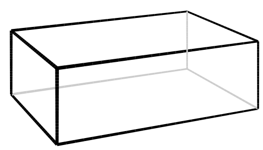
The Mystery of the Sealed Box

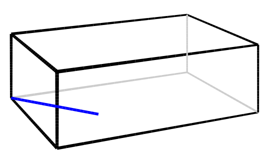
Frederick V. Henle   James M. Henle

athenahealth, Inc.   *Smith College*

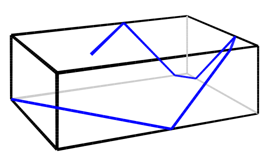
"Sealed" is a red herring. The mysteries here are all on the surface of the box; we don't care what's in the box or if it will ever be unsealed. This paper is all about what wonderful mathematics can lie hidden in a commonplace object. The commonplace object here is a box.



We investigate what happens when we start at one vertex and on one face of the box and move on the surface at angle of 45°from the edges.

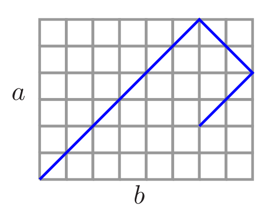


We keep going, bending whenever we hit an edge.

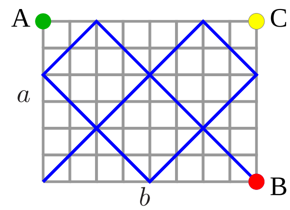


What will happen? Will the trail end at a vertex? If it does, which vertex? These are the mysteries of the sealed box.

We begin with the familiar problem of a billiard ball bouncing in a rectangle, striking the sides at 45°. Given a rectangle with dimensions (a, b), we start the ball at the lower-left corner.



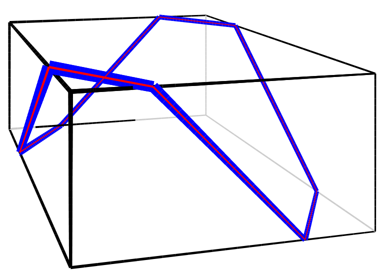
We watch it bounce until it eventually reaches a corner. Which corner it reaches depends on the dimensions a, b, of the rectangle. It's not hard to show that the ending corner is



- A, if a is more even than b,
- B, if b is more even than a,
- C, if a and b are equally even,

where the "evenness" of a number is measured by the number of 2's in its prime factorization. If the dimensions of the rectangle are not integers and if their ratio is not rational, then the path never ends.

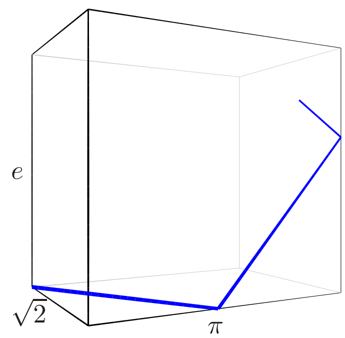
The generalization of this problem to a ray bouncing inside a box is not particularly interesting. Suppose, however, that instead of traveling through the interior of the box, we travel on the surface of the box. The path is a geodesic, by which we mean that it is the shortest path (restricted to the surface of the box) between any two points on the path that are sufficiently close. It's like a ribbon winding around the box.



As with the billiards problem on the rectangle, we can ask if the path reaches a vertex. If it does, we can ask which vertex it reaches. The ribbon problem is far more complicated and there are some surprises. In particular,

1. The pattern of vertices reached generates a fractal resembling a well-known gasket.

2. It's possible for a geodesic to reach a vertex even when the dimensions are not rationally related, for example, the path for the $\sqrt{2} \times \pi \times e$ box.



3. There are reasonably nice answers regarding the destinations of the geodesics and as in the case of the rectangle, evenness is involved.

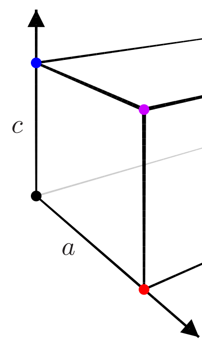
The connection to rectangles is underscored by the case where the height of the box is zero. Then the geodesic looks just like a ray bouncing in a rectangle—except that the box has a distinct top and bottom.

The question we investigate here is connected to many fields of research: billiards on polygons, geodesics on polyhedra, translation surfaces, cutting sequences, symbolic dynamics, Teichmüller theory, and generalized continued fractions, to name a few. Current research deals generally with infinite paths, infinite sequences, global issues. Geodesics are studied, but for geodesics on a box, for example, the focus is on paths that avoid vertices, that continue infinitely in both directions. Our interest here is on the local issue, on geodesics that start and end at vertices.

