

The Convergence of Difference Boxes

Antonio Behn, Christopher Kribs-Zaleta,
and Vadim Ponomarenko

1. INTRODUCTION. “Difference boxes,” also known as “diffy boxes,” are the basis for a simple mathematical puzzle that provides elementary-school students with subtraction practice. The idea’s original author is unknown, although among many Texas teachers it is traced back to Professor Juanita Copley of the University of Houston, who introduced it as a problem-solving activity in professional development sessions about twenty years ago (and who, in turn, cites her grandmother). Another source cites a World War II prisoner of war [4]. Diffy boxes have been used in numerous places as a mechanism for elementary teachers to provide arithmetic practice without the tedium of drill (their use as a problem-solving activity will be addressed in the final section of this paper).

One creates a difference box as follows:

1. Draw a (large) square, and label each vertex with a (real) number.
2. On the midpoint of each side write the (unsigned) difference between the two numbers at its endpoints.
3. Inscribe a new square in the old one, using these new numbers to label the vertices.
4. Repeat this process, and continue inscribing new boxes until reaching a square that has all four vertices labeled 0.

It is perhaps surprising that diffy boxes always tend to “converge” rather quickly, that is, it usually takes no more than a handful of iterations to get a box with all zeroes. Figure 1 shows a simple example, which converges to all zeroes after four iterations (on the fifth box).

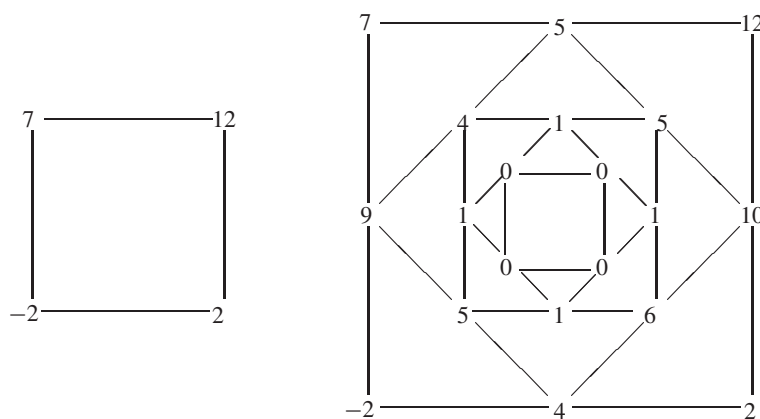


Figure 1. A simple diffy box (left) and its “descendants” (right).

The question we want to investigate is whether all diffy boxes really do converge to the zero box, and, if so, how quickly? It turns out that the answer is no if we require a

finite number of iterations; however, the essentially unique counterexample also converges to zero after a fashion. We approach the problem by considering the diffy box process as a map from the set of all possible four-tuples of real numbers (i.e., \mathbf{R}^4) into itself.

2. DEFINITIONS. We begin our analysis by establishing some notation. We denote by $[a\ b\ c\ d]$ the diffy box that contains the numbers a , b , c , and d on its upper left, upper right, lower right, and lower left corners, respectively (which makes the first box in Figure 1 [7 12 2 -2]).

Next we introduce notation to describe the diffy box iteration process, as well as some terms related to convergence:

Definition 1. The *child* of a given box $B = [a\ b\ c\ d]$ is $C(B) = [|b - a|\ |c - b|\ |d - c|\ |a - d|]$; B is a *parent* of $C(B)$. (We shall see that any given box has many parents.) The parent-child relation is signified by $B \triangleright C(B)$. We denote $C(C(B))$ by $C^2(B)$, etc. The box $C(B)$ is more commonly called the *iterate* of B .

Definition 2. A given box B *converges to zero in n generations* if $C^n(B) = [0\ 0\ 0\ 0]$ but $C^{n-1}(B) \neq [0\ 0\ 0\ 0]$; we also say in this case that B has *longevity n* .

Definition 3. The *range* of B , denoted $|B|$, is the largest difference between two (not necessarily adjacent) vertices of B .

Definition 4. A box B is *monotone* if its vertices are distinct and occur around the square in numerical order from least to greatest.

Note that we consider boxes such as [2 1 4 3] monotone, since the definition allows us to start with any vertex and proceed clockwise or counterclockwise from it. A couple of relatively quick results may help the reader begin to develop some intuition as to how and why diffy boxes tend to converge to zero.

Proposition 1. *The estimate $|C(B)| \leq |B|$ always holds, and the inequality is strict if the four numbers in B are distinct.*

Proof. Let $w \geq x \geq y \geq z$ be the numbers in B (not necessarily in order of appearance). We have $|B| = w - z$. The numbers in $C(B)$ all fall between 0 and $w - z$; hence $|C(B)| \leq (w - z) - 0 = |B|$. Furthermore, if w , x , y , and z are distinct, then $C(B)$ does not contain zero, whence $|C(B)| < |B|$. ■

Proposition 2. *Any nonmonotone diffy box converges to zero in six or fewer generations.*

Proof. Table 1 details the longevity of all diffy boxes whose vertices are nonmonotone. The proof is an exhaustive case analysis (a simpler proof follows from Figure 3 in section 3). To put a given diffy box into a form listed in the table, reorder the vertices (using reflection and/or rotation, isometry properties that we shall discuss in section 4) so that the smallest vertex is listed first, followed by the smaller of the two vertices adjacent to it. (If two copies of the smallest number occupy adjacent vertices, put them first and second, followed by the smaller of the two vertices adjacent to them.) This reordering does not affect longevity. ■

Table 1. Longevities for all nonmonotone diffy boxes
(here $a < b < c < d$).

