are reached when $turn(n) = -1$ right. The low digit is flipped and further digits rotated similar to lm but skipping digit 3.

Within a given expansion level, n is written in k many digits (with high 0s as necessary). If these end in state rm0 then a further 0 digit would go to state "unch". This means $other(n)$ is in the next level k+1 (across the join ahead in section 3.1).

The lm0 etc state δ, dz segments are located

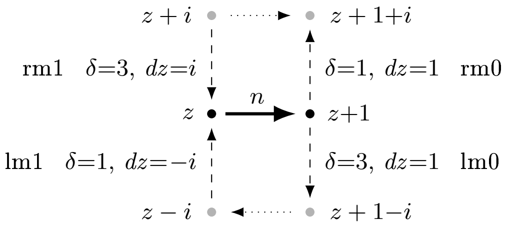
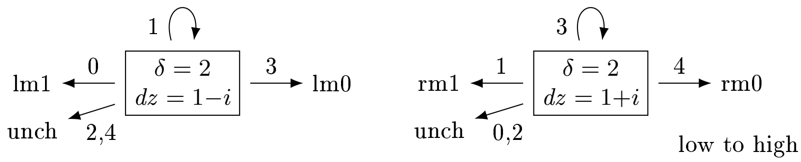


Figure 5

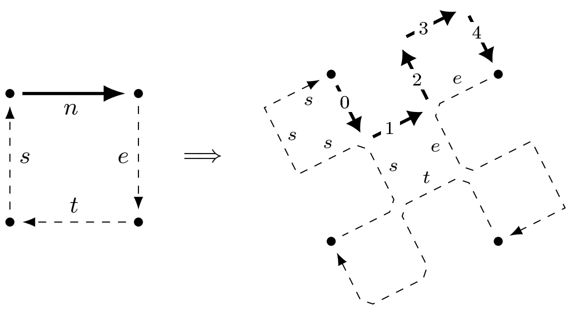
Starting from these states gives, from n, the segment numbers of those others. If such a segment is in an adjacent arm then the reversal is $5^k-1-output$ for segment rather than point.

Similar initial δ, dz can be used for other segments or points at locations relative to n. Bigger dz may extend further than just an adjacent arm, possibly reaching the 180° arm.

In figure 5, the bottom side shown dotted could be found by an lm0 start done twice (find the lm0 segment clockwise from n, then a second time finds the bottom segment). Starting directly δ=2, dz=1$-i$ goes to lm0 or lm1 according as lowest non-1 digit. Similarly the top side δ=2, dz=1+i.



Adjacent segment numbers can also be calculated by base-5 digits high to low. Suppose a segment n has segments s, t, e on its right. Expansion is a new low digit on n and the other segments



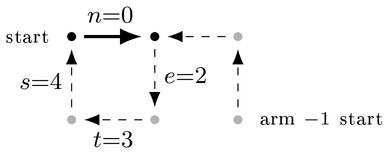
right
side
segments

The new adjacent s', t', u' follow from the new low digit of n,

n digit	s'	t'	u'
0	$s4$	$s3$	$s2$
1	$s1$	$t1$	$e1$
2	$e0$	$n4$	$n3$
3	$n2$	$e0$	$n4$
4	$n3$	$n2$	$e0$

(22)

Initial $s{=}4, t{=}3, e{=}2$ are segments in arm -1, on the right, directed towards the origin. Or start $s{=}0$ and an extra high 0 on n to step in (22) to $4, 3, 2$ (with initial t, e being unused by this).



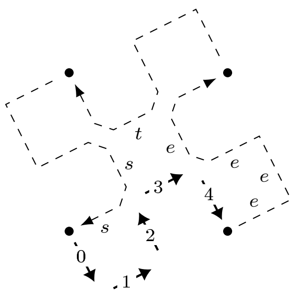
initial
$s{=}4, t{=}3, e{=}2$

A segment in arm -1 directed away from the origin is reversal $5^k{-}1{-}\mathit{output}$. After all digits of n are processed, an adjacent arm is identified by having high initial $2, 3, 4$, above the digits of n.

These right side segments give *other* of a left-turn n by one further expansion. A further low 1 digit or $100\ldots00$ sequence on n touches a corresponding low 2 or $200\ldots00$ of s. A further low 2 digit or $200\ldots00$ sequence on n touches a corresponding low 1 or $100\ldots00$ of e. (These $1{\leftrightarrow}2$ flips are per m0 low to high.)

Similar high to low holds for left side segments. Initial $s{=}4, t{=}2, e{=}1$ are the continuation of the curve and the arm to the right for s. The adjacent arm 1 is reversal $5^k{-}1 - \mathit{output}$ and is identified by high 4. One further expansion of these left segments give *other* of right turn n points at the starts of segments 3 and 4.

left
side
segments



n digit	s'	t'	u'
0	$s\,4$	$n\,2$	$n\,1$
1	$n\,0$	$s\,4$	$n\,2$
2	$n\,1$	$n\,0$	$s\,4$
3	$s\,3$	$t\,3$	$e\,3$
4	$e\,2$	$e\,1$	$e\,0$

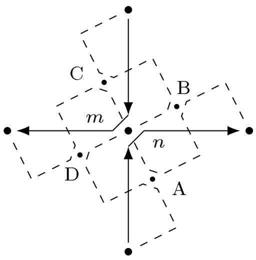
(23)

In tables (22),(23), some entries copy n for the new s', t', e'. This is where the output digits are to be n unchanged. It corresponds to somewhere at or above where the low to high of theorem 5 would be in "unch".

Theorem 6. *Differences $\Delta = \big|n - \mathit{other}(n)\big|$ which occur in R5 curve k are*

$$\begin{aligned} \Delta &= 4.\big(5^{k_0} \pm 5^{k_1} \pm 5^{k_1} \pm \cdots \pm 5^{k_t}\big) \\ &\quad \textit{distinct powers } k{-}2 \geq\ k_0 > k_1 > \cdots > k_t \geq 0 \\ &= 4, 16, 20, 24, 76, 80, 84, 96, 100, 104, 116, 120, \ldots \end{aligned} \tag{24}$$

Proof. The theorem is true trivially of $k = 0, 1$ where there are no $\mathit{other}(n)$ in the same curve arm. Suppose it is true in k. The segments at each double-visited point expand



$m = other(n)$ is the other visit to point n. On expansion they become $5n$ and $5m$ an difference 5Δ, all powers one bigger for one bigger k.

Point A is $5n{-}3$ and $5n{+}1$ for difference 4. Similarly point C before and after $5m$.

Point B is $5m{-}1$ and $5n{+}3$ so difference $5\Delta \pm 4$, with the sign according as m or n bigger. Point D similarly, but $5m{+}3$ and $5n{-}1$ so opposite ∓ 4.

The diagram shows a right turn n. The same new differences occur at a left turn. A single visited point (n alone) becomes new difference 4, which gives first difference 4 in $k{=}2$. □

Second Proof of Theorem 6. Differences can also be calculated from the *other* digit transformation of theorem 5. This is not as simple as the expansions, but shows where the difference powers fall in the *other* digit transformation.

In states lm0 and lm1, digit 1 loops and its output is 1 unchanged. Similarly digit 3 in rm0,rm1. Runs of other digits and their outputs are

Figure 6

	high	or 3,4 go to unch ↓ lm1	or 2,3 go to unch ↓ lm0	low
n	...	200...00	044...44	20...0
$other(n)$	...	100...00	200...00	10...0
		+4	−4	+4

all 1 digits unchanged

	high	or 1,2 go to unch ↓ rm1	or 0,1 go to unch ↓ rm0	low
n	...	400...00	244...44	40...0
$other(n)$	...	244...44	400...00	30...0
		+4	−4	+4

all 3 digits unchanged

m0 to lm0 is $2{\to}1$ which is difference -1 but take that as $+4$ and borrow -1 above. This is the low $+4$ shown in figure 6.

At lm0, $2{\to}3$ or $3{\to}4$ going to unch are $+1$ which cancels the borrow for a high 0 in Δ. So digit position $k{-}1$ must not have a power, making the highest power $k{-}2$.

lm0 $1{\to}1$ staying in lm0 and borrow is difference -1 but again take that as $+4$ and borrow above. $4{\to}0$ staying in lm0 and borrow is -5 which is 0 for Δ and continue borrow above.

lm0 $0{\to}2$ going to lm1 and borrow is difference $+1$. Or m0 to lm1 is $1{\to}2$ is likewise $+1$. But in both cases take that as -4 for Δ and carry $+1$ above in lm1. The middle -4 shown in figure 6 is lm0 to lm1.

At lm1, digits go similar to lm0 but carry $+1$ instead of borrow -1. lm1 $2{\to}3$ or $3{\to}4$ to unch are -1 which cancels the carry for 0 in Δ. $1{\to}1$ staying in lm1 and carry is $+1$ but again take that as -4 and carry $+1$ above. $0{\to}4$ staying in lm1 and carry is $+5$ which is 0 for Δ and continue carry above. $2{\to}0$ going to lm0 and carry is -1 and like m0 to lm0 take that as $+4$ for Δ and borrow -1 above.

Similar considerations on rm0 and rm1 gives the same $\pm 4, 0$ for Δ. $\square$

Nega-base-5 is base 5 using digits $0, \pm 1, \pm 2$ (instead of 0…4). (24) is equivalent to $\Delta \equiv 0 \bmod 4$, and $\Delta/4$ written in nega base-5 having digits $0, \pm 1$ only (no digits ± 2).

In plain base-5, the Δ powers are digits $0, 4$ with alternating pairs $3, 1$.

3,1 alternating and in pairs

$$\Delta = \begin{array}{c|c|c|c|c|c}
\scriptstyle\text{high} & & & & & \scriptstyle\text{low} \\
\hline
\cdots 0 \text{ or } 4 \cdots & 3 & \cdots 0 \text{ or } 4 \cdots & 1 & \cdots 0 \text{ or } 4 \cdots & \cdots \\
\hline
\geq 0 \text{ digits} & & \geq 0 \text{ digits} & & \geq 0 \text{ digits} &
\end{array}$$

Figure 7

$$\text{base-5} = 4, 31, 40, 44, 301, 310, 314, 341, 400, 404, 431, 440, \ldots$$

$5^{k_s} - 5^{k_t} = 044\ldots444$ in base-5 and multiplying 4 is $344\ldots441$. Taking just the positive Δ, any run of -4 powers has a $+4$ above it. Take k_t as the lowest of the -4s and k_s as the lowest $+4$ above it so digit pattern 34441. Other -4 in the runs of -4s cancel some of the 4s to 0s. Other $+4$ powers anywhere are 4 digits.

The number of distinct differences which occurs follows from the powers (24). Each k position is any of the three $0, \pm 4$, but not all 0s, and half to take only the positive Δ. Per the proof, the digit going to "unch" is a 0 in Δ so there are to be $k-1$ positions in curve k.

$$\underset{n,\, other(n) \,\leq\, 5^k}{\text{count distinct}} \ \big|n - other(n)\big| = \begin{cases} 0 & \text{if } k{=}0 \\ \frac{1}{2}(3^{k-1} - 1) & \text{if } k \geq 1 \end{cases}$$

Geometrically this corresponds to the join area J_k ahead at theorem 17.

Theorem 7. *The sum of distinct differences $\big|n - other(n)\big|$ in R5 curve k is*

$$\begin{aligned}
Odistinct_k &= \begin{cases} 0 & \textit{if } k{=}0 \\ \frac{2}{7}(15^{k-1} - 1) & \textit{if } k \geq 1 \end{cases} \\
&= 0, 0, 4, 64, 964, 14464, \ldots && \scriptstyle k \geq 2 \ \ 4\times\text{A135518} \\
&= 444\ldots44 \quad \textit{base 15, for } k{-}1 \textit{ digits}
\end{aligned}$$

Proof. Differences in k are those of $k-1$ with an additional low power, so 5 times each difference and 3 copies of each. The new low power can be $0, \pm 4$. Each $+4$ cancels with its corresponding -4, except $\Delta = 4$ that power alone.

So starting $Odistinct_1 = 4$ each successive k is a further base-15 digit 4.

$$Odistinct_k = 3.5.Odistinct_{k-1} + 4 \qquad k \geq 2$$ $\square$

Similar holds for a new high power. If 0 then $Odistinct_{k-1}$ total differences. If $+4$ then add all of $Odistinct_{k-1}$ but also subtract for the highest to be negative too, which cancels to just 4.5^{k-2} for two of each difference $\frac{1}{2}(3^{k-2} - 1)$, and also of all 0s below. The result 4.15^{k-2} is a new base-15 digit 4.

$$Odistinct_k = Odistinct_{k-1} + 4.5^{k-2}.(2\tfrac{1}{2}(3^{k-2} - 1) + 1) \qquad k \geq 2 \quad (25)$$

The mean distinct difference as a fraction of the curve length 5^k is

$$\frac{Odistinct_k}{5^k.\frac{1}{2}(3^{k-1}-1)} \to \frac{4}{35} = 0.114285\ldots \qquad \text{A021879}$$

1.5 Segments in Direction

Theorem 8. *The number of segments in directions $d \equiv 0,1,2,3 \bmod 4$ of R5 dragon curve k are*

$$\begin{aligned} S(k,d) &= \operatorname*{count}_{j=0}^{5^k-1} \Big(dir(j) \equiv d \bmod 4 \Big) \\ &= \tfrac{1}{4} \Big(5^k + (-i)^d b^k + \overline{(-i)^d b^k} + (-1)^d \Big) \qquad (26) \\ &= \tfrac{1}{4} \left| b^k + i^d \right|^2 - [0, \tfrac{1}{2}]_d \\ S(k,0) &= 1, 2, 5, 26, 153, 802, 3965, 19546, \ldots \\ S(k,1) &= 0, 2, 8, 30, 144, 762, 3928, 19670, \ldots \\ S(k,2) &= 0, 1, 8, 37, 160, 761, 3848, 19517, \ldots \\ S(k,3) &= 0, 0, 4, 32, 168, 800, 3884, 19392, \ldots \end{aligned}$$

Proof. When the curve expands, sub-parts 0 and 4 are in the same direction. Sub-parts 1 and 3 rotate $+90°$ which means two copies of the segments in direction d=3 move to direction d=0. Sub-part 2 rotates 180° which means the segments in direction d=2 move to direction d=0. So mutual recurrences

$$\begin{aligned} S(k{+}1,0) &= 2S(k,0) &&+ S(k,2) + 2S(k,3) \\ S(k{+}1,1) &= 2S(k,0) + 2S(k,1) &&+ S(k,3) \\ S(k{+}1,2) &= S(k,0) + 2S(k,1) + 2S(k,2) \\ S(k{+}1,3) &= S(k,1) + 2S(k,2) + 2S(k,3) \end{aligned}$$

which is

$$S(k{+}1,d) = 2S(k,d) + 2S(k,d{-}1) + S(k,d{-}2)$$

Substitutions or a little linear algebra gives a recurrence in a single d,

$$S(k{+}4,d) = 8S(k{+}3,d) - 22S(k{+}2,d) + 40S(k{+}1,d) - 25S(k,d)$$

which has characteristic polynomial

$$x^4 - 8x^3 + 22x^2 - 40x + 25 = (x-5)(x-1)(x-b)(x-\overline{b})$$

So $S(k,d)$ has a power form $W.5^k + X.b^k + Y.\overline{b}^k + Z$. Starting $S(0,0) = 1$ and others $S(0,1) = S(0,2) = S(0,3) = 0$ gives $W = \frac{1}{4}$, $X = \frac{1}{4}(-i)^d$, $Y = \frac{1}{4}i^d$, $Z = \frac{1}{4}(-1)^d$. ☐

In (26), the imaginary parts of the conjugate powers cancel out. The i^d factors select

$$(-i)^d b^k + \overline{(-i)^d b^k} = \begin{cases} 2\operatorname{Re} b^k & \text{if } d = 0 \\ 2\operatorname{Im} b^k & \text{if } d = 1 \\ -2\operatorname{Re} b^k & \text{if } d = 2 \\ -2\operatorname{Im} b^k & \text{if } d = 3 \end{cases}$$

The real and imaginary parts of b^k can be calculated by a Lucas sequence recurrence

$$v_{k+2} = 2v_{k+1} - 5v_k \tag{27}$$

$$\text{starting } v_0 = 1, v_1 = 1 \text{ for } \operatorname{Re} b^k = 1, 1, -3, -11, -7, 41, \ldots \qquad \text{A006495}$$

$$\text{starting } v_0 = 0, v_1 = 2 \text{ for } \operatorname{Im} b^k = 0, 2, 4, -2, -24, -38, \ldots \qquad \text{A006496}$$

Curve level k has 5^k segments in total. The b and constant parts cancel out in the total across $d = 0 \text{ to } 3$ so

$$S(k,0) \;+\; S(k,1) \;+\; S(k,2) \;+\; S(k,3) \;=\; 5^k$$

$|b| = \sqrt{5}$ so the b^k part in S grows roughly as a half power of 5 and so the segment counts in each direction are approximately $\frac{1}{4}5^k$ each, differing only by some fraction of the half power.

The half power is not periodic since, as from (15), the base $\arg b = \arctan 2$ angle is never a multiple of 2π.

A variation can be made by counting segments relative to the middle segment $n = \frac{1}{2}(5^k-1)$ which is direction $2k$ as from (8). The effect is to swap directions $0 \leftrightarrow 2$ and $1 \leftrightarrow 3$ at odd k. In the powers this becomes $-b$.

$$\begin{aligned}
SM(k,d) = S(k, 2k{+}d) &= \begin{cases} S(k,d) & \text{if } k \text{ even} \\ S(k,d{+}2) & \text{if } k \text{ odd} \end{cases} \\
&= \tfrac{1}{4}\left(5^k \,+\, (-i)^d(-b)^k \,+\, \overline{(-i)^d(-b)^k} \,+\, (-1)^d\right) \\
&= \tfrac{1}{4}\left|b^k + i^{d+2k}\right|^2 \,-\, [0, \tfrac{1}{2}] \\
SM(k,0) &= 1, 1, 5, 37, 153, 761, 3965, 19517, \ldots \\
SM(k,1) &= 0, 0, 8, 32, 144, 800, 3928, 19392, \ldots \\
SM(k,2) &= 0, 2, 8, 26, 160, 802, 3848, 19546, \ldots \\
SM(k,3) &= 0, 2, 4, 30, 168, 762, 3884, 19670, \ldots
\end{aligned}$$

Theorem 9. *Among the first n segments of the R5 dragon curve, the number of segments in direction $d \bmod 4$ is*

$$\begin{aligned}
SN(n,d) &= \operatorname*{count}_{j=0}^{n-1} \Big(\mathit{dir}(j) \equiv d \bmod 4 \Big) \\
&= \tfrac{1}{4}\Big(n + 2\operatorname{Re}(-i)^d \mathit{point}(n) \;+\; \big((-1)^d \;\textit{if } n \textit{ odd}\big)\Big) \qquad (28) \\
SN(n,0) &= 0, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 3, \ldots \\
SN(n,1) &= 0, 0, 1, 1, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, \ldots \\
SN(n,2) &= 0, 0, 0, 1, 1, 1, 1, 2, 2, 3, 3, 4, 4, 4, 4, \ldots \\
SN(n,3) &= 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 2, 2, 3, \ldots
\end{aligned}$$

Proof. These counts are cumulative sum of a direction predicate

$$DirPred(n,d) = \begin{cases} 0 & \text{if } dir(n) \equiv d \bmod 4 \\ 1 & \text{if not} \end{cases}$$

$$\begin{aligned}
DirPred(n,0) &= 1,0,0,0,1,0,0,0,0,0,0,0,1,\ldots \\
DirPred(n,1) &= 0,1,0,1,0,1,0,0,0,1,0,0,0,\ldots \\
DirPred(n,2) &= 0,0,1,0,0,0,1,0,1,0,1,0,0,\ldots \\
DirPred(n,3) &= 0,0,0,0,0,0,0,1,0,0,0,1,0,\ldots
\end{aligned}$$

A segment step $z = 1, i, -1, -i$ can be expressed as such a predicate by

$$\tfrac{1}{4}\Big(2 + 2\operatorname{Re}(-i)^d z - 2\big|\operatorname{Im}(-i)^d z\big|\Big) = \begin{cases} 1 & \text{if } z \text{ in direction } d \\ 0 & \text{if not} \end{cases}$$

Then applying that to steps $dpoint(n) = point(n{+}1) - point(n) = i^{\,dir(n)}$

$$\begin{aligned}
SN(n,d) &= \sum_{j=0}^{n-1} DirPred(n,d) \\
&= \tfrac{1}{4}\Bigg(\Big(\sum_{j=0}^{n-1} 2\Big) + \Big(2\operatorname{Re}(-i)^d \sum_{j=0}^{n-1} dpoint(j)\Big) && (29) \\
&\qquad\qquad - \Big(2\sum_{j=0}^{n-1}\big|\operatorname{Im}(-i)^d dpoint(j)\big|\Big)\Bigg) && (30)
\end{aligned}$$

The Re part (29) sum of *dpoint* is *point*(n). The $|\mathrm{Im}|$ part (30) is sum of terms 0 or 1 according as *dpoint* is horizontal or vertical, after rotation by d. The curve always turns left or right so segments are alternately horizontal and vertical so half each giving

$$2\sum_{j=0}^{n-1}\big|\operatorname{Im}(-i)^d dpoint(j)\big| \; = n - \big((-1)^d \text{ if } n \text{ odd}\big)$$

which subtracted from $\sum 2 = 2n$ is per (28). $\square$

Second Proof of Theorem 9. Segments alternate horizontal and vertical so total horizontals are $\lceil n/2 \rceil$ which is *SN* directions 0 plus 2. The difference of directions 0 and 2 is the net horizontal position Re *point*,

$$SN(n,0) + SN(n,2) = \lceil n/2 \rceil \qquad (31)$$
$$SN(n,0) - SN(n,2) = \operatorname{Re} point(n) \qquad (32)$$

(31)+(32) and (31)−(32) give

$$\begin{aligned}
SN(n,0) &= \tfrac{1}{2}\big(\lceil n/2 \rceil + \operatorname{Re} point(n)\big) \\
SN(n,2) &= \tfrac{1}{2}\big(\lceil n/2 \rceil - \operatorname{Re} point(n)\big)
\end{aligned}$$

Similarly for the verticals

$$\begin{aligned}
SN(n,1) + SN(n,3) &= \lfloor n/2 \rfloor \\
SN(n,1) - SN(n,3) &= \operatorname{Im} point(n)
\end{aligned}$$

$$
\begin{aligned}
SN(n,1) &= \tfrac{1}{2}\big(\lfloor n/2\rfloor + \operatorname{Im} point(n)\big) \\
SN(n,3) &= \tfrac{1}{2}\big(\lfloor n/2\rfloor - \operatorname{Im} point(n)\big)
\end{aligned}
$$

The $\pm\operatorname{Re},\operatorname{Im}$ parts are selected in (28) by $\operatorname{Re}(-i)^d point$, and the floor or ceil $n/2$ by the $(-1)^d$ offset part. ☐

A complete level k is 5^k segments $SN(5^k,d) = S(k,d)$. Its end $point(5^k) = b^k$ is the conjugate b^k parts of (26).

The *point* form of *SN* shows the mean number of segments by direction converges to $\frac{1}{4}$ each. A simple upper bound for $point(n)$ is to assume each of its digits is distance $|b^k| = \sqrt{5}^k$, and such a total grows slower than n.

$$
\begin{aligned}
base5len(n) &= \lceil \log_5 n{+}1\rceil \\
&= 0,1,1,1,1,2,2,2,2,2,2,2,2,\ldots \qquad \text{A110592}
\end{aligned}
$$

$$
\operatorname{Re} point(n) \le \sum_{k=0}^{base5len(n)-1} \sqrt{5}^k = \frac{\sqrt{5}^{base5len(n)} - 1}{\sqrt{5}-1} < \sqrt{5}^{base5len(n)}
$$

$$
\frac{1}{4} - \frac{2\sqrt{5}^{base5len(n)}+1}{n} < \frac{SN(n,d)}{n} < \frac{1}{4} + \frac{2\sqrt{5}^{base5len(n)}+1}{n}
$$

$$
\frac{SN(n,d)}{n} \to \frac{1}{4}
$$

2 Boundary

Theorem 10. *The boundary length of R5 dragon curve k is*

$$
\begin{aligned}
B_k &= 4.3^k - 2 \qquad \textit{boundary} \qquad (33) \\
&= 2,10,34,106,322,970,2914,\ldots \qquad k{\ge}1 \ \text{A079004}
\end{aligned}
$$

The curve is symmetric on each side so one side is

$$
\begin{aligned}
R_k &= \tfrac{1}{2}B_k = 2.3^k - 1 \qquad \textit{one-side boundary} \qquad (34) \\
&= 1,5,17,53,161,485,1457,\ldots \qquad \text{A048473}
\end{aligned}
$$

The length in a "U" part is

$$
\begin{aligned}
U_k &= 2.3^k + 1 \qquad \textit{“U” part boundary} \\
&= 3,7,19,55,163,487,1459,4375,\ldots \qquad \text{A052919, A100702}
\end{aligned}
$$

Proof. The curve consists of two R and one U. Points 1 and 2 are on the boundary since the two missing directions must be able to have curves added the to make the plane filling of theorem 2.

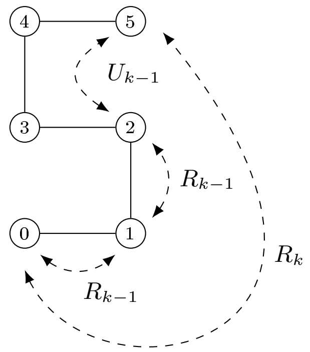


Figure 8: R expansion

$$R_k = 2R_{k-1} + U_{k-1} \tag{35}$$

$$R_0 = 1$$
$$U_0 = 3$$

The U part expands as follows. The two points marked with dots are on the boundary since the two missing directions must be able to have curves added the to make the plane filling of theorem 2.

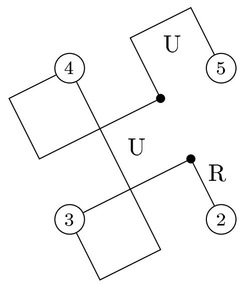


Figure 9: U expansion

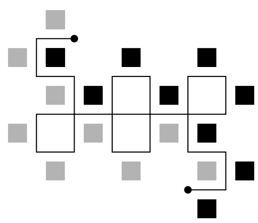
$$U_k = R_{k-1} + 2\,U_{k-1} \tag{36}$$

Using (35) for U_k and substituting into (36) gives recurrences

$$R_k = 4R_{k-1} - 3R_{k-2}$$
$$U_k = 4U_{k-1} - 3U_{k-2}$$

The characteristic polynomial is $x^2-4x+3 = (x-3)(x-1)$ so R and U have a power form $X.3^k + Y$. From initial $R_0 = 1$, $U_0 = 3$ and then $R_1 = 5$, $U_1 = 7$ by (35)(36) the powers are seen to be $2.3^k{-}1$ and $2.3^k{+}1$ respectively. □

2.1 Boundary Squares



$k = 2$

boundary squares

■ right $RQ_2 = 3^2 = 9$

■ left same

total $BQ_2 = 2.3^2 = 18$

Theorem 11. *The number of unit squares on the boundary of R5 dragon curve k is*

$$BQ_k = 2.3^k \qquad \textit{boundary squares} \tag{37}$$
$$= 2, 6, 18, 54, 162, 486, 1458, 4374, \ldots \qquad \text{A008776}$$

The curve is symmetric on each side so one side is

$$\begin{aligned} RQ_k = \tfrac{1}{2} BQ_k \; &= 3^k && \textit{one-side boundary squares} \\ &= 1, 3, 9, 27, 81, 243, 729, 2187, \ldots && \text{A000244} \end{aligned}$$

The "U" part is the same

$$UQ_k = RQ_k \; = 3^k \qquad \textit{"U" boundary squares}$$

Proof. Make the same breakdown into R and U parts as above. R and U meet at right angles so they do not have any boundary squares in common. The same recurrences (35) and (36) apply but U starts from a single initial square $UQ_0 = 1$.

$$RQ_k = 2\,RQ_{k-1} + \quad UQ_{k-1} \tag{38}$$

$$UQ_k = \quad RQ_{k-1} + 2\,UQ_{k-1} \tag{39}$$

$$\text{starting } RQ_0 = 1, \quad UQ_0 = 1$$

$RQ_0 = UQ_0$ and thereafter (38),(39) maintain $RQ_k = UQ_k$ so that $RQ_k = 3\,RQ_{k-1} = 3^k$. □

Second Proof of Theorem 11. The missing side of a U shape is an R

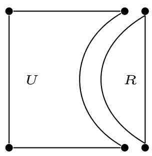


Figure 10:
U,R boundary squares
are opposites

Within the square all sides of all unit squares are traversed so the boundary squares of U overlap with the boundary squares of R, so that $UQ_k = RQ_k$. Then with the R expansion figure 8,

$$RQ_k = 3\,RQ_{k-1} \qquad \text{starting } RQ_0 = 1$$

□

The boundary squares have either 1 or 3 sides on the curve. There are no 2-side boundary squares since there are none in k=0 or k=1 and the breakdowns of figure 8 and figure 9 makes subsequent levels a multiple of the first two levels.

Theorem 12. *The number of 1-side and 3-side boundary squares of R5 dragon curve level k are*

$$\begin{aligned} BQ1_k &= 3^k + 1 && \text{A034472} \\ BQ3_k &= 3^k - 1 && \text{A024023} \\ RQ1_k &= \tfrac{1}{2}(3^k + 1) \\ &= 1, 2, 5, 14, 41, 122, 365, 1094, \ldots && \text{A007051} \\ UQ1_k &= \tfrac{1}{2}(3^k - 1) \\ &= 0, 1, 4, 13, 40, 121, 364, 1093, \ldots && \text{A003462} \\ RQ3_k &= UQ1_k && \textit{opposites} \\ UQ3_k &= RQ1_k \end{aligned}$$

Proof. In figure 10 the U and R boundary squares are opposed and every line segment is traversed so a 3-side of U is a 1-side of R so opposite $UQ3_k = RQ1_k$. Similarly a 1-side of U is a 3-side of R so $UQ1_k = RQ3_k$

The parts in figure 8 and figure 9 meet as the outside of right angles so do not change the number of sides of boundary squares in each part. The same recurrences (38),(39) apply to 1-side squares, but starting $RQ1_0 = 1$, and $UQ1_0 = 0$ (its single initial square is a 3-side). □

It also suffices to take just (38) applied to 1-side squares then use the UQ opposites and total $RQ_k = RQ1_k + RQ3_k$

$$
\begin{aligned}
RQ1_k &= 2RQ1_{k-1} + UQ1_{k-1} && \text{like (38)}\\
&= 2RQ1_{k-1} + RQ3_{k-1} && \text{opposites}\\
&= 2RQ1_{k-1} + RQ_{k-1} - RQ1_{k-1} && \text{total } RQ\\
&= RQ1_{k-1} + RQ_{k-1}\\
&= RQ1_0 + \sum_{j=0}^{k-1} RQ_j \quad = 1 + \frac{3^k - 1}{3 - 1}
\end{aligned}
$$

Likewise $RQ3_k$ but with $RQ3_0 = 0$ so that $RQ1_k = RQ3_k + 1$.

The R expansion figure 8 is shown as how a level k curve is comprised of level $k-1$ sub-curves. It can also be interpreted as how a line segment with a 1-side boundary square becomes two 1-sides and a 3-side when the line segment expands.

The U expansion figure 9 likewise as how 3 line segments with a 3-side boundary square become two 3-sides and a 1-side when the line segments expand.

The two interpretations are unfolding and segment expansion. The resulting recurrences are the same.

The boundary expansions R → R,R,U and U → R,U,U mean the number of sides of a boundary square is determined by numbering the squares in ternary, starting from m=0,

$$
\begin{aligned}
RQsides(m) &= \begin{cases} 1 & \text{if } TernaryLowestNon1(m) = 0\\ 3 & \text{if } TernaryLowestNon1(m) = 2 \end{cases}\\
&= TernaryLowestNon1(m) + 1\\
&= 1, 1, 3, 1, 1, 3, 1, 3, 3, 1, 1, 3, 1, 1, 3, 1, 3, 3, 1, 1, \dots
\end{aligned}
$$

$$
\begin{aligned}
TernaryLowestNon1(m) &= \text{ternary lowest non-1 digit of } m\\
&= 0, 0, 2, 0, 0, 2, 0, 2, 2, 0, 0, 2, 0, 0, 2, 0, 2, 2, 0, 0, \dots && 2{\times}\text{A116178}
\end{aligned}
$$

This number of sides gives a straightforward way to draw the boundary or part of the boundary. There are no straight ahead points on the boundary, since the remaining segments could not be traversed for plane filling without crossing or overlap. So the boundary is successive squares of $RQsides(m)$ with a left turn between each.

The left boundary of level k is a reversal of the right of k. For the left of the curve continued infinitely, the boundary expansions are reverse U → U,U,R and R → U,R,R starting from U. This is a swap R ↔ U of the right side above,

$$
\begin{aligned}
LQsides_{\infty}(m) &= \begin{cases} 3 & \text{if } \mathit{TernaryLowestNon1}(m) = 0 \\ 1 & \text{if } \mathit{TernaryLowestNon1}(m) = 2 \end{cases} \\
&= 3 - \mathit{TernaryLowestNon1}(m) \\
&= 4 - RQsides(m) \\
&= 3, 3, 1, 3, 3, 1, 3, 1, 1, 3, 3, 1, 3, 3, 1, 3, 1, 1, 3, 3, \ldots
\end{aligned}
$$

2.2 Boundary Segment Numbers

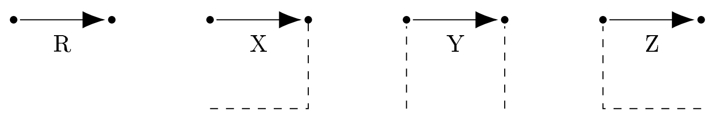
Theorem 13. *Number segments of the R5 dragon curve starting at n=0 for the first. The right boundary segments are characterized by*

$$
Rpred(n) = \begin{cases} 1 & \textit{if } n \textit{ in base-5 with digit } 1\textit{s deleted has no digit pair} \\ & \qquad\qquad\quad 22,\ \ 23,\ \ 24, \\ & \qquad\quad 30,\ \ 32,\ \ 33,\ \ 34, \\ & \qquad\quad 40 \\ 0 & \textit{otherwise} \end{cases}
$$

$$
= 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, \ldots
$$

$$
=1 \ \ at \ n{=} \begin{array}{llllll} \textit{decimal} & 0,1,2,3,4, & 5,6,7,8,9, & 10,11, & 16, & 21,22,23,24, \ \ 25,\ldots \\ \textit{base-5} & 0,1,2,3,4, & 10,11,12,13,14, & 20,21, & 31, & 41,42,43,44, \ \ 100,\ldots \end{array}
$$

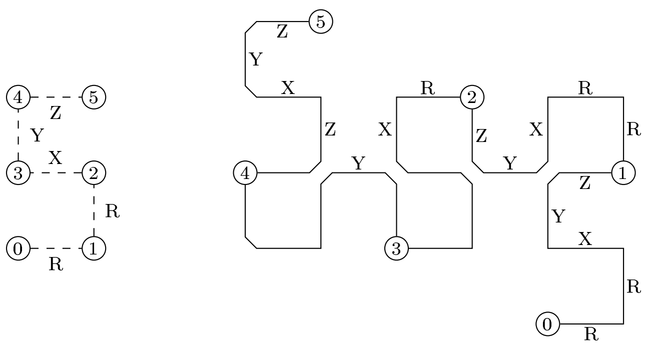
Proof. Take the boundary in four parts



R has no segments on its right. This is the right side of the full curve. X has a further curve at its end and so only some segments at its start are on the boundary. It also has a bottom curve, though that doesn't affect the result. Y has a curve at both start and end. Z has a curve at its start and below.

Let R_k, X_k, Y_k, Z_k be the set of segment numbers which are on the boundary in the respective configurations at level k. These numbers are in the range 0 to $5^k - 1$ and hence can be written with k many base-5 digits. The initial sets are a single 0 in each so $R_0 = X_0 = Y_0 = Z_0 = 0$ corresponding to a single line segment. These zeros are understood as 0 many digits.

The curve expands as



The R segment 0–1 expands to sub-parts 0.R, 1.R, 2.X, 3.Y, 4.Z. The number 0 to 4 is the high base-5 digit on top of the digits of the sub-part. Treating each part this way gives

$$\begin{array}{rlllll} R_k = & 0.R_{k-1}, & 1.R_{k-1}, & 2.X_{k-1}, & 3.Y_{k-1}, & 4.Z_{k-1} \\ X_k = & 0.R_{k-1}, & 1.X_{k-1} & & & \\ Y_k = & & 1.Y_{k-1} & & & \\ Z_k = & & 1.Z_{k-1}, & 2.X_{k-1}, & 3.Y_{k-1}, & 4.Z_{k-1} \end{array} \tag{40}$$

Taking base-5 digits from high to low this expansion is a state machine. For example in state R any digit is permitted and switch to state R, X, Y or Z according to the digit.