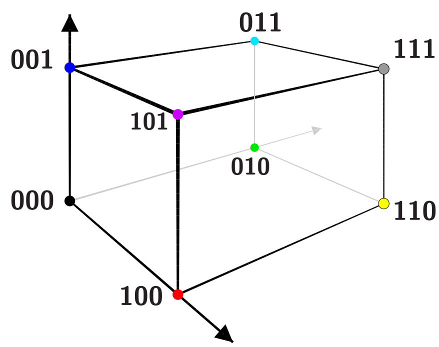
Our analysis of the destination of paths begins with the representation of box shapes as points in a triangle (section 1). In section 2 we reduce the problem to the class of "triangular" boxes. In section 3 we completely describe the destinations of paths on triangular boxes. Paths that never reach a vertex are treated briefly in section 4. We mention some wild ideas in section 5.

1 The Box Triangle

Imagine a box in the first octant of $\mathbb{R}^3$, with the starting vertex at the origin. Let (a, b, c) be the dimensions of the box.



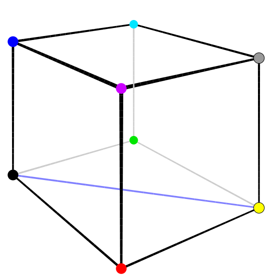
It will be convenient later to identify the eight vertices of the box as **000**, **001**, **010**, **011**, **100**, **101**, **110**, and **111**. The colors will also be useful.



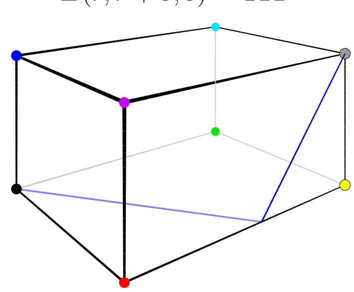
Our paths will always start at **000**, the origin on the xy-plane. If the path for dimensions (a, b, c) ends at a vertex, let $E(a, b, c)$ be that vertex, the "destination". Some simple examples:

Proposition 1 *For any* $r, s > 0$,

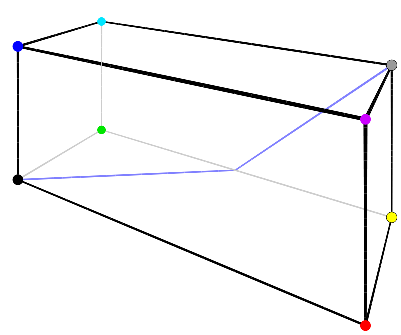
$E(r, r, s) =$ **110**



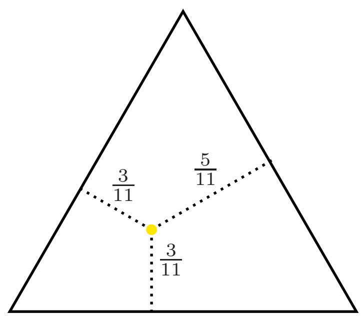
$E(r, r + s, s) =$ **111**



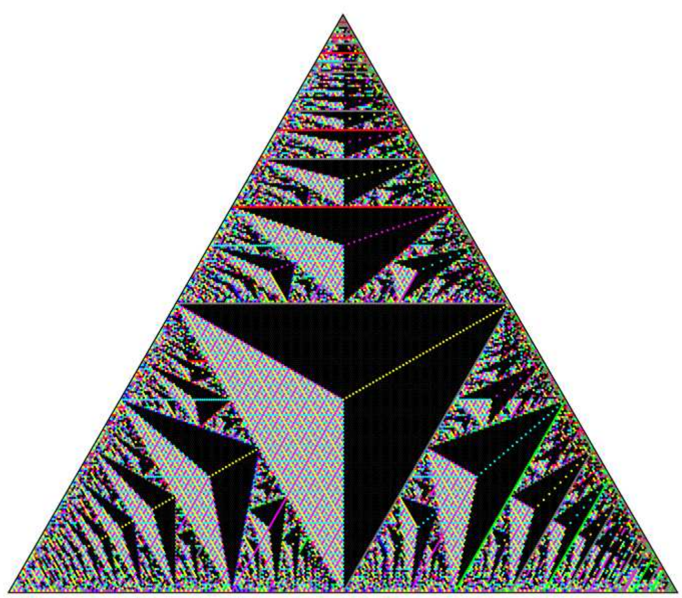
$E(r + s, r, s) =$ **111**



To get an idea of what is happening in general we interpret the dimensions of a rectangle, when normalized, as the barycentric coordinates of a point in an equilateral triangle of height 1. In such a triangle, the sum of the distances of any point to the three sides is always 1. The $3 \times 3 \times 5$ box, for example, corresponds to the point where the distances from the sides are $\left(\frac{3}{3+3+5}, \frac{3}{3+3+5}, \frac{5}{3+3+5}\right) = \left(\frac{3}{11}, \frac{3}{11}, \frac{5}{11}\right)$. Then, since $E(3,3,5) = \mathbf{110}$ and we colored $\mathbf{110}$ yellow, we color the point with barycentric coordinates $\left(\frac{3}{11}, \frac{3}{11}, \frac{5}{11}\right)$ yellow.