Longevity	Isometric form of diffy box
1	$[a a a a], a \neq 0$
2	$[a b a b]$ $[a b c b], b = (a + c)/2$
3	$[a a b b]$ $[a b b c], b = (a + c)/2$ $[a b d c], a - b = c - d$ $[a c b d], a - b = c - d$
4	$[a a a b]$ $[a a b c], b$ at least as close to c as to a $[a b a c]$ $[a b b b]$ $[a b b c], b \neq (a + c)/2$ $[a b c b], b \neq (a + c)/2$ $[a b c c], b$ at least as close to a as to c $[a b d c], a - b \neq c - d, (a + d)/2$ between b and c $[a c b c]$ $[a c b d], a - b \neq c - d$
6	$[a a b c], b$ closer to a than to c $[a b c c], b$ closer to c than to a $[a b d c], b$ and c on the same side of $(a + d)/2$

We remark that Proposition 2 includes any box whose vertex numbers are not all distinct. We will observe in section 4 that diffy boxes $[a b c d]$ whose vertices are monotone (with $a < b < c < d$ or $a > b > c > d$) have longevity five or greater.

Since, in view of Proposition 1, $|C(B)| = |B|$ only when the vertices of B are not all distinct, we can bound the longevity of any box that has *integer* vertices by observing that the range $|B|$ of any such box must decrease by at least 1 per iteration, until we reach a box C with at least one pair of identical vertices. At this point we compare in Table 1 the possible longevities of C (2, 3, 4, or 6) with the corresponding minimum possible range (1, 1, 1, or 3, respectively) and record the greatest difference. We therefore have the following result:

Corollary. *If B consists of integers, then the longevity of B is less than or equal to $|B| + 3$.*

We leave improvements on this specialized estimate as an exercise for the reader. For example, it can be sharpened from linear to logarithmic; namely, to $4(1 + \lceil \log_2 |B| \rceil)$.

3. CANONICAL FORM. A little experimentation with different sets of numbers quickly leads one to the observation that there are different boxes that behave the same way with respect to the iteration process. For example, adding the same number k to each of the vertices of a box B (written $B + k$) or changing the signs of all the vertices ($-B$) will produce other parents of $C(B)$, since each iteration records only differences between successive vertices; thus the map $B \mapsto C(B)$ is many-to-one. Furthermore, since our real interest is the history of families rather than of individuals, we observe three other types of changes that generate family histories parallel to the original. Multiplying each vertex of a box B by a positive constant k will produce

a box (kB) whose child is that same multiple of $C(B)$ (i.e., $kB \triangleright kC(B)$); B and kB therefore take the same number of iterations to reach the all-zero box. We may think of B and kB as “cousins” with the same family histories. Finally, rotating or reflecting the numbers on the vertices of a box B (call these $\text{rot}(B)$ and $\text{ref}(B)$) will create further cousins, boxes whose respective children are the rotated and reflected versions of $C(B)$ (i.e., $\text{rot}(B) \triangleright \text{rot}(C(B))$, $\text{ref}(B) \triangleright \text{ref}(C(B))$), and so on through successive iterations). These changes are merely cosmetic, since the numbers retain their positions relative to each other.

To simplify the analysis that follows, we define a set of equivalence classes that reduces the number of distinct boxes we must consider (and thereby reduce the dimension of the problem considerably, as we shall see).

Definition 5. The equivalence class of a box $B = [a \ b \ c \ d]$ consists of all boxes that can be obtained from B via a finite sequence of the following five elementary operations:

- (1) translation: $[a \ b \ c \ d] \mapsto [(a + \alpha) \ (b + \alpha) \ (c + \alpha) \ (d + \alpha)]$ ($\alpha \in \mathbf{R}$);
- (2) negation: $[a \ b \ c \ d] \mapsto [-a \ -b \ -c \ -d]$;
- (3) positive scaling: $[a \ b \ c \ d] \mapsto [(\alpha a) \ (\alpha b) \ (\alpha c) \ (\alpha d)]$ ($\alpha \in \mathbf{R}^+$);
- (4) rotation: $[a \ b \ c \ d] \mapsto [b \ c \ d \ a]$;
- (5) reflection: $[a \ b \ c \ d] \mapsto [d \ c \ b \ a]$.

We can consider the first three operations as field operations, and the last two as isometries. It is simple to verify that the relation \sim thus defined is indeed an equivalence relation on the set \mathcal{B} of all diffy boxes (i.e., it is reflexive, symmetric, and transitive). We also observe, following the same arguments given informally earlier, that if $B_1 \sim B_2$, then $C(B_1) \sim C(B_2)$, and B_1 and B_2 have the same longevity unless one of them is the zero box and the other is $[a \ a \ a \ a]$ ($a \neq 0$).

We would like to find a way to select a unique member from each equivalence class in \mathcal{B} so that we can concentrate our remaining analysis on a reduced domain. To this end, we define the canonical form for an equivalence class.

Definition 6. The *canonical form* for an equivalence class in \mathcal{B} is one of the following: (i) $[0 \ 0 \ 0 \ 0]$ for the class containing this (zero) box; (ii) $[0 \ 0 \ 1 \ 1]$ for the class containing this box; or (iii) the unique class member of the form $[0 \ 1 \ x \ y]$ for which (x, y) belongs to $S = \{(x, y) : x \geq 0, y \geq 1, x - 1 \leq y \leq x + 1\}$ otherwise.

Equivalence class (i) consists of all boxes $[a \ a \ a \ a]$ ($a \in \mathbf{R}$), each of which converges to zero in one iteration (if $a \neq 0$). Equivalence class (ii) consists of all boxes $[a \ a \ b \ b]$ and $[a \ b \ b \ a]$ ($a, b \in \mathbf{R}$), which are seen to converge to zero in three iterations ($[0 \ 0 \ 1 \ 1] \triangleright [0 \ 1 \ 0 \ 1] \triangleright [1 \ 1 \ 1 \ 1] \triangleright [0 \ 0 \ 0 \ 0]$). Identifying the canonical form for all other classes requires the notion of an *extreme element*, that is, an element a of a box that is either maximal (at least as big as each of the other elements) or minimal (at least as small as each of the other elements). Note that at least two of the four elements of each box must be extreme. The following result gives a procedure for determining a type (iii) canonical form, as well as a justification of its uniqueness.