In each case, digit 1 is followed by the same type sub-part, so no change of state. This means a run of 1 digits is permitted anywhere and does not change the state.

For other digits, it is seen that 0, when permitted, always goes to state R. 2 always to state X. 3 always to state Y. And 4 always to state Z. This means the state at any position is given by the preceding non-1 digit. A disallowed state transition is therefore a disallowed digit pair. So from part X not $22, 23, 24$, from part Y not $30, 32, 33, 34$, and from part Z not 40. □

Part Y_k has only a single sub-part $1.Y_{k-1}$. This means that once in state Y the only possible further digit is repeated 1s. Geometrically this is because when three curves are in a U arrangement the middle part always has just a single segment on the resulting boundary. This is the single middle square ahead in area theorem 16 second proof.

The parts can be shown as a state machine,

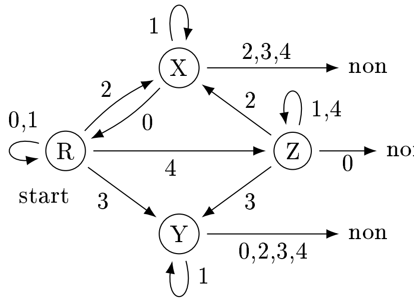


Figure 11: *Rpred* base-5 digits high to low

Going low to high is a little simpler. In the disallowed pairs, there are two cases. Digit 0 which is state r2 has 3 or 4 disallowed above. Digits 2,3,4 which is state r3 have digits 2,3 disallowed above. In all cases any digit 1 stays in the same state (like in high to low too).

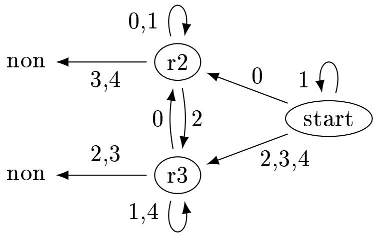


Figure 12: *Rpred* base-5 digits low to high

In state R of figure 11, the unit square on the right has just 1 segment. States X,Y,Z have 3 segments. So the state gives the number of sides of the unit square on the right of segment n,

$$
\begin{aligned}
\mathit{Rsides}(n) &= \begin{cases} 1 & \text{if } \mathit{Rpred} \text{ final state R} \\ 3 & \text{if } \mathit{Rpred} \text{ final state X,Y,Z} \\ 4 & \text{if } \mathit{Rpred} \text{ final state non} \end{cases} \\
&= 1, 1, 3, 3, 3, 1, 1, 3, 3, 3, 1, 3, 4, 4, 4, \ldots
\end{aligned}
\tag{41}
$$

States R,X,Y,Z are reached by digits $0, 2, 3, 4$ respectively, so lowest non-1 distinguishes 1 or 3 sides,

$$
\begin{aligned}
\mathit{Rsides}(n) &= \begin{cases} 1 & \text{if } \mathit{Rpred}(n) = 1 \text{ and } \mathit{LowestNon1}(n) = 0 \\ 3 & \text{if } \mathit{Rpred}(n) = 1 \text{ and } \mathit{LowestNon1}(n) \neq 0 \\ 4 & \text{if } \mathit{Rpred}(n) = 0 \end{cases} \\
\mathit{LowestNon1}(m) &= 0, 0, 2, 3, 4, 0, 0, 2, 3, 4, 0, 2, 2, 3, 4, \ldots
\end{aligned}
$$

Rsides can be expressed on digits low to high by finding the lowest non-1 and then the state machine of figure 12. In that state machine the lowest non-1 leaves start state, so *Rsides* would have one pair of r2,r3 for a 0 giving $\mathit{Rsides} = 1 \,\mathrm{or}\, 4$ and another pair for non-0 giving $\mathit{Rsides} = 3 \,\mathrm{or}\, 4$. In all cases "non" is 4 sides.

Theorem 14. *Number segments of the R5 dragon curve starting at n=0 for the first. The left boundary segments are characterized by*

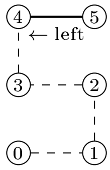
$$
\mathit{Lpred}_k(n) = \begin{cases} 1 & \textit{if } n \textit{ in } k \textit{ base-5 digits with 3s deleted has no digit pair} \\ & \qquad\qquad\qquad\qquad\quad 04, \\ & \qquad\quad 10,\quad 11,\quad 12,\quad 14, \\ & \qquad\quad 20,\quad 21,\quad 22 \\ 0 & \textit{otherwise} \end{cases}
$$

For the curve continued infinitely add a high 0 digit above the most significant non-zero so a most significant digit 4 becomes an 04 which is disallowed.

$$
\begin{aligned}
\mathit{Lpred}_\infty(n) &= \mathit{Lpred}_{k+1}(n) \quad \textit{for } k \textit{ with } 5^k > n \\
&= 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, \ldots \\
=1 \textit{ at } n= & \begin{array}{rl} \textit{decimal} & 0,1,2,3,\ 8,\ 13,14,\ 15,16,17,18,\ 43,\ldots \\ \textit{base-5} & 0,1,2,3,\ 13,\ 23,24,\ 30,31,32,33,\ 133,\ldots \end{array}
\end{aligned}
$$

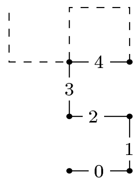
Proof. The curve is symmetric on its left and right sides so the left boundary segment numbers are the right segment numbers but numbered in reverse. This means digits 0,1,2,3,4 become 4,3,2,1,0. The digit pairs to exclude are the digit reversals of those in the right boundary pairs. The 1s deleted become 3s deleted.

For the left side of an infinite curve the reversal is from the final segment of an extra level of expansion. In the following diagram the 5–4 continues as 4–3 etc. This is an extra high 4 digit on the right boundary which reverses to be an extra high 0 digit for the left boundary.



□

The left digit pair 04 which is disallowed can be illustrated by the following diagram. Any 04 sub-part is closed off by the following "10", "11", "12".



Further it can be seen some of the "03" part will be closed off by the "13" curve blocking its end. In the theorem this follows from the "delete 3s".

For a finite curve digit 4 can be the high digit when n is written in k digits since there is nothing beyond it which would close it off.

For an infinite curve the extra high 0 digit can also be thought of as starting in state Z of theorem 13 rather than full R.

An *Lsides* follows from states similar to *Rsides* at (41), or by digit reversals,

$$\begin{aligned} \mathit{Lsides}_k(n) &= \mathit{Rsides}(5^k{-}1 - n) \\ \mathit{Lsides}_\infty(n) &= \mathit{Lsides}_{k+1}(n) \quad \text{for } k \text{ with } 5^k > n \\ &= 3, 3, 3, 3, 4, 4, 4, 4, 3, 4, 4, 4, 4, 3, 1, \ldots \end{aligned}$$

Segments on both the left and right boundary are those with none of either *Lpred* or *Rpred* disallowed pairs. This leaves allowed pairs

$$00, 01, 02, 03, \quad 13, \quad 31, \quad 41, 42, 43, 44$$

The only n which can be made with these pairs are fixed sets at start and end of the curve,

$$\begin{array}{lll} k{=}0 & n = 0 & \text{(all)} \\ k{=}1 & n = 0, 1, 2, 3, 4 & \text{(all)} \\ k{\ge}2 & n = 0\ldots\ 00, 01, 02, 03,\ 13,\ 31 & \text{base-5} \\ & \quad\ \ 4\ldots\ 44, 43, 42, 41,\ 13,\ 31 & \\ & \ \ = 0, 1, 2, 3, 8, 16 \ \text{ and } 5^k{-}1 - 0, 1, 2, 3, 8, 16 & \end{array} \tag{42}$$

For $k{=}2$ the 8,16 offsets in (42) are the same so there are 10 segments on both boundaries. For $k \ge 3$ there are a full 12 (6 at each end of the curve).

Theorem 15. *The lengths of sub-parts X, Y, Z from theorem 13 are*

$$\begin{aligned} X_k &= 3^k - k \\ &= 1, 2, 7, 24, 77, 238, 723, 2180, 6553, \ldots \end{aligned} \qquad \text{A024024}$$

$$
\begin{aligned}
Y_k &= 1 \\
Z_k &= 3^k + k \\
&= 1, 4, 11, 30, 85, 248, 735, 2194, 6569, \ldots \qquad \text{A104743}
\end{aligned}
$$

Proof. The expansions (40) are recurrences in the lengths

$$R_k = 2R_{k-1} + X_{k-1} + Y_{k-1} + Z_{k-1} \tag{43}$$
$$X_k = \quad R_{k-1} + X_{k-1} \tag{44}$$
$$Y_k = \qquad Y_{k-1}$$
$$Z_k = \qquad X_{k-1} + Y_{k-1} + 2Z_{k-1} \tag{45}$$

The initial lengths are a single segment each $X_0 = Y_0 = Z_0 = 1$.

$Y_k = Y_{k-1}$ means $Y_k = 1$ always.

Expanding X (44) repeatedly is cumulative R. So with the sum understood as empty when k=0, and R_k from (34),

$$X_k = X_0 + \sum_{j=0}^{k-1} R_j \;=\; 1 + \sum_{j=0}^{k-1} \left(2.3^j - 1\right) \;=\; 1 + 2\frac{3^k - 1}{2} - k \tag{46}$$

R, X and Y into (43) gives Z

$$
\begin{aligned}
Z_k &= R_{k+1} - 2R_k - X_k - Y_k \;=\; 3^k + k \\
&= (2.3^{k+1}-1) - 2(2.3^k-1) - (3^k-k) - 1
\end{aligned}
\qquad \square
$$

X, Y and Z are the three sides of U (and so (43) is the R expansion (35)).

$$X_k + Y_k + Z_k \;=\; U_k$$

Noticing this on the right of (45) gives Z as cumulative U, similar to X cumulative R at (46).

$$Z_k = Z_{k-1} + U_{k-1} \;= Z_0 + \sum_{j=0}^{k-1} U_j \;= 1 + \sum_{j=0}^{k-1} \left(2.3^j + 1\right) \;= 1 + 2\,\frac{3^k - 1}{3 - 1} + k$$

3 Area

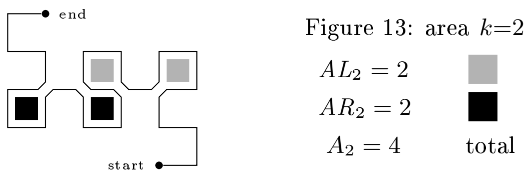
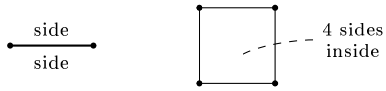


Figure 13: area k=2

Lemma 1. *Consider line segments on a square grid where any enclosed unit square has segments on all 4 sides. The enclosed area A and boundary B are related to total line segments N by*

$$4A + B = 2N \tag{47}$$

Proof. Count the sides of the line segments. There are N segments so total $2N$ sides. Each side is either a boundary or is inside.



There are B outside sides on the boundary. The inside sides are all in enclosed unit squares. Each area A square has 4 inside sides, so $4A$ and total $B + 4A = 2N$. ☐

Theorem 16. *The area enclosed by R5 dragon curve k is*

$$\begin{aligned} A_k &= \tfrac{1}{2}(5^k + 1) - 3^k && \textit{area} \qquad (48) \\ &= 0, 0, 4, 36, 232, 1320, 7084, 36876, \ldots \end{aligned}$$

Each side is symmetric so half area on each side

$$\begin{aligned} AR_k = AL_k &= \tfrac{1}{2}A_k && \textit{area one side} \\ &= \tfrac{1}{4}(5^k - 2.3^k + 1) \\ &= 0, 0, 2, 18, 116, 660, 3542, 18438, \ldots && \text{A007798} \end{aligned}$$

Proof. From the area–boundary relation (47) and boundary (33)

$$\begin{aligned} A &= \tfrac{1}{4}(2N - B) \\ A_k &= \tfrac{1}{4}(2.5^k - (4.3^k - 2)) \end{aligned}$$ ☐

Area can also be calculated from the boundary squares in the way Daykin and Tucker [5] use for the Heighway/Harter dragon curve. This applies to any curve where adjacent copies mesh perfectly.

Second Proof of Theorem 16. Arrange four R5 curves in a square

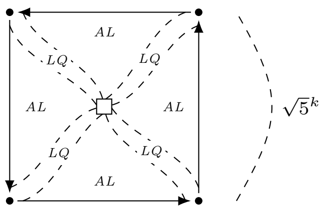


Figure 14: square 5^k
$4 \times AL_k$ area
$4 \times LQ_k$ boundary squares
with overlaps

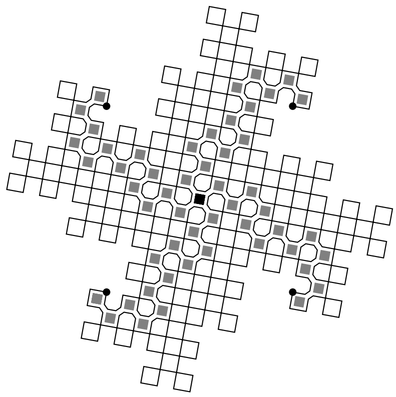
The curve endpoints are $(\sqrt{5})^k$ apart. The curves divide the plane into identical areas so the number of unit squares inside is $(\sqrt{5}^k)^2$. These squares are the left-side enclosed area AL_k and the left-side boundary squares LQ_k.

Every edge in the square is traversed so the curves mesh and the boundary squares overlap between adjacent curves. Along the diagonals the boundary squares are doubled. $LQ_k = 3^k$ is always odd so there is a middle unit square common to all four boundaries.

$$\begin{aligned} 5^k &= 4AL_k \\ &\quad + 4\tfrac{1}{2}(LQ_k - 1) && \text{sides, un-overlapped} \\ &\quad + 1 && \text{middle square} \end{aligned}$$

$$AL_k = \tfrac{1}{4}(5^k - 1) - \tfrac{1}{2}(3^k - 1)/2 \qquad \square$$

The actual joins of the square in figure 14 twist and turn. The following diagram shows an example. The grey boundary squares are overlapped twice. The middle black square is overlapped four times.



k=3
square area joins
and middle square

Third Proof of Theorem 16. The segments of each 1-side, 3-side and 4-side square expand into that square as

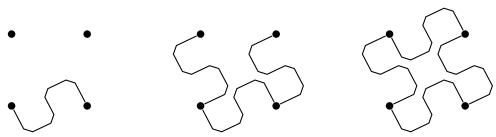


Figure 15:
$1, 3, 4$
side square
expansions

The number of unit squares enclosed by the expanded segments, on the side of the square, are 0, 2, 5 respectively. The 3-side squares are $BQ3_k$. The 4-side squares are the enclosed area A. So a recurrence for A_k giving a powered sum of $BQ3_k$.

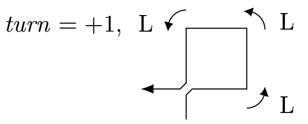
$$\begin{aligned}
A_k &= 5A_{k-1} \,+\, 2\,BQ3_{k-1} && (49)\\
&= 5^k A_0 \,+\, 2\sum_{j=0}^{k-1} 5^j\, BQ3_{k-1-j}\\
&= 2\sum_{j=0}^{k-1} 5^j\,(3^{k-1-j} - 1) && \text{since } A_0 = 0\\
&= 2\left(\frac{5^k - 3^k}{5-3} - \frac{5^k - 1}{5-1}\right) && (50)
\end{aligned}$$

$\square$

For the curve endpoints scaled to a unit length, the area limit is the coefficient of the 5^k power,

$$\frac{A_k}{5^k} \to \frac{1}{2} \qquad (51)$$

As from *TurnRun* in section 1.2, the curve turns go in runs of either 2 or 3 consecutive left or right. A run of 3 consecutive turns encloses a unit square.



three consecutive left turns,
is left-side enclosed unit square

The run lengths are pairs either 2,3 or 3,2. There is one 3 for each of the $5^{k-1}-1$ turns of the previous expansion level. So the number of runs of 3 turns in curve k is

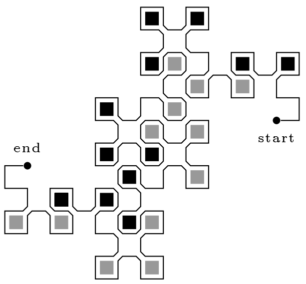
$$TurnRuns3_k = \begin{cases} 0 & \text{if } k{=}0 \\ 5^{k-1} - 1 & \text{if } k \geq 1 \end{cases} \tag{52}$$

$$= 0, 0, 4, 24, 124, 624, 3124, \ldots \qquad k{\geq}1 \text{ A024049}$$

The proportion of enclosed unit squares formed by 3-turns, out of the total area, is

$$\frac{TurnRuns3_k}{A_k} = \tfrac{2}{5} + \tfrac{2}{5}\frac{3^k-3}{A_k} \to \tfrac{2}{5}$$

This limit is approached from above since $3^k-3 > 0$ for $k \geq 2$ which is where $A_k > 0$. For example in k=3 the ratio is $\frac{2}{3}$,



k=3
LLL squares black
RRR squares grey

total
$TurnRuns3_3 = 24$
$A_3 = 36$

Some segments have these 3-turn squares on both sides. Such pairs are a sequence of turns LLLRRR. As from the turn expansion in figure 2, such consecutive 3-runs occur only as an LLRR plus L,R existing turns before and after. An L,R is then only the middle of an LLRR of preceding segment expansion. So there is one LLLRRR for each $k-2$ segment.

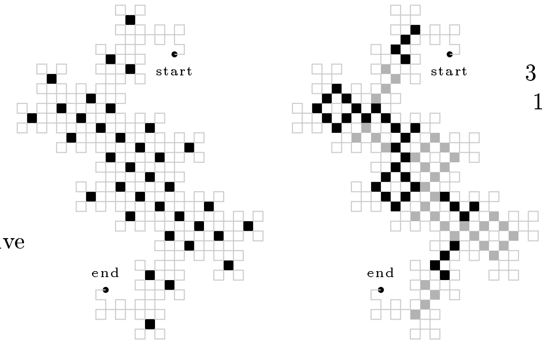
There are no RRRLLL pairs, since the Rs could only be an LLRR and existing R, but then LLR follows, not LLL.

$$TurnRuns3pairs_k = \begin{cases} 0 & \text{if } k = 0, 1 \\ 5^{k-2} & \text{if } k \geq 2 \end{cases}$$

The other enclosed unit squares are of two types, one has 3 sides consecutive and 1 separate, the other has 4 separate. This follows from the expansions of figure 15. The inner-most square has 4 separate, no matter what original sides. These are each preceding level enclosed unit square so A_{k-1} and located at a grid of b steps apart.

The outer squares in the expansions have 3 consecutive sides, and further side depending on whether the original sides were consecutive. Non consecutive original gives a non consecutive in the expansion. These can be counted by difference

$$A31_k = A_k - \mathit{TurnRuns3}_k - A_{k-1} \quad = \begin{cases} 0 & \text{if } k{=}0 \\ 2A_{k-1} & \text{if } k \geq 1 \end{cases}$$

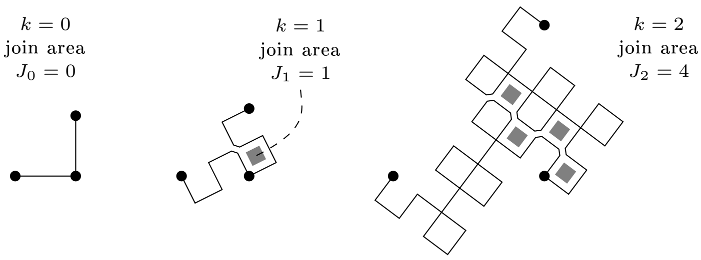


k=4
4 sides all
not consecutive

k=4
3 sides consecutive
1 not consecutive
left black
right grey

3.1 Join Area

When two copies of the R5 curve meet at right angles they touch and enclose new area. The simplest join is k=1 which encloses a single new unit square $J_1 = 1$ as shown in the following diagram.

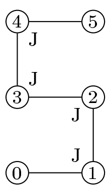


Theorem 17. *The area enclosed at the join of two R5 dragon curves at right angles is*

$$\begin{aligned} J_k &= \tfrac{1}{2}(3^k - 1) \qquad \textit{join area} \qquad (53) \\ &= \tfrac{1}{2}(RQ_k - 1) \\ &= 0, 1, 4, 13, 40, 121, 364, 1093, \ldots \qquad \text{A003462} \\ &= \textit{ternary } 0, 1, 11, 111, 1111, \ldots \qquad \text{A002275} \end{aligned}$$

Proof. In figure 14 the join area is the diagonal, so half the RQ_k excluding the middle 4-overlap square. □

Second Proof of Theorem 17. The join can also be calculated from the area as the extra which A_k has over its five A_{k-1} sub-copies.



joins

The horizontal sub-curves 0–1 and 2–3 don't touch since otherwise a square of four such sides would have its verticals touching too and so repeating some edge. Likewise 2–3 and 4–5 don't touch. So

$$A_{k+1} = 5A_k + 4J_k \tag{54}$$
$$AR_{k+1} = 5AR_k + 2J_k \qquad \text{one side} \qquad \square$$

Notice (54) is the same as area by 3-side boundary squares (49) since

$$J_k = \tfrac{1}{2} BQ3_k$$

Third Proof of Theorem 17. The join area is also the shortfall of boundary length B_{k+1} over five copies of B_k. There are 4 joins and each unit square enclosed by the joins reduces the boundary by 4 segments.

$$5B_k - B_{k+1} = 4J_k$$

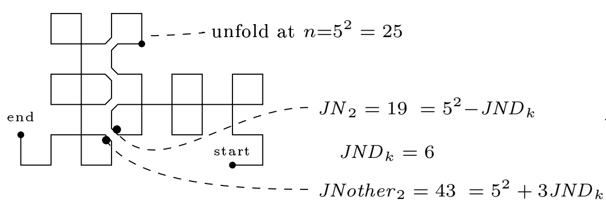
Similarly with the boundary squares. The shortfall of BQ_{k+1} over 5 copies of BQ_k is the 4 joins. Each join square is 2 overlapping boundary squares which are no longer on the boundary. At the end of the join the middle square from figure 14 is common to 3 adjacent curves and becomes a single boundary square. The effect of this middle square is to reduce by 1 for each of the four joins.

$$5BQ_k - BQ_{k+1} = 4\ (2J_k + 1) \qquad \square$$

In any non-overlapping self-similar curve the area is n copies of the previous level plus some amount extra which is where the copies join. The R5 curve is symmetric so there are four equal joins.

J_k is also the number of distinct differences $|n - other(n)|$ from theorem 7. The join is the only place new double-visited points occur on unfolding, so differences elsewhere are replications of differences in joins of previous levels.

The first point where a level k curve touches its unfolded copy is the smallest in the join and is the end-most part of the join area. The vertex numbers on each side can be calculated.



k=2
first and last
points where
unfold touches

Theorem 18. *The first vertex (smallest n) of a level k curve join is vertex number*

$$\begin{aligned} JN_k &= \tfrac{1}{4}\,(3.5^k + 1) && \textit{join } n \\ &= 1, 4, 19, 94, 469, 2344, \ldots && \text{A083065} \\ &= 33\ldots334 && \textit{base-5, k digits} \end{aligned}$$

and the opposing point it touches is

$$JNother_k = other(JN_k)$$

$$= \tfrac{1}{4}\,(7.5^k - 3) \qquad \textit{join n other}$$
$$= 1, 8, 43, 218, 1093, 5468, \ldots \qquad \text{A117617}$$
$$= 133\ldots33 \quad \textit{base-5, } k{+}1 \textit{ digits}$$

Or each measured back from the unfold point 5^k*,*

$$JN_k = 5^k - JND_k \qquad \textit{join n from unfold}$$
$$JNother_k = 5^k + 3\,JND_k$$
$$JND_k = \tfrac{1}{4}\,(5^k - 1)$$
$$= 0, 1, 6, 31, 156, 781, \ldots \qquad \text{A003463}$$
$$= 11\ldots11 \quad \textit{base-5, } k \textit{ digits repunit}$$

For k=0 there is nothing enclosed by the join and $JN_k = JNother_k = 5^k$. For $k \geq 1$ there is join area enclosed.

Proof. Join k comprises $k{-}1$ sub-curves

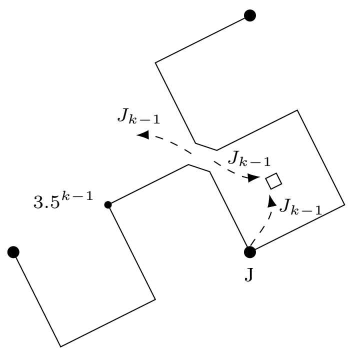


Figure 16

Join JN_k is at 3.5^{k-1} and further JN_{k-1},

$$JN_k = 3.5^{k-1} + JN_{k-1} \qquad \text{starting } JN_0 = 1$$

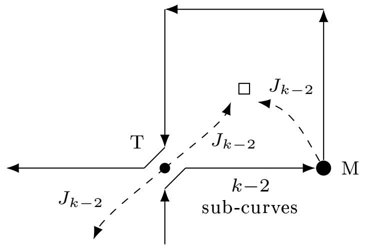
JND_k is back from point J by 5^{k-1} and further JND_{k-1} so

$$JND_k = 5^{k-1} + JND_{k-1} \qquad \text{starting } JND_0 = 0$$

$JNother_{k-1}$ would be relative to the 3.5^{k-1} point, so in k that plus the square 4.5^{k-1},

$$JNother_k = 7.5^{k-1} + JNother_{k-1} \qquad \text{starting } JNother_0 = 1 \qquad \square$$

The joins each way in figure 16 are a geometric interpretation of the high digit form of *Odistinct* at (25). The distinct differences in curve k are across join $k{-}1$ (as noted above). A $k{-}1$ join expands



join J_{k-1}
expansion

The join at M is the same set of differences as the previous level. Those at the new touch T are directed forwards and backwards. The differences there are offsets from difference 4.5^{k-2} around the big square of sub-curves shown. The differences forward and backward cancel out, leaving just 4.2^{k-4} for each point. There is 1 point each way for J_{k-2}, plus T itself. So

$$Odistinct_k = Odistinct_{k-1} + 4.5^{k-2}.\big(2J_{k-2}+1\big) \qquad k \geq 2$$

4 Points

4.1 Point Counts

The R5 dragon curve visits each point either once or twice. The number of single and double points can be calculated.

Lemma 2. *Consider a path on a square grid which does not repeat any segment and which always traverses all four sides of any enclosed unit square.*

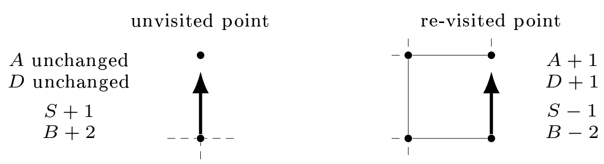
The number of single-visited points S, double visited points D, enclosed area A and boundary length B are related by

$$\begin{aligned} D &= A && \textit{double-visited = area} \\ S &= B/2+1 && \textit{single-visited and boundary} \end{aligned}$$

The starting point of the path can be revisited. If it is then that point is double-visited.

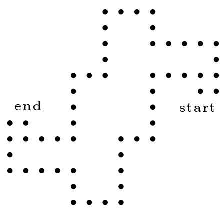
Proof. A path of no line segments is taken to be a single point at its start. The relations hold with $D = A = B = 0$ and $S = 1$.

When a further line segment is added to the end of the path it either goes to an unvisited point or it re-visits a point,



On going to an unvisited point a new single is added so S+1, and D unchanged. No new area is enclosed so A unchanged. The boundary increases by 2 (one on each side of the new line) so B+2. These new values satisfy the relations.

On re-visiting, a point which was single-visited becomes a double, so $S{-}1$ and $D{+}1$. A new unit square is enclosed so $A{+}1$. The boundary changes by -3 enclosed and $+1$ new outside so net $B{+}2$. These new values satisfy the relations. ☐



$k{=}3$
single-visited points
$S_3 = 54$

Theorem 19. *The number of single-visited points in the R5 dragon level k is*

$$S_k = \tfrac{1}{2}B_k + 1 \quad = 2.3^k$$
$$= 2, 6, 18, 54, 162, 486, 1458, 4374, \ldots \qquad \text{A008776}$$

Proof. The R5 dragon curve does not repeat any segment and always traverses all four sides of any unit square so lemma 2 applies. ☐

Second Proof of Theorem 19. When each line segment expands, its existing single and double visited points are unchanged. New points are made by the new vertices.

The expansions of figure 15 show the new vertices created within 1, 3 and 4-sided squares. 1-sided and 3-sided squares have 2 new single-visited points within the square. The 4-sided square has no new single-visited points (all are doubles).

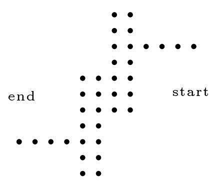
The 1-side and 3-side squares together are the boundary squares BQ_k from (37).

$$S_k = S_{k-1} + 2\,BQ_{k-1} \quad = S_0 + \sum_{j=0}^{k-1} 2\,BQ_j$$ ☐

S_k is even since $S + 2D =$ total points means it has the same parity as the total vertices $5^k + 1$. Any double point is made by taking away 2 singles from the total vertices.

If the curve did not touch itself then at the next level the single points would be $5S_k - 4$, for five copies with the endpoint meetings counted once each. This happens when $k{=}0$ as $S_1 = 5S_0 - 4$ but thereafter S_{k+1} is reduced by 2 for each point which touches. There is one such point for every join square (53) so

$$S_{k+1} = 5S_k - 4 \;-\; 8J_k$$



$k{=}3$
double-visited points
$D_3 = 36$

The double-visited points are the same as the area per lemma 2.

$$D_k = A_k$$

They can also be counted from the join area. Each join area square is formed by points on each side touching to make a double, so

$$D_k = 5D_{k-1} + 4J_{k-1}$$

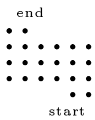
This is the same recurrence as (54) for A. Starting from the same $D_0 = A_0 = 0$ they maintain $D_k = A_k$.

Double-visited points can also be counted by the new vertices of the segment expansions in figure 15, as per the singles proof above. 1-side squares have no new doubles. 3-side squares have 2 new doubles. 4-side squares (ie. enclosed area) have 4 new doubles.

$$D_k = D_{k-1} + 2\,BQ3_{k-1} + 4\,A_{k-1}$$

When $D_{k-1} = A_{k-1}$ this is the same recurrence as area by 3-side boundary squares (49). They start $D_0 = A_0 = 0$ and thereafter remain equal.

For $k \le 3$ the number of single-visited points exceeds the double-visited but for $k \ge 4$ there are more double-visited. $D_k = A_k > S_k$ is when $5^k \ge 6.3^k$ which is $k \ge 4$.



k=2
distinct visited points
$P_2 = 22$

The number of distinct points visited by curve k is

$$\begin{aligned} P_k &= S_k + D_k && \text{distinct points} \\ &= 5^k + 1 - D_k && (55) \\ &= \tfrac{1}{2}\,(5^k + 1) + 3^k \\ &= 2, 6, 22, 90, 394, 1806, 8542, \ldots && 2\times \text{A146086} \end{aligned}$$

It can be noted P_k and doubles $D_k = A_k$ (48) differ only in the sign of the 3^k term. That term is $\frac{1}{2}S_k$.

$$D_k = \tfrac{1}{2}\,(5^k + 1 - S_k) \qquad \text{from } S + 2D = 5^k + 1 \qquad (56)$$

$$P_k = \tfrac{1}{2}\,(5^k + 1 + S_k) \qquad \text{and (55)} \qquad (57)$$

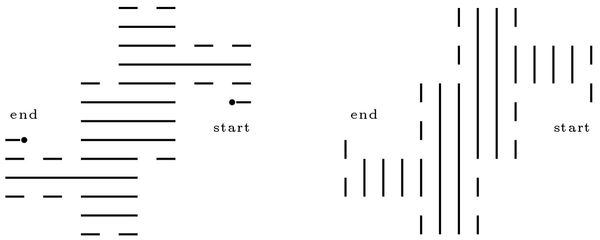
If there were no singles then it would be D doubles = P distinct = $\frac{1}{2}(5^k + 1)$ half the total points. Every 2 singles reduces the doubles by 1 and increases the distinct points by 1 (as +2 singles, -1 double).

As noted above S_k is the same parity as the total points $5^k + 1$ so S_k is even and (56),(57) are integers.

P_k is always even because the R5 curve is symmetric in 180° rotation about its midpoint $b^k/2$ which is the midpoint of the middle segment. Every vertex has a 180° rotated partner, making the total even.

4.2 Lines

Some line segments in the R5 dragon curve are consecutive and they can be considered in runs making horizontal and vertical lines.



k=3
total lines
$Lines_3 = 53$

$Horizontals_3 = 27$ lines $\quad$ $Verticals_3 = 26$ lines

The end of each line is a single-visited point. Since the curve always turns left or right 90°, every single-visited point is an end of a vertical and an end of a horizontal, except curve start and end. Curve start is only an end of a horizontal since the first segment is horizontal. Curve end is also only an end of a horizontal since the curve is identical in 180° rotation.

$$Horizontals_k = \tfrac{1}{2}S_k \qquad = 3^k \qquad \text{A000244}$$
$$Verticals_k = \tfrac{1}{2}S_k - 1 = 3^k - 1 \qquad \text{A024023}$$
$$Lines_k = Horizontals_k + Verticals_k = 2.3^k - 1 \qquad \text{A048473}$$

Join points where the curve touches on unfolding have both a horizontal and a vertical on each side so the two are reduced the same each. There is one join point for each join square J_k. The unfolding gives 3 copies the same direction or 180°, and 2 copies turned 90° so

$$\begin{aligned} Horizontals_k &= 3\,Horizontals_{k-1} + 2\,Verticals_{k-1} - 4J_{k-1} \\ Verticals_k &= 2\,Horizontals_{k-1} + 3\,Verticals_{k-1} - 4J_{k-1} \end{aligned}$$

4.3 Point Numbers

Single-visited and double-visited points n can be characterized by the $other(n)$ state machine from theorem 5. If "unch" is reached then there is another visit in the same arm and so a double point. This can be considered within k many digits for level k, or with infinite high 0 digits for the curve continued infinitely.

$$\begin{aligned} Dpred_k(n) &= 1 \text{ if } other \text{ reaches "unch" in } k \text{ digits} \\ Dpred_\infty(n) &= 1 \text{ if } other \text{ reaches "unch" with high 0s} \\ &= 0,0,0,0,1,0,0,0,1,1,0,0,1,1,0,0,1,1,\ldots \\ =1 \text{ at } n &= 4,8,9,12,13,16,17,19,20,21,23,24,28,29,\ldots \end{aligned}$$

The opposite is single-visited points

$$\begin{aligned} Spred_k(n) &= 1 - Dpred_k \\ Spred_\infty(n) &= 1 - Dpred_\infty \\ &= 1,1,1,1,0,1,1,1,0,0,1,1,0,0,1,1,0,0,\ldots \end{aligned}$$

$$=1 \text{ at } n{=}\, 0, 1, 2, 3, 5, 6, 7, 10, 11, 14, 15, 18, 22, 25, \dots$$

other is expressed by digits low to high. Usual state machine manipulations can go high to low for *Dpred* instead, but that is more complicated. Separating left turn doubles from right turn doubles simplifies a little.

Another approach for *Dpred* is to consider new doubles formed by expansion as from figure 15. An adjacent side in the preceding level gives a new double. In right boundary sides from theorem 13, states Y, Z and fully enclosed have a preceding adjacent side, so those with next digit 1 give double-visited points. Similarly states X, Y and fully enclosed with next digit 2. In both cases they are doubles with a left turn. Left side boundary states give doubles with right turns.

At a double-visited point the second visit is $other(n) > n$. In the *other* state machine this is when the transition to “unch” has a bigger output digit than its input digit. These are transitions out of lm1 and rm1.