When we do this for a great many trios of positive integers we get what we call the "box triangle."



The pattern of triangles resembles the Rauzy gasket, a figure homeomorphic to the Sierpinski gasket.[1] The Rauzy gasket is defined in terms of ternary "epis-

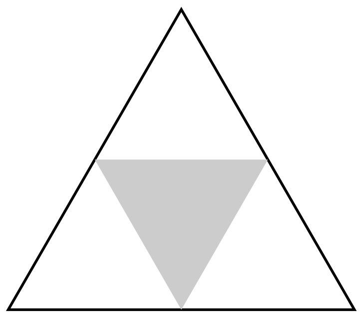
[1] We are indebted to Edmund Harriss for pointing this out to us.

turmian sequences," sequences with special complexity characteristics. Episturmian sequences are related to cutting sequences, sequences that chart, for a given geodesic, the order in which edges are crossed.

2 Triangularity

Definition 1 *A trio of numbers is* **triangular** *if none of them is greater than the sum of the other two. A box is triangular if its dimensions are triangular.*

Triangular boxes correspond to points in the center of the box triangle.



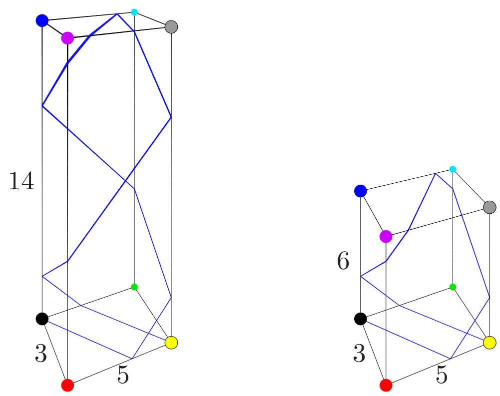
What happens there repeats infinitely in smaller, distorted versions, filling up the box triangle. Any trio of positive integers reduces to a triangular trio by the process of repeatedly subtracting the two smaller numbers from the largest. This action changes the destination of the path in the following way:

Proposition 2 *Let* $E(a, b, c) = \mathbf{xyz}$*. Then*

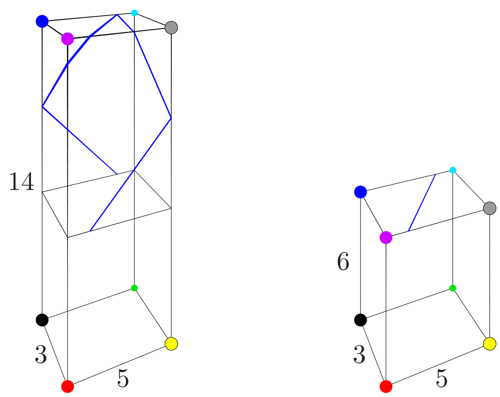
- $E(a + b + c, b, c) = \mathbf{x(y{\oplus}x)(z{\oplus}x)}$,
- $E(a, a + b + c, c) = \mathbf{(x{\oplus}y)y(z{\oplus}y)}$ *and*
- $E(a, b, a + b + c) = \mathbf{(x{\oplus}z)(y{\oplus}z)z}$,

where $\oplus$ *is the 'exclusive or' operation (sometimes written XOR):* $\mathbf{0} \oplus \mathbf{0} = \mathbf{1} \oplus \mathbf{1} \mathbf{=} 0$, $\mathbf{0} \oplus \mathbf{1} = \mathbf{1} \oplus \mathbf{0} = \mathbf{1}$.

Compare, for example, the paths on the $3 \times 5 \times 14$ box and the $3 \times 5 \times 6$ box.

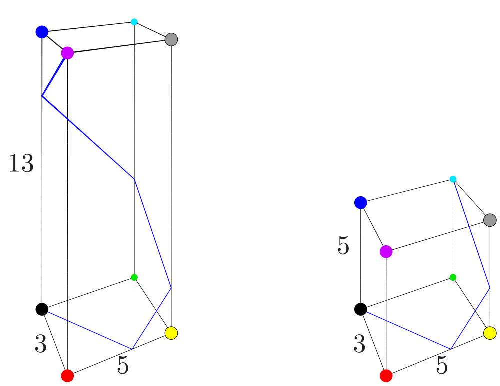


The destinations are the same, **110**. The additional height on the taller box simply takes the path on an extra run around the walls, matching the action on the top of the shorter box.



The key here is that **z** is **0** (that is, the destination is on the bottom). Operating $\oplus$ with **0** changes nothing, so **(1**$\oplus$**0)(1**$\oplus$**0)0**=**110**.