Theorem 1. Any equivalence class in \mathcal{B} of type (iii) (i.e., not including a box of the form $[a \ a \ a \ a]$ or $[a \ a \ b \ b]$) has a unique representative $[0 \ 1 \ x \ y]$ with (x, y) in $S = \{(x, y) : x \geq 0, y \geq 1, x - 1 \leq y \leq x + 1\}$.

Proof. We first provide an algorithm for finding the representative, using our elementary operations. We begin with an arbitrary diffy box $[a \ b \ c \ d]$ and operate on it as follows (relabeling the results $[a \ b \ c \ d]$).

1. Rotate (operation (4)) until a box is obtained for which $|d - a|$ is maximal among $|a - b|$, $|b - c|$, $|c - d|$, $|d - a|$.
2. (a) Observe that either a or d must be extreme. If necessary, reflect $[a \ b \ c \ d]$ (operation (5)) to ensure that a is extreme.
 - (b) If a and d are both extreme, it's possible that $|a - b| < |c - d|$. If necessary, reflect (operation (5)) to make $|a - b| \geq |c - d|$.
3. If a is maximal, use negation (operation (2)) to make a minimal.
4. Use translation (operation (1)) to make $a = 0$.
5. Use positive scaling (operation (3)) to make $b = 1$.

Observe that properties created at any step are preserved in subsequent steps. The last step is always possible since steps 2(b) and 4 together imply that $c = d$ when $b = 0$, and the only two such cases correspond to equivalence classes of type (i), $c = d = 0$, or (ii), $c = d \neq 0$. Otherwise $b > 0$ (from steps 3 and 4), so positive scaling can be used to normalize b .

We now have a class representative of the form $[0 \ 1 \ c \ d]$ (from steps 4 and 5). Since $a = 0$ is minimal, c and d are nonnegative. From steps 1, 4, and 5, we infer that $d \geq 1$. On the basis of 2(b) we have that $|c - d| \leq 1$, so $c - 1 \leq d \leq c + 1$. Thus (c, d) lies in S .

It remains to show that the canonical form (iii) is unique: if $[0 \ 1 \ a \ b] \sim [0 \ 1 \ c \ d]$, where (a, b) and (c, d) belong to S , then $(a, b) = (c, d)$. This can be seen first by observing that 0 is a unique minimal number (i.e., $a, b, c, d > 0$ except for $[0 \ 1 \ 0 \ 1]$, which has no other $[0 \ 1 \ c \ d]$ representation), and second on a case-by-case basis by taking a or b to be maximal and showing that each of the seven transformations that would transform 0, 1, a , or b to 0 and normalize either of the numbers adjacent to it results in a box with (c, d) not in S . As the details are simple but technical, we leave them as an exercise for the reader (we likewise leave as an exercise the proof that the equivalence classes containing $[0 \ 0 \ 0 \ 0]$ and $[0 \ 0 \ 1 \ 1]$ have no type (iii) canonical form representation). ■

4. A TWO-DIMENSIONAL MAP. We can now focus our remaining analysis on what happens to equivalence classes in \mathcal{B} of the form $[0 \ 1 \ x \ y]$ with (x, y) in S . We first observe that in the special case $y = 1$ the children are of type (i) or (ii) (since $[0 \ 1 \ x \ 1] \triangleright [1 \ |x - 1| \ |x - 1| \ 1]$), hence convergence to zero occurs in two ($x = 0, 2$) or four generations.

We now consider the diffy box process as a continuous map from $S \setminus \{(x, y) : y = 1\}$ into S that calculations show to be given by

$$(u, v) = f(x, y) = \begin{cases} \left(1 + \frac{x}{y-1}, \frac{x}{y-1}\right) & \text{if } 0 \leq x \leq 1, 1 < y \leq x+1 \ ((x, y) \in S_1); \\ \left(1 + \frac{2-x}{y-1}, \frac{x}{y-1}\right) & \text{if } x > 1, x < y \leq x+1 \ ((x, y) \in S_2); \\ \left(1 + \frac{2-x}{y-1}, 2 + \frac{2-x}{y-1}\right) & \text{if } x, y > 1, x-1 \leq y \leq x \ ((x, y) \in S_3). \end{cases}$$

Figure 2 illustrates the three sets S_1 , S_2 , and S_3 into which this definition decomposes S .

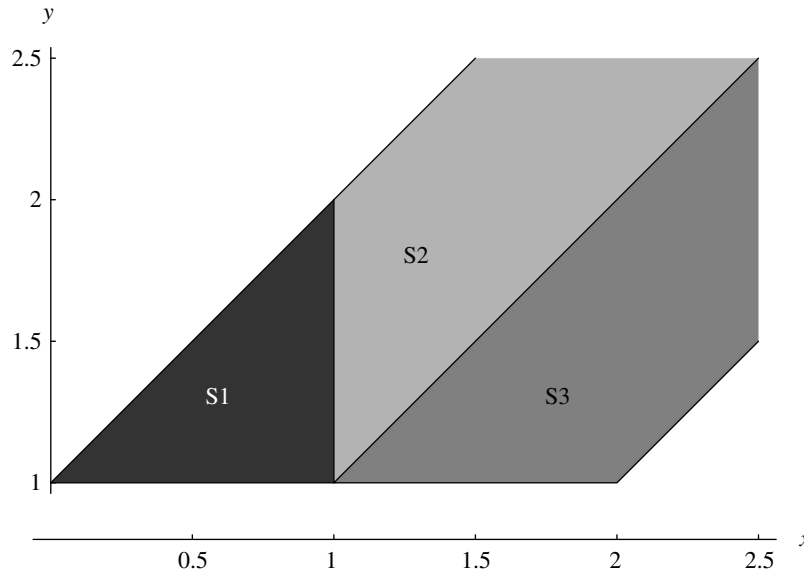


Figure 2. Subdivision of S into regions corresponding to the three branches of f .