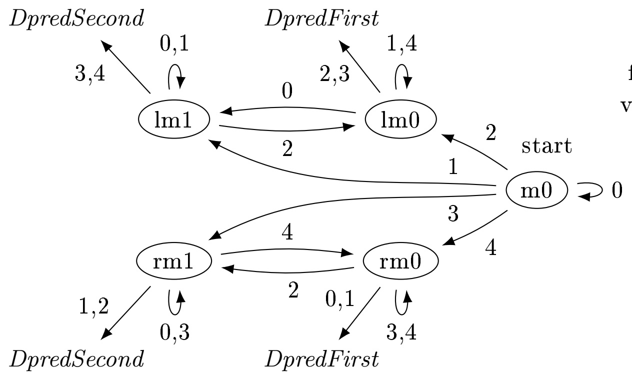


Figure 17:
first and second
visits to a double

$$\begin{aligned}
\mathit{DpredFirst}_\infty(n) &= 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, \dots \\
&=1 \text{ at } n{=}\, 4, 9, 12, 17, 19, 20, 23, 24, 29, 34, 37, 42, \dots \\
\mathit{DpredSecond}(n) &= 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, \dots \\
&=1 \text{ at } n{=}\, 8, 13, 16, 21, 28, 33, 38, 39, 40, 41, 43, 46, \dots
\end{aligned}$$

The number of visits to a location is 1 for a single-visited or 2 for double-visited,

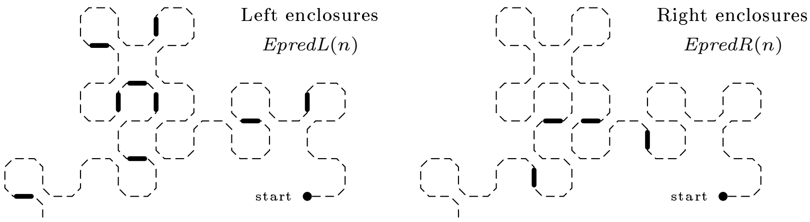
$$\begin{aligned}
\mathit{Visits}_k(n) &= \mathit{Dpred}_k(n) + 1 \qquad (58) \\
\mathit{Visits}_\infty(n) &= \mathit{Dpred}_\infty(n) + 1 \\
&= 1, 1, 1, 1, 2, 1, 1, 1, 2, 2, 1, 1, 2, 2, 1, 1, 2, 2, 1, \dots
\end{aligned}$$

Or which visit each n is, 1 for the first visit or 2 for the second,

$$\begin{aligned}
\mathit{VisitNum}(n) &= \mathit{DpredSecond}(n) + 1 \\
&= 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 2, 1, 1, 2, 1, 1, \dots
\end{aligned}$$

4.4 Enclosure Sequence

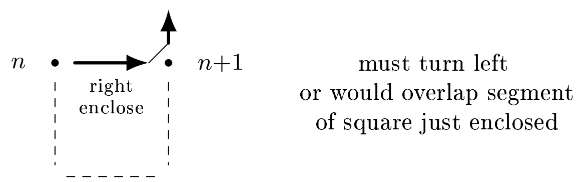
As each segment is successively appended to the R5 dragon curve it may enclose a new unit square on the right or left of the curve, or not.



A segment encloses a new unit square when its end is the second visit to that point.

$$\begin{aligned} &\mathit{Epred}(n) = \mathit{DpredSecond}(n{+}1) \\ &\quad =1 \text{ at } \ n = 7, 12, 15, 20, 27, 32, 37, 38, 39, 40, \ldots \end{aligned}$$

The unit square may be on the left or right of the segment. The turn at $n{+}1$ must be away from the square, or the next segment would overlap the square.



$$\begin{aligned} &\mathit{EpredR}(n) = \mathit{DpredSecond}(n{+}1) \text{ and } \mathit{turn}(n{+}1) = 1 \text{ (left)} \\ &\quad =1 \text{ at } \ n = 15, 20, 40, 45, 65, 70, 75, 76, 79, 80, \ldots \\ &\mathit{EpredL}(n) = \mathit{DpredSecond}(n{+}1) \text{ and } \mathit{turn}(n{+}1) = -1 \text{ (right)} \\ &\quad =1 \text{ at } \ n = 7, 12, 27, 32, 37, 38, 39, 42, 52, 57, \ldots \end{aligned}$$

Testing $n{+}1$ with the *Dpred* state diagram at figure 17 means the low digit is reduced, similar to next *turn* from figure 1. For *Epred* low digit 0 or 1 of n is 1 or 2 of $n{+}1$ and goes to lm. Similarly rm. A run of low 4 digits in n increments to low 0 digits in $n{+}1$ so digit 4 loops in state m0. The other states are unchanged.

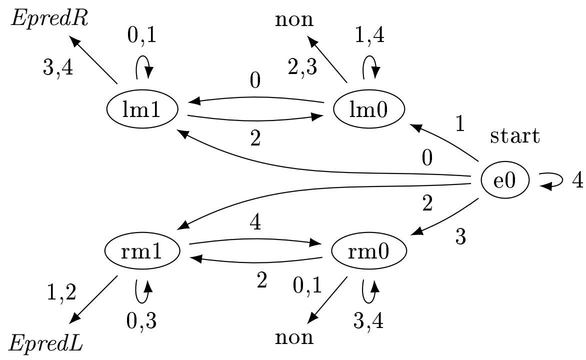
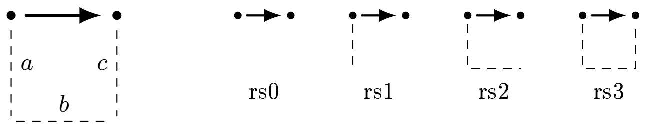


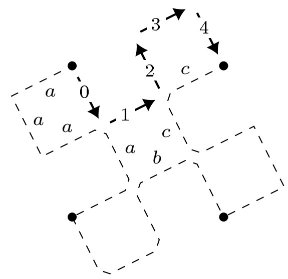
Figure 18:
Epred
low to high

$$\begin{aligned} \mathit{EpredR}(n) &= 1 \text{ iff reach } \mathit{EpredR} \\ \mathit{EpredL}(n) &= 1 \text{ iff reach } \mathit{EpredL} \end{aligned}$$

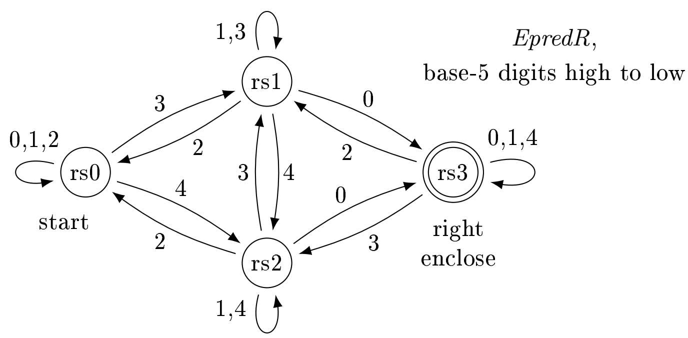
Instead of going from *Dpred*, the states can be found directly by considering segments beside. This is similar to the boundary segment numbers but for the enclosure sequence any later segment is not considered, just those preceding.



rs0 has no side segments. This is initial segment n=0. *rs3* has 3 segments on its right and encloses a new unit square. The line segments expand



This expansion gives the following transitions between configurations when a new low digit is added to n. For example *rs1* has the a side segment and with it the expanded segment 0 is an *rs3*. So *rs1* with a new 0 digit goes to *rs3*.



rs0,*rs1*,*rs2* *rs3* are the only combinations of side segments which occur in the expansions. c cannot occur without b since it would require the curve to curl around on the left to reach c, and that would overlap 4-arm plane filling. Likewise b cannot occur without a.

Usual DFA state machine manipulations can reverse to match digit strings low to high, and that is the pair of states lm0 and lm1 in figure 18. For *EpredL* similar configurations and transitions are made.

In low to high figure 18, for *EpredR* any 1 digit stays in lm0 or in lm1. n with 1 digits ignored can be treated as starting in lm0 since transitions are the same as from m0, with low $2, 3$ being not *Epred*. So digit runs (similar to the second proof of theorem 6)

lm1		*lm0*		*lm1*		*lm0*	
00…00	0	44…44	2	00…00	0	44…44	*EpredR* *DeleteOnes*(n)
3, 4 to enc		2, 3 to not-enc		3, 4 to enc		2, 3 to not-enc	

Likewise *EpredL* digit 3 does not change the *rm* state. So with all 3s ignored

rm1		*rm0*		*rm1*		*rm0*	
00…00	2	44…44	4	00…00	2	44…44	*EpredL* *DeleteThrees*(n)
1, 2 enc		0, 1 not-enc		1, 2 enc		0, 1 not-enc	

Theorem 20. *n which is an enclosure is given in terms of digits by*

$$\mathit{EpredR}(n) = \begin{cases} \textit{take } n \textit{ in base-5, delete all 1-digits} \\ 1 \quad \textit{if pair } 30 \textit{ or } 40 \textit{ with no } 3 \textit{ below} \\ \quad\;\; \textit{and any } 2s \textit{ below are in pairs } 20 \\ 0 \quad \textit{otherwise} \end{cases} \tag{59}$$

$$\mathit{EpredL}(n) = \begin{cases} \textit{take } n \textit{ in base-5, delete all 3-digits} \\ 1 \quad \textit{if pair } 10, 12, 20, 22 \textit{ with no } 1 \textit{ below} \\ \quad\;\; \textit{and any } 0s \textit{ at and below are in runs } 00\ldots02 \\ 0 \quad \textit{otherwise} \end{cases}$$

As an example, $n = 2009$ is base-5 "31014". For *EpredR* delete 1s to 304. It has 30 and no 2, 3 below it to consider, so $\mathit{EpredR}(2009) = 1$.

For *EpredL*, a 0 in the pairs is included in the 00…02 requirement. For example, $n = 2144$ is base-5 "32034" and delete 3s to 204. It has a 20 pair, but its 0 is 04 which is not an 0…02 run, so $\mathit{EpredL}(2144) = 0$.

Proof. For *EpredR*, enclosing is reached only by 3, 4 from lm1. That state is distinguished by having a 0 below for pairs 30, 40. The 0 is either from the run in lm1 or the transition to lm1.

Consider the lowest 30, 40 pair. Any 3 below it is in lm0 so is not enclosing. Any 2 below it is either in lm0 not enclosing or is the transition from lm1 to lm0. The latter is distinguished by having a 0 of lm1 or transition below it (lm1 being all 0s below the lowest 30, 40 pair).

For *EpredL*, enclosing is reached only by 1, 2 from rm1. That state is distinguished by having a 0 below from the rm1 run, or 2 below from the transition to it.

Consider the lowest 10, 12, 20, 22 pair. Any 1 below it is in rm0 so is not enclosing. Any 0 below it could be rm1 or rm0. The former is distinguished by being an rm1 run 00…02 (any rm1 below the lowest pair is all 0s). ☐

EpredR requirement 30, 40 can be compared to *Rpred* from theorem 13. *Epred* must have 30, 40 and *Rpred* must not, because an n which encloses on its right is certainly not right boundary.

Segments which are not *EpredR* and also not *Rpred* are going to be in an enclosed square, but one completed later. This will be before the next boundary segment since after a right boundary segment the curve cannot reach back to the right side of anything earlier. (If it did so on the right then the right boundary segment would not be on the boundary. If it did so by the left then it would overlap the four-arm plane filling.)

In *EpredR*, flipping n pairs $20 \leftrightarrow 40$ allows the digit conditions to be expressed in terms of lowest 2 or 3.

$$\mathit{EpredR}(n) = \begin{cases} 1 & \mathit{Tperm}(\mathit{DeleteOnes}(n)) \text{ lowest 2 or 3 is 20 or 30} \\ 0 & \text{otherwise} \end{cases} \tag{60}$$

$$\begin{aligned} \mathit{Tperm}(n) &= \text{flip base-5 pairs } 20 \leftrightarrow 40 \\ &= 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 20, 11, \ldots \end{aligned}$$

Conditions (59) are already lowest 3 must be 30. After flipping, the lowest 2 is either 20 from a flipped 40, or is an unchanged 2 because it didn't have a 0 below. The latter is non-enclose.

Total enclosures n=0 to 5^k-1 inclusive is the area AR_k on each side of the curve from theorem 16.

$$\sum_{n=0}^{5^k-1} \mathit{EpredR}(n) \;=\; \sum_{n=0}^{5^k-1} \mathit{EpredL}(n) \;=\; AR_k \;=\; AL_k$$

This *EpredR* sum can be calculated from the digit conditions. Doing so is more complicated than theorem 16 but is a combinatorial interpretation for the area. It's convenient to use *Tperm* form (60). An *EpredR* is digit 2 or 3, possible 1s, then 0 and no other 2 or 3 below.

$$\begin{array}{rccccc} k \geq 2 \text{ digits} & & & & & \\ k = h+1+s+1+t & \boxed{\ldots\text{any}\ldots} & \boxed{2,3} & \boxed{1\ldots1} & \boxed{0} & \boxed{\ldots0,1,4\ldots} \\ \text{combinations} & 5^h & 2 & 1^s & & 3^t \end{array}$$

$$\begin{aligned} 2\sum_{h=0}^{k-2} 5^h \sum_{t=0}^{k-2-h} 3^t &= 2\sum_{h=0}^{k-2} 5^h \, \frac{3^{k-1-h}-1}{3-1} \\ &= 3\frac{5^{k-1}-3^{k-1}}{5-3} - \frac{5^{k-1}-1}{5-1} = AR_k \end{aligned} \tag{61}$$

The powers in (61) are similar to area form (50). Adding 5^{k-1} to both terms (positive and negative) in (61) raises them to 5^k.

EpredR can enclose up to 3 unit squares consecutively. There cannot be 4 or more consecutive *EpredR* or that would be 4 left turns and segments would overlap. Similarly *EpredL*.

Some state machine manipulations can test whether n+1 is also the respective enclosure, then intersection $n, n{+}1, n{+}2$ for a triple. Taking that low to high shows they are the original digit forms with extra low.

$$\mathit{EpredR3} = \overset{\text{high}}{\boxed{\mathit{EpredR}}}\boxed{0 \text{ or } 4}\overset{\text{low}}{\boxed{4}} \qquad \mathit{EpredL3} = \overset{\text{high}}{\boxed{\mathit{EpredL}}}\overset{\text{low}}{\boxed{2}}$$

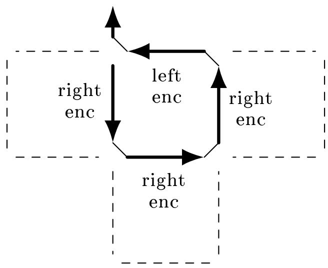
$$\mathit{EpredR3}(n) = \mathit{EpredR}(n) \text{ and } \mathit{EpredR}(n{+}1) \text{ and } \mathit{EpredR}(n{+}2)$$

$=1$ at $n = 379, 399, 504, 524, 1004, 1024, 1129, 1149, \ldots$

$$\mathit{EpredL3}(n) = \mathit{EpredL}(n) \text{ and } \mathit{EpredL}(n{+}1) \text{ and } \mathit{EpredL}(n{+}2)$$

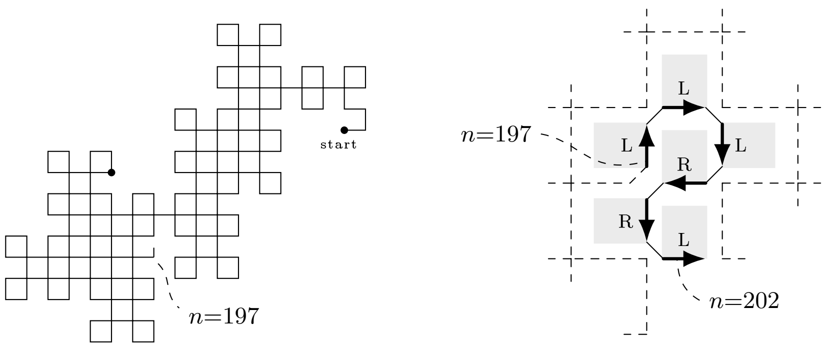
$=1$ at $n = 37, 62, 137, 162, 187, 192, 197, 212, \ldots$

When 3 consecutive *EpredR* occur, the next segment is always a left enclosure, since there were 3 left turns. Likewise 3 consecutive *EpredL* is always followed by a right enclosure. These are the 3 turns per *TurnRuns3* from (52).



3 right enclosures
are 3 left turns
so next segment
is a left enclosure

Runs of right and left enclosures can occur. For example at n=197 there is a run of 6 consecutive enclosures LLL,RRL.



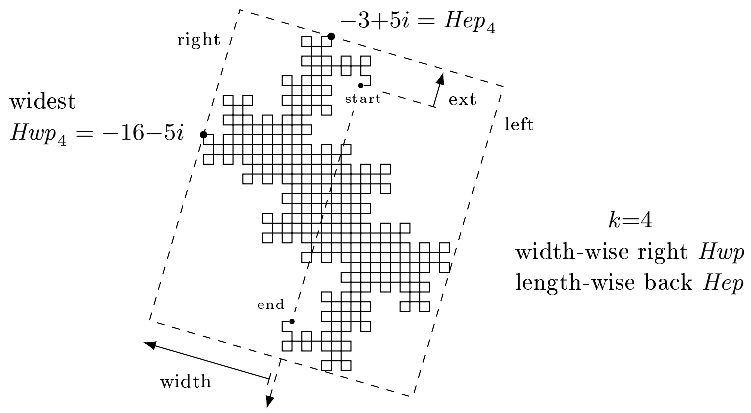
There are no runs longer than 6. This can be seen by some state machine manipulations on *Epred* to test whether n+1, n+2 etc are enclosing. The intersection of *Epred* on 7 terms n through n+6 inclusive is empty.

State machine manipulations on the 6 intersection shows it is *EpredL* with extra low digits,

	high	low
EpredSix =	*EpredL*	4 2

The last point of a level is single-visited so the last segment is non-enclosing and a run of enclosures does not extend across levels. From the digits, the number of runs of 6 in level k is simply AL_{k-2}, for $k \geq 2$.

5 Width and Length



Theorem 21. *The widest point on the right side of R5 dragon k, meaning furthest on the right side perpendicular to a line start to end, is located at*

$$\begin{aligned} Hwp_k &= b^{k-1} + ib^{k-2} + i^2 b^{k-4} + i^3 b^{k-5} + i^4 b^{k-7} + \cdots + i^{some} b^0 \\ &= \sum_{d=1}^{k-\lfloor ad\rfloor \geq 0} i^{d-1}\, b^{k-\lfloor ad\rfloor} \quad \textit{and always } b^0 \textit{ term when } k>0 \qquad (62) \\ &= 0,\ 1,\ 1{+}3i,\ -6{+}5i,\ -16{-}5i,\ -6{-}38i,\ 71{-}50i, \ldots \end{aligned}$$

$$\textit{where}\ \ a = \frac{\pi}{2\arctan 2} = \frac{\pi/2}{\arg b} = 1.418776\ldots$$

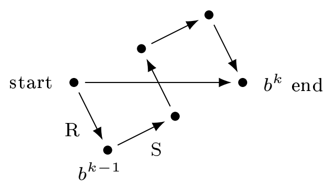
b *exponents* 1,2, 4,5, 7,8,9, 11,12, 14,15, 17,18,19, ...

The furthest point length-wise back in direction parallel to start to end is

$$\begin{aligned} Hep_k &= b^{k-2} + ib^{k-4} + i^2 b^{k-5} + i^3 b^{k-7} + \cdots + i^{some} b^0 \\ &= \sum_{d=2}^{k-\lfloor ad\rfloor \geq 0} i^{d-2}\, b^{k-\lfloor ad\rfloor} \quad \textit{and always } b^0 \textit{ term when } k>0 \qquad (63) \\ &= 0,\ 0,\ 1,\ 1{+}3i,\ -3{+}5i,\ -14{-}i,\ -12{-}30i, \ldots \end{aligned}$$

In (62),(63) the sums are some powers of b, with successive powers of i. The b powers range from $k{-}1$ to 0 inclusive. The powers are spread and some skipped by $\lfloor ad\rfloor$. There is at most 1 skip between included powers. The lowest term b^0 is always included, even when not among the $k{-}\lfloor ad\rfloor$. The terms in Hwp and Hep are the same but Hep starts d=2 so omits the first.

Proof. Curve k comprises 5 sub-curves $k{-}1$.



For *Hwp*, sub-curve R is the furthest on the right among the sub-curves parallel to R. Similarly S is furthest to the right among sub-curves parallel to S. The point furthest to the right must be in either R or S.

R and S in turn comprise sub-curves $k-2$.

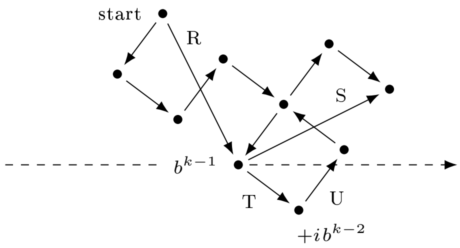


Figure 19: $k-2$

Among the sub-curves of the two orientations, T and U are the furthest to the right. T is the first sub-curve of S so its direction is $+90°$ as compared to the first segment of R, so an extra factor i for $+ib^{k-2}$.

Similarly in the next expansion $+i^2 b^{k-3}$ the sub-curves of U. The segments turn by $+\arctan\frac{1}{2}$ each time. At $k-4$ they turn enough that the furthest on the right is instead in the first sub-curve V,

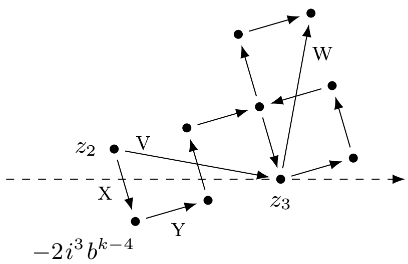


Figure 20: $k-4$ angle enough that sub-curves of first are furthest right

Offset $-2i^3 b^{k-4}$ is back from z_3 to the end of X. The offset has extra factor i since its direction is parallel to the first sub-curve of W, which is $+90°$ as compared to V.

X and Y are in sub-curve V so their directions do not gain a factor of i. On expansion the next term is then $+i^3 b^{k-5}$, the same power of i as the -2 offset here.

This procedure repeats. Each time there are two sub-curves at right angles and on expansion whichever new two sub-curves are furthest right are taken. Sometimes the angle is enough to make it the first original, like figure 20.

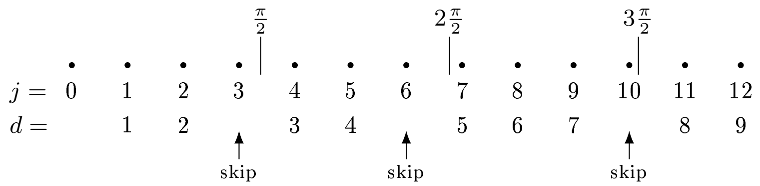
For computer calculation this can be followed as an algorithm (the powers of b calculated exactly). Formula (62) is written by noting that a step and a following -2 step back can be combined

$$+i^2 b^{k-3} \quad -2i^3 b^{k-4} \qquad = +i^2 b^{k-4}$$

In figure 20, this means don't go to z_3 for V–W but instead straight from z_2 to the end of X by $+i^2 b^{k-4}$. Done this way, each term always has a new factor i, which is index d in (62).

The skipped powers b^{k-j} are where $j.\gamma$ precedes a multiple of $\frac{\pi}{2}$, where

$$\begin{aligned} \gamma = \tfrac{\pi}{2} - \arg b \ = \arctan\tfrac{1}{2} \ &= 26.565051° \ldots \\ &= 0.463647\ldots \quad \text{radians} \qquad \text{A073000} \end{aligned}$$



Multiplying through by $1/\gamma$ becomes skip integers j preceding a multiple of

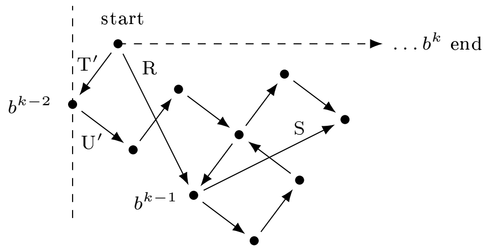
$$\begin{aligned} h &= \tfrac{\pi}{2}/\gamma \;= \frac{\pi}{2\arctan\frac{1}{2}} = 3.387909\ldots \\ \text{skip}\;\; j &= \lfloor mh \rfloor \text{ for integer } m \\ &= 0, 3, 6, 10, 13, 16, 20, 23, 27, 30, 33, 37, \ldots \end{aligned}$$

$\lfloor mh \rfloor$ is a Beatty sequence. Its complement is the b powers to include. Per Rayleigh [10] and Beatty [3], this is $\lfloor ad \rfloor$ for integer d and factor a from h by

$$\frac{1}{a} + \frac{1}{h} = 1$$

A skip is to be done only when the next step back will cancel it out. At b^0 there is no lower term which will step back. So b^0 should not be skipped, but included, with its next i power. Hence the rule to always include term b^0 in sums (62),(63).

For *Hep* extending back from the start, a similar argument is made. The first point extending back past the curve start is b^{k-2},



T′ and U′ are the same angles relative to the vertical as T and U are to their horizontal in figure 19. So further terms of *Hep* follow in the same pattern as in *Hwp*. The effect in (63) is to start at d=2. □

The sub-curves show point Hwp_k is unique for $k \geq 1$. For k=0, the curve is 1 segment and its two points $z = 0, 1$ are both width 0 away. $Hwp_0 = 0$ is chosen so that the b powers summed range from 0 to $k{-}1$ in all cases, and that range understood as empty when k=0.

For the curve scaled to endpoints a unit length, the limits are

$$\frac{Hwp_k}{b^k} \to Hwpf = \sum_{d=1}^{\infty} \frac{i^{d-1}}{b^{\lfloor ad \rfloor}} \;= 0.385314\ldots - 0.576988\ldots i \qquad (64)$$

$$\frac{Hep_k}{b^k} \to Hepf = \sum_{d=2}^{\infty} \frac{i^{d-2}}{b^{\lfloor ad \rfloor}} \;= -0.176988\ldots - 0.185314\ldots i$$

$$= -i\left(Hwpf - \frac{1}{b}\right)$$

The special handing of the low b^0 does not apply since its offset becomes ever smaller with k.

The width part of $Hwpf$ can be written with factor i put through (64) to give a Re sum. The extension part of $Hepf$ similarly with i^2 for the same sum but starting d=2.

$$\operatorname{Im} Hwpf = -\operatorname{Re}\sum_{d=1}^{\infty} \frac{i^d}{b^{\lfloor ad \rfloor}} = -0.576988\ldots \qquad \text{width (down)}$$

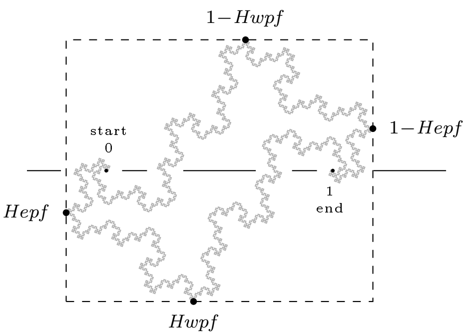
$$\operatorname{Re} Hepf = -\operatorname{Re}\sum_{d=2}^{\infty} \frac{i^d}{b^{\lfloor ad \rfloor}} = -0.176988\ldots \qquad \text{extension (back)}$$

The terms in these sums are all positive since they are steps to successive widest points. If just magnitudes are taken and skipped powers included then a simple upper bound is

$$|Hwpf| < \sum_{d=1}^{\infty} \frac{1}{\sqrt{5}^d} = \frac{1+\sqrt{5}}{4} = 0.809016\ldots \qquad \text{A019863}$$

This is also a bound on omitted terms if stopping at some $1/b$ power.

Width and extension give the area of the bounding rectangle around the curve and aligned to the curve endpoints,



width and extension bounding rectangle

$$HewA_k = 5^k \left(-2\operatorname{Im}\frac{Hwp_k}{b^k}\right)\left(1 - 2\operatorname{Re}\frac{Hep_k}{b^k}\right)$$
$$= 0, 4, \tfrac{806}{25}, \tfrac{21306}{125}, \tfrac{574454}{625}, \ldots$$

$$\frac{HewA_k}{5^k} \to HewAf = \big(-2\operatorname{Im} Hwpf\big)\big(1 - 2\operatorname{Re} Hepf\big)$$
$$= 1.562458\ldots$$

Point numbers n of locations Hwp and Hep follow from the segment steps. Each step taken is digit 1 and each skip is digit 0.

$$Hdigit(j) = \begin{cases} 1 & \text{if } (j \bmod h) < h{-}1 \\ 0 & \text{otherwise} \end{cases} \tag{65}$$

$$= \left\lfloor (j{+}1)\tfrac{1}{a} \right\rfloor - \left\lfloor j\tfrac{1}{a} \right\rfloor \tag{66}$$

At (65), $j \bmod h$ means remainder r from division by h, so $j = qh + r$ with integer q and $0 \le r < h$. j immediately preceding a multiple of h is a skip, which is $r > h{-}1$. Conversely steps taken are $r < h{-}1$. Since b is never a multiple of a full circle, per (15), h is irrational and never have remainder exactly $r = h{-}1$.

This is the "characteristic word" of $1/a$ per floors at (66), or Sturmian word (homogeneous since no offset to the multiples), or Christoffel word (which for irrational slope is a Sturmian word).

$$\frac{1}{a} = \frac{\arctan 2}{\pi/2} = 0.704832\ldots$$

The geometric interpretation is 1 when power b^{j+1} crosses into a new quadrant from where b^j was. The floors at (66) are the usual interpretation of a Sturmian word as when a straight line of slope $1/a$ crosses up to a new integer part.

In *Hwp*, the lowest b^0 is never skipped, so always a 1 digit, when $k > 0$.

$$\begin{aligned} \mathit{HwN}_k &= \text{base-5} \begin{cases} \text{no digits} & \text{if } k{=}0 \\ \mathit{Hdigit}(1\ldots k{-}1),\ 1 & \text{if } k \ge 1 \end{cases} \\ &= 0,\ 1,\ 11,\ 111,\ 1101,\ 11011,\ 110111,\ 1101101, \ldots \ \text{base-5} \\ &= 0, 1, 6, 31, 151, 756, 3781, 18901, \ldots \end{aligned}$$

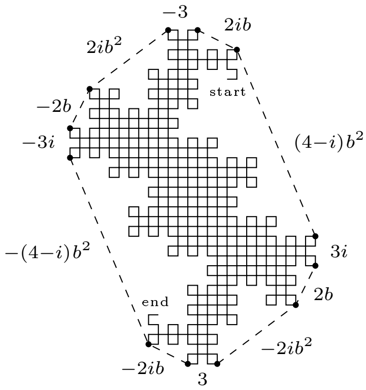
For *Hep*, the same but the high $j{=}1$ digit is 0.

$$\begin{aligned} \mathit{HeN}_k &= \begin{cases} 0 & \text{if } k \le 1 \\ \mathit{HwN}_k - 5^{k-1} & \text{if } k \ge 2 \end{cases} \\ &= 0,\ 0,\ 1,\ 11,\ 101,\ 1011,\ 10111,\ 101101, \ldots \ \text{base-5} \\ &= 0, 0, 1, 6, 26, 131, 656, 3276, \ldots \end{aligned}$$

For the curve scaled to endpoints a unit length, the limit fractional location along the curve is all of *Hdigits*. The low digit k exception does not apply. So *HwNf* is the characteristic word as base-5 fractional digits.

$$\begin{aligned} \frac{\mathit{HwN}_k}{5^k} \to \mathit{HwNf} = \sum_{j=1}^{\infty} \frac{\mathit{Hdigit}(j)}{5^j} &= 0.110110111\ldots && \text{base-5} \\ &= 0.241935\ldots && \text{decimal} \\ \frac{\mathit{HeN}_k}{5^k} \to \mathit{HeNf} = \sum_{j=2}^{\infty} \frac{\mathit{Hdigit}(j)}{5^j} &= 0.010110111\ldots && \text{base-5} \\ &= 0.041935\ldots && \text{decimal} \\ &= \mathit{HwNf} - \tfrac{1}{5} \end{aligned}$$

6 Convex Hull



k = 4
convex hull,
12 sides

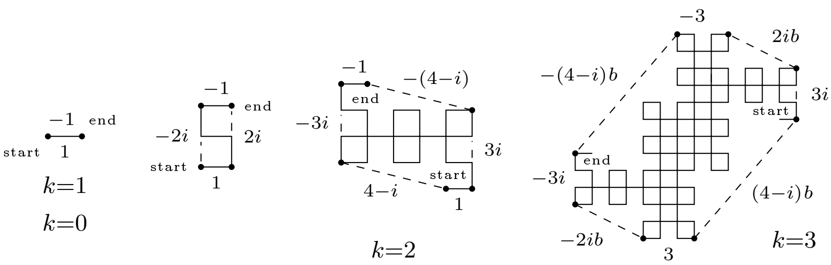
Theorem 22. *The sides of the convex hull around R5 dragon curve k are a set of complex number offsets (to be taken anti-clockwise in* arg *order),*

$$
\begin{aligned}
Hsides_0 &= \pm 1 && k=0\\
Hsides_1 &= \pm 1,\ \pm 2i && k=1\\
Hsides_2 &= \pm 1,\ \pm 3i,\ \pm(4-i) && k=2\\
Hsides_k &= \begin{cases} 3i^d & d=0 \text{ to } 3 \qquad k\ge 3\\ 2i^d b^j & d=0 \text{ to } 3 \text{ and } j=1 \text{ to } k-3\\ \pm 2ib^{k-2},\ \pm(4-i)b^{k-2} \end{cases} && \qquad (67)
\end{aligned}
$$

so number of sides

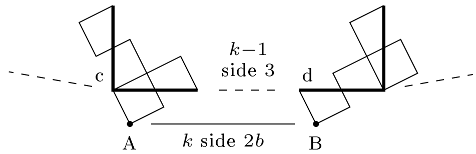
$$
\begin{aligned}
HnumSides_k &= \begin{cases} 2,4,6 & \text{if } k=0,1,2\\ 4k-4 & \text{if } k\ge 3 \end{cases}\\
&= 2,4,6,8,12,16,20,\dots \qquad 2\times \text{A004278}
\end{aligned}
$$

Proof. Hull sides for $k \le 3$ can be calculated explicitly.



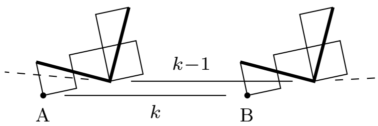
For $k \ge 4$, suppose the theorem is true of $k-1$, and further that the hull vertices are all at corners of two segments (not curve start or end), which is so of $k-1 = 3$.

The $3i^d$ sides of $k-1$ expand as follows

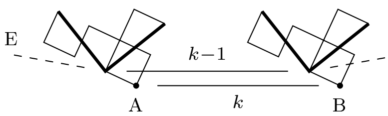


The $k-1$ segments are shown thick. They are the segments of furthest extent perpendicular to their hull side. So in the manner of theorem 21, the expanded A and B points are then furthest extent in k and thus a hull side of k. Those points are offsets from points c, d which are 2 apart in $k-1$ so expansion factor b gives sides $2i^d b$ in k. These are $j=1$ at (67).

The segments of $k-1$ are a square grid so those four $3i^d$ sides are the only sides aligned to the segments. The other sides have segments at some angle to the side. Their corner segments expand as follows

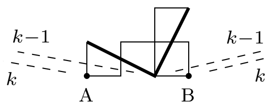


or



These corners are the maximum extent in $k-1$ perpendicular to the hull side. Again in the manner of theorem 21, points A and B are then new maximum extents so they are a side of the k hull. These points are at offset -2 in the left form, or $+1$ in the right form (according as the angle between the corner segments and the $k-1$ side). A and B are the same offset, so a further factor b on the $k-1$ side. This is sides $(4-i)b^{k-2}$ and $2b^j$ for $j \geq 2$.

At a vertex of $k-1$, it's possible to have a further new side in k. This happens when an offset -2 is followed by offset $+1$. This gives a new side $3i^d$.



The k segments are a square grid so there are 4 orientations this might happen. If A–B is co-linear with an adjacent side of k then it is not a new further side (only a continuation of another). But the sides $2b^j$ in k are not co-linear with A–B since from (15) the b powers never fall on the axes.

The sides $(4-i)b^{k-2}$ are not co-linear with A–B either, since $(4-i)b^{k-2}$ always has non-zero real and imaginary parts (in fact the same sequence as b powers).

$$\begin{aligned} &\operatorname{Re}(4-i)b^{k-2} \equiv \operatorname{Re} b^{k-2} \not\equiv 0 \mod 5 \qquad k \geq 3 \\ &\operatorname{Im}(4-i)b^{k-2} \equiv \operatorname{Im} b^{k-2} \not\equiv 0 \mod 5 \end{aligned} \qquad \square$$

The general case at (67) could combine the four-way 3 and $2b^j$ cases by reckoning coefficient 3 when $j=0$ and coefficient 2 when $j \geq 1$. The high $k-2$ case could be combined too by reckoning coefficient $4-i$ when $j = k-2$ and d even.