On the other hand, if **z** is **1** (the destination is on the top), as it is in the $3 \times 5 \times 13$ and $3 \times 5 \times 5$ boxes,



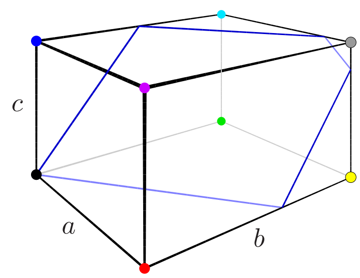
the destination changes from **011** to **101** because the additional height takes the path halfway around the walls. Operating $\oplus$ with **1** changes everything and we get **(0$\oplus$1)(1$\oplus$1)1=101**.

A complete proof of Proposition 2 can be found at [4].

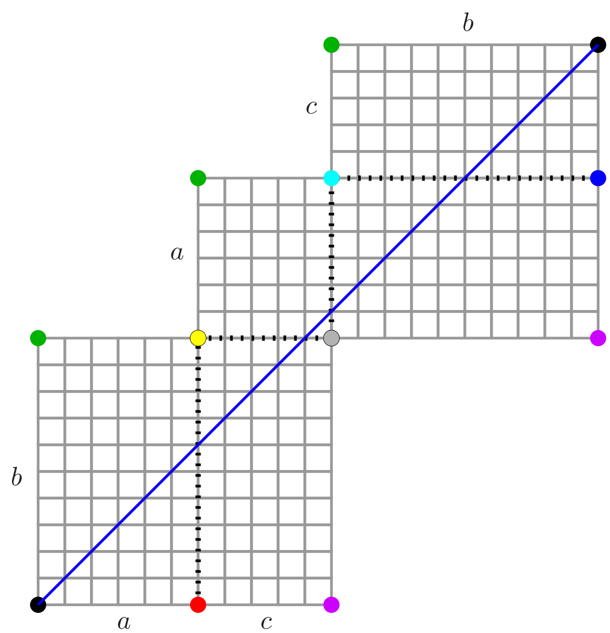
In view of Proposition 2, we can restrict our attention to triangular boxes.

3 Triangular Boxes

If we follow the path in the first example where the dimensions were $(6, 10, 5)$, the path returns to the original vertex.



We can see how this works by laying the path out in the plane. We call the following diagram the "unfolding" of the box.

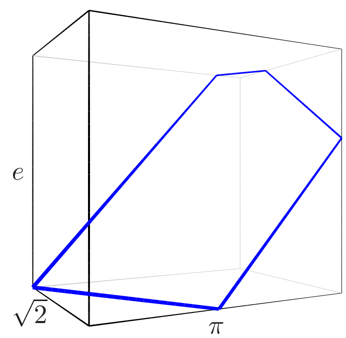


What makes this path miss certain vertices and return to the starting point are the inequalities $a + c > b$ and $b > a, c$. This gives us:

Proposition 3 *For any triangular a, b, c with either $a > b, c$ or $b > a, c$,*

$$E(a, b, c) = \mathbf{000}$$

Proposition 3 holds not just for natural numbers but for all positive real numbers. In particular, it gives us the fact mentioned earlier that a path from a vertex might reach another vertex even if the dimensions of the box are mutually incommensurable.



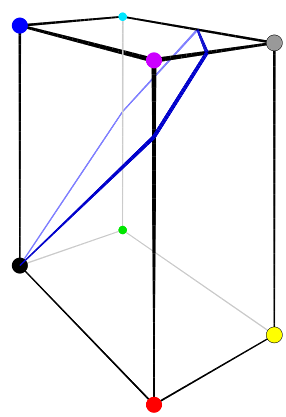
Another consequence of Proposition 3 is that at least two-thirds of the box triangle is black, the color of vertex **000**.

The case where $c \geq a, b$ is more complicated. Like billiards on rectangles, the solution involves evenness.

Proposition 4 *For any triangular a, b, c with $c \geq a, b$,*

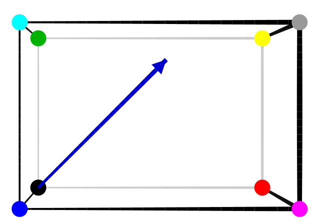
$$\begin{array}{ll} E(a,b,c) = \mathbf{101} & \textit{if } c-a \textit{ is more even than } c-b \\ E(a,b,c) = \mathbf{011} & \textit{if } c-b \textit{ is more even than } c-a \\ E(a,b,c) = \mathbf{110} & \textit{if } c-a \textit{ and } c-b \textit{ are equally even.} \end{array}$$

Proof: Note first that in view of Proposition 3, any path starting on one of the four walls of the box (instead of the bottom or top) must end at the vertex where it began, traveling on the adjacent wall.



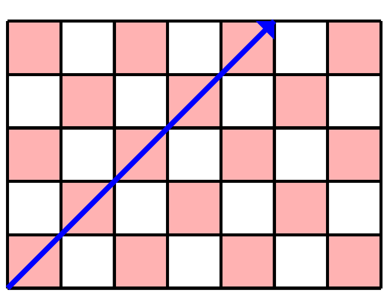
Every path that ends on a wall is part of a loop that begins on a wall, so no path that ends on a wall can begin on the bottom or top. Thus, a path starting on the bottom face must end running along the bottom face or end on the top face.