By inspection we can see that the first and third branches of this function are infinite-to-one mappings: the first branch sends all points to the line $v = u - 1$ (with $u \geq 2$, since $y \leq x + 1$ implies that $x/(y - 1) \geq 1$), while the third branch sends all points to the line $v = u + 1$ (with $u \geq 0$, since $y \geq x - 1$ implies that $(2 - x)/(y - 1) \geq -1$). On the first branch (where $x \leq 1$), $f(x_1, y_1) = f(x_2, y_2)$ holds when

$$\frac{y_1 - 1}{x_1 - 0} = \frac{y_2 - 1}{x_2 - 0},$$

that is, when (x_1, y_1) and (x_2, y_2) lie on the same line segment from $(0,1)$. Likewise, on the third branch (where $y \leq x$), $f(x_1, y_1) = f(x_2, y_2)$ occurs when

$$\frac{y_1 - 1}{x_1 - 2} = \frac{y_2 - 1}{x_2 - 2},$$

that is, when (x_1, y_1) and (x_2, y_2) lie on the same line segment from $(2,1)$. On these branches, f compresses two-dimensional regions into rays on the boundary of S .

The second branch, however, is one-to-one: for (x_1, y_1) and (x_2, y_2) both in S_2 , $f(x_1, y_1) = f(x_2, y_2)$ implies that $(x_1, y_1) = (x_2, y_2)$ (we can see this by rewriting

$$\frac{x_1}{y_1 - 1} = \frac{x_2}{y_2 - 1} \quad \text{as} \quad y_2 - 1 = \left(\frac{y_1 - 1}{x_1}\right) x_2$$

and substituting this into

$$1 + \frac{2 - x_1}{y_1 - 1} = 1 + \frac{2 - x_2}{y_2 - 1}.$$

We can also see that the first and third branches have no fixed points inside their respective domains: the first branch sends all points to the line $v = u - 1$, which isn't in S_1 , and branch 3 sends all points to the line $v = u + 1$, which isn't in S_3 . (In fact, there is no possible periodicity in these regions, either, since for (x, y) in S_1 or S_3 , $(c, d) = f(f(x, y))$ has $d = 1$.) Therefore any fixed points must lie in S_2 . Indeed, straightforward calculations show that S_2 contains the unique fixed point $(q(q - 1), q) \approx (1.5437, 1.8393)$, where

$$q = \frac{1}{3} \left(1 + \sqrt[3]{19 + 3\sqrt{33}} + \sqrt[3]{19 - 3\sqrt{33}} \right)$$

is the unique real solution to $q^3 - q^2 - q - 1 = 0$. Here $C([0 \ 1 \ q(q - 1) \ q]) \sim [0 \ 1 \ q(q - 1) \ q]$ (i.e., the child is in the same equivalence class as the parent, and therefore takes just as long to converge—in other words, it has infinite longevity). The only other equivalence class for which this is true is $[0 \ 0 \ 0 \ 0]$ (but properly speaking this class has no longevity).

We should note, however, that the foregoing does *not* imply that $B = C(B)$ for members of the class containing $[0 \ 1 \ q(q - 1) \ q]$. In fact, in an \mathbf{R}^4 -norm sense, as well as that of Definition 3, members of this class do approach the zero box under iteration—they just take infinitely long to get there. For example, calculation of a few iterations of the diffy box process beginning with $[0 \ 1 \ q(q - 1) \ q]$ shows that after the first step the four entries gradually diminish in size (by a factor of $q - 1$) as they cycle around counterclockwise, in keeping with Proposition 1.

We now determine whether the fixed point $(q(q - 1), q)$ is stable, in the sense of other nearby points' images approaching it under repeated application of f . To determine the stability of a fixed point of a complex map f , one looks at its Jacobian matrix J_f . This is derived from the linearization of the map and consists of the map's partial derivatives, evaluated at a given fixed point. For $f = (f_1(x, y), f_2(x, y))$, the Jacobian is given by

$$J_f = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix}.$$

This (second) branch of f has

$$J_f(x, y) = \begin{bmatrix} -\frac{1}{y-1} & \frac{x-2}{(y-1)^2} \\ \frac{1}{y-1} & -\frac{x}{(y-1)^2} \end{bmatrix},$$

$$J_f(q(q-1), q) \approx \begin{bmatrix} -1.1915 & -0.6478 \\ 1.1915 & -2.1915 \end{bmatrix}.$$

The matrix on the right has eigenvalues $\lambda \approx -1.6915 \pm 0.7224i$. Because the eigenvalues have magnitude greater than 1, the fixed point is unstable. Because the imaginary components are nonzero, we see that points near $(q(q - 1), q)$ spiral away from it under (repeated) application of f . (To read more about stability analysis of fixed points, see [2] or [3].) Because this is the only fixed point, one might expect that further applications of f will eventually move any other point to the boundary of the

domain, and then out of it entirely. Therefore, we might expect that those diffy boxes that take longest to converge to zero correspond to those points in S closest to the fixed point $(q(q-1), q)$. As we shall see later, these intuitive notions turn out to be correct.

If we begin a case-by-case analysis of the successive applications of f in S , we notice the domain subdividing into regions beginning along the boundaries and working in toward the fixed point.

Example 1. Any box $[0 \ 1 \ x \ y]$ with $x \leq 1$ converges to zero within four generations (three if $y = x + 1$, two if $(x, y) = (0, 1)$). We calculate

$$\begin{aligned} [0 \ 1 \ x \ y] &\triangleright [1 \ (1-x) \ (y-x) \ y] \triangleright [x \ (y-1) \ x \ (y-1)] \\ &\triangleright [|y-x-1| \ |y-x-1| \ |y-x-1| \ |y-x-1|] \triangleright [0 \ 0 \ 0 \ 0] \end{aligned}$$

and observe convergence to zero one or two generations sooner in the aforementioned special cases. (This corresponds to region S_1 in Figure 2.)