The hull boundary length follows from the sides. The powers of b are $\sqrt{5}$ lengths. $4-i$ is length $\sqrt{17}$.

$$\begin{aligned} HB_k &= \sum_{z \in Hsides_k} |z| \\ &= \begin{cases} 2,\ 6 & k=0,1 \\ \left(6+2\sqrt{5}+2\sqrt{17}\right)\sqrt{5}^{\,k-2}\ +2-2\sqrt{5} & k{\ge}2 \end{cases} \\ &= 2,\ 6,\ 8{+}2\sqrt{17},\ 12{+}4\sqrt{5}{+}2\sqrt{85},\ 32{+}8\sqrt{5}{+}10\sqrt{17},\ldots \end{aligned}$$

For curve endpoints scaled to a unit length, the limit set of sides is countably infinite

$$\begin{aligned} Hsidesf &= Hsides_k/b^k \qquad \text{as } k \to \infty \\ &= \begin{cases} 2i^d/b^j & \text{for } d=0 \text{ to } 3 \text{ and } j \ge 3 \\ \pm 2i/b^2,\ \pm(4{-}i)/b^2 \end{cases} \end{aligned}$$

and the hull boundary length limit is

$$\frac{HB_k}{\sqrt{5}^k} \to HBf = \frac{6+2\sqrt{5}+2\sqrt{17}}{5} \ = \frac{6}{5}+\frac{2}{\sqrt{5}}+\frac{2}{5}\sqrt{17} \ = 3.743669\ldots$$

Theorem 23. *The area of the convex hull around R5 dragon curve k is*

$$HA_k = \begin{cases} 0,2 & k=0{,}1 \\ 17.5^{k-2}-1+\displaystyle\sum_{j=1}^{k-2} \begin{array}{l} \left(3.5^{k-2-j}-1\right) HAgrow\left(2b^j\right) \\ +\ 2.5^{k-2-j}\ \ HAgrow\left((4{-}i)b^j\right) \end{array} & k \ge 2 \end{cases} \tag{68}$$

$$= 0, 2, 16, 106, 578, 2954, 15064, \ldots$$

$$HAgrow(z) = abs1or2\left(\operatorname{Im}\ z \,.\, i^{-quadrant\left((2+i)z\right)}\right) \tag{69}$$

$$abs1or2(x) = |x| \,.\begin{cases} 1 & \textit{if } x \ge 0 \\ 2 & \textit{if } x < 0 \end{cases}$$

$$quadrant(z) = \left\lfloor \frac{\arg z}{\pi/2} \right\rfloor \qquad quadrant \equiv 0 \text{ to } 3 \bmod 4 \ \textit{containing } z$$

The effect of *HAgrow* is to take either Re or Im of z and a factor 1 or 2 according to where z falls in the following sectors.

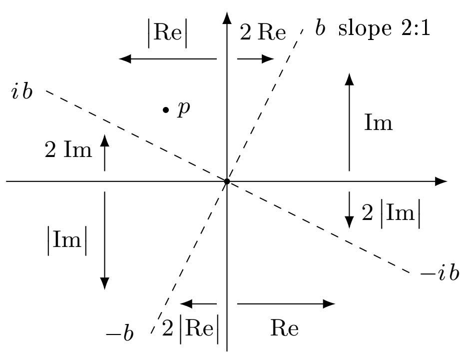


Figure 21:
HAgrow sectors
Re or Im
and factor 1 or 2

For example point p shown is in the $|\mathrm{Re}|$ sector so $HAgrow(p) = |\mathrm{Re}\, p\,|$.

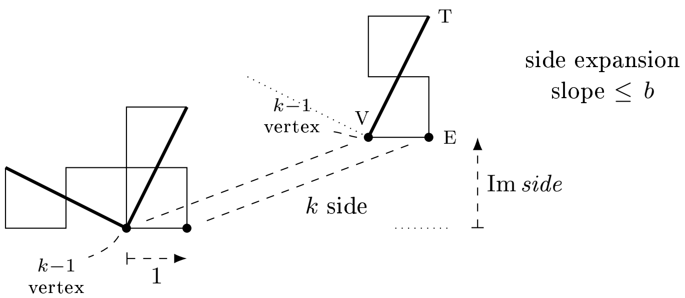
All measures are ≥ 0. The b line in the first quadrant is slope 2:1 so the adjacent sectors Im below or $2\,\mathrm{Re}$ above are the same result. Similarly the other b lines.

(69) is written with *quadrant* finding the quarter between b lines which contains z. Factor $2{+}i = -i\overline{b}$ rotates $-ib$ up to the x axis so quarter 0 starts at $-ib$. The resulting power of i rotates z to between $-ib$ (inclusive) and b (exclusive), and there Im or $2|\mathrm{Im}|$. As noted, the result on the b lines is the same from either adjacent quarter, so *quadrant* can take the x, y axes as part of either of their adjacent quadrants.

Proof of Theorem 23. Areas of hulls $k \leq 2$ can be calculated explicitly.

For $k \geq 3$, take the curve with its first segment directed East, per *Hsides*. The approach is to consider what extra area the k hull has over the $k{-}1$ hull expanded by factor b.

There is a bottom-most horizontal side in k. This is side $+3$ in Hsides_k. If it is followed by a side of slope $\leq b$ then the expansions from $k{-}1$ are



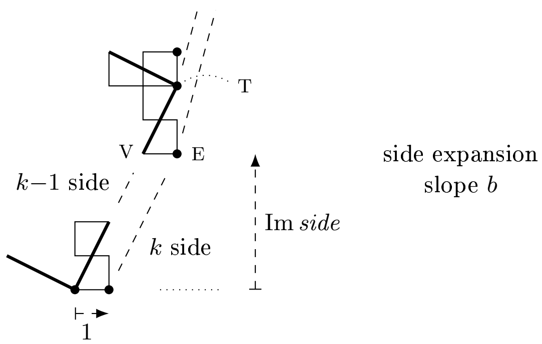
The $k{-}1$ segments are shown thick, then their resulting expansions for k. The $+3$ bottom side of k is the bottom left.

The $k{-}1$ hull points have factor b in the expansion, which is $\times|b| = \sqrt{5}$ lengths so $5\times$ area of the $k{-}1$ hull.

The k side is, from the segment expansions, 1 segment right of the hull side of $k{-}1$. The additional area over $k{-}1$ is thus a vertical rectangle width 1 which has been sheared to the right. Its height gives area $\mathrm{Im}\,\mathit{side}_k = \mathit{HAgrow}(\mathit{side}_k)$.

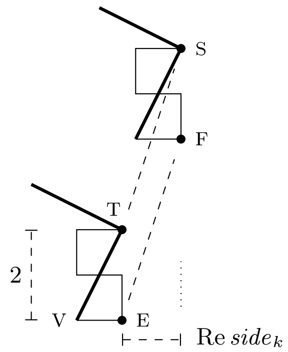
Further sides of slope $< b$ (if any) in k are sheared 1 segment rectangles too.

When the k side is slope b, the $k{-}1$ segments were co-linear



In this case, the $k{-}1$ hull vertex is at T rather than V. The rectangle wanted is up to V. Basing the sheared rectangle on E in the k side accomplishes this.

The next side is slope $> b$ as follows,



side expansion
slope $> b$

Area T–E upwards to S–F is a horizontal rectangle height 2 which has been sheared upwards. Its width gives area $2 \operatorname{Re} side_k = HAgrow(side_k)$.

When this side immediately follows a slope $\le b$, the area taken by the previous side extends only to V–E so triangle V–E–T is additional area 1 in that case.

These slopes $> b$ (if any) continue until a vertical side in hull k. There the pattern of side slopes and expansion repeats starting from the horizontal etc above, all turned $+90°$. Then likewise turned $+180°$ and $+270$ around the hull. So area increase is $HAgrow$ of each $Hsides_k$, plus 4 triangles area 1 each.

$$HA_k = 5 HA_{k-1} + 4 + \sum_{z \in Hsides_k} HAgrow(z) \qquad k \ge 3 \tag{70}$$

Expanding (70) down to $k{-}1 = 2$ gives $HAgrow$ on all $Hsides_k$ down to $Hsides_3$, with successive powers of 5. The set of $2b^j$ sides in each are

$$\begin{array}{rllllll}
 & i^d 2b & i^d 2b^2 & i^d 2b^2 & \cdots & i^d 2b^{k-3} & i^d 2b^{k-2} & \pm 2b^{k-2} \\
5\times & i^d 2b & i^d 2b^2 & i^d 2b^2 & \cdots & i^d 2b^{k-3} & \pm 2b^{k-2} \\
5^2\times & i^d 2b & i^d 2b^2 & i^d 2b^2 & \cdots & \pm 2b^{k-3} \\
\cdots \\
5^{k-5}\times & i^d 2b & i^d 2b^2 & \pm 2b^3 \\
5^{k-4}\times & i^d 2b & \pm 2b^2 \\
5^{k-3}\times & \pm 2b
\end{array}$$

The columns are $j{=}1$ through to $j = k{-}2$. Terms in a column are 4 directions i^d except the bottom-most only $\pm$. Since $HAgrow(z) = HAgrow(i.z)$, these directions do not change $HAgrow$.

So a column is $HAgrow(2b^j)$ with power of 5 factors and multiplicity 4 or 2. The total is per (68),

$$4.1 + 4.5 + 4.5^2 + \cdots + 4.5^{k-j-3} + 2.5^{k-j-2} = 3.5^{k-j-2} - 1$$

Each $Hsides$ has 2 sides $(4{-}i)b^j$. Its power of 5 is 2.5^{k-2-j} which is again per (68). The four sides $3i^d$ have $HAgrow = 0$. Constant 4 in (70) is powers $4.5^0 + \cdots + 4.5^{k-3} = 5^{k-2} - 1$. The starting $k{-}1 = 2$ is further $5^{k-2}.HA_2 = 16.5^{k-2}$, for total $17.5^{k-2} - 1$ in (68). □

Scaled to curve endpoints a unit length, the hull area has limit

$$\frac{HA_k}{5^k} \to HAf = \frac{17}{25} + \frac{1}{25}\sum_{j=1}^{\infty} \Big(6\,HAgrow\big((\tfrac{b}{5})^j\big) + 2\,HAgrow\big((4-i)(\tfrac{b}{5})^j\big)\Big)$$

$$= \frac{17}{25} + \frac{1}{25}\sum_{j=1}^{\infty} \left(6\,HAgrowf\Big(\frac{1}{b^j}\Big) + 2\,HAgrowf\Big(\frac{4+i}{b^j}\Big)\right) \qquad (71)$$

$$= 0.976164\ldots$$

$$HAgrowf(z) = HAgrow(\overline{z}) \;= abs2or1\Big(\mathrm{Im}\; z\,.\,i^{-quadrant(b.z)}\Big)$$

$$abs2or1(x) = |x|\,.\begin{cases} 2 & \text{if } x \geq 0 \\ 1 & \text{if } x < 0 \end{cases}$$

At (71), $b/5 = \overline{1/b}$, and *HAgrowf* accounts for the conjugate by taking the pattern of figure 21 in vertical mirror image.

Theorem 24. *A simple bound on terms of the HAf sum* (71) *is*

$$\sum_{j=k}^{\infty} HAtermf_j \;<\; \frac{1}{\sqrt{5}^k} \qquad (72)$$

where $HAtermfZ(z) = \frac{1}{25}\Big(6\,HAgrowf(z) + 2\,HAgrowf((4+i)z)\Big)$ *and* $HAtermf_j = HAtermfZ(1/b^j)$ *so that* $HAf = \frac{17}{25} + \sum_{j=1}^{\infty} HAtermf_j$.

Proof. *HAtermf* has z and $(4+i)z$ and working through the sectors they fall in, and consequent $|\mathrm{Re}\,z|$ or $2\,|\mathrm{Im}\,z|$ in *HAgrowf*, shows bound

$$HAtermfZ(z) \;\leq\; \frac{1}{25}\,\frac{26}{\sqrt{5}}\,|z|$$

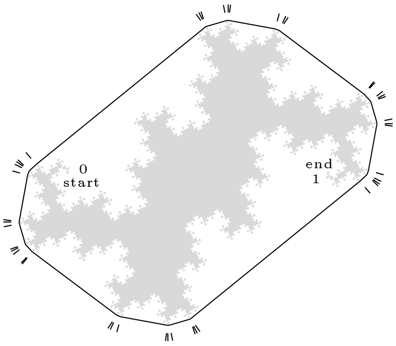
with equality when z is in direction $2+i$. Magnitudes $|1/b^j| = 1/\sqrt{5}^j$ are a geometric sum with total $\frac{1}{4}(5+\sqrt{5})\,\sqrt{5}^k$ so that together

$$\sum_{j=k}^{\infty} HAtermf_j \;\leq\; \frac{1}{25}\,\frac{26}{\sqrt{5}}\,\frac{(5+\sqrt{5})}{4}\,\frac{1}{\sqrt{5}^k}$$

and the factor is $\frac{13}{50}(1+\sqrt{5}) = 0.841377\ldots < 1$. □

Using *HAtermf* with its related z and $(4+i)z$ means the final factor is < 1 so is convenient to round up so the simple (72). Taking z and $(4+i)z$ separately in *HAgrowf* would be a factor bigger than 1. A slightly tighter bound could be made by using related directions of for instance consecutive $1/b^j$ and $1/b^{j+1}$.

HAf can be compared to the area limit $\frac{1}{2}$ of the curve it surrounds. Just under half is empty.

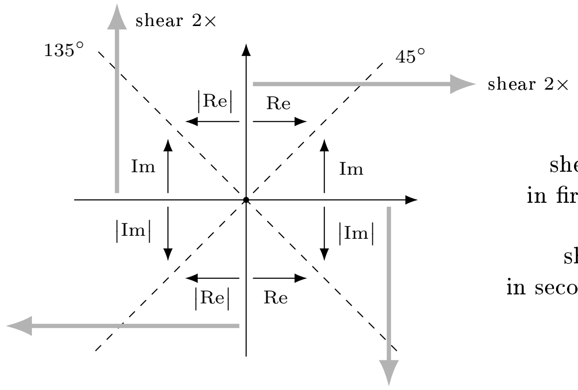


R5 hull limit
area 0.976...
boundary 3.743...
countably infinite
vertices

Another approach for *HAgrow* is to take the factor 2 in *abs1or2* as a factor on the real part in the first and third quadrants, and the imaginary part in the second and fourth, as a kind of quadrant-dependent shear.

$$HAgrow(z) = \min\Big(\,|\mathrm{Re}\,z'|, |\mathrm{Im}\,z'|\,\Big)$$

$$\text{where sheared } z' = \begin{cases} 2(\mathrm{Re}\,z) + (\mathrm{Im}\,z)i & \text{if Re, Im same signs} \\ (\mathrm{Re}\,z) + 2(\mathrm{Im}\,z)i & \text{if Re, Im different signs} \end{cases}$$



shear horizontally for z
in first and third quadrants

shear vertically for z
in second and fourth quadrants

The effect of the shear is to move the slope 2:1 line b or $i\,b$ etc in figure 21 to become 45° here. The shear applies the desired factor 2, and also allows min(Re, Im) to select between the horizontal or vertical measure.

7 Centroid

The R5 dragon is symmetric in 180° rotation so the centroid of the segments, points or area are all the curve midpoint $\frac{1}{2}b^k$. Some measures can be made on just one side of the curve.

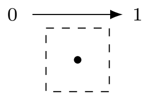
Theorem 25. *The centroid of the right boundary squares of R5 dragon curve k is*

$$\begin{aligned} GRQ_k &= \left(\tfrac{2}{5} - \tfrac{1}{5}i\right) b^k \;+\; \left(\tfrac{1}{10} - \tfrac{3}{10}i\right)\left(\tfrac{1}{3}\right)^k \\ &= \tfrac{1}{2} - \tfrac{1}{2}i,\;\; \tfrac{5}{6} + \tfrac{1}{2}i,\;\; -\tfrac{7}{18} + \tfrac{13}{6}i,\;\; -\tfrac{259}{54} + \tfrac{25}{18}i,\;\; -\tfrac{1231}{162} - \tfrac{443}{54}i,\;\; \dots \end{aligned}$$

$b = 1 + 2i$

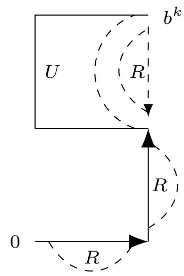
Proof. For k=0 the curve is a single line segment with a single boundary square.

k=0



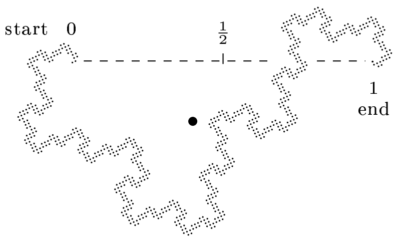
$$GRQ_0 = \frac{1}{2} - \frac{1}{2}i$$

As in figure 10, the boundary squares in a U part are a reversal of the R part, so the centroid is the mean of the three copies of R in the previous level.



$$\begin{aligned}
GRQ_k &= \tfrac{1}{3}\big(GRQ_{k-1} \ + \ b^{k-1} + i\,GRQ_{k-1} \ + \ b^k - i\,GRQ_{k-1}\big) \\
&= \tfrac{1}{3}\,GRQ_{k-1} \ + \ \tfrac{1}{3}(2{+}2i)\,b^{k-1} \\
&= GRQ_0 . \left(\tfrac{1}{3}\right)^k \ + \ \tfrac{1}{3}(2{+}2i)\sum_{j=0}^{k-1}\left(\tfrac{1}{3}\right)^j b^{k-1-j} \\
&= \left(\tfrac{1}{2} + \tfrac{1}{2}i\right)\left(\tfrac{1}{3}\right)^k \ + \ \tfrac{1}{3}(2{+}2i)\frac{\left(\tfrac{1}{3}\right)^k - b^k}{\tfrac{1}{3} - b}
\end{aligned}$$ $\square$

Scaling the endpoint b^k to a unit length, the limit GRQ_k/b^k as $k \to \infty$ is the coefficient of the b^k term. Notice this is not the middle horizontally, but a little towards the start at $\frac{2}{5}$.



right boundary
centroid limit
$$\frac{2}{5} - \frac{1}{5}i$$

This limit is the same as the whole Heighway/Harter dragon curve centroid (its whole curve, not just the boundary squares).

7.1 Centroid of Join

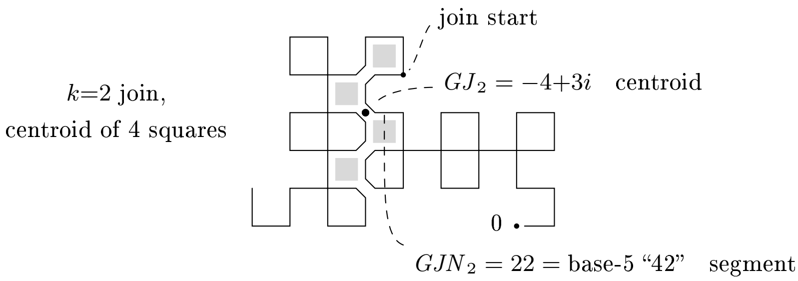


Figure 22: Join area centroid sample

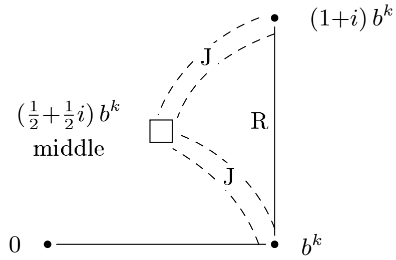
Theorem 26. *For $k \geq 1$ the join between two level k R5 dragon curves encloses some unit squares. The centroid of those squares is*

$$GJ_k = (\tfrac{9}{10}+\tfrac{3}{10}i)\,b^k + (\tfrac{2}{5}-\tfrac{1}{5}i)\frac{b^k-1}{3^k-1} \qquad k \geq 1 \tag{73}$$

$$= \tfrac{1+5i}{2},\ \tfrac{-32+24i}{8},\ \tfrac{-247-131i}{26},\ \tfrac{64-1904i}{80},\ \tfrac{11697-5323i}{242},\dots$$

$$J_k.GJ_k = \tfrac{1+5i}{2},\ -16+12i,\ \tfrac{-247-131i}{2},\ 32-952i,\ \tfrac{11697-5323i}{2},\dots$$

Proof. The right boundary squares are two joins and a middle square in between, as per the square arrangement of figure 14. So with suitable weights and rotations the centroid of those two joins and the middle square sum to the right boundary squares centroid GRQ_k.

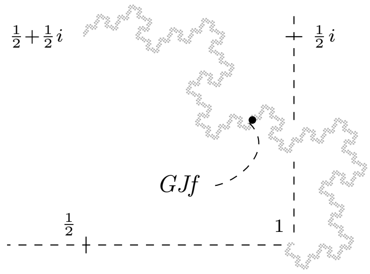


$$J_k.GJ_k + (\tfrac{1}{2}+\tfrac{1}{2}i)b^k + J_k.\big(b^k + i\,GJ_k\big) \;=\; RQ_k.\big((1+i)b^k - i\,GRQ_k\big)$$

$$GJ_k = \frac{RQ_k\Big((1+i)b^k - i.GRQ_k\Big) \;-\; \big(\tfrac{1}{2}+\tfrac{1}{2}i\big)\,b^k \;-\; J_k b^k}{(1+i)\,J_k} \qquad \square$$

With the curve scaled to the endpoint b^k a unit length, the limit is the coefficient of the b^k term in (73).

$$\frac{GJ_k}{b^k} \to GJf = \frac{9}{10} + \frac{3}{10}i$$



Theorem 27. *Centroid GJ_k of the join area squares is located in one of those squares. For k=2 it is at a corner where two squares touch, otherwise entirely within.*

Counting squares from 0 at the start of the join, this square is number

$$\begin{aligned} GJQ_k &= \tfrac{1}{4}\Big(3^k - 2 + (-1)^k\Big) && k \ge 1 \\ &= 0, 2, 6, 20, 60, 182, 546, 1640, 4920, \ldots && k \ge 1 && \text{A081251} \\ &= \textit{ternary } 202020\ldots \textit{ for } k{-}1 \textit{ digits} \end{aligned}$$

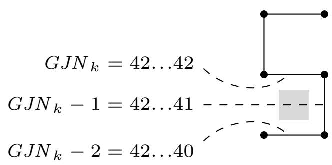
This is the boundary square on the left of segment number

$$\begin{aligned} GJN_k &= \tfrac{1}{12}\Big(11.5^k - 9 - 2(-1)^k\Big) && k \ge 1 \\ &= 4, 22, 114, 572, 2864, 14322, 71614, \ldots && k \ge 1 && 2\times\text{A037490} \\ &= \textit{base-5 } 424242\ldots \textit{ for } k \textit{ digits} && && 2\times\text{A037495} \end{aligned}$$

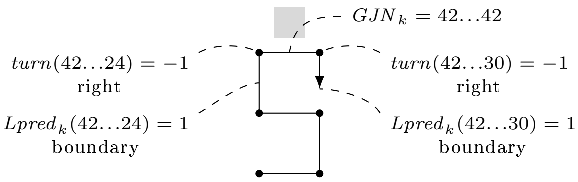
For k even this boundary square is 3-sided and GJN_k is the last, the other two sides being $GJN_k - 1$ and $GJN_k - 2$.

Proof. That GJN is on the boundary can be verified by $Lpred_k(GJN_k)$ from section 2.2, using the base-5 digits of GJN. Digits 4242… do not have any of the *Lpred* disallowed pairs.

For k even, $GJN_k = 42\ldots42$ so it is the 2 segment of the base figure and so is in the left-side U part and is a 3-side boundary square with the preceding $GJN_k - 1$ and $GJN_k - 2$ also on the boundary.



For k=1, $GJN_1 = 4$ which is a 1-side boundary square. For k odd ≥ 3, $GJN_k = 42\ldots424$ and its preceding $GJN_k{-}1 = 42\ldots423$ and following $GJN_k{+}1 = 42\ldots430$ are also boundary segments. $turn(GJN_k) = turn(GJN_k{-}1) = -1$ so it is a 1-side boundary square. Geometrically the preceding turn is the base figure, and the following turn is after 2 in the level above.



Location $point(GJN(k))$ can be calculated from its base-5 digits per section 1.4. Form (14) becomes a sum with coefficients repeating in a period-4 pattern.

$$\begin{aligned} point(GJN(k)) &= 2ib^{k-1} + (1{+}i)b^{k-2} - 2ib^{k-3} - (1{+}i)b^{k-4} \quad + + - - \ \cdots \\ &= (\tfrac{9}{10}{+}\tfrac{3}{10}i)b^k \quad + [-\tfrac{9}{10}{-}\tfrac{3}{10}i,\ -\tfrac{3}{10}{-}\tfrac{1}{10}i,\ \tfrac{9}{10}{+}\tfrac{3}{10}i,\ \tfrac{3}{10}{+}\tfrac{1}{10}i] \qquad (74) \end{aligned}$$

Total (74) can be formed by taking terms in the pattern separately (or odd and even). But it's convenient to go by generating functions for a cumulative sum with descending pattern of coefficients, in this case a cumulative sum of powers b^k. The coefficients are in the numerator of the factor.

$$\begin{aligned} gGJNpos(x) &= x\,\frac{2i + (1{+}i)x - 2ix^2 - (1{+}i)x^3}{1 - x^4} \cdot \frac{1}{1{-}bx} \\ &= (\tfrac{9}{10} + \tfrac{3}{10}i)\frac{1}{1{-}bx} \ - \ \tfrac{1}{10}\,\frac{(9{+}3i) + (3{+}i)x}{1 + x^2} \qquad (75) \end{aligned}$$

Factor x shifts the indexing down since $point(GJN_k)$ starts with term b^{k-1}. Partial fractions (75) give a power of b and the period-4 repeating part $1{+}x^2$ per (74).

The direction of $dir(GJN)$ is found from its 2-digits either East or West in another 4-period repeating E,E,W,W as $k \equiv 0$ to $3 \bmod 4$. Using this with offset $\frac{1}{2} + \frac{1}{2}i$ goes to the middle of the unit square on the left of the segment.

$$\begin{aligned} GJmid_k &= point(GJN(k)) + (\tfrac{1}{2} + \tfrac{1}{2}i)i^{dir(GJN(k))} \qquad \text{square middle} \\ &= (\tfrac{9}{10}{+}\tfrac{3}{10}i)b^k \quad + [-\tfrac{2}{5}{+}\tfrac{1}{5}i,\ \tfrac{1}{5}{+}\tfrac{2}{5}i,\ \tfrac{2}{5}{-}\tfrac{1}{5}i,\ -\tfrac{1}{5}{-}\tfrac{2}{5}i] \end{aligned}$$

$(\frac{9}{10}{+}\frac{3}{10}i)b^k$ is the same as in GJ. The periodic part of $GJmid$ shows this is at most $\frac{2}{5}$ away from the centre of the square horizontally or vertically. For $k < 6$ it can be verified explicitly that the further term of GJ_k does not push it outside the square, only to the upper left corner for $k{=}2$. For $k \ge 6$ the further term of GJ_k has smaller real or imaginary part than the $\frac{1}{10}$ distance to the edge of the square,

$$\begin{aligned} |\mathrm{Re}| \text{ or } |\mathrm{Im}|\left((\tfrac{2}{5}{-}\tfrac{1}{5}i)\frac{b^k - 1}{3^k - 1} \right) \ &< \ \frac{2}{5} \cdot \frac{\sqrt{5}^k + 1}{3^k - 1} \\ &< \ \frac{2}{5} \cdot \frac{\sqrt{5}^6 + 1}{3^6 - 1} \cdot \frac{\sqrt{5}^{k-6}}{3^{k-6}} \ = \ \frac{9}{130}\left(\frac{\sqrt{5}}{3}\right)^{k-6} < \frac{1}{10} \end{aligned}$$

For GJQ_k, the number of boundary squares after GJN_k is determined by its base-5 digits. Each 4-digit is the end of the sub-part and there are no further boundary squares beyond it. Each 2-digit is the end of the first left-side U part and there are two RQ sides after. So

$$\begin{aligned} GJQ_k &= 2RQ_{k-2} + 2RQ_{k-4} + \cdots \\ &= 2.3^{k-2} + 2.3^{k-4} + \cdots \quad = \text{ternary } 2020\ldots \end{aligned} \qquad \square$$

For k=1, the theorem is trivial in that there is only a single join area unit square so its centroid is its midpoint.

For k=2, the sample figure 22 shows segment GJN_2 to the right of the centroid GJ_2. The adjacent boundary square is below the segment. The squares $GJQ_2 = 2$ are the two above which are nearer the join start.

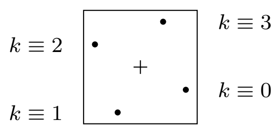
In *GJmid* the period-4 offset from $(\frac{9}{10}+\frac{3}{10}i)b^k$ to the centre of its contained square arises because the real and imaginary parts of b^k mod 10 are period-4. Per (27) those parts are a recurrence. Any modulo recurrence is eventually period since there are only so many values for so many recurrence terms. In this case they repeat from k=1 onwards.

$$\operatorname{Re}(b^k) \bmod 10 \equiv 1,\ \ 1,7,9,3,\ \ 1,7,9,3,\ldots \qquad {\scriptstyle k\geq 1\ \text{A001903}}$$
$$\operatorname{Im}(b^k) \bmod 10 \equiv 0,\ \ 2,4,8,6,\ \ 2,4,8,6,\ldots \qquad {\scriptstyle k\geq 1\ \text{A000689}}$$

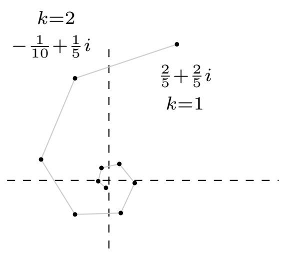
With factor 9+3i they become purely periodic. Notice the imaginary values are the same as the real but one k earlier.

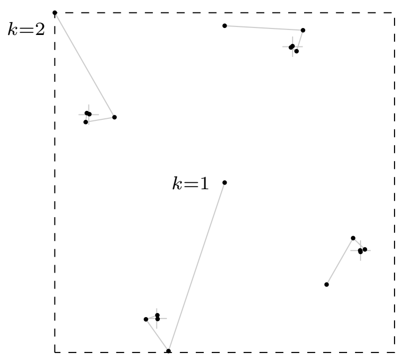
$$\operatorname{Re}\big((9{+}3i)b^k\big) \bmod 10 \equiv 9,3,1,7,\ \ 9,3,1,7,\ldots \qquad {\scriptstyle \text{A001903}}$$
$$\operatorname{Im}\big((9{+}3i)b^k\big) \bmod 10 \equiv 3,1,7,9,\ \ 3,1,7,9,\ldots \qquad \text{same}$$

So these determine the offsets (as tenths) to round $(\frac{9}{10}+\frac{3}{10}i)\,b^k$ to an integer. Or the distance to 5 goes to a $\frac{1}{2}$ integer which is the centre of the unit square containing $(\frac{9}{10}+\frac{3}{10}i)\,b^k$.



In GJ term $(\frac{2}{5}-\frac{1}{5}i)\frac{b^k-1}{3^k-1}$ approaches 0 since $\sqrt{5}/3 < 1$. Its effect is to spiral anti-clockwise towards $(\frac{9}{10}+\frac{3}{10}i)b^k$. The period-4 position of the latter within the containing unit square gives spirals every 4th term towards those positions. Notice how the k=1 offset gives the centre of the square and k=2 gives the top left corner.





8 Moment of Inertia

The mass moment of inertia $I = \sum mr^2$ of a rigid body rotating around a given axis is the ratio of torque to angular acceleration, similar to the way mass is the ratio of force to linear acceleration.

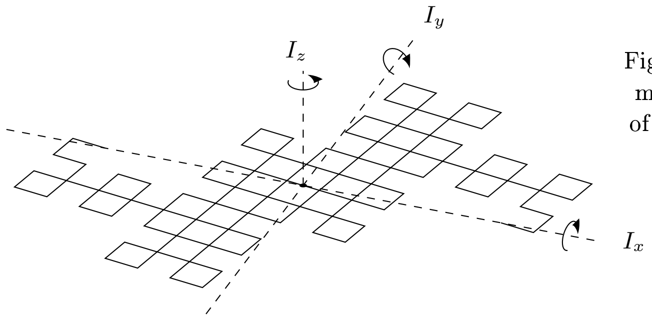


Figure 23: moment of inertia

Theorem 28. *Consider the R5 dragon curve to have unit length line segments and a total mass 1 distributed as point masses $1/5^k$ at the midpoint of each line segment.*

The centre of mass is the centre of the curve. With the x axis aligned to the curve endpoints the moment of inertia tensor about the centre of gravity is

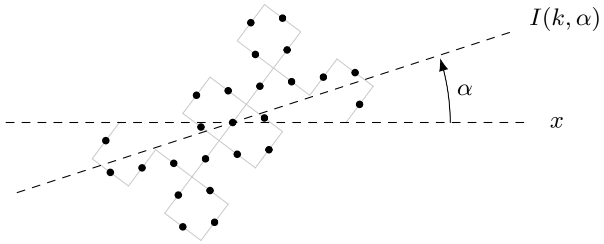
$$\begin{pmatrix} I_x & -I_{xy} & 0 \\ -I_{xy} & I_y & 0 \\ 0 & 0 & I_z \end{pmatrix} \qquad \begin{array}{ll} I_x = \sum y^2 & I_{xy} = \sum xy \\ I_y = \sum x^2 & I_z = \sum x^2 + y^2 \end{array}$$

where

$$\begin{aligned} I_x(k) &= \tfrac{54}{820}\,5^k + \tfrac{3}{328}\operatorname{Re}\left(\tfrac{b}{5}\right)^{2k} + \tfrac{14}{328}\operatorname{Im}\left(\tfrac{b}{5}\right)^{2k} - \tfrac{3}{40} \\ &= 0, \tfrac{13}{50}, \tfrac{981}{625}, \tfrac{254903}{31250}, \tfrac{16048266}{390625}, \tfrac{4017923233}{19531250}, \dots \\ I_y(k) &= \tfrac{69}{820}\,5^k - \tfrac{3}{328}\operatorname{Re}\left(\tfrac{b}{5}\right)^{2k} - \tfrac{14}{328}\operatorname{Im}\left(\tfrac{b}{5}\right)^{2k} - \tfrac{3}{40} \\ &= 0, \tfrac{17}{50}, \tfrac{1269}{625}, \tfrac{326347}{31250}, \tfrac{20514234}{390625}, \tfrac{5134420517}{19531250}, \dots \\ I_{xy}(k) &= \tfrac{7}{164}\,5^k - \tfrac{7}{164}\operatorname{Re}\left(\tfrac{b}{5}\right)^{2k} + \tfrac{3}{328}\operatorname{Im}\left(\tfrac{b}{5}\right)^{2k} \\ &= 0, \tfrac{11}{50}, \tfrac{667}{625}, \tfrac{166721}{31250}, \tfrac{10420662}{390625}, \tfrac{2605159031}{19531250}, \dots \\ I_z(k) &= I_x(k) + I_y(k) \\ &= \tfrac{3}{20}\left(5^k - 1\right) \\ &= 0, \tfrac{3}{5}, \tfrac{18}{5}, \tfrac{93}{5}, \tfrac{468}{5}, \tfrac{2343}{5}, \tfrac{11718}{5}, \tfrac{58593}{5}, \dots \qquad \tfrac{1}{5}\,\text{A125833} \end{aligned}$$

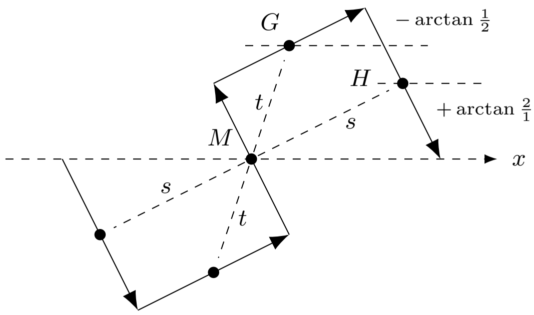
I_x, I_y and I_z are the moments of inertia rotating about the x,y,z axes of figure 23. They can be combined with I_{xy} in the usual way for inertia about an axis at angle α in the plane

$$I(k,\alpha) \;=\; I_x(k).\cos^2\alpha \;-\; 2\,I_{xy}(k).\cos\alpha\sin\alpha \;+\; I_y(k).\sin^2\alpha$$



Proof. For k=0 the curve is a single line segment and its single midpoint is inertia all zeros $I_x(0) = I_{xy}(0) = I_y(0) = 0$ which is per the formulas.

For $k \geq 1$ the inertia is calculated by rotations and the parallel axis theorem from the 5 copies of level $k{-}1$.