To see what is going on when we start along the bottom, we view the box from above.

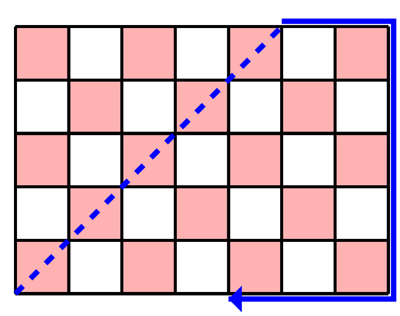


We draw a checkerboard pattern on the bottom and top faces of the box to see which vertices are possible destinations and which are not. We use the same pattern for top and bottom with the lower-left vertex pink. As an example, consider a box with dimensions $(7, 5, 10)$.

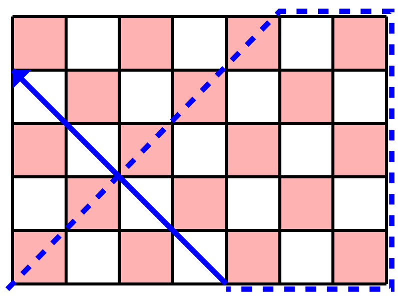
Seen from above, the path starts like this on the floor,



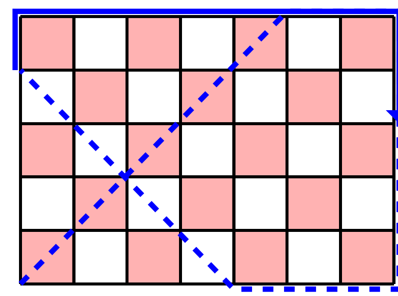
then continues like this as it climbs up the walls of the box.



The path continues on the top of the box,



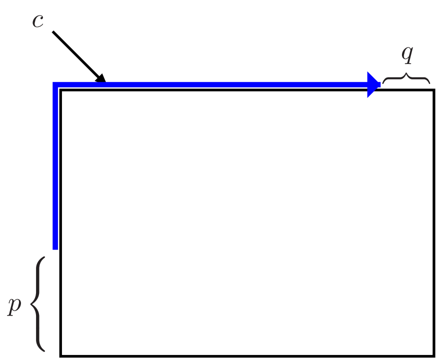
and then down the walls of the box.



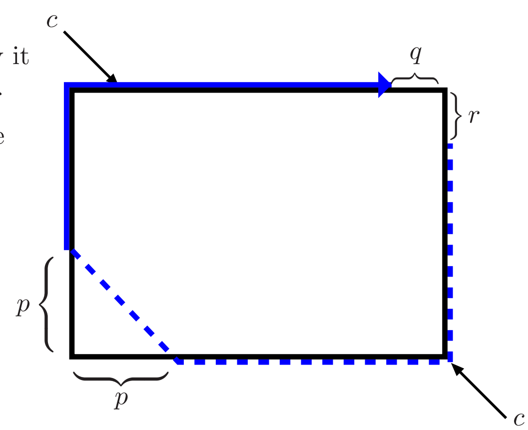
Notice that on the bottom the path moves SW-NE $\nearrow$ or NE-SW $\swarrow$ and on the top the path moves SE-NW $\nwarrow$ or NW-SE $\searrow$.

Claim 1 *Seen from the top, as the path moves on the walls it always passes exactly two vertices.*

Proof of Claim 1: It must pass at least one vertex, since $c > a, b$. It can't pass three because $c \leq a + b$. But if it ever passed just one vertex.



then by symmetry it would have earlier passed by just one vertex (note that on the left $p + c + q = a + b$, and on the right, $p + c + r = a + b$, so $q = r$).



Thus there can't be a first time to pass just one vertex.

$\blacksquare_{\text{Claim}}$

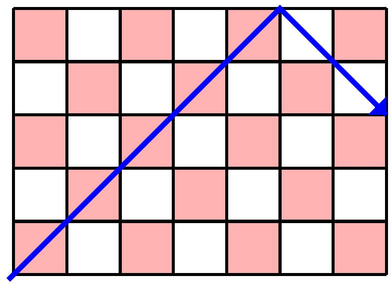
Claim 2 *Seen from the top, the path always moves SW-NE or NE-SW on the bottom and SE-NW or NW-SE on the top.*

Proof of Claim 2: At the start the path is on the bottom and moving SW-NE. From the previous claim, when the path climbs the walls, it ends up on the other wall of the box and so the direction of the path on the top will be SE-NW and appear perpendicular (looking down) to the path on the bottom. This pattern continues. The paths on the bottom are always perpendicular to the paths on the top, establishing the claim.