Example 2. Any box $[0 \ 1 \ x \ y]$ with $x \geq y$ and $x \geq 2$ converges to zero within four generations (three if $y = x - 1$, two if $(x, y) = (2, 1)$). We compute

$$\begin{aligned} [0 \ 1 \ x \ y] &\triangleright [1 \ (x-1) \ (x-y) \ y] \triangleright [(x-2) \ (y-1) \ (2y-x) \ (y-1)] \\ &\triangleright [(y-x+1) \ (y-x+1) \ (y-x+1) \ (y-x+1)] \triangleright [0 \ 0 \ 0 \ 0], \end{aligned}$$

again observing the quicker convergence for the special cases.

Example 3. Any box $[0 \ 1 \ x \ y]$ with $1 < y \leq x < 2$ converges to zero within six generations. Here we find that

$$\begin{aligned} [0 \ 1 \ x \ y] &\triangleright [1 \ (x-1) \ (x-y) \ 2] \triangleright [(2-x) \ (y-1) \ (2y-x) \ (y-1)] \\ &\triangleright [|x+y-3| \ (y-x+1) \ (y-x+1) \ |x+y-3|] \triangleright [p \ 0 \ p \ 0] \\ &\triangleright [p \ p \ p \ p] \triangleright [0 \ 0 \ 0 \ 0], \end{aligned}$$

where $p = ||x+y-3| - y+x-1|$. (Examples 2 and 3 together comprise region S_3 in Figure 2.)

Note that Examples 1, 2, and 3, together with the case $y = 1$ discussed earlier and the type (i) and (ii) classes, cover all nonmonotone classes of diffy boxes.¹ (Monotone classes have $0 < 1 < x < y$.) Further calculations show that monotone classes have longevity at least five. Figure 3 demonstrates how S is subdivided into regions of different longevities (the lighter the region, the greater the longevity). The only two equivalence classes not depicted are (i) and (ii); here (ii) can be considered as the point at infinity (by which is meant the point that must be adjoined to \mathbf{R}^2 in order to compactify it). The black dot in the center is the fixed point, and the white region around it represents all equivalence classes of longevity ten or more generations. Note that where two regions of different longevity meet, the boundary between them belongs to the region of lower longevity.

To determine the longevity of monotone classes in full detail, we change our approach from the sort of increasingly detailed calculations in the foregoing examples to a consideration of preimages under f . We will need to invoke the following result regarding the invertibility of the map f .

¹They thus provide an alternate proof of Proposition 2.

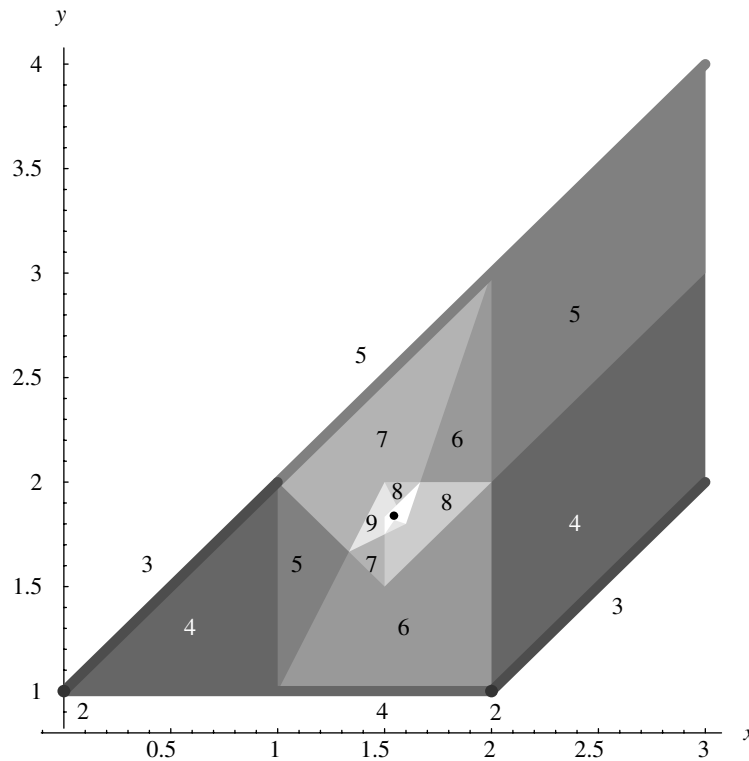


Figure 3. Subdivision of S into regions colored by longevity.

Theorem 2. *The map $f|_{S_2}$ has an inverse g that is defined on the set*

$$A = \text{int } S \cup \{(x, 1) : 0 < x < 2\},$$

maps the interior of S into the interior of S_2 , and preserves line segments.

Proof. We have already seen that of the three branches in the definition of f , only the second is one-to-one. Since the images of the first and third branches lie on the left and right boundaries of S , the inverse map $g = f|_{S_2}^{-1}$ should be well-defined on the set A . We can invert the expression for f on the second branch to find that

$$g(x, y) = \left(\frac{2y}{x + y - 1}, \frac{x + y + 1}{x + y - 1} \right)$$

for (x, y) in A . We observe that g maps the interior of S into the interior of S_2 : for $g(x, y) = (u, v)$, $x > 0$, $y > 1$, and $x - 1 < y < x + 1$ together imply that $u > 1$ and $u < v < u + 1$.