The 1st, 3rd and 5th copies have the x axis at $+\arctan\frac{2}{1}$ relative to those copies. The 3rd is the opposite direction but the curve is identical forward and back. The axes are turned by a matrix of rotation in the usual way

$$R = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} & 0 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{rotate axes by } +\arctan\tfrac{2}{1}$$

The 2nd and 4th copies have the x axis at $-\arctan\frac{1}{2}$ relative to those copies. For them take the inverse of a rotation

$$S = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{rotate axes by } +\arctan\tfrac{1}{2}$$

Distance s is the $k{-}1$ curve ends so $s = (\sqrt{5})^{k-1}$. This is at $+\arctan\frac{1}{2}$ to the axes which is rotation S^{-1} again to shift the centre of mass by the parallel axis theorem. Distance t is a diagonal across half curve lengths so $t = s/\sqrt{2}$ at a further 45° angle from s so a rotation matrix.

$$T = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{3}{\sqrt{5}} & 0 \\ \frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{rotate axes by } S + 45°$$

$$I(k) = \tfrac{3}{5}\, R^{-1}.I(k{-}1).R \qquad\qquad \text{1st, 3rd, 5th}$$

$$
\begin{array}{ll}
+\,\tfrac{2}{5}\,S.I(k{-}1).S^{-1} & \text{2nd, 4th} \\
+\,2.\tfrac{1}{5}\,(\quad\sqrt{5}^{k-1})^2\;S.\left(\begin{smallmatrix}0&0&0\\0&1&0\\0&0&1\end{smallmatrix}\right).S^{-1} & ms^2\text{, 1st, 5th} \\
+\,2.\tfrac{1}{5}\,(\tfrac{1}{\sqrt{2}}\sqrt{5}^{k-1})^2\;T.\left(\begin{smallmatrix}0&0&0\\0&1&0\\0&0&1\end{smallmatrix}\right).T^{-1} & mt^2\text{, 2nd, 4th}
\end{array}
$$

Multiplying through gives mutual recurrences

$$
\begin{aligned}
I_x(k) &= \tfrac{11}{25}I_x(k{-}1) \;-\; \tfrac{4}{25}I_{xy}(k{-}1) \;+\; \tfrac{14}{25}I_y(k{-}1) \;+\; \tfrac{13}{50}\,5^{k-1} \\
I_y(k) &= \tfrac{14}{25}I_x(k{-}1) \;+\; \tfrac{4}{25}I_{xy}(k{-}1) \;+\; \tfrac{11}{25}I_y(k{-}1) \;+\; \tfrac{17}{50}\,5^{k-1} \\
I_{xy}(k) &= \tfrac{2}{25}I_x(k{-}1) \;-\; \tfrac{3}{25}I_{xy}(k{-}1) \;-\; \tfrac{2}{25}I_y(k{-}1) \;+\; \tfrac{11}{50}\,5^{k-1} \\
I_z(k) &= I_z(k{-}1) \;+\; \tfrac{3}{5}.5^{k-1}
\end{aligned}
$$

Expanding I_z repeatedly down to $I_z(0){=}0$ is $I_z(k) = \sum_{j=0}^{k-1}\tfrac{3}{5}.5^j$ per its formula.

$I_z = I_x{+}I_y$ is true of any plane figure. Substituting $I_y = I_z - I_x$ gives three equations in I_x and I_{xy}. The first and last are negations leaving two in two unknowns which are a second-order recurrence for I_x

$$
\begin{aligned}
&I_x(k) = -\tfrac{6}{25}I_x(k{-}1) - \tfrac{1}{25}I_x(k{-}2) + \tfrac{216}{125}\,5^{k-2} - \tfrac{12}{125} \qquad\qquad k \ge 2 \\
&\qquad\text{starting } I_x(0) = 0 \text{ and } I_x(1) = \tfrac{13}{50}
\end{aligned}
$$

Usual recurrence or generating function manipulation then gives I_x as powers of the roots $5, 1, (\frac{b}{5})^2 = -\frac{3}{25}{+}\frac{4}{25}i$ and its conjugate. The imaginary parts cancel out and the real parts can be expressed by real and imaginary parts of the single power $(\frac{b}{5})^{2k}$. I_x then gives I_{xy} and I_z. □

The power $(\frac{b}{5})^{2k}$ is similar to that for segments in direction counts (26) and like there it is not a multiple of 2π so the inertia components are not periodic in k.

For the curve scaled to a unit length, and unit mass (which is twice its area), the limit for the inertia tensor is the high coefficients in I_x etc.

$$
\frac{I(k)}{5^k} \;\to\; \begin{pmatrix} \frac{54}{820} & -\frac{7}{164} & 0 \\ -\frac{7}{164} & \frac{69}{820} & 0 \\ 0 & 0 & \frac{123}{820} \end{pmatrix}
$$

An inertia matrix is real and symmetric so can be diagonalized with a suitable matrix of rotation turning to the eigenvectors which are its principal axes. The physical significance of this is that rotation about those axes is perfectly balanced with no torque exerted on the mounting points.

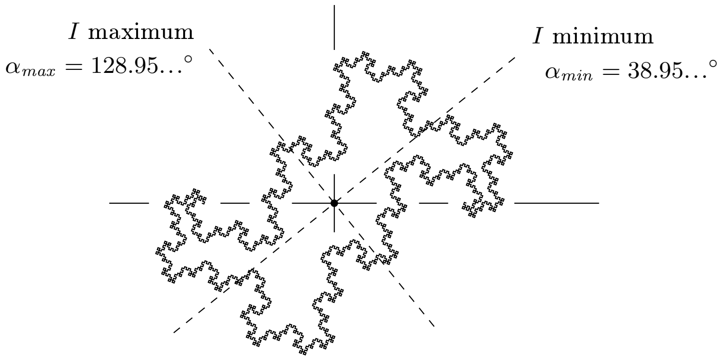
In the usual way for a 2×2 matrix, the eigenvectors are complex number direction d and then angle α,

$$
\begin{aligned}
d^2 &= \big(I_x(k) - I_y(k)\big) \;-\; 2I_{xy}(k)\,i \\
\alpha &= \tfrac{1}{2}\arctan\frac{-2\,I_{xy}(k)}{I_x(k) - I_y(k)} \;+ (0 \,\text{or}\, \tfrac{\pi}{2}) \\
&= \tfrac{1}{2}\arctan\left(\tfrac{14}{3} + \epsilon_k\right) \;+ (0 \,\text{or}\, \tfrac{\pi}{2})
\end{aligned}
$$

$$\text{where } \epsilon_k = \frac{\operatorname{Im}\left(\frac{b}{5}\right)^{2k}}{\frac{9}{205}5^k - \frac{9}{205}\operatorname{Re}\left(\frac{b}{5}\right)^{2k} - \frac{42}{205}\operatorname{Im}\left(\frac{b}{5}\right)^{2k}}$$

$\left|\left(\frac{b}{5}\right)^2\right| = \frac{1}{5}$ so $\epsilon_k \to 0$ as $k \to \infty$ and limits for the principal axes are

$$\begin{aligned}
\alpha_{min} &= \tfrac{1}{2}\arctan\tfrac{14}{3} \;+ (\text{terms in } \epsilon) \qquad \text{arctan first quadrant} \\
&\to 38.95262^\circ\ldots \\
\alpha_{max} &= \alpha_{min} + 90^\circ \\
&\to 128.95262^\circ\ldots
\end{aligned}$$

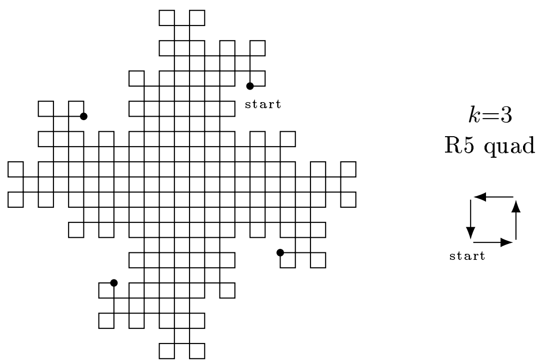


For the curve scaled to unit length, unit mass (which is twice its area), and rotated to principal axes, the limit for the inertia tensor is

$$\begin{pmatrix} \frac{3}{40} - \frac{\sqrt{205}}{328} & 0 & 0 \\ 0 & \frac{3}{40} + \frac{\sqrt{205}}{328} & 0 \\ 0 & 0 & \frac{3}{20} \end{pmatrix}$$

9 R5 Quad

Four R5 dragon curves can be arranged in a square "quad". k=0 is a single unit square.



Each expansion is

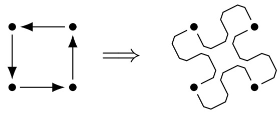


Figure 24:
R5 quad
expansion

Each unit square expands to 5 squares, so starting 1 in k=0 is total 5^k inside. The outside unit squares are four right sides $4AR_k$, for total area

$$\begin{aligned} QA_k &= 5^k + 4AR_k \\ &= 2.5^k - 2.3^k + 1 \\ &= 1, 5, 33, 197, 1089, 5765, \ldots \end{aligned}$$

Scaled to unit length endpoints for the component curves, the area limit is $QA_k/5^k \to 2$, which is $4\times$ curve area limit $\frac{1}{2}$ from (51).

The convex hull around the quad follows in a similar way to the hull around the plain curve in theorem 22. For the quad, there are only sides $3i^d$ and $2i^d b^j$ (no $4-i$) and the general case can start at k=1.

$$QHsides_0 = \; i^d \qquad d = 0 \text{ to } 3 \qquad\qquad\qquad \text{for } k{=}0$$

$$QHsides_k = \begin{cases} 3i^d & d = 0 \text{ to } 3 \qquad\qquad\qquad \text{for } k \ge 1 \\ 2i^d b^j & d = 0 \text{ to } 3 \text{ and } j = 1 \text{ to } k{-}1 \end{cases}$$

which is number of sides

$$\begin{aligned} QHnumSides_k &= \begin{cases} 4 & \text{if } k = 0 \\ 4k & \text{if } k \ge 1 \end{cases} \\ &= 4, 4, 8, 12, 16, 20, 24, \ldots \end{aligned}$$

The quad hull boundary length is then

$$\begin{aligned} QHB_k &= \sum_{z \,\in QHsides_k} |z| \\ &= 2\left(1{+}\sqrt{5}\right)\sqrt{5}^k \;+ 2\left(1{-}\sqrt{5}\right) \\ &= 2\left(5^{\lceil k/2 \rceil}{+}1\right) \;+\; 2\left(5^{\lfloor k/2 \rfloor}{-}1\right)\sqrt{5} \\ &= 4,\; 12,\; 12{+}8\sqrt{5},\; 52{+}8\sqrt{5},\; 52{+}48\sqrt{5}, \ldots \end{aligned}$$

rational dup $2\times$A034474, irratational dup $2\sqrt{5}\times$A024049

Scaled to unit length component curves, the hull boundary length limit is

$$\frac{QHB_k}{\sqrt{5}^k} \to 2 + 2\sqrt{5} \;\; = 6.472135\ldots \qquad 2\times\text{A134945}$$

The area of the quad hull follows in a similar way to the curve hull in theorem 23. The equivalent of recurrence (70) applies for $k \ge 2$ with $k-1 = 1$ being 4 sides $3i^d$. The powers of 5 on each $2b^j$ are $1 + 5 + 5^2 + \cdots + 5^{k-1-j} = (5^{k-j} - 1)/4$ so with 4 of each factor $5^{k-j} - 1$. The constant 4 and initial QHA_1 become $2.5^k{-}1$.

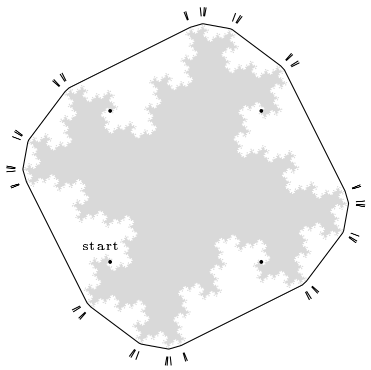
$$QHA_k = 2.5^k - 1 \,+ \sum_{j=1}^{k-1} \left(5^{k-j}{-}1\right) HAgrow(2b^j)$$

$$= 1, 9, 65, 369, 1905, 9697, 48985, \ldots$$

Scaled to unit length component curves, the hull area limit is

$$\frac{QHA_k}{5^k} \to 2 + \sum_{j=1}^{\infty} 2\,HAgrowf\left(\tfrac{1}{b^j}\right) \qquad = 3.153977\ldots$$

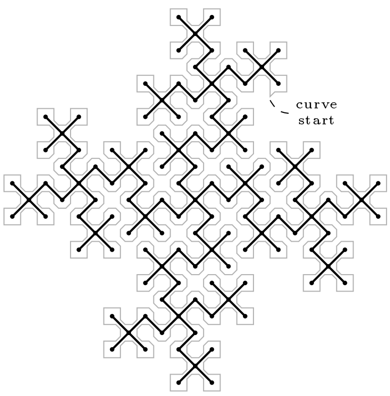
This can be compared to area limit 2 for the quad it surrounds, so over a third of the hull is empty.



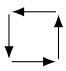
R5 quad hull limit
area 3.153...
boundary 6.472...
countably infinite
vertices

9.1 R5 Quad Area Tree

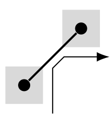
When corners of an R5 quad are chamfered off, the inside unit squares are connected through the resulting gaps. Call this an area tree.



k=3 R5 quad area tree,
vertex each unit square
inside R5 quad



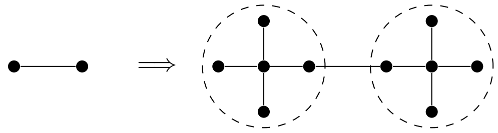
An equivalent definition is to connect unit squares which are on the left of consecutive curve segments.



edge between unit squares
on left sides of consecutive
curve segments

When the curve turns to the right the unit squares on the left of the segments are distinct. A turn is always left or right (never straight ahead) so those connections are though corners of the squares

Per figure 24, each unit square expands to 5 new unit squares in a star pattern. So a bottom-up expansion rule is a repeated star replacement. Starting from a single vertex for k=0, each vertex is replaced by a 5-star and existing edges are between arms of the stars.



star
replacement

Theorem 29. *Vertices of the R5 quad area tree are degrees 1,2,4 after the initial degree-0 in k=0. The number of each degree in level k are*

$$QADegCount(k,0) = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \geq 1 \end{cases} \qquad \text{A000007}$$

$$QADegCount(k,1) = \begin{cases} 0 & \text{if } k{=}0 \\ 2.5^{k-1} + 2 & \text{if } k \geq 1 \end{cases}$$

$$= 0, 4, 12, 52, 252, \ldots \qquad 2{\times}\text{A034474}$$

$$QADegCount(k,2) = \begin{cases} 0 & \text{if } k = 0 \text{ to } 2 \\ 2.5^{k-1} - 2 & \text{if } k \geq 3 \end{cases} \tag{76}$$

$$= 0, 0, 8, 48, 248, \ldots \qquad 2{\times}\text{A024049}$$

$$QADegCount(k,4) = \begin{cases} 0 & \text{if } k{=}0 \\ 5^{k-1} & \text{if } k \geq 3 \end{cases} \tag{77}$$

$$= 0, 1, 5, 25, 125, \ldots$$

Proof. Degree-4 vertices arise from the star replacement, so each of the 5^{k-1} vertices in level $k-1$ becomes a degree-4 in level k for (77).

Degree-2 vertices arise from the star replacement as two in each previous edge. There are $5^{k-1} - 1$ edges in level $k-1$, so (76).

The star replacement leaves only degree 1,2,4 vertices so the remainder of the total 5^k in level k are degree-1. Or alternatively the star replacement gives 3 degree-1 for each previous degree-1 and 2 for each previous degree-2 so

$$QADegCount(k{,}1) = 3\,QADegCount(k{-}1,1) + 2\,QADegCount(k{-}1,2) \quad k{\geq}2 \qquad \square$$

A further approach for degree-1 is that they occur when the curve makes 3 consecutive left turns. The total number of 3 consecutive turns is *TurnRuns3* from (52) in section 3. The curve is the same forward and reverse so the rights and lefts are half each. The end of the sub-curves making up the quad has a left turn to the next and this makes a further 3 consecutive lefts when $k \geq 1$, so

$$QADegCount(k,1) = 4.\tfrac{1}{2}\,TurnRuns3_k + 4 \qquad k \geq 1$$

Theorem 30. *The diameter of R5 quad area tree k is*

$$QAdiameter_k = RQ_k - 1 \tag{78}$$

$$= 3^k - 1 \qquad \text{A024023}$$

It is attained by the left boundary squares of the curve (inside the quad) and various other paths. The total number of paths attaining the diameter is

$$QAdiameterCount_k = \begin{cases} 1 & \text{if } k{=}0 \\ 6\,.\,9^{k-1} & \text{if } k \ge 1 \end{cases}$$
$$= 1, 6, 54, 486, 4374, \ldots \qquad \text{A092810}$$

The number of diameter endpoints, and total number of vertices on some diameter are

$$QAdiameterEnds_k = \begin{cases} 1 & \text{if } k{=}0 \\ 4.3^{k-1} & \text{if } k \ge 1 \end{cases}$$
$$= 1, 4, 12, 36, 108, \ldots \qquad \text{A003946}$$

$$QAdiameterVertices_k = 4k\,.\,3^{k-1} + 1$$
$$= 1, 5, 25, 109, 433, \ldots$$

Proof. For any path in level $k{-}1$, the star replacement inserts 2 further edges into it for level k, so $3\times$ the length. A path between any of those new vertices is shorter. If a path in $k{-}1$ ends at a degree-1 vertex then the star there is new leaf vertices attached.

A diameter must be between degree-1 vertices (otherwise could be extended). So, with diameter 0 for the single vertex of $k{=}0$,

$$QAdiameter_k = 3\,QAdiameter_{k-1} + 2 \qquad \text{starting } QAdiameter_0 = 0 \tag{79}$$

There are 3 new leaves at the end of the diameter path. The single vertex of $k{=}0$ is trivially the left boundary squares. Choosing the boundary square of the left side on each expansion gives the diameter as left boundary squares, hence (78). The 3 choices are then, once the diameter is not 0,

$$QAdiameterEnds_k = 3\,QAdiameterEnds_{k-1} \quad \text{starting } QAdiameterEnds_1 = 4$$

and combinations of the 3 new at each end is 9 new paths for each existing one

$$QAdiameterCount_k = 9\,QAdiameterCount_{k-1} \qquad QAdiameterCount_1 = 6$$

For total vertices of diameters, on star replacement each existing diameter vertex has 2 new vertices towards the middle of the tree, except at the middle vertex itself. The new $QAdiameterEnds_k$ outer vertices are immediately adjacent to existing diameter vertices. So

$$QAdiameterVertices_k = 3\,QAdiameterVertices_{k-1} - 2 + QAdiameterEnds_k$$
$$\text{starting } QAdiameterVertices_0 = 1 \qquad \square$$

In $k = 1, 2$ all the degree-1 vertices are diameter endpoints, but in $k \ge 3$ some degree-1 are not diameter endpoints. The degree-1 vertices grow as 5^k whereas the diameter endpoints grow only as 3^k.

$$QAdiameterEnds_k = QADegCount(k, 1) \qquad k = 1, 2$$

$$QAdiameterEnds_k < QADegCount(k,1) \qquad k \ge 3$$

A top-down definition of the tree is to take the expansion of figure 24 as a level k square comprising 5 level $k{-}1$ squares connected by new edges between each.

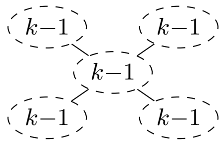


Figure 25: area tree k as 5 copies of $k{-}1$ and new edges between

The connections are at vertices which are diameter endpoints, since that is necessary to give (79) as three diameters of level $k{-}1$ plus 2 edges between.

The vertices attaining the diameter are in the outer sub-trees. Not all pairs of *QAdiameterEnds* endpoint vertices make a diameter, only those between different outer sub-trees. By symmetry there are $\frac{1}{4}$ vertices in each, and there are binomial $\binom{4}{2} = 6$ pairs of outer trees, so

$$QAdiameterCount_k = 6\left(\tfrac{1}{4}\,QAdiameterEnds_k\right)^2 \quad k \ge 1$$

A top-down approach to *QAdiameterVertices* can be made by noting the middle $k{-}1$ has only paths going middle to outer connections, and the outers have only path out to their middle, nothing in the quarter tree at the connection. So 12 outer quarter trees and diameter or half diameters across. A constant offset is applied for middles counted twice or not by these diameters.

$$QAdiameterVertices_k = \tfrac{12}{4}\,QAdiameterVertices_{k-1} + (2+\tfrac{4}{2})\,QAdiameter_{k-1} + 2$$

The Wiener index is a measure of total distance between pairs of vertices in a graph.

$$Wiener\ index = \tfrac{1}{2} \sum_{\text{vertices } u,v} distance(u,v)$$

Factor $\frac{1}{2}$ has the effect of taking distance between a pair u,v in just one direction, not also its reverse v,u.

Theorem 31. *The Wiener index of R5 quad area tree k is*

$$\begin{aligned} QAW_k &= 5^k\left(\tfrac{2}{7}\,15^k - 5^{k-1} - \tfrac{3}{35}\right) \\ &= 0, 16, 1480, 117400, 8962000, \ldots \end{aligned}$$

Proof. Per figure 25, sub-parts connect at a degree-1 vertex C which is a diameter endpoint. Let $QAwC$ be the total path length from there to all other vertices. Total length C in its own sub-tree is $QAwC_{k-1}$. To reach the middle tree is the diameter plus 1 edge for each of the 5^{k-1} destination vertices in the middle part. To reach the far outer 3 parts is twice that distance. In each of those parts the further distance to the vertices is $QAwC_{k-1}$, for total 5 of those.

$$\begin{aligned} QAwC_k &= 5\,QAwC_{k-1} + 5^{k-1}(1+3.2)\big(\,QAdiameter_{k-1}{+}1\,\big) \\ &= \tfrac{7}{2}5^{k-1}(3^k - 1) \qquad\qquad \text{starting } QAwC_0 = 0 \end{aligned}$$

$$= 0, 7, 140, 2275, 35000, \ldots$$

Total paths within the sub-trees is QAW_{k-1} each. Then paths from each of the 4 outer trees to the middle is $QAwC_{k-1}$ for vertices to the connecting edge, times 5^{k-1} destinations, and the same from the middle connection vertex into the middle part. The connecting edge is crossed by $5^{k-1}.5^{k-1}$ many paths between the parts.

Similarly for the 6 combinations of paths between outer sub-trees, except instead of a single connecting edge to cross there is the middle sub-tree diameter plus an edge each side.

$$\begin{aligned} QAW_k &= 5\,QAW_{k-1} + 4\Big(5^{k-1}.2.\,QAwC_{k-1} + 25^{k-1}\Big) \\ &\quad + 6\Big(5^{k-1}.2.\,QAwC_{k-1} + 25^{k-1}(QAdiameter_{k-1}{+}2)\Big) \end{aligned} \qquad \square$$

Second Proof of Theorem 31. Suppose a given edge has x many vertices on one side and y many on the other, so that for the paths in QAW it is crossed xy times,

$$QAW_k = \sum\nolimits_{\text{edges}} x\,y$$

Star replacement means each edge becomes 3

$$x \;\text{—}\; y \quad\Longrightarrow\quad x \;\text{—}\; a \;\text{—}\; b \;\text{—}\; y$$

The middle a–b has $5x$ vertices on one side and $5y$ on the other. Edges x–a and b–y have one more and less each side. So total crossings

$$5x.5y + (5x{+}1)(5y{-}1) + (5x{-}1)(5y{+}1) \;=\; 75xy - 2$$

There are new edges in $k{+}1$ to its leaf vertices. Those edges have 1 vertex on one side and $5^{k+1}{-}1$ on the other. So

$$\begin{aligned} QAW_{k+1} &= 75\,QAW_k - 2.(5^k{-}1) \;+\; QADegCount(k{+}1,1).1.(5^{k+1}{-}1) \\ &= 75\,QAW_k + 10.25^k + 6.5^k \end{aligned} \qquad \square$$

Wiener index divided by number of vertex pairs is a mean distance between vertices. Such a mean is usually taken over vertex pairs in one direction (like the Wiener index) and excluding a vertex to itself, so number of pairs is binomial $\binom{5^k}{2} = \frac{1}{2}5^k(5^k-1)$. The mean can be expressed as a fraction of *QAdiameter*. The limit of that fraction as $k{\to}\infty$ follows from the high coefficients of the terms.

$$\frac{QAW_k}{\frac{1}{2}5^k(5^k-1)\,.\,QAdiameter_k} \;\to\; \frac{4}{7} \quad = 0.571428\ldots \qquad (80)$$

For interest, a similar repeated star replacement can be made for stars of $s{\ge}3$ vertices ($s{=}5$ is the quad area tree). This has the same *QAdiameter*, and working through the forms above for the Wiener index gives

$$StarRepW(s,k) = s^{k-1}\left(\tfrac{1}{2}\,\frac{(s{-}1)(3s{-}5)}{3s{-}1}\,(3s)^k \;-\; \tfrac{1}{2}(s{-}3)\,s^k \;-\; \frac{s{+}1}{3s{-}1}\right) \qquad s{\ge}3$$

$$
\begin{aligned}
s{=}3 \quad &= 0, 4, 120, 3276, 88560, 2391444, \ldots \\
s{=}4 \quad &= 0, 9, 516, 25872, 1258560, 60674304, \ldots \\
s{=}5 \quad &= QAW \\
s{=}6 \quad &= 0, 25, 3390, 389700, 42928920, 4666565520, \ldots
\end{aligned}
$$

Case $s{=}3$ is a middle and 2 outer vertices. Its repeated replacement is simply a path of 3^k vertices.

The limit for mean distance between distinct vertices is then

$$
\begin{aligned}
\frac{StarRepW(s,k)}{\frac{1}{2}s^k(s^k-1) \,.\, QAdiameter_k} \;\to\; \frac{(s{-}1)(3s{-}5)}{s\,(3s{-}1)} &= 1 - \frac{5}{s} + \frac{8}{3s{-}1} \\
&= \tfrac{1}{3}, \tfrac{21}{44}, \tfrac{4}{7}, \tfrac{65}{102}, \tfrac{24}{35}, \tfrac{133}{184}, \tfrac{88}{117}, \ldots \qquad s \ge 3
\end{aligned}
$$

From the formula, this approaches 1 as $s \to \infty$. Roughly speaking, as s increases, more paths become diameters. The number of diameter paths is $\binom{s-1}{2}((s{-}2)^2)^k$ out of all paths $\frac{1}{2}s^k(s^k{-}1)$ so ratio approaches 1.

Gutman, Furtula and Petrović[7] consider a terminal Wiener index which is distances between pairs of terminal vertices (ie. leaf nodes, degree 1).

Theorem 32. *The terminal Wiener index of R5 quad area tree k is*

$$
\begin{aligned}
QATW_k &= \tfrac{8}{175}75^k + \tfrac{38}{125}25^k + \tfrac{69}{175}5^k - 1 \\
&= 0, 12, 456, 24084, 1565424, \ldots
\end{aligned}
$$

Proof. Make a calculation similar to QAW theorem 31 above. The connection vertices in $k{-}1$ are degree-1 but on joining are no longer, so adjust to exclude them. The recurrences can begin at $k{=}2$ so that the connection vertices in the $k{-}1$ sub-trees are distinct.

Let $QAtwC_k$ be the total path lengths from the connecting vertex to all others in area tree k.

$$
\begin{aligned}
QAtwC_k &= 5\,QAtwC_{k-1} \; - 4\,QAdiameter_{k-1} \qquad k \ge 2 \qquad (81) \\
&\quad + (QADegCount(k{-}1,1) - 4)\,(QAdiameter_{k-1}{+}1) \\
&\quad + (QADegCount(k{-}1,1) - 1)\,3.2.(QAdiameter_{k-1}{+}1) \\
&\qquad\qquad \text{starting } QAtwC_{0,1} = 0, 6 \\
&= \begin{cases} 0 & \text{if } k = 0 \\ \tfrac{7}{25}15^k + \tfrac{14}{25}5^k - 1 & \text{if } k \ge 1 \end{cases} \\
&= 0, 6, 76, 1014, 14524, \ldots
\end{aligned}
$$

$QAtwC_k$ comprises the five $k{-}1$ sub-trees plus the extra distance to reach the connection at the middle and 3 far outers. The first $QAtwC_{k-1}$ should not include the path to its connection to the middle since that vertex is no longer a leaf. The middle $QAtwC_{k-1}$ similarly should not include the path to its connections to the 3 outer, hence the 4 subtracted at (81).

$$
QATW_k = 5\,QATW_{k-1} \; - 8\,QAtwC_{k-1} + 6\,QAdiameter_{k-1} \qquad k \ge 2 \qquad (82)
$$

$$+4\begin{pmatrix}(QADegCount(k{-}1,1)-4)\ QAtwC_{k-1}\\ +(QADegCount(k{-}1,1)-1)\,(QAtwC_{k-1}-3QAdiameter_{k-1})\\ +(QADegCount(k{-}1,1)-4)\,(QADegCount(k{-}1,1)-1)\end{pmatrix} \tag{83}$$

$$+6\begin{pmatrix}2\,(QADegCount(k{-}1,1)-1)\ QAtwC_{k-1}\\ +(QADegCount(k{-}1,1)-1)^2\,(QAdiameter_{k-1}{+}2)\end{pmatrix} \tag{84}$$

$$\text{starting } QATW_{0,1}=0,12$$

$QATW_k$ comprises the five $k{-}1$ sub-trees at (82), less paths from the 8 vertices which are connections into their respective parts. In the middle part there are 4 such and subtracting $QAtwC$ of all removes paths both ways between, hence adding back $\binom{4}{2}=6$ diameters.

(84) is paths between the 4 outer sub-trees. (83) is paths between the middle sub-tree and the 4 outer, firstly from the middle to the outers, and then from the outers to the middle. For the latter the $QAtwC$ has paths to the other middle connection vertices subtracted. □

Second Proof of Theorem 32. In a similar manner to the second proof of QAW theorem 31, consider an edge with x vertices on once side and y on the other (all vertices, not just terminals). Star replacement expands it to 3 edges. Each vertex on expands to 3 leaf if degree-1, 2 if degree-2 or none if degree-4. New leaves are thus $4-deg$. Total of all degrees is $2.edges$ in the usual way, so $4x-2(x{-}1)=2x{+}1$ expanded leaf vertices. Crossings of the expanded edges for $QATW$ are then

$$3(2x{+}1)(2y{+}1)=12xy+6(x{+}y)+3$$

Total xy is the full Wiener index QAW. Each $x{+}y$ is simply all vertices 5^k. It and $+3$ are over all $5^k{-}1$ edges. Then with further edges at each leaf crossed to the other $k{+}1$ leaves,

$$\begin{aligned}QATW_{k+1}&=12QAW_k+(6.5^k+3)(5^k-1)\\&\quad+QADegCount(k{+}1,1).1.(QADegCount(k{+}1,1)-1)\\&=12QAW_k+10.25^k+3.5^k-1\end{aligned}$$ □

The limit mean distance between distinct degree-1 vertices is the same as all vertices from (80)

$$\frac{QATW_k}{\binom{QADegCount(k,1)}{2}\,QAdiameter_k}\to\frac{4}{7}\qquad\text{same as } QAW$$

The star replacement turns each existing vertex into a degree-4, with 3 edges between them instead of 1. So a Wiener index among degree-4 vertices is $3\,QAW_{k-1}$.

Theorem 33. *In the R5 quad area tree, take the root as the vertex at the quad curve start. The width of the tree (number of vertices) at a given depth (distance from the root, starting d=0 for the root) is*

$$QAwidth(d)=3^{CountTernaryTwos(d)} \tag{85}$$

$$= 1, 1, 3, 1, 1, 3, 3, 3, 9, 1, 1, 3, \ldots$$

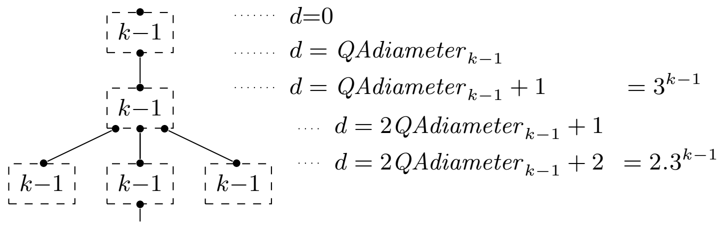
$$\mathit{CountTernaryTwos}(n) = \textit{number of digit 2s in } n \textit{ written in ternary}$$
$$= 0, 0, 1, 0, 0, 1, 1, 1, 2, 0, 0, 1, \ldots \qquad \text{A081603}$$

Generating functions

$$\mathit{gQAwidth}(x) = \prod_{j=0}^{\infty} \left(1 + x^{3^j} + 3x^{2.3^j}\right) \tag{86}$$

$$\mathit{gCountTernaryTwos}(x) = \frac{1}{1{-}x} \sum_{j=0}^{\infty} \frac{x^{2.3^j}}{1 + x^{3^j} + x^{2.3^j}} \tag{87}$$

Proof. As from the top-down above, sub-trees connect across diameters, so there is no overlap between parts in the descent.



$\mathit{QAdiameter}_{k-1} = 3^k - 1$ means the middle sub-tree begins at depth $d = 3^{k-1}$. It is widths of tree $k{-}1$ unchanged. The three low sub-trees are a further diameter across the middle plus one edge so they begin at depth $d = 2.3^{k-1}$. Depths from there to $3^k - 1$ are a digit 2 of d written in ternary and are a factor 3 on the widths of tree $k{-}1$, hence the power (85).

Generating function $\mathit{gQAwidth}$ at (86) follows by considering how to add a new low ternary digit to d. Existing terms spread to $3d$ by substituting x^3. Terms at $3d{+}1$ are to be the same as $3d$, and terms at $3d{+}2$ a further factor 3. So $\mathit{gQAwidth}_k(x)$ good for k many digits is

$$\mathit{gQAwidth}_k(x) = (1 + x + 3x^2)\ \mathit{gQAwidth}_{k-1}(x^3)$$
$$\text{starting } \mathit{gQAwidth}_0(x) = 1$$

Generating function $\mathit{gCountTernaryTwos}$ at (87) has each term giving coefficients 1 when there is a ternary digit 2 at position j, for $j{=}0$ the least significant digit. Each term is 1s where $n \bmod 3.3^j$ in the range $2.3^j \le n < 3.3^j$.

It's also possible to consider a new low ternary digit on $\mathit{gCountTernaryTwos}$. This is a spread and 3 copies of the existing values, then $+1$ at $n \equiv 2 \bmod 3$.

$$\mathit{gCountTernaryTwos}_k(x) = (1{+}x{+}x^2)\ \mathit{gCountTernaryTwos}_{k-1}(x^3) + \frac{x^2}{1 - x^3}$$
$$\text{starting } \mathit{gCountTernaryTwos}_0(x) = 0$$

Expanding repeatedly is

$$\mathit{gCountTernaryTwos}(x) = \sum_{j=0}^{\infty} \frac{x^{2.3^j}}{1 - x^{3^{j+1}}} \prod_{l=0}^{j-1} \left(1 + x^{3^l} + x^{2.3^l}\right)$$

Each term of this sum has the product cancelling into denominator $1-x^{3^{j+1}}$ leaving (87). $\square$

The non-overlap of k with deeper expansion means that sum of widths up to the diameter is the total 5^k vertices.

$$\sum_{d=0}^{QAdiameter_k} QAwidth(d) \quad = 5^k \qquad \text{vertices in tree } k$$

An independent set in a graph is a set of vertices which have no edges between them. The empty set or one vertex set are always independent. Bigger sets are restricted to no adjacent vertices.

Theorem 34. *The number of independent sets in R5 quad area tree k is*

$$\begin{aligned} QAindsets_k &= \begin{cases} 2 & \text{if } k{=}0 \\ h_k(0) & \text{if } k \geq 1 \end{cases} \\ &= 2,\ 17,\ 540497,\ 17578416511181776606206459281, \ldots \end{aligned}$$

where for $n = 0$ to 4

$$h_k(n) = \sum_{v_{1,2,3,4}=0,1} h_{k-1}\Big(\sum_{l=1}^{4} v_l\Big) \prod_{l=1}^{4} \begin{cases} h_{k-1}(t{+}1) & \text{if } v_l = 0 \\ h_{k-1}(t) - h_{k-1}(t{+}1) & \text{if } v_l = 1 \end{cases} \tag{88}$$

where each $t{=}1$ if $l \leq n$ or $t{=}0$ if $l > n$

starting $h_1(0 \text{ to } 4) = 17, 9, 5, 3, 2$

$$\begin{aligned} h_k(1) &= 9,\ 290281,\ 9440829864910747123331181609, \ldots \qquad k \geq 1 \\ h_k(2) &= 5,\ 155905,\ 5070380968700781122208928001, \ldots \\ h_k(3) &= 3,\ \ 83737,\ 2723146538559765117772726809, \ldots \\ h_k(4) &= 2,\ \ 44977,\ 1462518717280970760659330417, \ldots \end{aligned}$$

Proof. $h_k(n)$ is the number of independent sets in area tree k which have n many of the connection vertices absent from the sets (the rest either present or absent). h starts at $k{=}1$ where there are 4 distinct connection vertices. For example $h_1(4) = 2$ is where all 4 connection vertices are absent, so just the centre present or absent.