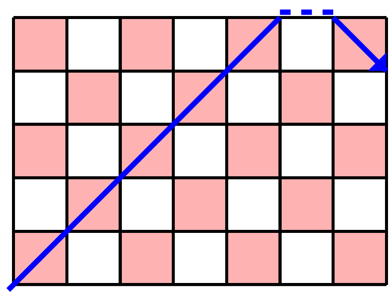
$\blacksquare_{\text{Claim}}$

Claim 3 *The path on the bottom always moves on pink squares. The path on the top always moves on pink squares if c is odd; it always moves on white squares if c is even.*

Proof of Claim 3: If c were 0, the parity (pink/white) would change between bottom and top.



If c were 1, and we're not near a vertex, the parity wouldn't change.



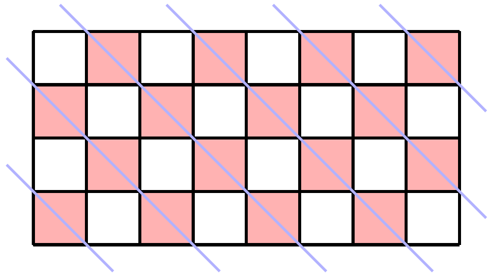
This pattern continues for $c = 2, 3, 4, \ldots$. The only difficulty is that going around a vertex changes whether or not the parity shifts. But by Claim 1, the path always passes exactly two vertices, removing the effect of vertices on parity.

$\blacksquare_{\text{Claim}}$

The proof of Proposition 4 now proceeds by cases.

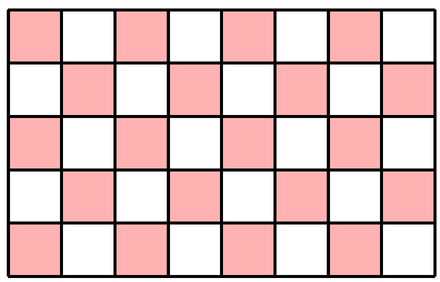
Case 1: a and b even We consider the subcases, c even and c odd, separately. If c is even, we can divide the dimensions by two. The resulting box has the same properties as the original box and reduces to this or another case, but on a smaller box.

If c is odd, consider the top of the box. On the top the path will travel on pink squares, going SE-NW and NW-SE.



There is no vertex it can reach traveling this way, so the destination must be on the bottom. On the bottom the path travels on pink squares SW-NE and thus can only end at **110**. This is as the Proposition states ($c - a$ and $c - b$ are odd, hence equally even).

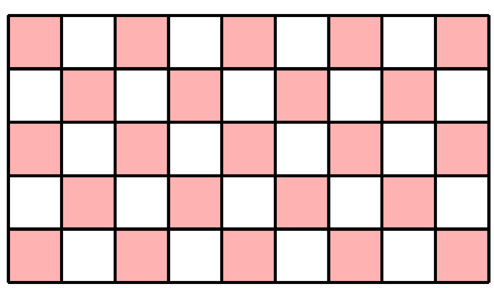
Case 2: One of a, b is even and one is odd Suppose, for example, that a is odd and b is even (a similar argument works for a even and b odd).



There is no vertex on the bottom that the path can reach traveling SW-NE. If c is even, the path will travel on white squares on top. The only vertex possible is the SE vertex, **101**, as predicted by the Proposition ($c - a$ is even, $c - b$ is odd).

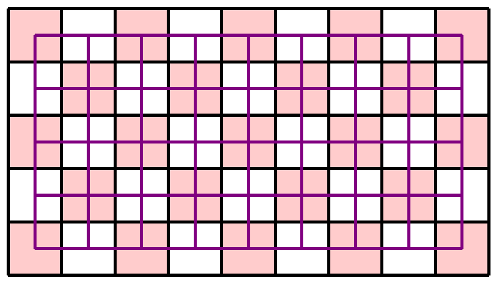
If c is odd, the path will travel on pink squares on top and can only end at the NW vertex, **011**, and again agreeing with the Proposition.

Case 3: a and b both odd

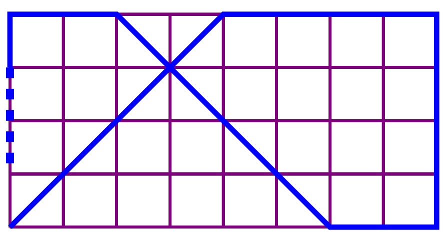


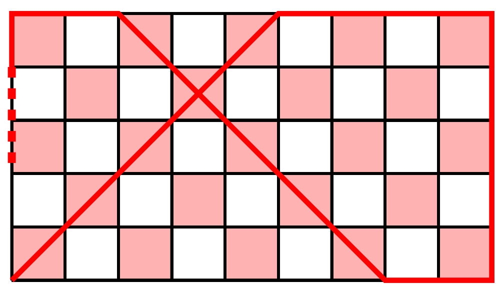
On the bottom, the path can end at the NE vertex (**110**). If c is even the path will travel on white squares on top and no vertex is possible so the destination must be **110**. This agrees with the proposition.