Straightforward calculation shows that g preserves line segments in the interior of S and that f preserves line segments in S_2 . ■

We define a sequence of sets T_n ($n > 1$) inductively, as follows. Let $T_2 = S$, and for $n > 1$ let $T_{n+1} = f^{-1}(T_n)$. Because the backward map g is not defined on the left and right boundaries of S , we shall consider the first few examples individually, until we arrive at a T_n such that $T_n \subset \text{int } S$. We also need to consider the type (ii) class (the

point at infinity), because when $0 < x < 2$,

$$[0 \ 1 \ x \ 1] \triangleright [1 \ |x - 1| \ |x - 1| \ 1] \sim [0 \ 0 \ 1 \ 1] \triangleright [0 \ 1 \ 0 \ 1],$$

that is, the diffy box process sends points $(x, 1)$ on the boundary ($0 < x < 2$) to the type (ii) class and then sends the type (ii) class to the point $(0, 1)$ in S .

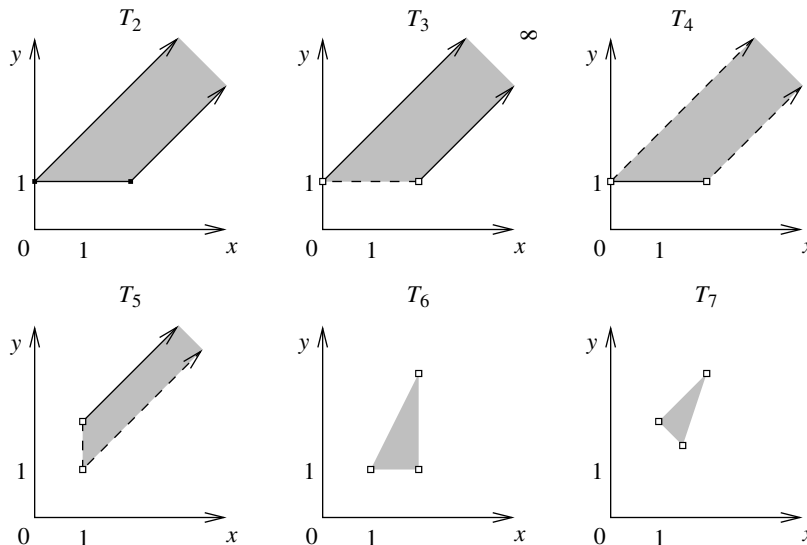


Figure 4. The sets T_n for $2 \leq n \leq 7$.

Four further calculations are necessary before we can make use of g . Excluding the point at infinity from S , we find that $T_3 = f^{-1}(S)$ is S without its lower boundary $\{(x, 1) : 0 < x < 2\}$ but with the point at infinity (see Figure 4), since the only points in S that do not have images in S are on the lower boundary $y = 1$. (The previous paragraph addresses the point at infinity.) Next, $T_4 = f^{-1}(T_3)$ is S without its left and right boundaries (i.e., $T_4 = A$), as the preimage of the interior of S is the interior of S_2 , the preimage of the left boundary of S is $S_3 \cap \text{int } S$, the preimage of the right boundary of S is $S_1 \cap \text{int } S$, and the preimage of the point at infinity is the lower boundary of S (endpoints excluded). The left and right boundaries of S (endpoints excluded), which we exclude from T_4 , are the preimage of the lower boundary of S (endpoints included), which we excluded from T_3 . (The corners are not preimages of anything in S .) Next we find that $T_5 = f^{-1}(T_4) = S_2$; the included boundary is the preimage of the lower boundary of S , endpoints excluded. Finally, we find that $T_6 = f^{-1}(T_5)$ is the interior of the triangle with vertices $(1,1)$, $(2,1)$, and $(2,3)$, again by excluding the preimages of the parts of S excluded from T_5 (see Figure 4 for sketches of all these). T_6 lies in $\text{int } S$.

If we continue in this way, we discover that the T_n for $n > 6$ are also interiors of triangles. When $n \geq 8$ the vertices of these triangles are in the interior of S , which allows us to keep track of the T_n via their vertices (Table 2 provides a partial list, and Figure 5 shows boundaries of some of the T_n ($n \leq 10$) superimposed upon each other).

The utility of the T_n derives from the fact that all equivalence classes of longevity n ($n > 1$) are represented in T_n . It is simple enough to check this for the first few examples; thereafter the result follows by induction. It is also worth noting that, although T_n does not contain all classes of longevity $n + 1$, it does contain all classes

Table 2. Vertices of T_n for $6 \leq n \leq 10$.

n	Vertex 1	Vertex 2	Vertex 3
6	(1, 1)	(2, 1)	(2, 3)
7	(2, 3)	(1, 2)	$(\frac{3}{2}, \frac{3}{2})$
8	$(\frac{3}{2}, \frac{3}{2})$	(2, 2)	$(\frac{3}{2}, 2)$
9	$(\frac{3}{2}, 2)$	$(\frac{4}{3}, \frac{5}{3})$	$(\frac{8}{5}, \frac{9}{5})$
10	$(\frac{8}{5}, \frac{9}{5})$	$(\frac{5}{3}, 2)$	$(\frac{3}{2}, \frac{11}{6})$