Recurrence (88) is area tree k comprising five copies of $k{-}1$ with new edges between. Index l is the outer parts numbered 1 to 4. $v_l = 1$ when the outer part has its connection vertex present in the set, or $v_l = 0$ when absent.

$h_{k-1}(\sum v_l)$ is sets in the middle part. It must have absent connection vertices where each v_l outer is present, or anything where the outer is absent.

The product terms are sets in the outer parts. One connection vertex from each outer part becomes the connection vertices of tree k. This is t as 0 or 1 to make total n many outer parts have a connection absent. If $v_l = 0$ then a second connection vertex absent too so $h_{k-1}(t{+}1)$. If $v_l = 1$ then the second connection vertex must be present. This is formed by count arbitrary less count absent. $\square$

See join area tree section 9.2 for *QAindsets* by quarters.

The independence number of a graph is the maximum number of vertices in any independent set.

Theorem 35. *The independence number of R5 quad area tree k is*

$$QAindnum_k = \begin{cases} 1, 4 & \textit{if } k = 0, 1 \\ 16.5^{k-2} + 1 & \textit{if } k \geq 2 \end{cases} \tag{89}$$
$$= 1, 4, 17, 81, 401, 2001, \ldots$$
$$\textit{base-5} \quad = 1, 4, 32, 311, 3101, 31001, \ldots$$

The number of independent sets of this size is

$$QAindnumCount_k = \begin{cases} 1 & \textit{if } k{=}0 \\ 2^{\frac{1}{2} QADegCount(k-1,\, 2)} & \textit{if } k \geq 1 \\ \quad = 2^{5^{k-2}-1} \quad k{\geq}2 \end{cases} \tag{90}$$
$$= 1, 1, 1, 16, 16777216, \ldots$$

k=0 is a single vertex. k=1 is the 4 outer vertices of the star. k=2 is 4 vertices adjacent to the 4 outer degree-4s, and the centre degree-4,

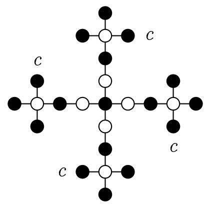


Figure 26: k=2 sole independent set attaining $QAindnum_2 = 17$

For $k \geq 3$ the sub-trees are this k=2 form but at each connection between sub-trees omit 1 of the 2 adjacent vertices across the new edge. The choice of which at each is the power (90)*.*

Proof. The theorem can be verified explicitly for $k = 0, 1, 2$. For k=2, additionally if some number n of the 4 connection vertices marked c are required to be omitted from the set then the maximum size is reduced to $QAindsets_2 - n$.

Suppose the theorem and the additional n absent condition is true of some $k{-}1 \geq 2$. For k a set can be made as described above by omitting one of the two vertices across each new connection edge, which gives (89)

$$QAindnum_k = 5\, QAindnum_{k-1} - 4 \qquad k \geq 3$$
$$QAindnumCount_k = 2^4 . QAindnumCount_{k-1}^5$$

If any set bigger than this $QAindnum_k$ is independent then one of its sub-trees has an additional vertex, making it bigger than the maximum of the induction hypothesis, for any choice of which sub-trees have the vertex omitted at the connections.

Similarly when some number n of the 4 connection vertices of k are omitted. A set bigger than $QAindnum_k - n$ in k would again be at least one sub-tree with a set bigger than the induction hypothesis for any choice at the $k{-}1$ connections.

As a remark, induction cannot begin from k=1 since the 5-star there has $QAindnum_1 = 4$ but if n=4 connection vertices are omitted then there is a set of size 1, not $QAindnum_1 - 4 = 0$ for the argument. In figure 26 this is the centre vertex included in the set there. This also means for k=2 no choice of vertices at the connections of middle star to outer stars, only the configuration shown.

The connection vertices where the choice is to be made are new degree-2s in k. All degree-2s arise like this, except those of the initial k=1 and its replications have no choice. These non choices can be reckoned as 1 fewer star replacements, so in k there are $QADegCount(k{-}1, 2)$ connection vertices where a choice of one or the other is to be made, hence (90). $\square$

The independence ratio of a graph is the ratio of independence number to number of vertices. For the R5 quad area tree this is

$$\frac{QAindnum_k}{5^k} = \begin{cases} 1, \frac{4}{5} & \text{if } k = 0, 1 \\ \frac{16}{25} + \frac{1}{5^k} & \text{if } k \geq 2 \end{cases} \rightarrow \frac{16}{25} = 0.64 \quad \text{as } k \to \infty$$

An independent edge set is a set of edges with no end vertices in common, also called a matching since it is vertices in matched pairs with edge between. The match number is the number of pairs in the largest matching.

Theorem 36. *The match number of R5 quad area tree k is*

$$QAmatchnum_k = \begin{cases} 0, 1 & \textit{if } k = 0, 1 \\ 9\,.\,5^{k-2} - 1 & \textit{if } k \geq 2 \end{cases} \tag{91}$$
$$= 0, 1, 8, 44, 224, 1124, \ldots \qquad k{\geq}2 \text{ A198768}$$
$$\textit{base-5} \quad = 0, 1, 13, 134, 1344, 13444, \ldots$$

The number of matchings of this size is

$$QAmatchnumCount_k = \begin{cases} 1,\ 4,\ 513 & \textit{if } k = 0 \text{ to } 2 \\ \dfrac{3^4.13^2}{2^8.5^2} \cdot \left(2^{12}.3^8.5^2.7.13^2\right)^{5^{k-3}} & \textit{if } k \geq 3 \end{cases} \tag{92}$$
$$= 1, 4, 513, 1699990476912, \ldots$$

Factor $3^4.13^2/(2^8.5^2) = 13689/6400$ *and power* $2^{12}.3^8.5^2.7.13^2 = 794794291200$.

Proof. $QAmatchnum$ can be verified explicitly for $k \leq 2$. For k=2 additionally if all 4 connection vertices are required to be unmatched then the match number is still $QAmatchnum_2 = 8$.

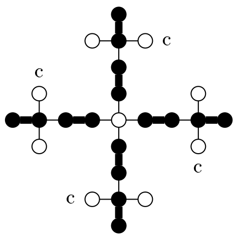


Figure 27:
k=2 matching
$QAmatchnum_2 = 8$,
connection vertices unmatched

Suppose the $QAmatchnum$ formula at (91) and no reduction for connections unmatched is true of some $k{-}1 \geq 2$. Tree k comprises 5 copies of $k{-}1$. Their unmatched connection vertices form new pairs across the new edges,

$$QAmatchnum_k = 5\,QAmatchnum_{k-1} + 4 \qquad \text{starting } QAmatchnum_2 = 8$$

For unmatched connections in k, those connection vertices are some of the outer $k{-}1$ tree connections which can also be unmatched without reducing the match number.

The *QAmatchnumCount* number of such matchings can be verified explicitly for $k \le 1$. $k{=}0$ is the empty matching.

The match number construction above means there is always a pair across the connection vertices of $k \ge 2$, so there are $k{=}2$ sub-parts separated by fixed pairs at those connections. Star replacement means those $k{=}2$ parts correspond to vertices in tree $k{-}2$. The degree of those vertices is the number of connections paired at each, and thus required to be absent for matchings within that $k{=}2$ part.

For $k{=}2$ itself, the $k{-}2 = 0$ tree is a single vertex with no connections. In figure 27 the outermost pairs can be in 3 orientations each, or if one of the inner edges moves in to the middle then 4 orientations in its outer. So

$$QAmatchnumCount_2 = 3^4 + 4.4.3^3 = 513$$

Let $QAmatchnumCount_k(n)$ be the number of matchings in tree k with $n = 0 \,\text{to}\, 4$ of the connection vertices left unmatched. For $k{=}2$ this reduces n many of the outer pair orientations from 3 to 2, or 4 to 3. The following formula reckons outers as 1 to 4 with the first n of them having connections absent. The sum over j is for the respective inner pair shifted towards the middle. Its 4 orientations reduce to 3 when $j \le n$ for connection required to be absent.

$$\begin{aligned} QAmatchnumCount_2(n) &= 2^n.3^{4-n} + \sum_{j=1}^{4} \begin{cases} 3.2^{n-1}.3^{4-n} & \text{if } j \le n \\ 4.2^n \quad .3^{3-n} & \text{if } j > n \end{cases} \\ &= 513, 351, 240, 164, 112 \qquad n = 0 \,\text{to}\, 4 \end{aligned}$$

The total number of matchings is then product of these counts according as number of vertices of each degree d in $k{-}2$,

$$QAmatchnumCount_k = \prod_{d=0,1,2,4} QAmatchnumCount_2(d)^{QADegCount(k-2,\,d)}$$

Each *QADegCount* is a power 5^{k-1}, giving 5^{k-3} in (92). Factor 2 in degree $d = 1, 2$ means corresponding $QAmatchnumCount_2$ is squared so base

$$\begin{aligned} &QAmatchnumCount_2(1)^2.QAmatchnumCount_2(2)^2.QAmatchnumCount_2(4) \\ &\quad = 2^{12}.3^8.5^2.7.13^2 \; = 794794291200 \end{aligned}$$

The constant offsets ± 2 in $d = 1, 2$ become factor

$$\frac{QAmatchnumCount_2(1)^2}{QAmatchnumCount_2(2)^2} = \frac{3^4.13^2}{2^8.5^2} = \frac{13689}{6400}$$

In each case the factorizations of $QAmatchnumCount_2(1, 2, 4)$ give the factorizations of the result. $\square$

The number of matchings in k when n connection vertices required absent follow for $k \geq 3$ by changing n many degree-1 vertices to degree-2 in the $k-2$, giving extra absent outer connections.

$$\begin{aligned}
&\frac{QAmatchnumCount_2(2)}{QAmatchnumCount_2(1)} = \frac{80}{117} \qquad \text{so} \\
&QAmatchnumCount_k(n) = QAmatchnumCount_k \, . \, \left(\tfrac{80}{117}\right)^n \qquad k \geq 3 \\
&\qquad n=1 \quad = 1162386650880, \ldots \qquad k \geq 3 \\
&\qquad n=2 \quad = \ \ 794794291200, \ldots \\
&\qquad n=3 \quad = \ \ 543449088000, \ldots \\
&\qquad n=4 \quad = \ \ 371589120000, \ldots
\end{aligned}$$

A top-down form for the match count is k comprising a $k-1$ middle with 4 connections absent and outers n with 2 absent and the rest 1 absent.

$$\begin{aligned}
QAmatchnumCount_k(n) = \ & QAmatchnumCount_{k-1}(4) \qquad k \geq 3 \\
& . \, QAmatchnumCount_{k-1}(2)^{\,n} \\
& . \, QAmatchnumCount_{k-1}(1)^{\,4-n}
\end{aligned}$$

A dominating set in a graph is a set of vertices for which all other vertices of the graph are adjacent to one or more in the set. The domination number is the number of vertices in the smallest dominating set of a graph.

Theorem 37. *The domination number of R5 quad area tree k is*

$$\begin{aligned}
QAdomnum_k &= \begin{cases} 1 & \text{if } k{=}0 \\ 5^{k-1} & \text{if } k \geq 5 \end{cases} \\
&= 1, 1, 5, 25, 125, 625, \ldots \qquad k \geq 1 \ \text{A000351}
\end{aligned}$$

which is uniquely attained by the middle vertex of each previous level star replacement.

Proof. The middles of the star replacements are degree-4. This is the maximum degree in the tree, and they and their dominated neighbours are disjoint so this is the minimum possible set. □

Second Proof of Theorem 37. Cockayne, Goodman, and Hedetniemi[4] show the domination number of a tree is obtained by starting at leaf vertices and dominating them by requiring their neighbours present in the dominating set. Those leaves and required vertices can then be removed. Remaining vertices they dominated which become leaves or isolated can be removed too, then the procedure repeated.

In the R5 quad area tree, leaf vertices in k occur in the star replacements of degree 1 and 2 vertices of $k-1$. Their neighbours are the centres of those stars. The centre and leaves are removed. The other vertices of the star are then dominated leaves so removed too.

The remaining vertices are the stars from degree-4 vertices in $k-1$. Repeated star replacement means these degree-4s are not adjacent, so the 1,2 star removals make those stars disconnected from each other. Their centre vertices are adjacent to leaves, so domination number is all the centres, which is all the 5^{k-1} vertices of tree $k-1$. □

The domination ratio is the ratio of domination number to number of vertices in a graph. For the R5 quad area tree this is

$$QAdomRatio_k = \frac{QAdomnum_k}{5^k} = \begin{cases} 1 & \text{if } k{=}0 \\ \frac{1}{5} & \text{if } k \geq 1 \end{cases}$$

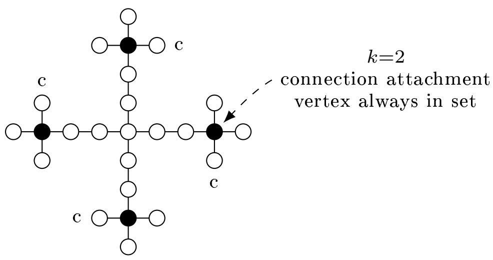
A *total dominating set* in a graph is a set of vertices for which all graph vertices are adjacent to one or more in the set. This differs from an ordinary dominating set in that a vertex in the set does not dominate itself, it must have some neighbour.

Theorem 38. *The number of total dominating sets in R5 quad area tree k is*

$$\begin{aligned} QAtotdomsets_k &= \begin{cases} 0, 15 & \textit{if } k = 0, 1 \\ 820815^{5^{k-2}} & \textit{if } k \geq 2 \end{cases} \\ &= 0, 15,\ 820815,\ 372585905337900111961365759375, \ldots \\ & \textit{with } 820815 = 3.5.54721 \end{aligned}$$

Proof. $k{=}0$ is a single isolated vertex so has no neighbour to dominate it, hence no sets $QAtotdomsets_0 = 0$.

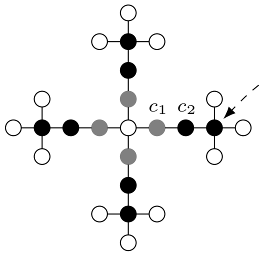
The theorem can then be verified explicitly for $k = 1, 2$. In $k{=}2$ the connection vertices are at a star replacement of a degree-1, so it is one of 3 leaves on the connection attachment. That connection attachment must be present in the set to dominate those other leaves.



When $k{=}2$ trees join there could be additional sets arising by an undominated c which is dominated across the join edge. But this does not give any additional sets since the connection attachment is always present so each c is always dominated already. So sets are simply the product of those in the preceding level.

$$QAtotdomsets_k = 5\, QAtotdomsets_{k-1} \qquad k \geq 3$$

As a remark, it can be noted this is not true for $k{=}2$ from copies of $k{=}1$. The middle $k{=}1$ does not need to have its centre vertex if all its surrounding connections are dominated by the outer sub-trees.



k=2
connection attachment
vertex always in set

c_1 is dominated by c_2 and likewise in the other arms. The centre of the middle is dominated by having at least one of its four c_1 etc neighbours in the set. That is $2^4-1 = 15$ combinations, and the outer leaves can each be present or absent so

$$QAtotdomsets_2 = QAtotdomsets_1^5 + 15.2^{12} \qquad \square$$

The total domination number is the size of the smallest total dominating set of a graph.

Theorem 39. *The total domination number of R5 quad area tree k is*

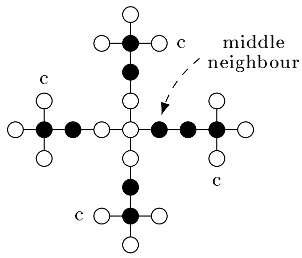
$$QAtotdomnum_k = \begin{cases} none & \text{if } k{=}0 \\ 2 & \text{if } k{=}1 \\ 9.5^{k-2} & \text{if } k \geq 2 \end{cases}$$
$$= none, 2, 9, 45, 225, 1125, \ldots \qquad k{\geq}2 \text{ A189274}$$

and the number of sets of this size is

$$QAtotdomnumCount_k = \begin{cases} 0, 4 & \text{if } k = 0, 1 \\ 4^{5^{k-2}} & \text{if } k \geq 2 \end{cases}$$
$$= 0, 4, 4, 1024, 1125899906842624, \ldots \qquad k{\geq}2\ 4{\times}\text{A185981}$$
$$= 0, 2^2, 2^2, 2^{10}, 2^{50}, 2^{250}, 2^{1250}, \ldots \qquad \text{exponent } k{\geq}2 \text{ A020729}$$

k=0 is a single isolated vertex so has no neighbour to dominate it, hence no value for $QAtotdomnum_0$ and count $QAtotdomnumCount_0 = 0$.

Proof. The theorem can be verified explicitly for $k = 0, 1$. For k=2 the leaf attachments must be in the set, as from theorem 38. To dominate them taking their inner neighbour (rather than a leaf) dominates 2 vertices so gives a smaller set. The centre of the tree can be dominated by any 1 of its 4 neighbours. So total domination number 9 in 4 ways.



k=2
total dominating set
attaining
$QAtotdomnum_2 = 9$

Again as from theorem 38, for $k \geq 3$ there is no cross-domination between sub-trees to reduce set size or increase how many sets. So sum total domination numbers of the sub-parts and product of their counts,

$$\begin{aligned} QAtotdomnum_k &= 5\, QAtotdomnum_{k-1} \\ QAtotdomnumCount_k &= QAtotdomnumCount_{k-1}^5 \end{aligned} \qquad \square$$

The total domination ratio is the ratio of total domination number to number of vertices in a graph. For the R5 quad area tree this is

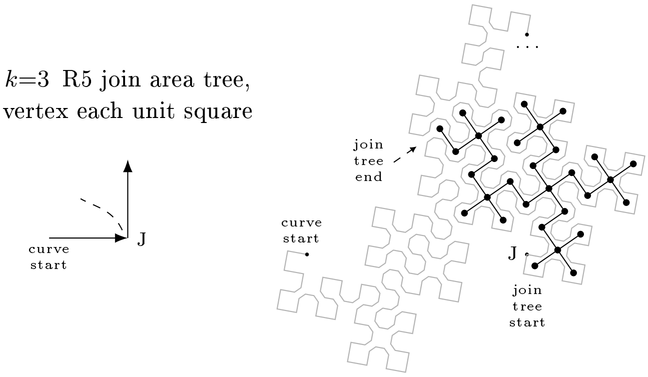
$$QAtotdomRatio_k = \frac{QAtotdomnum_k}{5^k} = \begin{cases} -1, \frac{2}{5} & \text{if } k = 0, 1 \\ \frac{9}{25} = 0.36 & \text{if } k \geq 2 \end{cases}$$

9.2 R5 Join Area Tree

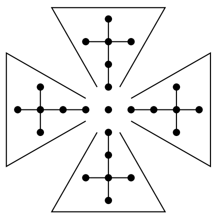
In the join from section 3.1, a tree can be formed by taking the enclosed unit squares on the left of the curve from points JN_k to $JNother_k$.

Like the quad area tree, each vertex is an enclosed unit square and edges are between those consecutive in the curve, or equivalently joined between gaps when the corners of the curve are chamfered off.

k=3 R5 join area tree,
vertex each unit square



This is a quarter of the quad area tree. The unit square at J is an outermost connection vertex in the quad area tree. The innermost end is the centre of the quad area tree. The centre vertex of the quad area tree is not part of the join area tree.



R5 quad area tree
as 4 join area trees

Or conversely the join area tree is descending size quad area trees,

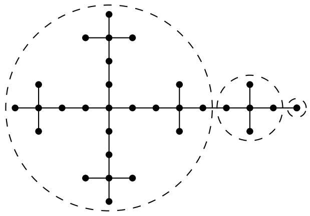


Figure 28:
k=3 join area tree
as descending
quad area trees

This follows from the quad area tree growing by further outer copies of itself (figure 25). Or alternatively, the join area growth in theorem 18 has the existing join area expanding (so star replacement) and further unit square at the end of each expansion

The number of vertices is the quarter of the quad area tree

$$JAV_k = \tfrac{1}{4}(5^k - 1)$$

Or vertices by degree similarly (with the quarter changing a degree-2 from the quad tree to a degree-1 in the join tree),

$$JADegCount(k,0) = \begin{cases} 1 & \text{if } k=1 \\ 0 & \text{otherwise} \end{cases} \qquad \text{A000007}$$

$$\begin{aligned} JADegCount(k,1) &= \begin{cases} 0 & \text{if } k \le 1 \\ \frac{1}{2}(5^{k-1}+3) & \text{if } k \ge 2 \end{cases} \\ &= 0,0,4,14,64,\ldots \qquad k{\ge}2 \text{ A132079} \end{aligned}$$

$$\begin{aligned} JADegCount(k,2) &= \begin{cases} 0 & \text{if } k \le 1 \\ \frac{1}{2}(5^{k-1}-3) & \text{if } k \ge 2 \end{cases} \\ &= 0,0,1,11,61,\ldots \qquad \text{A137410} \end{aligned}$$

$$\begin{aligned} JADegCount(k,4) &= \begin{cases} 0 & \text{if } k{=}0 \\ \frac{1}{4}(5^{k-1}-1) & \text{if } k \ge 3 \end{cases} \\ &= 0,0,1,6,31,\ldots \qquad \text{A003463} \end{aligned}$$

The centroid of a tree is the vertex or vertices for which all its attached sub-trees are $\le \frac{1}{2}$ total tree vertices. Per Jordan the centroid is either 1 vertex or 2 adjacent vertices.

For the quad area tree the centroid is the tree middle. Its neighbours are $\frac{1}{4}$ of the tree each. Its neighbour (or any other vertex) has sub-tree towards the middle with $\ge \frac{3}{4}$ of the tree. Or by symmetry if any other vertex was a centroid then there would be 4 or more equivalent vertices also centroids, contrary to at most 2.

Theorem 40. *The centroid of join area tree k is the middle of its biggest component quad area tree.*

Proof. The biggest quad middle has attached sub-trees of JAV_{k-1} and $2JAV_{k-1}$ so $\le \frac{1}{2}$ the whole tree $5JAV_{k-1}+1$. The neighbours of the quad middle have $3JAV_{k-1}+1$ or $4JAV_{k-1}+1$ towards the middle, so $> \frac{1}{2}$ the whole. □

Or by symmetry vertices of the three sub-trees not with the descending further quads are equivalent so if any were centroids then there would be 3, contrary to at most 2 centroids.

Theorem 41. *The diameter of R5 join area tree k is*

$$JAdiameter_k = \sum_{j=1}^{k-1} QAdiameter_j + 1 \tag{93}$$

$$= \begin{cases} none & \text{if } k{=}0 \\ \frac{1}{2}(3^k - 3) & \text{if } k \geq 1 \\ \quad = \frac{1}{2} QAdiameter_k - 1 & \end{cases}$$

$$= 0, 0, 3, 12, 39, 120, \ldots \qquad k{\geq}1 \text{ A029858}$$

$$ternary \quad = 0, 0, 10, 110, 1110, 11110, \ldots \qquad k{\geq}1 \text{ A105279}$$

The number of paths attaining the diameter (reckoning the empty graph as no paths) is

$$JAdiameterCount_k = \begin{cases} 0 & \text{if } k{=}0 \\ 3^{k-1} & \text{if } k \geq 1 \end{cases} \qquad \text{A140429}$$

The number of diameter endpoints, and total number of vertices on some diameter are

$$JAdiameterEnds_k = \begin{cases} 0, 1 & \text{if } k = 0, 1 \\ 3^{k-1} + 1 & \text{if } k \geq 2 \end{cases}$$

$$= 0, 1, 4, 10, 28, \ldots \qquad \text{A103457}$$

$$JAdiameterVertices_k = \begin{cases} 0, 1 & \text{if } k = 0, 1 \\ k.3^{k-1} & \text{if } k \geq 2 \end{cases}$$

$$= 0, 1, 6, 27, 108, \ldots \qquad \text{A027471}$$

Proof. Take the join tree as descending component quad trees per figure 28. A path entirely within a single component quad tree is at most its diameter. A path starting and ending in two quad trees is at most the sum of those, the trees between, and the edge between each, since the quad trees are joined at connection vertices, which are a diameter distant.

The sum at (93) is then all quad trees in the join tree. It has +1 for the edge between each, and extends down only to $j{=}1$ so there is no +1 after the final level 0 quad tree. Quad tree 0 is a single vertex $QAdiameter_0 = 0$ so its omission from the sum does not change the result.

The end quad tree $k{=}0$ is 1 vertex so is the end of all diameter paths. The starts are vertices in the $k{-}1$ component quad tree which are a diameter away from the connection vertex. The $k{=}1$ quad in join tree $k{=}2$ has 3 such vertices (all its leaves). Then in the manner of theorem 30, star replacement make 3 new vertices as ends of those paths, so that

$$JAdiameterCount_k = 3\, JAdiameterCount_{k-1}$$
$$\text{starting } JAdiameterCount_2 = 3$$

The 3 parts of the end-most quad which are away from its connection give join area tree endpoints and total vertices.

$$\begin{aligned} \mathit{JAdiameterEnds} &= \tfrac{3}{4}\,\mathit{QAdiameterEnds}_{k-1} + 1 && k \ge 2 \\ \mathit{JAdiameterVertices} &= \tfrac{3}{4}\big(\mathit{QAdiameterVertices}_{k-1} - 1\big) && k \ge 1 \\ &\quad + \mathit{JAdiameter}_{k-1} - \tfrac{1}{2}\,\mathit{QAdiameter}_{k-1} + 1 && \square \end{aligned}$$

Theorem 42. *The number of independent sets in R5 join area tree k is*

$$\begin{aligned} \mathit{JAindsets}_k &= aa_k \\ &= 1, 2, 26, 11050330, \ldots \end{aligned}$$

where

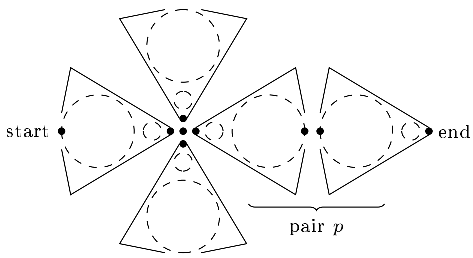
$$\begin{aligned} aa_k &= aa_{k-1}^3 paa_{k-1} &&+ az_{k-1}^3 paz_{k-1} \qquad k \ge 2 \\ az_k &= aa_{k-1}^3 paz_{k-1} &&+ az_{k-1}^3 pzz_{k-1} \\ za_k &= za_{k-1}\, aa_{k-1}^2 paa_{k-1} &&+ zz_{k-1}\, az_{k-1}^2 paz_{k-1} \\ zz_k &= za_{k-1}\, aa_{k-1}^2 paz_{k-1} &&+ zz_{k-1}\, az_{k-1}^2 pzz_{k-1} \end{aligned} \tag{94}$$

$$\begin{aligned} paa_k &= aa_k^2 - wa_k^2 \qquad & wa_k &= aa_k - za_k \\ pzz_k &= az_k^2 - wz_k^2 \qquad & wz_k &= az_k - zz_k \\ paz_k &= aa_k\, az_k - wa_k\, wz_k \end{aligned}$$

$$\textit{starting} \quad aa_1 = 2, \quad za_1 = za_1 = zz_1 = 1$$

$$\begin{aligned} az_k &= 1, 1, 17,\ 7186721, \ldots \qquad & paa_k &= 1, 3, 532, 95941084541604, \ldots \\ za_k &= 1, 1, 14,\ 5934794, \ldots \qquad & paz_k &= 1, 2, 346, 62396495941658, \ldots \\ zz_k &= 1, 1,\ \ 9,\ 3859769, \ldots \qquad & pzz_k &= 1, 1, 225, 40580349121537, \ldots \end{aligned}$$

Proof. The join area tree can be taken as 5 join trees $k-1$. The biggest component quad tree is $k-1$ and it comprises 4 join trees $k-1$.



aa, za, az, zz are counts of independent sets,

$$\begin{aligned} aa &= \text{start and start and end each either present or absent} \\ za &= \text{start required to be absent} \\ az &= \text{end required to be absent} \\ zz &= \text{both start and end vertices required to be absent} \end{aligned}$$

The p forms are for the pair of join trees shown at the right. They have a new edge between two $k-1$ start vertices. So paa is sets aa^2 except not both across the edge so subtract wa^2. wa is number of sets with the start vertex present. It is the unrestricted aa less za which is those without start vertex. Similarly the other p forms, and using wz which is start present and end absent.

The vertex at the centre of the 4 sub-trees is in addition to those trees. The term at (94) etc are without the centre (allowing any end of the joined parts) and with the centre (so the ends must be absent). The pair p is symmetric so a pza for it is the same as its paz. The recurrence applies for $k \geq 2$ where the sub-trees are not empty. □

These aa, az, za, zz are recurrences in 4 quantities rather than 5 for $QAindsets$. The quad area tree as 4 join area parts gives the latter as, again the centre without or with,

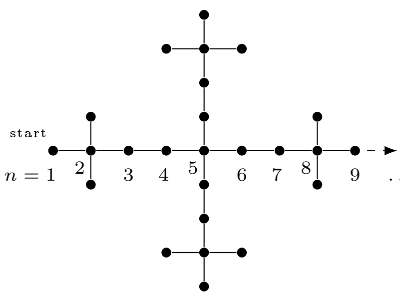
$$QAindsets_k = aa_k^4 + az_k^4$$

Another approach for both join and quad area trees is to work along a diameter. The sub-trees branching off are join area trees of appropriate level. For computer calculation or similar this is less efficient, but the formulas are simpler.

Theorem 43. *Consider a tree comprising the first n many vertices across a diameter of the quad area tree (of level big enough to have diameter $\geq n$), and the branches off those vertices.*

Let da_n be the number of independent sets in this tree, and dz_n the number which do not include vertex n. They are given by recurrences

$$\begin{aligned}
da_n &= dz_n + dz_{n-1}\, dz^{\,2}_{TernaryLowOnes(n-1)} && \textit{starting } da_0 = 1\\
dz_n &= \phantom{dz_n + {}} da_{n-1}\, da^{\,2}_{TernaryLowOnes(n-1)} && dz_0 = 1\\
da_n &= 1, 2, 9, 17, 26, 22489, 40065, 62554, 290281, 540497, \ldots\\
dz_n &= 1, 1, 8,\ \ 9, 17, 17576, 22489, 40065, 250216, 290281, \ldots\\
TernaryLowOnes(n) &= n \textit{ in ternary keep only low run of 1 digits}\\
&= \tfrac{1}{2}\left(3^{TernaryCountLowOnes(n)} - 1\right)\\
&= 0, 1, 0, 0, 4, 0, 0, 1, 0, 0, 1, 0, 0, 13, \ldots\\
TernaryCountLowOnes(n) &= 0, 1, 0, 0, 2, 0, 0, 1, 0, 0, 1, 0, 0, 3, \ldots && \text{A253786}
\end{aligned}$$



R5 quad area tree
vertices along a diameter

Proof. The recurrences are again the usual way to combine counts for sub-trees at a given vertex. dz_n is without vertex n so it is product of all combinations in its attached sub-trees. da_n adds sets which include the vertex at n. For them the sub-trees must not have their end vertex so dz_{n-1} etc.

The empty tree is $a_0 = 1$ since the empty set is independent. $z_0 = 1$ likewise, reckoned as "without" vertex n. It has no vertices at all and remains independent if the next vertex $n{+}1$ is present.

The sub-trees at n are its preceding $n{-}1$, and trees branching above and below. The tree pattern means these are determined by the low 1 digits of $n{-}1$ in ternary.

Star replacement in the tree turns each n into new vertices $n' = 3n{-}2, 3n{-}1, 3n$. For example $n{=}1$ becomes $1, 2, 3$. The first and last of these new vertices have empty trees above and below. They are $n'{-}1 \equiv 0, 2 \bmod 3$ which is no low ternary 1-digits. In the recurrences $\mathit{TernaryLowOnes}(n'{-}1) = 0$ is those empty trees.

The middle new vertex $n' = 3n{-}1$ has the existing branches above and below n. It is $n'{-}1 \equiv 1 \bmod 3$ so an additional low ternary 1-digit over what $n{-}1$ had. Hence on repeated star replacement $\mathit{TernaryLowOnes}(n{-}1)$ is the sub-tree attaching above and below vertex n. □

A full width diameter is

$$\mathit{QAindsets}_k = da_{3^k}$$

The recurrences only re-use previous values of da, dz at

$$\begin{aligned} n &= \mathit{TernaryRunOnes}(m) = m \text{ many ternary 1 digits} \\ &= \tfrac{1}{2}(3^m - 1) && \text{A003462} \end{aligned}$$

These re-used values are join area trees. Each n is a centre vertex and the branches above and below are the appropriate level join tree for that centre. (Per the star replacement construction in the proof.) So $\mathit{JAindsets}$ and the az from theorem 42 are

$$\begin{aligned} \mathit{JAindsets}_k &= da_{\,\mathit{TernaryRunOnes}(k)} \\ az_k &= dz_{\,\mathit{TernaryRunOnes}(k)} && \text{start of JA not in set} \end{aligned}$$

Or the recurrences can be written with az and $aa = \mathit{JAindsets}$ and count of low 1s

$$\begin{aligned} da_n &= dz_n + dz_{n-1}\, az^{\,2}_{\mathit{TernaryCountLowOnes}(n-1)} \\ dz_n &= \phantom{dz_n +{}} da_{n-1}\, aa^{\,2}_{\mathit{TernaryCountLowOnes}(n-1)} \end{aligned}$$

Theorem 44. *The independence number of R5 join area tree k is*

$$\mathit{JAindnum}_k = \begin{cases} 0, 1 & \text{if } k = 0, 1 \\ 4.5^{k-2} & \text{if } k \ge 2 \end{cases}$$

The number of independent sets of this size is

$$\mathit{JAindnumCount}_k = \begin{cases} 1, 1, 2 & \text{if } k \le 2 \\ 2^{\frac{1}{2}(\mathit{JADegCount}(k-1,2)+3)} & \text{if } k \ge 3 \end{cases} \tag{95}$$

$$= 1, 1, 2, 4, 128, 4294967296, \ldots$$

$$\log_2 \mathit{JAindnumCount}_k = 0, 0, 1, 2, 7, 32, 157, 782, \ldots \qquad k{\ge}2 \ \text{A047850}$$

Proof. $k{=}0$ is empty which has a single empty independent set.

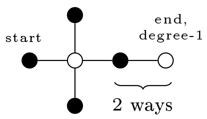
$k{=}1$ is a single vertex so its biggest independent set is only that single vertex.

Suppose then the formula is true of some $k{-}1 \geq 1$ and further that if the start vertex (connection in the biggest component quad tree) is required to be absent then the biggest set is $JAindnum - 1$. This is so for $k{-}1 = 1$.

Join area tree k has a quad area tree $k{-}1$ attached at the start of join area tree $k{-}1$. Either the start vertex of the join tree or the connection vertex of the quad tree must be omitted. As from theorem 35, omitting one vertex from it reduces to $QAindnum - 1$ there too. So

$$JAindnum_k = JAindnum_{k-1} + QAindnum_{k-1} - 1 \qquad k \geq 2$$

There are 2 choices for which connecting vertex is omitted. In the manner again of theorem 35, these choices become, for $k \geq 3$, all the degree-2 vertices except the lowest level, but with the end-most degree-1 counted too, and the degree-2 pair between quads 1 and 2, hence (95).