But if c is odd, there will be three possible vertices: **110** on the bottom, **011** and **101** on the top. Since all three dimensions are odd, let $a = 2k+1$, $b = 2n+1$, $c = 2m+1$. To analyze the situation, we consider a smaller box, a box with dimensions $(2k, 2n, 2m)$.

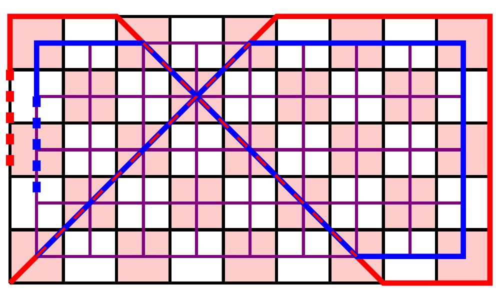


The paths on the two boxes have the same form, that is, the same sequence of diagonal, vertical and horizontal segments.

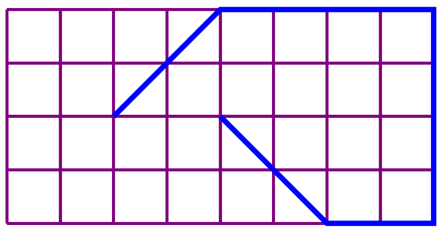


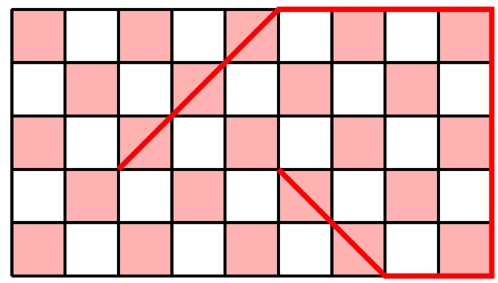


This is easier to see with one picture on top of the other.



The only tricky part is noting the distance traveled on the sides of the box (the edge of the rectangle). On the smaller box the path travels on the outside a distance of $2m$ blocks, while on the larger it is $2m + 1$ blocks.

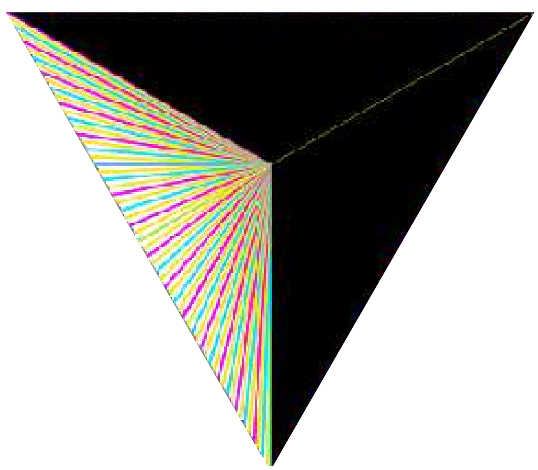




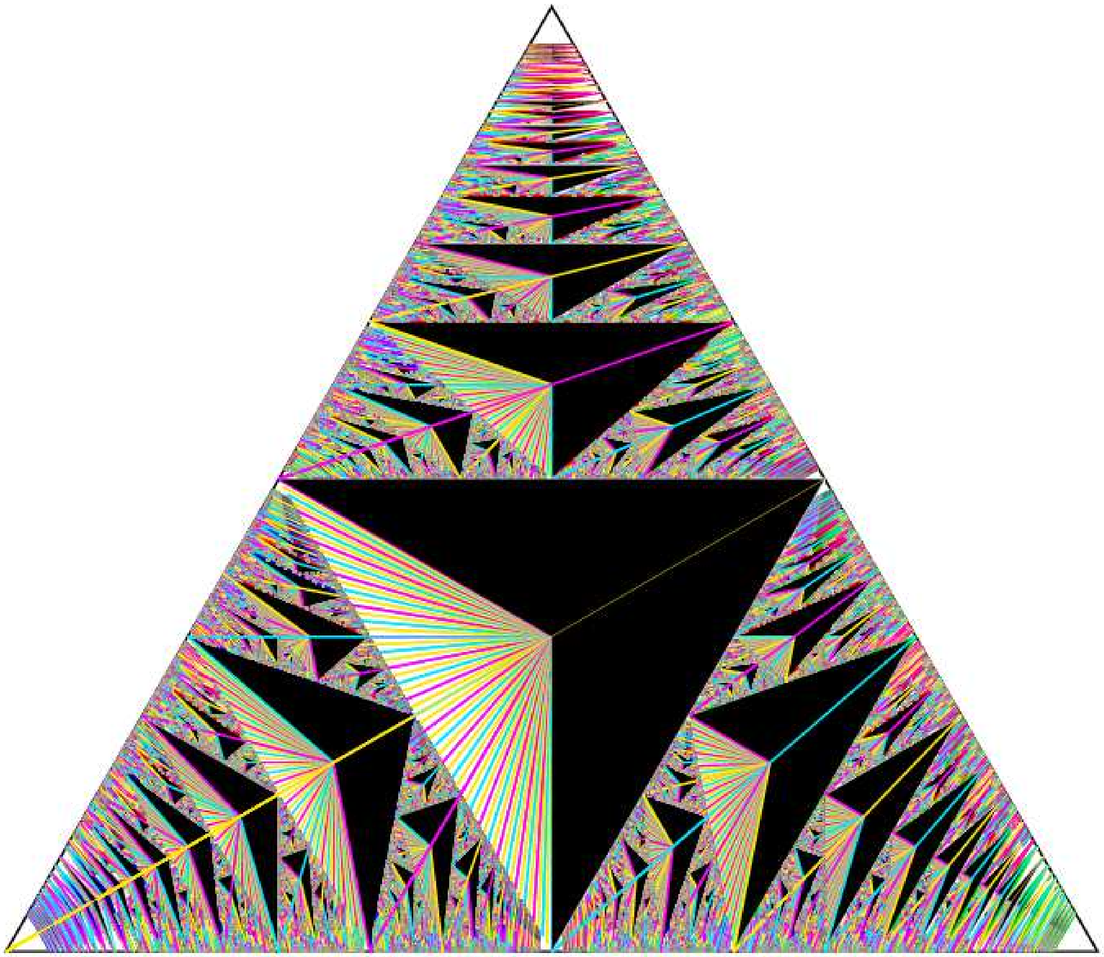
The difference is made up by the fact that the outer path is slowed by exactly two vertices. Thus, Case 3 reduces to the smaller box, which falls under Case 1. Note that the differences, $c - b$ and $c - a$, are the same for the two boxes and hence there is no change in the relative evenness.