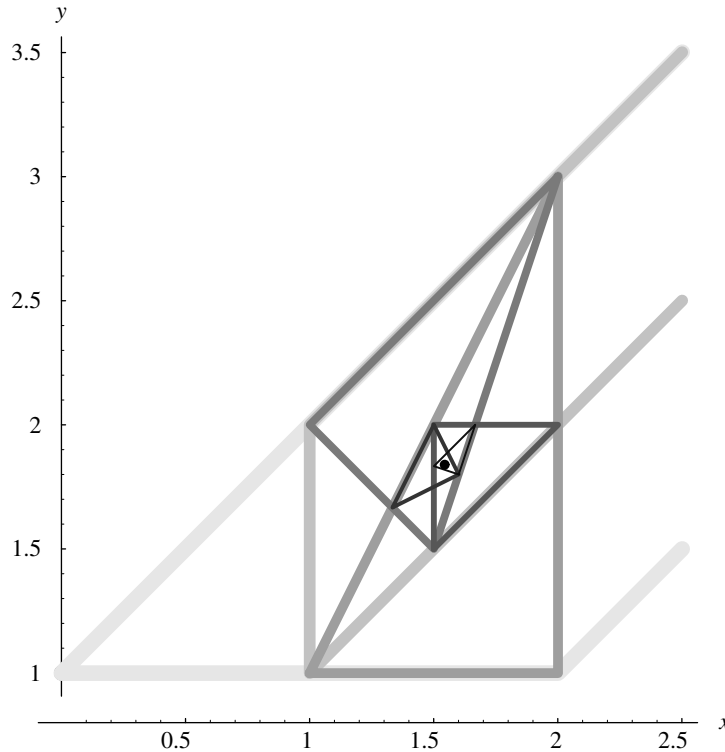


Figure 5. T_n ($n \leq 10$) superimposed upon each other.

of longevity $n + 2$ or greater. Furthermore, we observe (again by induction, starting with $n = 2$) that $T_{n+2} \subset T_n$, so that $T_{n+3} \subset T_{n+1}$, and we can classify the set of all equivalence classes of longevity n ($n > 1$) as precisely $T_n \setminus (T_{n+1} \cup T_{n+2})$. That is, an equivalence class has longevity n if and only if its canonical representative is in T_n but not T_{n+1} or T_{n+2} .

At this point the question arises of how to test where a given point (i.e., equivalence class) falls relative to the sequence of triangles T_n ($n \geq 6$). Arguably the simplest is just to plot it on a graph containing (enough of) the T_n . There are also several simple algebraic approaches, however, to test whether a point falls within a given triangle. One is to write the points involved in vector form. First let the interior of each triangle be written as the set of points whose coordinates are a weighted average of the coordinates of the three vertices V_1 , V_2 , and V_3 :

$$\hat{T} = \{(x, y) : (x, y) = rV_1 + sV_2 + (1 - r - s)V_3; \quad r, s > 0, r + s < 1\}.$$

Now, to test whether a point P is inside \hat{T} , calculate the corresponding “coordinates” r and s :

$$\begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} \overrightarrow{V_3V_1} & \overrightarrow{V_3V_2} \end{bmatrix}^{-1} \overrightarrow{V_3P},$$

and check whether $r > 0$, $s > 0$, and $r + s < 1$. For example, T_6 is the set of (x, y) such that $(x, y) = r(1, 1) + s(2, 1) + (1 - r - s)(2, 3)$ for some r and s satisfying $r > 0$, $s > 0$, and $r + s < 1$, and the fixed point P is $(q(q - 1), q) \approx (1.5437, 1.8393)$, making

$$\overrightarrow{V_3V_1} = (-1, -2), \quad \overrightarrow{V_3V_2} = (0, -2), \quad \overrightarrow{V_3P} \approx (-0.4563, -1.1607),$$

and

$$\begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -2 & -2 \end{bmatrix}^{-1} \begin{bmatrix} -0.4563 \\ -1.1607 \end{bmatrix} = (0.4563, 0.124),$$

verifying that P lies in T_6 .

The only drawback to an algebraic approach is that it is inescapably recursive, and the number of calculations required to continue testing whether a given point falls inside each T_n is comparable to the number of calculations required simply to take the diffy box process toward its eventual end. A graphical approach merely requires plotting a sufficient number of T_n so that the point falls outside two consecutive triangles.

We close this section with one more way to look at the domain S . Figure 6 divides S into three disjoint invariant regions by shades of gray: that is, each shade (light, medium, or dark) represents a sequence of images and preimages under f , jumping around and toward the fixed point. The fading of the colors near the fixed point indicates increasing longevity.

5. CONCLUSIONS, APPLICATIONS, AND EXTENSIONS. We now return to our original question: Does every diffy box converge to the zero box in a finite number of generations, and, if so, how many generations will it take? We can now reinterpret the results of our analysis on equivalence classes in terms of boxes as four-tuples.

The answer to our question as phrased is no, for the diffy boxes belonging to the class containing the box $[0 \ 1 \ q(q - 1) \ q]$ associated with the fixed point of the map f require an infinite number of iterations to reach zero: any diffy box in this class has entries (vertex numbers) that become smaller and smaller but never actually reach zero. However, the counterexample class is unique, and all other diffy boxes converge to zero in finitely many generations (the number depends on how close their canonical forms are to the fixed point). In fact, for any specified longevity, there are classes of diffy boxes that take precisely that long to converge to zero. Nonmonotone boxes converge especially quickly (in no more than six generations), while monotone boxes have longevity at least five. To determine the longevity of a monotone diffy box, it is simplest to put the box in canonical form and compare its coordinates (x, y) with a graph of the regions of various longevity identified in the previous section. We might also make the observation that (as seen by the regions into which S is subdivided in Figure 3) the use of “complicated” numbers such as radicals or transcendentals does not really prolong convergence much, since within a couple of generations the differences have propagated through the four vertices and get subtracted out.