$k{=}2$ R5 join area tree
independence number
$JAindnum_2 = 4$ vertices
$JAindnumCount_2 = 2$ ways

□

The independence ratio of the R5 join area tree is then, for $k \geq 1$ where it is not empty,

$$\frac{JAindnum_k}{JAV_k} = \begin{cases} 1 & \text{if } k=1 \\ \frac{16}{25}\left(1 - \frac{1}{5^k - 1}\right) & \text{if } k \geq 2 \end{cases} \to \frac{16}{25} \quad \text{same as quad area tree}$$

Theorem 45. *The match number of R5 join area tree k is*

$$\begin{aligned} JAmatchnum_k &= \begin{cases} 0 & \textit{if } k \leq 1 \\ \frac{1}{4}\left(9\,.\,5^{k-2} - 1\right) & \textit{if } k \geq 2 \end{cases} \\ &= 0, 0, 2, 11, 56, 281, 1406, \ldots \qquad {\scriptstyle k \geq 2 \ \text{A198769}} \\ \textit{base-5} \quad &= 0, 0, 2, 21, 211, 2111, 21111, \ldots \end{aligned}$$

Proof. Take the join area tree as connected descending quad area trees. From theorem 36, $QAmatchnum$ is not reduced by requiring connection vertices absent, so a new pair between each quad, and the end-most part of the join area tree considered explicitly gives

$$JAmatchnum_k = 2 + \sum_{j=2}^{k-1} QAmatchnum_j + 1$$

□

Theorem 46. *The domination number of R5 join area tree k is*

$$\begin{aligned} JAdomnum_k &= \begin{cases} 1 & \textit{if } k{=}0 \\ \frac{1}{4}(5^{k-1} + 3) & \textit{if } k \geq 5 \end{cases} \\ &= 0, 1, 2, 7, 32, 157, \ldots \qquad {\scriptstyle k \geq 1 \ \text{A047850}} \end{aligned}$$

Proof. In the same way as theorem 37, the smallest dominating set is all the degree-4 vertices, except at the end-most degree-1 either it or its degree-2 neighbour is required too.

$$JAdomnum = JADegCount(k, 4) + 1 \qquad k \geq 1$$ □

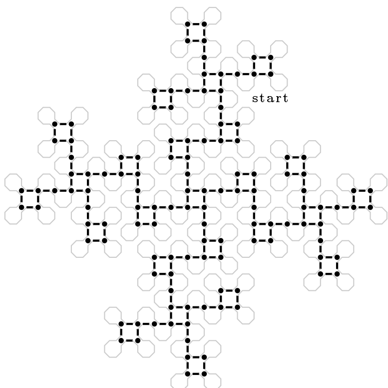
The domination ratio of the R5 join area tree is then, for $k \geq 1$ where it is not empty,

$$JAdomRatio_k = \frac{JAdomnum_k}{JAV_k} = \frac{1}{5}\left(1 + \frac{14}{5^k - 1}\right) \to \frac{1}{5} \quad \begin{array}{c}\text{same as}\\ \text{quad area tree}\end{array} \qquad k \geq 1$$

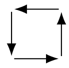
9.3 R5 Quad Turns Graph

When the R5 dragon or quad revisits a location, the second visit is the same turn as the first. This is so for any non-crossing closed curve, or curve continuing infinitely and not encircling its start, since an opposite turn would enclose either the end or the start,

In an R5 quad, locations with right turns and the segments between them form a graph which is a tree of cycles,



k=3 R5 quad turns graph, vertex at each right turn, edges segments between



This has the same structure as the area tree, but the deepest 5-stars are instead 4-cycles.

The sub-curve expansion in figure 24 shows that the connections between sub-graphs are at adjacent arms of the sub-graphs. For example c to j in the following

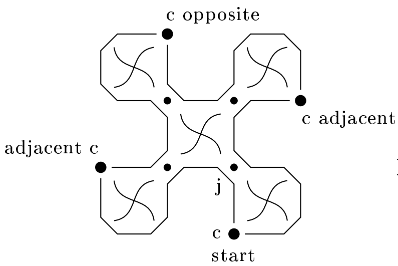


Figure 29: R5 quad turns graph connections

The quad area tree above has adjacent sub-connections like this too, but for it there is always a single centre vertex so arm order doesn't matter. Here the centre of each sub-graph is a 4-cycle so positions around it do matter.

The expansions show the turns graph is a kind of 4-cycle replacement of the $k{-}1$ area tree. Each vertex is replaced by a 4-cycle and existing edges go between those cycles with an extra vertex in between. At a degree-2 vertex the new edges are at adjacent vertices of the cycle, since they are curve starts and ends. But at a degree-4 vertex the edge order must follow the geometry.

The number of vertices follows from 5 copies of $k{-}1$ with a new vertex between each of the 4 connections. Similarly edges, with 2 new edges between each of the 4 connections.

$$\begin{aligned} QTV_k &= 5QTV_k + 4 \quad \text{starting } QTV_1 = 4 \\ &= 5^k - 1 \qquad\qquad \text{number of vertices} \qquad\qquad \text{A024049} \\ QTE_k &= 5\,QTE_{k-1} + 8 \quad \text{starting } QTE_1 = 4 \\ &= \begin{cases} 0 & \text{if } k{=}0 \\ 6.5^{k-1} - 2 & \text{if } k \ge 1 \end{cases} \qquad \text{number of edges} \\ &= 0, 4, 28, 148, 748, 3748, \ldots \qquad\qquad 2\times \text{ A198762} \end{aligned}$$

At each of the 4 connections the new vertex is degree-2, and two degree-2 vertices on each side become degree-3 so the number of each are

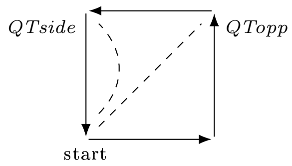
$$\begin{aligned} QTV2_k &= 5QTV2_{k-1} - 4 \quad \text{starting } QTV2_1 = 4 \\ &= \begin{cases} 0 & \text{if } k{=}0 \\ 3.5^{k-1} + 1 & \text{if } k \ge 1 \end{cases} \qquad \text{degree 2} \\ &= 0, 4, 16, 76, 376, 1876, \ldots \qquad\qquad \text{A199214} \\ QTV3_k &= 5QTV3_{k-1} + 8 \quad \text{starting } QTV3_1 = 0 \\ &= \begin{cases} 0 & \text{if } k{=}0 \\ 2.5^{k-1} - 2 & \text{if } k \ge 1 \end{cases} \qquad \text{degree 3} \\ &= 0, 0, 8, 48, 248, 1248, \ldots \qquad\qquad 2\times\text{A024049} \\ QTV_k &= QTV2_k + QTV3_k \quad \text{total} \end{aligned}$$

Theorem 47. *The diameter of R5 quad turns graph k is*

$$\begin{aligned} QTdiameter_k &= \begin{cases} 0 & \text{if } k = 0 \\ 3^k + 2k - 3 & \text{if } k \ge 1 \end{cases} \\ &= 0, 2, 10, 30, 86, 250, \ldots \end{aligned} \tag{96}$$

This is attained by 2 paths, between opposite corners of the quad.

Proof. Connections between sub-trees are at the start and ends of the component curves. Let *QTside* be the distance through the graph from one connection vertex to another along a side, and *QTopp* to the opposite.



There are two side connections, one on the left and one on the right. By symmetry their distance is the same. From figure 29 the sub-graph side or opposite parts are, plus 2 edges at each of 2 connections,

$$\begin{aligned}
QTside_k &= 3QTside_{k-1} + 4 \quad \text{starting } QTside_1 = 1 \\
&= \begin{cases} 0 & \text{if } k = 0 \\ 3^k - 2 & \text{if } k \geq 1 \end{cases} \\
&= 0, 1, 7, 25, 79, 241, \ldots && \text{A058481} \\
QTopp_k &= 2QTside_{k-1} + QTopp_{k-1} + 4 \quad \text{starting } QTopp_1 = 2 \\
&= 3^k - 1 && \text{A024023}
\end{aligned}$$

Expanding repeatedly shows $QTopp$ differs from $QTside$ by 1

$$QTopp_k = QTside_k + 1 \qquad k \geq 1$$

In the graph the adjacent and opposite arms are the same by symmetry, but to reach the opposite one is 1 extra edge at the middle cycle.

Let $QTecc$ be the eccentricity of a connection vertex. Going $QTside$ to the middle $k-1$ sub-graph, then $QTopp$ across it, and $QTecc$ there, plus 2 edges at each of the 2 connections is

$$\begin{aligned}
QTecc_k &= QTside_{k-1} + QTopp_{k-1} + QTecc_{k-1} + 4 \qquad k \geq 2 \\
&= \begin{cases} 0 & \text{if } k = 0 \\ 3^k + k - 2 & \text{if } k \geq 1 \end{cases} \qquad \text{connection eccentricity} \\
&= 0, 2, 9, 28, 83, 246, 733, 2192, \ldots
\end{aligned}$$

This is greater than going into one of the side graphs across the middle (by 1 edge, being $QTside$ instead of $QTopp$). From the formulas it is greater than staying in the originating $k-1$ sub-graph ($QTecc_{k-1}$ unchanged), or staying in the middle ($QTside_{k-1} + 2 + QTecc_{k-1}$).

Eccentricity 2 in k=1 is uniquely attained across to the opposite side of the 4-cycle, so all subsequent k are unique too.

A path going opposite connections of the middle and eccentricity into the outer sub-graphs is the diameter (96),

$$QTdiameter_k = 2QTecc_{k-1} + QTopp_{k-1} + 4 \qquad k \geq 2$$

This is greater than going adjacent connections across the middle, again by 1 edge. From the formulas it is greater than staying in one $k-1$ sub-graph ($QAdiameter_{k-1}$ unchanged), or between just two sub-graphs ($2QTecc_{k-1} + 2$).

There are 2 cross pairs of outer sub-graphs, and since the eccentricity into them is then uniquely attained there are just 2 paths attaining the diameter. □

Theorem 48. *The Wiener index of the R5 quad turns graph is*

$$QTW_k = \begin{cases} 0 & \textit{if } k{=}0 \\ 5^k\left(\frac{2}{7}15^k + (\frac{1}{5}k - \frac{17}{20})5^k + \frac{79}{140}\right) & \textit{if } k \geq 1 \end{cases}$$
$$= 0, 8, 1340, 116700, 9021000, \ldots$$

Proof. Let $QTwC_k$ be the total path lengths in graph k from a connection vertex to all other vertices. The $k{-}1$ sub-graphs give

$$\begin{aligned} QTwC_k &= 5\,QTwC_{k-1} \\ &\quad + QTV_{k-1}\begin{pmatrix} QTside_{k-1} + 2 \\ 2\big(2QTside_{k-1} + 4\big) \\ QTside_{k-1} + QTopp_{k-1} + 4 \end{pmatrix} && (97) \\ &\quad + QTside_{k-1} + 1 && (98) \\ &\quad + 2\big(2QTside_{k-1} + 3\big) \\ &\quad + QTside_{k-1} + QTopp_{k-1} + 3 \\ &\qquad \text{starting } QTwC_1 = 4 \\ &= \tfrac{1}{10}5^k\big(7.3^k + 2k - 17\big) + 1 \qquad \text{if } k \geq 1 \\ &= 0, 4, 126, 2226, 34876, \ldots \end{aligned}$$

(97) is distances to the middle and 3 outer sub-graphs, multiplied by QTV many vertices at each. The further terms (98) are path lengths to the new vertices at the connections.

The Wiener index is then a recurrence

$$\begin{aligned} QTW_k &= 5\,QTW_{k-1} && (99) \\ &\quad + 10.2.\,QTV_{k-1}.\,QTwC_{k-1} && (100) \\ &\quad + QTV_{k-1}^2\begin{pmatrix} 4.2 \\ +\,4\,(QTside_{k-1}{+}4) \\ +\,2\,(QTopp_{k-1}{+}4) \end{pmatrix} && (101) \\ &\quad + 4\left(5QTwC_{k-1} + QTV_{k-1}\begin{pmatrix} 2.1 \\ +\,2\,(QTside_{k-1}{+}3) \\ +\,2\,(QTopp_{k-1}{+}3) \end{pmatrix}\right) && (102) \\ &\quad + 4\,(QTside_{k-1}{+}3) + 2\,(QTopp_{k-1}{+}3) && (103) \end{aligned}$$

(99) is paths within the 5 sub-graphs. Paths from one sub-graph to another must go to the connection vertex of the sub-graph. This is total $QTwC$, multiplied by the number of destinations QTV. Likewise at the other end so this twice, and there are 10 pairs of the 5 sub-graphs so (100).

The paths then have a distance between their sub-graph connection vertices. Middle to outer is distance 2 each for 4 pairs. Adjacent outers is $QTside + 4$ each for 4 pairs. Opposite outers is $QTopp + 4$ each for 2 opposing pairs. Those distances are between QTV many vertices on each side so that number squared for (101).

(102) is distance from the 4 new vertices to everything else, and (103) is

between those new vertices. $\square$

The limit for mean distance between distinct vertices is the same as the area tree (80). The number of vertices differs by a constant, the diameter by a linear term, and the high coefficient in QTW is the same.

$$\frac{QTW_k}{\binom{QTV_k}{2}\, QTdiameter_k} \to \frac{4}{7} \qquad \text{same as } QAW$$

Theorem 49. *The independence number of R5 quad turn graph k is*

$$\begin{aligned} QTindnum_k &= \begin{cases} 1 & \text{if } k{=}0 \\ \frac{1}{2}\,QTV_k & \text{if } k \ge 1 \end{cases} \\ &= 1, 2, 12, 62, 312, 1562, \dots \qquad\qquad k{\ge}1 \text{ A125831} \end{aligned}$$

Proof. $k{=}0$ is the empty graph and $k{=}1$ is a 4-cycle. They can be verified explicitly.

Let $QTindnum_k(n)$ be the size of the biggest set with n many of the 4 connection vertices required to be absent. For $n{=}2$ there are two configurations, either adjacent vertices required absent or opposites required absent.

Suppose then the theorem is true of some $k{-}1 \ge 1$ and further that there

$$QTindnum_{k-1}(n) = QTindnum_{k-1} - \begin{cases} 0 & \text{if } n = 0, 1 \\ 0 & \text{if } n = 2 \text{ opposite} \\ 1 & \text{if } n = 2 \text{ adjacent} \\ 1 & \text{if } n = 3 \\ 2 & \text{if } n = 4 \end{cases} \tag{104}$$

This holds for the 4-cycle level $k{-}1 = 1$.

Graph k then comprises 5 of $k{-}1$ with join vertices j between.

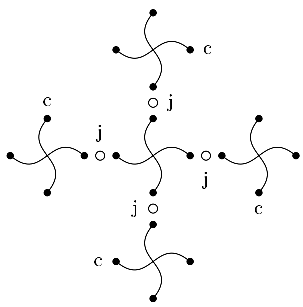


Figure 30:
R5 quad turns graph
5 copies and
join vertices between

Each j vertex can be present only when both its adjacent connection vertices are absent. For 2 join vertices, the middle can be an $n{=}2$ opposite and their outers $n{=}1$ absent. These parts are all no reduction at (104), so

$$QTindnum_k = 5\,QTindnum_{k-1} + 2 \qquad k \ge 2 \tag{105}$$

To see this is the maximum, for 3 or 4 join vertices the middle must have $n{=}3$ or $n{=}4$ absent and per (104) this reduces the biggest there by 1 or 2 so net the same $+2$ as (105).

For 0 or 1 join vertices, the middle and outers have no reduction in (104), but only +0 or +1 join vertices so less than (105).

To show reductions (104) hold in k, the connection vertices in k are the outer vertices marked c in figure 30. Take some of them absent as necessary for the n of k. If its corresponding j vertex is present then the maximum set in the outer is n=2 adjacent (reduction 1), or if j absent then n=1 (no reduction).

The set of j present or absent determines the type of the middle part. Enumerating the 16 possible j present or absent gives the reductions of (104) as the biggest sets in k too. □

The independence ratio of the turns graph exists, for $k \ge 1$ which is where it is not empty, is

$$\frac{QTindnum_k}{QTV_k} = \frac{1}{2} \qquad k \ge 1 \quad \text{independence ratio}$$

In the construction at (105), $\frac{1}{2}$ of the join vertices are included in k and the sub-parts are $\frac{1}{2}$ their existing vertices, so that ratio is maintained.

10 Fractional Locations

The location of a point $0 \le f \le 1$ along the R5 dragon curve fractal is a limit

$$fpoint(f) = \lim_{k \to \infty} \frac{point(\lfloor f.5^k \rfloor)}{b^k} \qquad \text{fractional point}$$

$n = \lfloor f.5^k \rfloor$ is the first k digits below the radix point of f written in base-5. The location is then a change to powers $\pm b^j$ high to low as per (14).

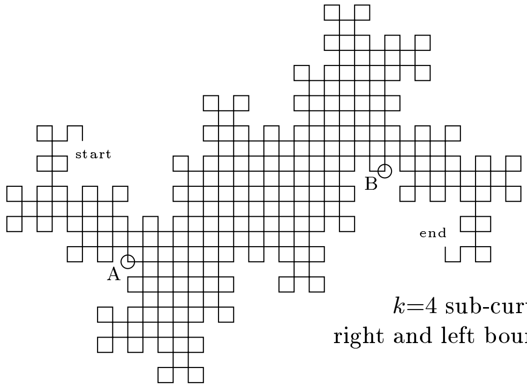
When f is rational, its digits are an initial fixed part then a repeating periodic part (of length at most denominator $-$ 1). The b powers are then likewise periodic and give a location as some $x+iy$ with rational x, y.

When taking successive digits of f written in base-5, this is the fractional position within a sub-fractal. For rational f the denominator is unchanged (hence eventually repeating).

10.1 Fractional Boundary

Theorem 50. *The only points on both left and right boundary of the R5 dragon fractal are curve start and end $f = 0, 1$.*

Proof. k=4 sub-curves are as follows



k=4 sub-curves for
right and left boundary boxes

Take bounding boxes of width *Hwpf* and extension *Hepf* from section 5. Put bounding boxes around the left boundary segments. The curve is non-crossing so all left boundary locations are within these boxes.

The right boundary is within corresponding bounding boxes around right boundary segments. The boxes from A around to B are disjoint from the left boundary boxes. So the spiralling and curling within those right boundary parts never reaches the left boundary.

On expansion to k=5, the right boundary parts from start to A expand to the same form as k=4, leaving only sub-curves through to a new smaller A as possible both boundary. Repeating this excludes points at arbitrarily small distance from the start, leaving only the start as both left and right boundary.

Right boundary parts end through B likewise on expansion. ☐

Boundary locations A and B are chosen based on bounding boxes. The box size means they must be after the points which are an offset 1 segment distance between left and right (above A and above right B).

A tighter bound on the sub-curve extents can show no overlap there either. The 10-side hull bound ahead in theorem 52, drawn in figure 31, is one possibility. It allows a bigger A–B, or smaller k, for the same result.

Theorem 51. *The R5 dragon fractal has no cut points, ie. is a topological disc.*

Proof. If a cut point separates start and end then it is on both left and right boundary, but from theorem 50 there are no such points.

Suppose a cut point separates a lobe from the boundary. If this point is somewhere within a sub-curve then it separates start and end of that sub-curve, but again no such point exists.

Otherwise the point is at the start or end of some sub-curve. The curve would go to the lobe and back again so such a point would be double-visited. But by plane filling, a double-visited is not on the boundary, so not a cut point. ☐

Theorem 52. *Fractional f on the boundary of the R5 dragon fractal are*

> *f written in base-5 fractional digits,*
> *$fRpred(f) = 1$ iff f sans 1s none of the Rpred disallowed digit pairs*
> *$fLpred(f) = 1$ iff f sans 3s none of the Lpred disallowed digit pairs*
> *$fBpred(f) = fRpred(f)$ or $fLpred(f)$*

Proof. A sub-curve which is *Rpred* boundary has outside points at most $\frac{1}{2}\sqrt{5}^k$ away. Any f which has strings of initial digits always satisfying *Rpred* is therefore an arbitrarily small distance from the outside and so on the boundary. Such an f has none of the *Rpred* disallowed pairs.

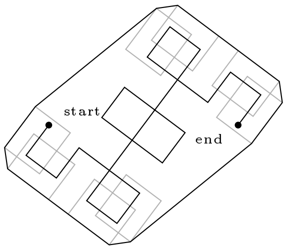
A sub-curve which does not satisfy *Rpred* has 3 enclosing sub-curves on its right side. Since they have no cut points, they enclose all of that side except start and end.

A double-visited start is not on the boundary, since the 4 sub-curves enclose that point per plane filling.

A single-visited left turn is on the right boundary since it has an absent sub-curve on its right. *Rpred* accepts this already too for the same reason. A single-visited right turn is not on the right boundary by plane filling.

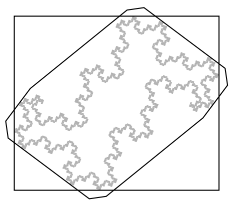
Similarly *fLpred*. □

Second Proof of Theorem 52. An upper bound for the convex hull around the fractal can be formed by taking k=2 segments, putting *Hwpf*, *Hepf* width and extent boxes around them (section 5), and taking the convex hull around those boxes. The result is a 10-side polygon.



10-side hull around
width and extent boxes
of k=2 segments

It can be noted this hull bound is wider and extends further than the plain *Hwpf* box. Both the box and the bound contain the curve. Roughly speaking, the box corners are mostly empty so having them as hull vertices goes out beyond actual curve.



This hull bound shows that curves 1 apart and offset do not touch. Level k=2 boxes are used in the bound for this reason.

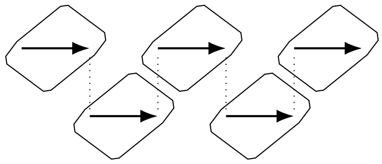


Figure 31:
offset 1 apart
segments

A sub-curve m has its hull bound touched or overlapped by the hull bounds of the following surrounding sub-curves,

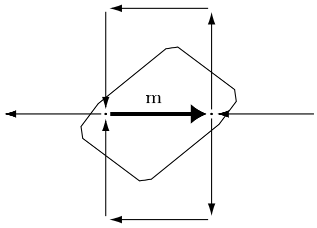


Figure 32:
surrounding segments
whose hull bounds
touch or overlap
bound of m

If m has all the segments of figure 32 surrounding it then it is non-boundary since, by construction, it does not touch or overlap the hull of any absent outside. Conversely, if m has one or more of the segments of figure 32 absent, then that is some part of the hull bound of m which is outside the curve and therefore some of m possibly on the boundary.

Segments and their hull bounds beyond figure 32, so not touching m, can be illustrated

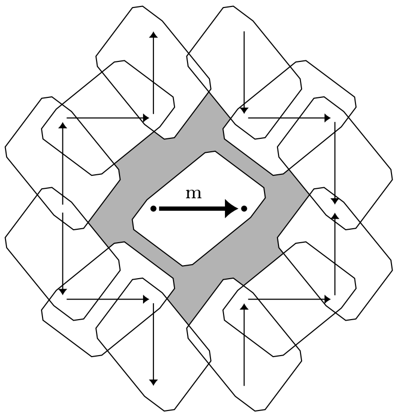


Figure 33:
non-touching
hulls

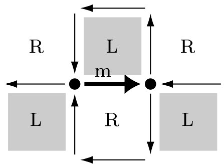
When m is surrounded by all segments of figure 32, the grey area here is a minimum amount of filled region surrounding m. The outer hulls shown are those necessary to delimit the grey. Actually the grey may be bigger, since the curve continues at the left and right segment ends in figure 32 (the curve does not start or end this way), so there is at least one vertical there also surrounding m. But knowing that is not necessary.

A given sub-curve m has some of figure 32 surrounding segments. The initial single segment $k{=}0$ has none. On expansion, m is five new segments and there are other new segments around. On expansion there are new segments around the three new sub-curves. The segments of figure 32 suffice to determine the corresponding set of segments around each new segment. A finite set of configurations arise and give a state machine traversed by base-5 digits of f.

A fully surrounded configuration expands to fully surrounded for any next digit 0 to 4. So if the digits of f ever reach fully surrounded then it remains so always. If f never reaches fully surrounded then that is an absent sub-curve at an arbitrarily small distance, and hence f is boundary.

$$fBpred(f) = \begin{cases} 0 & \text{if ever reach fully surrounded} \\ 1 & \text{if never fully surrounded} \end{cases}$$

To distinguish right and left boundary, segments of the curve always turn left or right and so divide the plane into alternating left or right side squares (eg. as previously for area in figure 13). The actual sub-curves are curling spiralling shapes, but they divide into logical squares.



possible boundary squares,
left and right sides

Squares are shown with just the segments of figure 32. If a square has at least 1 of the sides segments shown, but not all of them, then this is some of its R or L as boundary for m.

A configuration with no R squares expands to no R again for any next digit 0–4. Similarly L.

$$fRpred(f) = \begin{cases} 0 & \text{if ever reach no } 1,2 \text{ side R triangles} \\ 1 & \text{if always a } 1,2 \text{ side R triangle} \end{cases}$$

$$fLpred(f) = \begin{cases} 0 & \text{if ever reach no } 1,2 \text{ side L triangles} \\ 1 & \text{if always a } 1,2 \text{ side L triangle} \end{cases}$$

Total 18 configurations arise. There are 5 with R fully enclosed and 5 with L fully enclosed. 1 configuration is common to these, being the full set of segments.

Some usual state machine comparison shows R side is the same as the finite iterations *Rpred*. Likewise L the same as *Lpred*, and their union *Bpred*. ☐

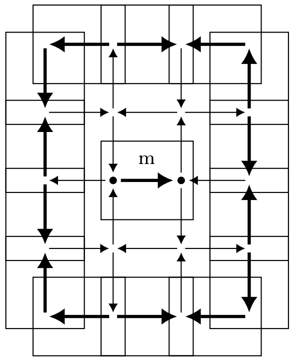
The finite *Bpred* considers only the 6 segments above and below m, not the 2 horizontals before and after m in figure 32. A full set of 6 is *Bpred*, which means start and end of m have 3 segments. But that means the 2 horizontals are also present, since the curve must arrive and leave twice (curve start and end are not double-visited). So the presence or absence of each horizontal in figure 32 is determined by the other segments.

The *Rpred* state machine considers only the 3 segments below m. If the 3 segments below are all present then the further R squares top left and top right can only have 0 or 2 segments, which mean not boundary, so that those squares add nothing. Similarly left L squares.

The hulls state machine does use such considerations of 3 segments at a point, but rather this arises from the way the expansions make new configurations.

The state machine also does not use theorem 50 no points on both left and right boundary. That follows mechanically from the state machine by taking the intersection of *Rpred* and *Lpred*. State machine manipulations show the only arbitrarily long strings matched are $f{=}0$ digits .000… and $f{=}1$ digits .444….

The surrounding segments in figure 33 can also be chosen further out, by some bigger upper bound on the convex hull. For example just $\mathit{Hwpf}, \mathit{Hepf}$ rectangles,



surrounding boxes
not touching
or overlapping m

The result of a bigger set of surrounding segments is more configurations. Some may be "eventually enclosed" in the sense that more digits from f, no matter what value, will eventually reach enclosed. Such configurations can be treated as enclosed since f always has further digits (trailing 0s or 4s if an otherwise terminating exact fraction $/5^k$). In all cases the f bit patterns matched by the resulting state machine are the same.

The number of f which are *fNonRpred* is uncountably infinite, since once reaching "non", further base-5 digits of f can be an arbitrary real.

The number of f which are boundary *fRpred* is uncountably infinite too. That can be seen in the *Rpred* state machine (figure 11) where there are various different ways digits of f can loop among R,X,Z so as to always stay away from "non". (But not Y since once there only 1s avoid non.) For example digits 20 is R$\rightarrow$X returning to R, and 210 is R$\rightarrow$X$\rightarrow$X returning to R. The bits of an arbitrary real can be coded to base-5 digits as $0 \rightarrow 20$, $1 \rightarrow 210$ so there are at least as many *fRpred* as reals. The same argument holds for *fLpred*, and then union for *fBpred*.

Theorem 53. *A location in the R5 dragon fractal can be visited 1 to 4 times. The number of visits to the location of a given f is*

$$fVisits(f) = \begin{cases} Visits_k(n) & \text{if } f = n/5^k \text{ for integer } n,k \\ Rsides(n) & \text{if } f = (n+\frac{1}{4})/5^k \\ Lsides_k(n) & \text{if } f = (n+\frac{3}{4})/5^k \\ & \text{and otherwise} \\ \quad 2 & \text{if fNonBpred but sub-curve fBpred} \\ \quad 1 & \text{otherwise} \end{cases}$$

The $fVisits = 2$ case is where f is not on the whole curve boundary, but its digits at some digit position and below are *fBpred* so as to be on the boundary of the sub-curve there.

Proof. An exact fraction $f = n/5^k$ is a vertex of curve k and the visits there are the same as $Visits_k$ from (58). By plane filling, those visits enclose the point so no other sub-curves visit.

An exact fraction $f = (n+\frac{3}{4})/5^k$ is repeating base-5 digits 3 and so always a side of the join end square per theorem 18. The number of sub-curve sides around that square is $Lsides_k(n)$ and each of them has a corresponding $\frac{3}{4}$ visit.

Similarly $f = (n+\frac{1}{4})/5^k$ is repeating base-5 digits 1 and is the middle square on the right and so $Rsides(n)$.

The claimed cases for *fBpred* whole curve boundary or not, and sub-curve eventually or never *fBpred* are

	whole curve	
	fBpred	*fNonBpred*
sub-curve eventually *fBpred*	1	2
sub-curve never *fBpred*	no such	1

An f which is on the boundary of some sub-curve, meaning its digits at some digit position and below are *fBpred*, might have an adjacent sub-curve like

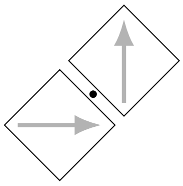


Figure 34:
f on sub-curve boundary
and adjacent other sub-curve

If it has this further sub-curve then by plane filling and no cut points the two enclose the location so visits are only those arising from the two.

If no such further sub-curve then the visits are only those arising from the f sub-curve itself. An f which is on a sub-curve boundary like this has only 1 visit because any other would be, for suitable yet smaller sub-curves, an adjacent enclosing further sub-curve like figure 34 and so not on the boundary. So in the table the first row cases are sub-curve boundary as 1 or 2 visits according to whole curve boundary or not.

An f which is *fNonBpred* non-boundary, and its digits at all positions below are also *fNonBpred*, is never on the boundary of any sub-curve and so always a non-zero distance away from any other sub-curve and so just 1 visit. $\square$

fNonBpred of a non-$/5^k$ means somewhere one of the forbidden *Rpred* digit pairs (theorem 13) and also somewhere one of the *Lpred* digit pairs (theorem 14). Pair 22 is common to these so 22 anywhere is *fNonBpred*.

The $fVisits = 2$ case is at least 1 each of the *Rpred* and *Lpred* disallowed, so as to be non-boundary, but only finitely many of one of them so eventually on a sub-curve boundary.

The $fVisits = 1$ case is the converse. Either none of one of *Rpred* or *Lpred* so as to be *fBpred* whole curve boundary, or infinitely many of both of them so never *fBpred* in a sub-curve.

The latter case, infinitely many of both, can be either rational or irrational. Requisite pairs in a repeating pattern is rational, or a non-repeating pattern is irrational. The simplest rational is $f = .222\ldots = \frac{1}{2}$ which is in the middle of the curve, then middle of the middle sub-curve, and so on, always a non-zero distance away from any other f digits.

It can be noted *fVisits* is not decided by initial digits of f. Further digits $222\ldots$ is $fVisits = 1$. An exact $/5^k$ point can be a double $Visits = 2$. And an exact $+\frac{1}{4}$ can be an $Rsides = 4$ visits.

An $fVisits = 3$ can be excluded by initial digits of f, since it is only a $+\frac{1}{4}$ or $+\frac{3}{4}$ into a 3-side boundary square. An f sub-curve must be a boundary segment to ensure such squares are possible.

Theorem 54. *For $fVisits(f) = 2$ by eventually sub-curve right boundary, its other visit $fOther(f)$ is all 1 digits unchanged and flip runs of non-1s by*

non-1 digits (106)

	high	*fRpred* disallowed pair		low
f	...	344...442	00...000	44...442	...
$fOther(f)$	...	400...000	44...442	00...000	...
		+4	−4	+4	

Runs are alternating 4442 *and* 0000. *fOther flips to the opposite form* $4442 \leftrightarrow 0000$. *Each run is* ≥ 1 *digit. Runs begin with the low digit of the lowest fRpred disallowed pair. The high of the pair is incremented when above run* 4442 *or decremented when above* 0000.

For $fVisits(f) = 2$ by eventually sub-curve left boundary, the same but digit patterns reversed $0 \leftrightarrow 4$ and starting from the lowest fLpred disallowed pair

non-3 digits (107)

	high	*fLpred* disallowed pair		low
f	...	100...002	44...444	00...002	...
$fOther(f)$	...	044...444	00...002	44...444	...
		−4	+4	−4	

Proof. For the right side, the 3 sub-curves on the right are calculated high to low as per table (22). There are 3 segments on the right. A 1 digit of n is the same s, t, e and 1 for all of them, so 1 digits unchanged and ignored for the purpose of the segments.

For other digits, only one of the 3 segments is used. n digit 0 uses only s and digits $2, 3, 4$ use only e. So at a sub-curve f the next lower digit of f determines which sub-curve s or e is wanted. So pairs of non-1 digits of f get the s or e digit as output at the high position of the pair,

R, non-1 digits of f (108)

f pair	00	02	03	04	20	22	23	24	30	32	33	34	40	42	43	44
output	4	2	2	2	0	$f3$	$f3$	$f3$	$f2$	$f4$	$f4$	$f4$	$f3$	0	0	0

When table (22) has an "n" it is a copy of f for the output. This is shown as output f at (108) here. It occurs for the *fRpred* disallowed pairs so that *fOther* is unchanged above such a pair.

At the lowest *fRpred* disallowed pair, following the pairs there onwards in f and the output digits in (108) gives the run forms (106).

For the left side similarly, with the pairs being

L, non-3 digits of f

f pair	00	01	02	04	10	11	12	14	20	21	22	24	40	41	42	44
output	4	4	4	$f1$	$f0$	$f0$	$f0$	$f2$	$f1$	$f1$	$f1$	4	2	2	2	0

□

The left side is $0 \leftrightarrow 4$ digit reversals of the right patterns and outputs. This is since the curve is the same in 180° reverse, so that $1-f$ measures back from the end and then $1-fOther(1-f)$ measures again from the start. $1-f$ is a $0\leftrightarrow 4$ reversal.

Theorem 55. *Differences $|f - fOther(f)|$ which occur for f eventually sub-curve boundary or exact $f = n/5^k$ are*

$$\big|f - fOther(f)\big| = \frac{4}{5^{k_0}} \pm \frac{4}{5^{k_1}} \pm \frac{4}{5^{k_2}} \pm \cdots \tag{109}$$
$$\text{where } 0 > k_0 > k_1 > k_2 > \cdots \text{ distinct powers}$$

Proof. Differences $|f - fOther(f)|$ which occur for runs (106),(107) are marked there as ± 4.

4442 to 0000 with borrow -1 from below is difference $+4$ and a carry $+1$ above. Conversely 0000 to 4442 with carry $+1$ from below is difference -4 and borrow -1 above. The runs are infinite so there is always a carry or borrow from below. At the highest run the carry or borrow going above suitably changes the high digit of the disallowed pair.

If there is a 1 digit within 4442 then the carry into that digit would be difference $+1$. But instead take it as -4 and continue the carry above. Conversely if a 1 digit within 0000 then borrow there would be difference -1 but instead take $+4$ and continue borrow -1 above.

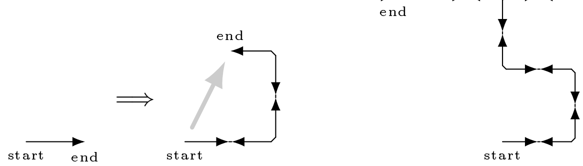
For exact $f = n/5^k$, theorem 6 scaled $/5^k$ has the same sum of ± 4 powers, in similar manner, but finite terms. ∎

All differences (109) occur by choosing suitable run lengths in f. The fractional base-5 digits are the $0, 4$ and alternating $3, 1$ of figure 7.

11 Half Curve

Jeffrey Ventrella[12] gives an R5 variant which is half the R5 curve, ie. that part of the curve through to the midpoint $\frac{1}{2}b^k$ (half the middle curve segment).

Figure 35:
5 segments
for half curve



In the expansion, the replacement directions are by rotating the base figure. Each pair of segments directed towards each other is a single segment of the R5 curve. This correspondence is maintained by the expansion of such a pair, and the expansion of the final unpaired segment maintains the final unpaired segment.

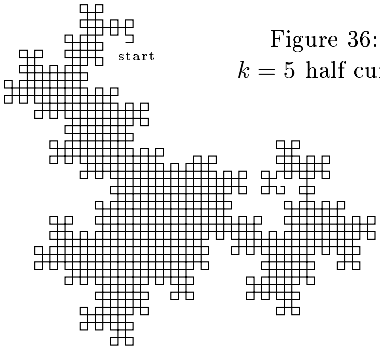


Figure 36:
$k = 5$ half curve

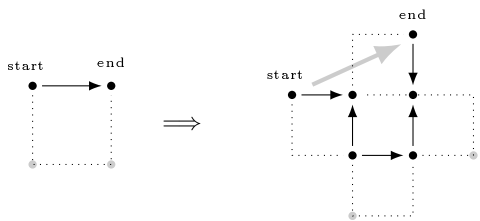
The half curve here is taken with the start as the full curve start. It could also be taken the other way around so the full curve middle is the half curve start. In figure 35 the expansion has the last segment forward, the same as the original, so on repeated expansion each level has the preceding as a prefix. In this form two half curve arms directed at 180° fill the plane.

Per the expansion in figure 35, going middle as start has the end (the full curve start) at $-1-2i$ which is $-116.56°$ around. Successive expansions are multiples of that angle. In figure 36 it can be seen starting from the middle the curve spirals clockwise, and faster than the usual R5 spiral so that it makes about $1+\frac{5}{8}$ rotations relative to the initial half segment.