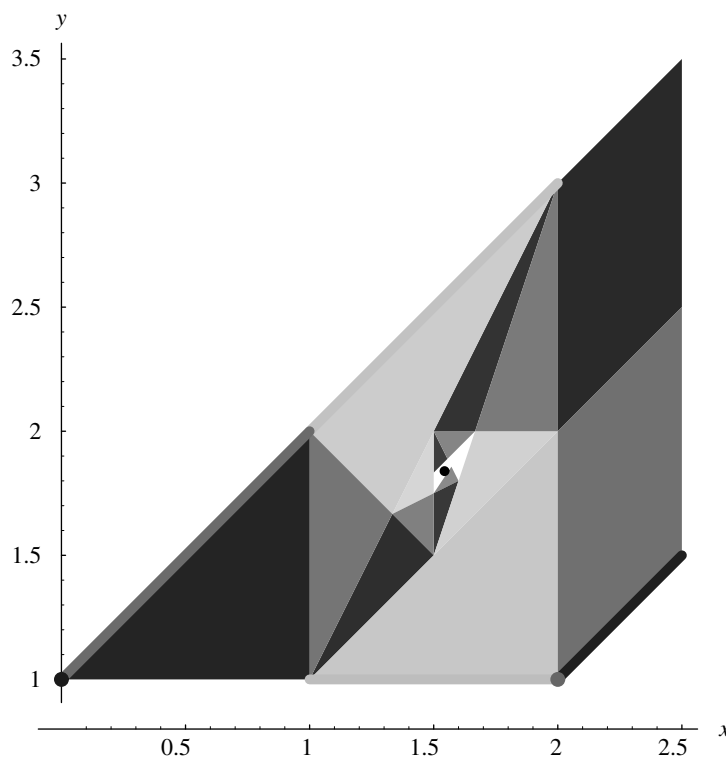


Figure 6. Subdivision of S into 3 invariant regions (by color).

As mentioned in the introduction, diffy boxes can be used as problem-solving contexts for elementary grades students (see [1]). After working through several diffy boxes, children can group them according to longevity and begin to observe some patterns in the forms of boxes that converge in one, two, three, and possibly four generations. They can also observe the properties we used in Definition 5 to define equivalence, as well as the effects (or lack thereof) of using numbers other than whole numbers.

A natural extension of the diffy-box problem that we leave to the reader is the generalization from squares to other polygons. For example, a quick investigation of “diffy triangles” reveals a peculiar chasing pattern and the surprising(?) result that *no* “diffy triangles” ever converge to the all-zero triangle, except for those with all three numbers the same (to convince yourself of this, try to construct the parent of such a diffy triangle). From this observation one might try to classify the types of possible behavior of diffy triangles, or else move to a larger scale and perhaps consider convergence of “diffy polygons” with varying numbers of edges. (If we place “diffy polygons” in the context of graph theory, we see that any simple generalization to a more general class of graphs is prevented by the fact that only the cycle graphs C_n (i.e., polygons) have line graphs isomorphic to themselves. Considering polygons as cycle graphs, however, does allow us to include the trivial example C_2 , the two-sided polygon, which converges to zero in two steps for any two starting values.) The reader interested in the “diffy n -gon” may wish to take a hint from Winkler [4, chap. 2], who observes (p. 17) that, for the special case where all vertices are integers, “A little analysis via polynomials over the integers modulo 2 shows that the salient issue is whether n is a power of 2.” Winkler’s book also uses integers modulo 2 to establish

the logarithmic bound suggested at the end of section 2 for integer-valued diffy boxes (using the maximum entry rather than $|B|$).

Another possible extension is a change in the distance function used to calculate the vertices of a given box's child. We have used the symmetric "one-dimensional" norm $f(a, b) = |a - b|$, but we might instead have used a "two-dimensional" norm $f(a, b) = \sqrt{a^2 - b^2}$, or an asymmetric one-norm $f(a, b) = |Aa - Bb|$ for fixed weights A and B with $A + B = 2$ that places more emphasis on one vertex than the other (in this case we would clearly have to identify vertices by orientation, for example, b as clockwise from a).

It is also interesting to note that the irrational number q involved in the fixed-point diffy box class is also associated with sequences of numbers called *tribonacci numbers*. Similar to the notion of Fibonacci numbers, tribonacci numbers are a sequence of numbers t_n that obey the recursive relation $t_n = t_{n-1} + t_{n-2} + t_{n-3}$. (The sequence typically begins with $t_1 = 1$, $t_2 = 1$, and $t_3 = 2$.) Like Fibonacci numbers, any sequence of tribonacci numbers tends toward a geometric increase, with the ratio of any two successive numbers in the sequence approaching a fixed constant. For tribonacci numbers that constant is q . In fact, beginning as indicated, $t_n \sim q^n$ as $n \rightarrow \infty$. (If we look for geometric solutions $t_n = a^n$ to the defining recursive relation, we see that we must have $a^n = a^{n-1} + a^{n-2} + a^{n-3}$ or $a^3 = a^2 + a + 1$, the same equation we solved to obtain q .)

It is remarkable how mathematically rich such a simple notion can be. We invite the reader to explore further.

ACKNOWLEDGMENT. The authors thank George Christ for introducing them to diffy boxes.

REFERENCES

1. Charles A. Dana Center, *TEXTEAMS Rethinking Elementary School Mathematics, Part I*, University of Texas, Austin, 2002.
2. J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1983.
3. S. H. Strogatz, *Nonlinear Dynamics and Chaos*, Addison-Wesley, New York, 1994.
4. P. Winkler, *Mathematical Puzzles: A Connoisseur's Collection*, AK Peters, Natick, MA, 2003.

ANTONIO BEHN received a Ph.D. in mathematics from the University of Wisconsin. He held a postdoctoral position at Texas A&M University before returning to Chile to teach mathematics at the Universidad de Chile. His research interests include ring theory.
Universidad de Chile, Santiago, Chile

CHRISTOPHER KRIBS-ZALETA is a professor of mathematics and curriculum & instruction at the University of Texas at Arlington. He received undergraduate degrees in electrical engineering and mathematics from Duke University, a master's degree in electrical engineering from Georgia Tech, and a master's and Ph.D. in mathematics from the University of Wisconsin. His research interests include mathematical biology and mathematics education.
University of Texas at Arlington, Arlington, TX 76019-0408
kribz@uta.edu

VADIM PONOMARENKO received an undergraduate degree in mathematics and computer science from the University of Michigan and a master's degree in computer science and Ph.D. in mathematics from the University of Wisconsin. He is a professor of mathematics at Trinity University. His diverse research interests lie primarily within combinatorics.
Trinity University, San Antonio, TX 78212
vadim@trinity.edu