12 Quartet

12.1 Quartet Curve

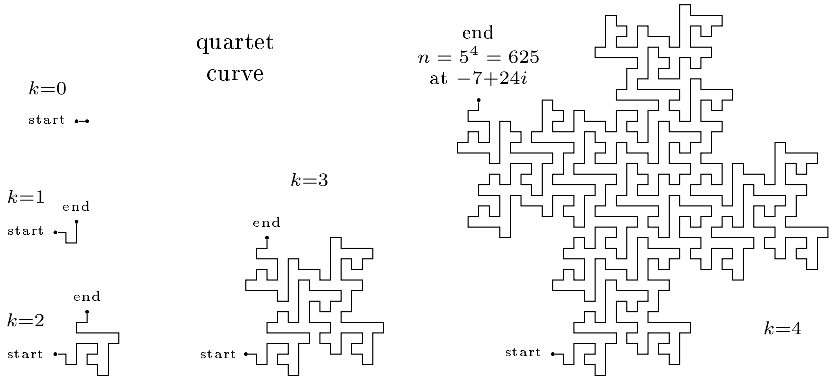
Mandelbrot [9] defines a "quartet curve" by an expansion going along sides of unit squares of a complex base $2+i$.



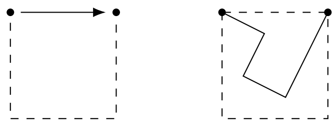
quartet curve
segment expansion

The arrows in the expansion are orientations for subsequent expansion by reversing the base figure. A segment represents the square on its right. The second segment in the expansion is reverse so its right is the middle square. The fourth segment likewise reverse for the top square. Mandelbrot conceives these squares as four "players" around a middle "table", hence the name quartet though there are 5 total parts and calculations are conveniently made in base-5.

Mandelbrot doesn't specify start and end. The start is chosen here so that successive expansion levels extend the previous. With the initial segment horizontal the curve end goes as powers of base $2+i$.

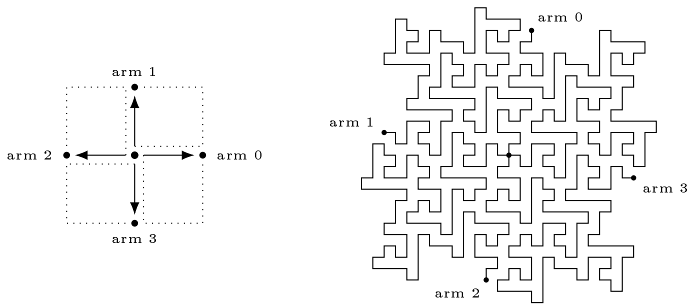


The segment expansion is entirely within its original square, touching only at the original start and end.



So starting from any set of segments with distinct squares on the right (in direction of expansion) gives a non-touching curve. The expanded segments are distinct squares too so repeated such expansions are non-touching.

A single initial segment has this form. Four arms at 90° likewise. Four arms are plane filling since each $2{\times}2$ expands to at least 1 unit square bigger.



Segments in opposite directions represent the squares on either side. Such segments start overlapping but after the first expansion they are a closed non-touching curve.

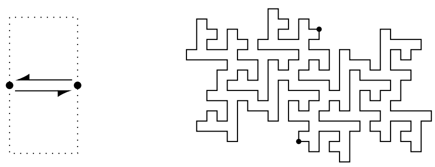


Figure 37:
$k = 3$
quartet curves
back-to-back

Theorem 56. *Number segments of the quartet curve starting $n{=}0$ for the first. Those which are "reverse" for the purpose of sub-curve expansion (or for having their unit square on the left instead of right) are characterized by*

$$QuartetRev(n) = \begin{cases} 0 & \textit{if } LowestNon2(n) = 0 \text{ or } 3 \\ 1 & \textit{if } LowestNon2(n) = 1 \text{ or } 4 \end{cases}$$
$$= 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, \ldots$$

$$LowestNon2(n) = \textit{base-5 lowest non-2 digit of } n$$
$$\textit{with a high } 0 \textit{ understood above } n \textit{ if necessary}$$
$$= 0, 1, 0, 3, 4, 0, 1, 1, 3, 4, 0, 1, 0, 3, 4, \ldots$$

$$gQuartetRev(x) = \sum_{k=0}^{\infty} \frac{x^{(3.5^k-1)/2}\,(1 + x^{3.5^k})}{1 + x^{5^k} + x^{2.5^k} + x^{3.5^k} + x^{4.5^k}} \tag{110}$$

$$gLowestNon2(x) = \sum_{k=0}^{\infty} \frac{x^{(3.5^k-1)/2}\,(1 + x^{5^k})\,(1 - x^{5^k} + 4x^{2.5^k})}{1 - x^{5^{k+1}}} \tag{111}$$

Proof. The base figure has sub-parts 0,2,3 forward and parts 1,4 reverse.

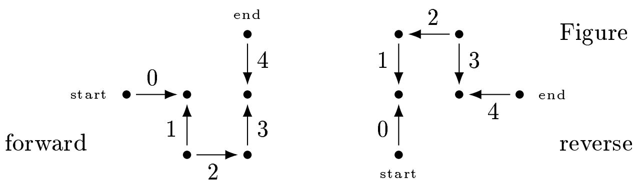


Figure 38

In reverse, parts $1, 4$ are again reverse and $0, 3$ forward, but part 2 remains reverse. So going high to low, digits 0,3 are always forward, 1,4 are always reverse, and 2 is unchanged from above. Hence *LowestNon2*.

Each term k in *gQuartetRev* sum (110) is those n with k many low base-5 digit 2s, which is $(5^k-1)/2$, and coefficient 1 at digit or 1 or 4 above them.

$$gQuartetRev(x) = \sum_{k=0}^{\infty} \frac{x^{(5^k-1)/2}\,(x^{5^k} + x^{4.5^k})}{1 - x^{5^{k+1}}}$$

In (110), exponent $(3.5^k-1)/2$ is $n = 122\ldots22$ with is the digit 1 above, and then add 3 more for the 4 above.

gLowestNon2 is similarly k many low base-5 digit 2s, then above those a non-2 digit $0, 1, 3, 4$. The numerator factorizes for (111).

$$gLowestNon2(x) = \sum_{k=0}^{\infty} \frac{x^{(5^k-1)/2}\,(x^{5^k} + 3x^{3.5^k} + 4x^{4.5^k})}{1 - x^{5^{k+1}}} \qquad \square$$

Theorem 57. *The number of reverse segments in quartet curve level k is*

$$QuartetRevA_k = \sum_{n=0}^{5^k-1} QuartetRev(n) \;= \tfrac{1}{2}(5^k - 1)$$

Proof. Curve k comprises 3 sub-curves forward, and 2 sub-curves reverse. The latter have $5^{k-1} - QuartetRevA_k$ reverse segments each,

$$QuartetRevA_k = 3\,QuartetRevA_{k-1} + 2\,(5^{k-1} - QuartetRevA_k)$$

$$= QuartetRevA_{k-1} + 2.5^{k-1} \tag{112}$$

(112) also follows from a new low base-5 digit on each n. A low 2 is no change so $QuartetRevA_{k-1}$, and low $1, 4$ are always reverse. $\square$

The area inside the back-to-back two curves in figure 37 is unit squares on the left of each component curve. These are the squares associated to the reverse segments. The curves mesh perfectly so the area inside is $2.QuartetRevA_k = 5^k - 1$.

Theorem 58. *Number points of the quartet curve starting $n=0$ at the origin so the first turn is at $n=1$. The turn at point n is determined by base-5 digits*

$$n = \begin{array}{l} {\scriptstyle high} \qquad \neq 2 \qquad\qquad \neq 0 \qquad {\scriptstyle low} \\ \boxed{\;\ldots\;} \boxed{r} \boxed{22\ldots22} \boxed{t} \underbrace{\boxed{00\ldots00}}_{l \text{ zeros}} \end{array} \qquad \text{base-5,} \;\; n \geq 1$$

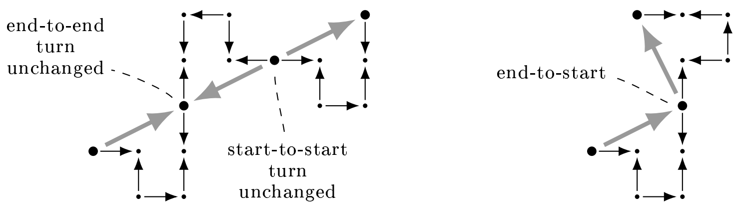
$$QuartetTurn(n) = \begin{array}{c} t = \\ \\ \\ \\ \\ \end{array} \begin{array}{rrrrl} 1 & 2 & 3 & 4 & \\ \hline -1 & +1 & +1 & 0 & \quad r = 0 \text{ or } 3 \text{ and } l=0 \\ -1 & +1 & 0 & 0 & \quad r = 0 \text{ or } 3 \text{ and } l \geq 1 \\ 0 & -1 & -1 & +1 & \quad r = 1 \text{ or } 4 \text{ and } l=0 \\ 0 & 0 & -1 & +1 & \quad r = 1 \text{ or } 4 \text{ and } l \geq 1 \end{array} \tag{113}$$

$$= -1, +1, +1, 0, -1, 0, -1, -1, +1, +1, -1, +1, +1, 0, \ldots$$

l is the number of low 0 digits and t is the lowest non-0 digit. Above there, r is the *LowestNon2*, with a high 0 understood for r if otherwise entirely 2s.

Proof. When l=0 so no low zero digits, t is the turns in the base figure, but reversed when r is $0, 3$, per $QuartetRev(\lfloor n/5 \rfloor)$. Reversal means digit t taken reversed 4 to 1, and the turn negated. Hence the two l=0 rows of (113).

When $l{\neq}0$, the further expansion replaces the directed segment with the base figure forward or reverse.



The diagram at the left is sub-curves meeting end-to-end or start-to-start. The turn in both cases is unchanged by expansion and both become start-to-start so are unchanged by further expansions too.

In the base figure, the only end-to-start is at segments 2 to 3 which is a left turn. The diagram at the right shows that changes to straight, so table (113) has t=3 left +1 when l=0 but straight 0 when $l \geq 1$, and corresponding reversal when $r = 1 \text{ or } 4$. The new straight is start-to-start so is unchanged by further expansions. $\square$

A little arithmetic can be used for the reversal r if desired. This emphasises the reversal both negating the turn of t and reversing the order.

$$
\mathit{QuartetTurn}(n) = \mathit{rev}\,.\begin{cases} -1, +1, +1, 0 & \text{for } t' = 1 \text{ to } 4 \text{ if } l = 0 \\ -1, +1, \ 0, 0 & \text{for } t' = 1 \text{ to } 4 \text{ if } l \ge 1 \end{cases}
$$

$$
\mathit{rev} = \begin{cases} +1 & \text{if } r = 0, 3 \\ -1 & \text{if } r = 1, 4 \end{cases} \qquad t' = \begin{cases} t & \text{if } \mathit{rev} > 0 \\ 5{-}t & \text{if } \mathit{rev} < 0 \end{cases}
$$

A generating function for *QuartetTurn* can be formed by a sum over l many low 0s and k many 2s above t. Coefficients are $+1$ or -1 as appropriate for digit combination r and t. The extra term outside the l sum is the extras for l=0.

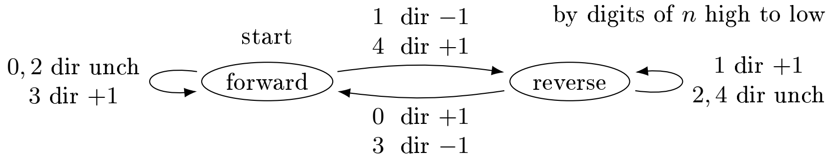
$$
\begin{aligned}
&g\mathit{QuartetTurn}(x) \\
&= \sum_{k=0}^{\infty} \left(\begin{aligned} &\frac{x^{5.(5^k-1)/2}}{1 - x^{5^{k+2}}} \left(x^3 \left(1 + x^{3.5^{k+1}}\right) - x^2 \left(x^{5^{k+1}} + x^{4.5^{k+1}}\right)\right) \\ &+ \sum_{l=0}^{\infty} \frac{x^{5^{l+1}.(5^k-1)/2}}{1 - x^{5^{k+l+2}}} \left(\begin{aligned} &(-x^{5^l} + x^{2.5^l})\,(\ 1 + x^{3.5^{k+l+1}}) \\ + &(-x^{3.5^l} + x^{4.5^l})\,(x^{5^{k+l+1}} + x^{4.5^{k+l+1}}) \end{aligned} \right) \end{aligned} \right) \\
&= \sum_{k=0}^{\infty} \left(\begin{aligned} &\frac{x^{3+5.(5^k-1)/2}\,(1 + x^{3.5^{k+1}})\,(1 - x^{5^{k+1}-1})}{1 - x^{5^{k+2}}} \\ &- \sum_{l=0}^{\infty} \frac{x^{5^l.(1+5.(5^k-1)/2)}\,(1 + x^{3.5^{k+l+1}})\,(1 + x^{5^l.(2+5^{k+1})})\,(1 - x^{5^l})}{1 - x^{5^{k+l+2}}} \end{aligned} \right)
\end{aligned}
$$

Number segments starting n=0 for the first. The net direction of segment n is sum of the turns preceding it, with the empty sum as $\mathit{QuartetDir}(0) = 0$,

$$
\begin{aligned}
\mathit{QuartetDir}(n) &= \sum_{j=0}^{n} \mathit{QuartetTurn}(j) \qquad n \ge 0 \\
&= 0, -1, 0, 1, 1, 0, 0, -1, -2, -1, 0, -1, 0, 1, 1, \ldots
\end{aligned}
$$

Theorem 59. *QuartetDir can be calculated from the base-5 digits of n,*

Figure 39: *QuartetDir* by digits of n high to low



Proof. Forward or reverse state is per *QuartetRev*. In forward state, the amount added to *QuartetDir* for a new digit is the segment direction $0, -1, 0, 1, 1$ in the base figure along the direction of the curve.

In reverse state, the amount is the reversed base figure as in figure 38. These $1, 1, 0, -1, 0$ are the same as forward, but read last to first. □

A generating function for *QuartetDir* is cumulative *QuartetTurn* using a factor $1/(1-x)$ in the usual way. The direct interpretation would be terms putting $+1, 0, -1$ for the digit t at position l and forward or reverse according to r.

$$gQuartetDir(x) = \frac{1}{1-x}\, gQuartetTurn(x)$$

The minimum *QuartetDir* in curve level k is by taking the -1 case each time in figure 39. This is alternating base-5 digits $1, 3$.

$$\begin{aligned} \min_{n=0}^{5^k-1} QuartetDir(n) &= -k, \quad \text{at } n = \tfrac{1}{3}\big(5^k - [5,1]\big) \\ &\qquad\qquad = 1313\ldots \text{ base-5} && \text{A037582} \\ &\qquad\qquad = 0, 1, 8, 41, 208, 1041, \ldots && \text{A037577} \end{aligned}$$

The maximum *QuartetDir* in curve level k is by taking the $+1$ case each time in figure 39. There are two digits with $+1$ in each state. In forward, either $3, 4$ (and 4 goes to reverse). In reverse, either $0, 1$ (and 1 goes back to forward). So repeating digit runs of the form (114).

$$\max_{n=0}^{5^k-1} QuartetDir(n) = k$$

$$\text{at } n = \overset{\text{high}}{\underset{\geq 0 \text{ 3s}}{\boxed{33\ldots334}}}\ \underset{\geq 0 \text{ 1s}}{\boxed{11\ldots110}}\ \overset{\text{low}}{\boxed{\text{repeat}}} \quad \text{base-5} \tag{114}$$

$$\begin{aligned} &= 0 && k = 0 \\ &= 3, 4 && k = 1 \\ &= 18, 19, 20, 21 && k = 2 \\ &= 93, 94, 95, 96, 103, 104, 105, 106 && k = 3 \end{aligned}$$

The first n in direction k is base-5 all 3s staying in forward state, since a digit 4 to go to reverse is bigger than staying.

$$\begin{aligned} \text{first } n \text{ with } QuartetDir(n) = k &= \tfrac{3}{4}\,(5^k - 1) \\ &= 0, 3, 18, 93, 468, 2343, \ldots && \text{A125833} \\ &= 0, 3, 33, 333, 3333, 33333, \ldots \text{ base-5} && \text{A002277} \end{aligned}$$

12.2 Quartet Tree

Mandelbrot gives a variation on the quartet curve which instead forms a tree. The change is to the segment representing the middle square in the expansion. This expansion is still entirely within its original square etc, so non-overlapping etc.

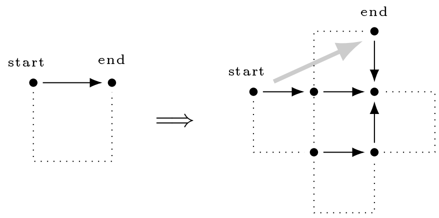
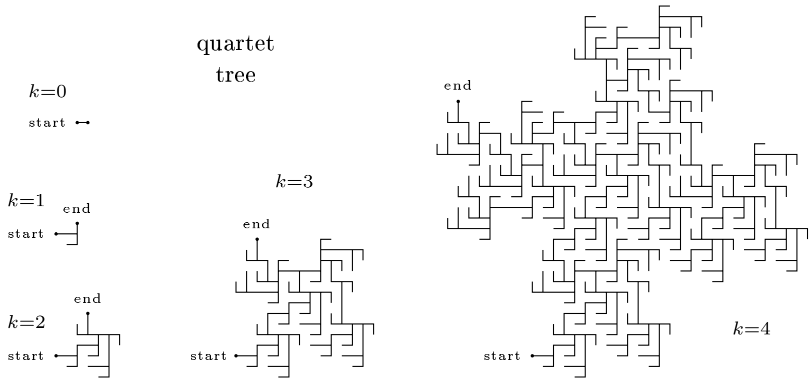


Figure 40: quartet tree expansion



Theorem 60. *The diameter of quartet tree k and eccentricities of its start and vertices are*

$$QuartetDiameter_k = \begin{cases} 1 & \textit{if } k{=}0 \\ 4.3^{k-1} & \textit{if } k \geq 1 \end{cases}$$

$$= 1, 4, 12, 36, 108, 324, 972, \ldots \qquad \text{A003946}$$

$$QuartetEccS_k = QuartetDiameter_k$$

$$QuartetEccE_k = 3^k$$

Proof. In k=0, start to end is a single edge distance 1. Expansion replaces each edge with 3 new edges, so a factor 3 each level and so distance 3^k start to end.

This is *QuartetEccE* too since at the middle fork in the base figure, the branch to the start is the same as the branch away from the start (including both directed towards the middle).

For *QuartetEccS*, the branch towards the end is the same as the branch away from it, except the latter has an extra 3^{k-1} start to end sub-tree. Thus

$$QuartetEccS_k = 3.3^{k-1} + QuartetEccE_{k-1} \;= 4.3^{k-1} \qquad k \geq 1$$

This is the diameter too since all sub-trees are directed inward toward the middle, so any path stopping in one of the non-terminal sub-trees would be lengthened by going to the same place in the outermost sub-tree. $QuartetEccS_k$ is greater than $QuartetDiameter_{k-1}$ which would be any path entirely within a sub-tree. □

For k=0, the two vertices are both centres (middles of the diameter) and both centroid vertices.

For $k \geq 1$, the diameter is even so unicentral and by symmetry the centre is the vertex which is the degree-3 meeting of $k-1$ sub-trees. This is also the tree centroid, since the number of vertices in each direction are $\frac{1}{5}, \frac{2}{5}, \frac{2}{5}$ of the total (excluding that centroid itself).

Theorem 61. *The Wiener index of quartet tree k is*

$$\begin{aligned} \mathit{QuartetW}_k &= 5^k \left(\tfrac{13}{49}15^k + \tfrac{7}{20}5^k + \tfrac{3}{35}k + \tfrac{377}{980}\right) \\ &= 1, 31, 1725, 117475, 8531625, \ldots \end{aligned}$$

Proof. Let WS_k be the total distances from the start to all other vertices, and similarly WE_k end to other vertices.

$$\begin{aligned} WS_k &= \sum\nolimits_v \mathit{distance}(\mathit{start}, v) \\ &= 1, 13, 179, 2599, 38549, \ldots \\ WE_k &= \sum\nolimits_v \mathit{distance}(\mathit{end}, v) \\ &= 1, 11, 147, 2117, 31317, \ldots \end{aligned}$$

Tree k as sub-trees $k{-}1$ gives mutual recurrences in respective directions, and distances 3^{k-1} to each part and 5^{k-1} vertices there (that count excluding the first vertex of each sub-part).

$$\begin{aligned} WS_k &= 2\,WS_{k-1} + 3\,WE_{k-1} \; + (1{+}2.2{+}3)\,3^{k-1}.5^{k-1} \\ WE_k &= \quad WS_{k-1} + 4\,WE_{k-1} \; + (2.1{+}2.2)\,3^{k-1}.5^{k-1} \end{aligned}$$

Some substituting or generating function manipulations gives

$$\begin{aligned} WS_k &= \tfrac{53}{70}15^k + \tfrac{7}{20}5^k - \tfrac{3}{28} \\ WE_k &= \tfrac{43}{70}15^k + \tfrac{7}{20}5^k + \tfrac{1}{28} \end{aligned}$$

Wiener index total distances between vertices are then 5 within the sub-trees and pair-wise between sub-trees. Paths between sub-trees are multiples of 3^{k-1} distances to the respective start or end, times 5^{k-1} destinations for each. There are just 5^{k-1} destinations, since paths to the common vertex are included in $\mathit{QuartetW}_{k-1}$ for adjacent parts, or the preceding pair-wise for those further apart.

$$\begin{aligned} \mathit{QuartetW}_k = 5\;\mathit{QuartetW}_k \; &+ 5^{k-1}\left(2\,WS_{k-1} + 18\,WE_{k-1}\right) \\ &+ 5^{k-1}.5^{k-1}\,(1{+}1{+}2{+}1{+}1)\,3^{k-1} \end{aligned} \qquad \square$$

The start and end vertices are not equivalent for $k \geq 1$ and they give different WS, WE. The difference between the two is

$$\begin{aligned} WS_k - WE_k &= \tfrac{1}{7}\left(15^k - 1\right) \\ &= 0, 2, 32, 482, 7232, \ldots \qquad \text{2×A135518} \end{aligned}$$

Second Proof of Theorem 61. Suppose a given edge has x many vertices on one side and y many on the other, so that for $\mathit{QuartetW}$ the number of crossings of that edge is product xy. Then total

$$\mathit{QuartetW}_k = \sum\nolimits_{\text{edges}} x\,y \tag{115}$$

Segment replacement means each edge becomes 3 edges across and a 2 edge branch. In an edge of tree $k{-}1$, take x as the number of vertices at the start and y at the end, in its direction of expansion.

$$\begin{array}{ccc}
\begin{tikzcd} x \arrow[r] & y \end{tikzcd} & \Longrightarrow & \begin{tikzcd} X \arrow[r] & A \arrow[r] & B & Y \arrow[l] \\ C \arrow[r] & D \arrow[ur] & & \end{tikzcd} \\
\text{tree } k{-}1 & & \text{tree } k
\end{array}$$

Figure 41:
segment expansion

Segment replacement means x many vertices preceding in $k{-}1$ become $X = 5x{-}4$ in k. Similarly $Y = 5y{-}4$ following. So for example edge $X{-}A$ is crossed $(5x{-}4)(5y{-}4+4)$ times for its vertex count on each side. Total crossings then, with $N = 5^k{+}1 = X{+}Y$ vertices in k and $n = 5^{k-1}{+}1 = x{+}y$ in $k{-}1$,

$$\begin{aligned}
&(5x{-}4).5y + (5x{-}3)(5y{-}1) + (5x)(5y{-}4) + 1.(N{-}1) + 2.(N{-}2) \\
&\quad = 75xy - 25x - 35y + 3N - 2 \\
&\quad = 75xy + 10x - 35n + 3N - 2
\end{aligned} \tag{116}$$

To sum over all edges, the product term xy in (115) is $\mathit{QuartetW}$, but the separate x requires a further calculation. Let WX_k be its sum over edges,

$$\mathit{WX}_k = \sum\nolimits_{\text{edges}} x$$

For an edge with x preceding vertices, the replacement segment directions inward in figure 41 give for the new 5 edges and their directions

$$(5x{-}4) + (5x{-}3) + \big(N{-}(5x{-}4)\big) + 1 + 2 \;=\; 5x + N - 4$$

so sum over all such

$$\begin{aligned}
\mathit{WX}_k &= 5\,\mathit{WX}_{k-1} + (5^k{+}1-4).5^{k-1} \qquad \text{starting } \mathit{WX}_0 = 1 \\
&= 5^k\left(\tfrac{1}{4}5^k - \tfrac{3}{5}k + \tfrac{3}{4}\right) \\
&= 1, 7, 145, 3775, 96625, \dots
\end{aligned}$$

So sum of (116) over all edges is

$$\begin{aligned}
\mathit{QuartetW}_k &= 75\,\mathit{QuartetW}_{k-1} + 10\,\mathit{WX}_{k-1} \\
&\qquad + \big({-35}(5^{k-1}{+}1) + 3(5^k{+}1) - 2\big).5^{k-1}
\end{aligned} \qquad \square$$

Similar to page 76, the mean distance between distinct pairs of vertices as a fraction of the diameter is, with limit from the high coefficients,

$$\frac{\mathit{QuartetW}_k}{\frac{1}{2}(5^k+1)\,5^k.\mathit{QuartetDiameter}_k} \to \frac{39}{98} = 0.397959\dots$$

For a rooted tree, the width of the tree at a given depth d is the number of vertices at that depth. Depth $d{=}0$ is the root itself (width 1). For the quartet tree take the root as the start. Vertices at depths can be illustrated by drawing as follows. The top horizontal is the spine continuing infinitely, with finite branches from it. Those branch depths variously overlap.

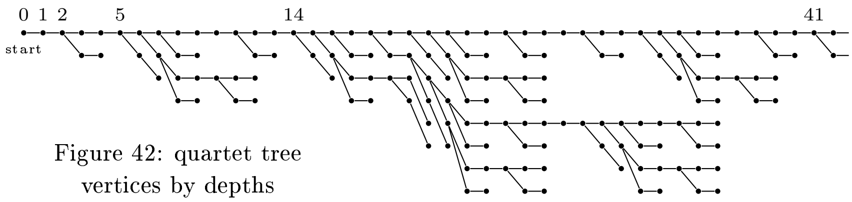


Figure 42: quartet tree vertices by depths

Widths follow by considering the expansion sub-parts from figure 40. Let $QuartetWidthS_k(d)$ be width of tree k at depth d down from the start vertex, and $QuartetWidthE_k(d)$ similarly but from the end vertex. Then mutual recurrences,

$$\begin{aligned}
QuartetWidthS_k(d) &= QuartetWidthS_{k-1}(d) + QuartetWidthS_{k-1}(d-3^{k-1}) \\
&\quad + 2\,QuartetWidthE_{k-1}(d-2.3^{k-1}) + QuartetWidthE_{k-1}(d-3.3^{k-1}) \\
&\quad - \begin{cases} 1 & \text{if } d=3^{k-1} \text{ or } 3.3^{k-1} \\ 2 & \text{if } d=2.3^{k-1} \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{117}$$

$$\begin{aligned}
QuartetWidthE_k(d) &= QuartetWidthS_{k-1}(d) \\
&\quad + 2\,QuartetWidthE_{k-1}(d-3^{k-1}) + 2\,QuartetWidthE_{k-1}(d-2.3^{k-1}) \\
&\quad - \begin{cases} 2 & \text{if } d=3^{k-1} \text{ or } 2.3^{k-1} \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{118}$$

$$\begin{aligned}
QuartetWidthS_0(d) &= 1,1 \\
QuartetWidthS_1(d) &= 1,1,1,2,1 \\
QuartetWidthS_2(d) &= 1,1,1,2,2,1,2,3,4,4,1,2,2
\end{aligned}$$

Widths of depths $d<0$ arising are understood as 0. The constants subtracted at (117),(118) are where sub-trees have a vertex in common. Those vertices are to be counted just once each.

Successive k adds further vertices below the end vertex at depth 3^k. So a depth $d \le 3^k$ has width unchanging with bigger k. This is the width of the tree continued infinitely,

$$\begin{aligned}
QuartetWidthS_\infty(d) &= QuartetWidthS_k(d) \ \text{ for any } k \text{ with } 3^k \ge d \\
&= 1,1,1,2,2,1,2,3,4,4,2,3,4,2,1,2,3,4,4,3,4,6,6,8,8,4,\dots
\end{aligned}$$

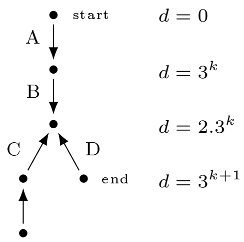
All widths are ≥ 1 since the tree continues infinitely. In figure 42 the labelled depths $0,1,2,5,\dots$ have width 1 vertex.

Theorem 62. *In the quartet tree extended infinitely, depths with just a single vertex $QuartetWidthS(\infty,d)=1$ are at d equal to*

$$\begin{aligned}
QuartetDOne(m) &= \begin{cases} 0 & \textit{if } m{=}0 \\ \frac{1}{2}\left(3^{m-1}+1\right) & \textit{if } m \ge 1 \end{cases} \\
&= 0,1,2,5,14,41,122,365,1094,3281,\dots && {\scriptstyle m\ge 1\ \text{A007051}} \\
&= \textit{ternary } 111\ldots112 \textit{ for } m{-}1 \ge 1 \textit{ many digits} && {\scriptstyle m\ge 1\ \text{A047855}}
\end{aligned}$$

Proof. The theorem can be verified explicitly for depths $d \leq 3$. Suppose the theorem is true of depths in the range $0 \leq d \leq 3^k$ for $k{\geq}1$ and consider then range $3^k < d \leq 3^{k+1}$.

k sub-trees



Depths in the range $2.3^k < d \leq 3^{k+1}$ have parallel identical sub-trees C,D so all widths there ≥ 2.

Depths in the range $3^k < d \leq 2.3^k$ are sub-tree B, but it is overlapped by some of A and those overlaps likewise have width ≥ 2. Sub-tree A extends down to $d = \mathit{QuartetEccS}_k = 4.3^{k-1}$. So in sub-tree B seek a depth $4.3^{k-1} < d \leq 2.3^k$ which is width 1 in that B sub-tree, so $d = \mathit{QuartetDOne}(m) + 3^k$ for some m.

The desired range is ternary $110\ldots00 < d \leq 200\ldots00$. The only suitable *QuartetDOne* in that range is $m{=}k{+}1$ giving

$$d = \mathit{QuartetDOne}(m{+}1) = \mathit{QuartetDOne}(m) + 3^k$$

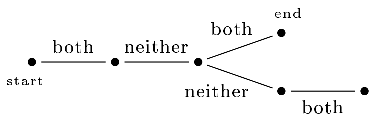
an extra high ternary 1 digit for $\mathit{QuartetDOne}(m{+}1)$. □

A perfect matching is a matching (as from page 82) which has all vertices in some pair.

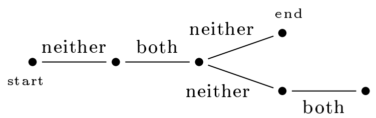
Theorem 63. *The quartet tree has a perfect matching.*

Proof. Suppose the theorem is true of some k, and further that if the start and end vertices are omitted then that reduced tree also has a perfect matching. This is so of the 2 vertices $k{=}0$.

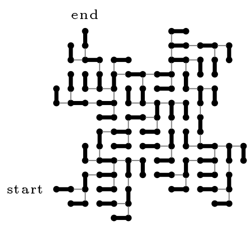
Tree $k{+}1$ then has a perfect matching by taking alternating both/neither sub-trees start to end and into the branch. And tree $k{+}1$ with start and end vertices omitted likewise but opposite start to end.



both start, end



neither start, end □



quartet tree $k = 3$
perfect matching

Any tree with a perfect matching has independence number half its vertices, by starting at any pair of vertices and working outwards taking successive vertices present or absent for an independent set.

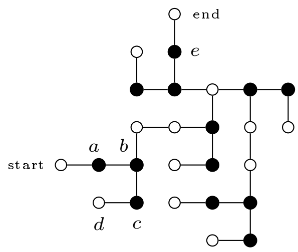
Theorem 64. *The total domination number of quartet tree k is*

$$\begin{aligned} \mathit{QuartetTotdomnum}_k &= \begin{cases} 2, 3 & \text{if } k = 0, 1 \\ 12.5^{k-2} + 1 & \text{if } k \geq 2 \end{cases} \\ &= 2, 3, 13, 61, 301, 1501, \ldots \end{aligned}$$

The number of total dominating sets of this size is

$$\begin{aligned} \mathit{QuartetTotdomnumCount}_k &= \begin{cases} 1 & \text{if } k \leq 2 \\ 12^{\frac{1}{4}(5^{k-2}-1)} & \text{if } k \geq 2 \end{cases} \\ &= 1, 1, 1, 12, 2985984, \ldots \end{aligned}$$

Proof. The theorem can be verified explicitly for $k \leq 2$. The single total dominating set for k=2 is (black vertices in the set),



$\mathit{QuartetTotdomnum}_2 = 13$
unique set

If the start vertex is allowed to be undominated then there is still a single smallest set, being the full set less attachment vertex a. Similarly if the end vertex is allowed to be undominated then solely e omitted. If both start and end then solely both a and e omitted.

Suppose the theorem is true of some $k{-}1 \geq 2$ and further that, like k=2, if the start and/or end are allowed to be undominated then the total domination number reduces by 1 each and the sets are different but count of sets unchanged.

Tree k comprises five $k{-}1$ sub-trees which have one vertex common to 3 sub-trees and two vertices each common to 2 sub-trees. The common vertex can be dominated by just one of those sub-trees each, which is start or end undominated in the others,

$$\mathit{QuartetTotdomnum}_k = 5\,\mathit{QuartetTotdomnum}_{k-1} - 4 \qquad k \geq 3$$

The number of such sets is by choosing which sub-tree dominates the 3-branch and two 2-branches. So 1 of 3 and two places 1 of 2 for 3.2.2 = 12 choices and product of counts of sets within the sub-trees

$$\mathit{QuartetTotdomnumCount}_k = 12\,\mathit{QuartetTotdomnumCount}_{k-1}^{5}$$

These sets are all with the common vertices absent. To see that any common vertex present would be bigger, if the start vertex of k=2 is present then it dominates a. This could allow b to be omitted, but if it is then d must be present to dominate c. So the start is added and number of others unchanged, giving a bigger set. Similarly at the end vertex. This increase with start or end likewise holds for $k > 2$. □

References

[1] Jörg Arndt, "Matters Computational: Ideas, Algorithms, Source Code", section 1.31.5 "Dragon Curves Based on Radix-R Counting".
http://www.jjj.de/fxt/fxtpage.html

[2] Jörg Arndt, Online Encyclopedia of Integer Sequences (ed. N. J. A. Sloane) entry A175337, April 2010.
http://oeis.org/A175337

[3] Samuel Beatty, Problem 3173, American Mathematical Monthly, volume 33, number 3, March 1926, page 159. Answers volume 34, number 3, March 1927, pages 159–160.

[4] Cockayne, Goodman, and Hedetniemi, "A Linear Algorithm for the Domination Number of a Tree", Information Processing Letters, volume 4, number 2, November 1975, pages 41–44.

[5] D. E. Daykin and S. J. Tucker, "Sequences from Folding Paper", unpublished manuscript, January 1975. Reproduced in Online Encyclopedia of Integer Sequences (ed. N. J. A. Sloane) entry A003229.
http://oeis.org/A003229

[6] F. M. Dekking, "Paperfolding Morphisms, Planefilling Curves, and Fractal Tiles", Theoretical Computer Science, volume 414, issue 1, January 2012, pages 20–37.
http://arxiv.org/abs/1011.5788 (preprint version 2, January 2011)

[7] Ivan Gutman, Boris Furtula and Miroslav Petrović, "Terminal Wiener Index", Journal of Mathematical Chemistry, volume 46, 2009, pages 522–531.
DOI 10.1007/s10910-008-9476-2

[8] Donald E. Knuth, "Selected Papers on Fun and Games" addendum to reprint of "Number Representations and Dragon Curves", 2010, pages 571–614.
http://www-cs-faculty.stanford.edu/~uno/fg.html

[9] Benoit B. Mandelbrot, "The Fractal Geometry of Nature", W. H. Freeman and Co., 1983, section 7, "Harnessing the Peano Monster Curves", pages 72–73.

[10] Rayleigh, "The Theory of Sound", volume 1, second edition, section 92a, pages 122–123.

[11] Tiling Search database entry AL04, and with enlarged square part L3010.
http://tilingsearch.org/HTML/data182/AL04.html
http://tilingsearch.org/HTML/data149/L3010.html

[12] Jeffrey Ventrella, "Brain-Filling Curves", 2012, section 4, "$\sqrt{5}$ Family", pages 84–90.
`http://www.fractalcurves.com`
`http://archive.org/BrainfillingCurves-AFractalBestiary/BrainFilling.pdf` (about 245mb)

Index

OEIS A-Numbers