■

By Proposition 3 the center of the triangle is two-thirds black. By Proposition 2, the entire triangle is two-thirds black. Since destinations are preserved when a box is magnified equally in all directions, Proposition 4 tells us that in the non-black sector of the central triangle, the only possible destinations are $E(a, b, c) =$ **101**, **011**, or **110**. In fact, we can show ([4]) that all points in the non-black sector on a straight line from the center of the triangle to the edge have the same destination (and hence the same color).



Putting all this information together, we have a pretty detailed picture of the box triangle.



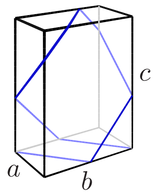
4 Infinite Paths

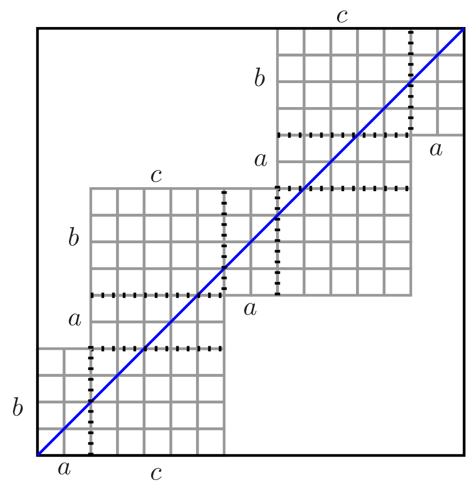
Propositions 2, 3 and 4 combine to show that for all boxes with positive integral dimensions, the path from a vertex always reaches a vertex. With irrational dimensions, there can be infinite paths.

Proposition 5 *If a, b, and c are linearly independent as vectors over the field of rational numbers and $a, b < c < a + b$, then the path never reaches a vertex.*

Proof:

Consider the unfolding of a box with a finite path.





The diagram is contained in a square. Each side of the square is a linear combination of a, b, and c. In the case above, the square is $3a + 2c$ wide and $3b + 2a$ high.

The sums of the coefficients in the linear expressions note the number of times each dimension of the box is traversed. For example, the sum of the coefficients of a is 5 and indeed the short side a is crossed five times in the path. The sums of the coefficients of a and b are odd; the sum of the coefficients of c is even and so the destination of the path is **110**.

Definition 2 *Call a triple* (a, b, c) balanced *if in the unfolding, the coefficients of* a*,* b*, and* c *in the expressions for the lengths of the sides of the unfolding are identical.*

Suppose that for some linearly independent a, b, c, the path reaches a vertex. Consider the unfolding. The sides of the square containing the unfolding will each be linear combinations of a, b, and c, and as sides of a square, they will be equal. Since a, b, c are linearly independent, we must have that the coefficients for the two sides are identical, so (a, b, c) is balanced. That means that the sums of the two coefficients for each of a, b and c are even. Hence the destination of the path is **000**.

Claim 1 *The set of points in the color triangle whose barycentric coordinates are balanced is open.*

Proof of Claim 1: If (a, b, c) is balanced, then sufficiently slight changes in a, b, c will not change the topology of the path in the unfolding. Imagine the effect on the unfolding of making a slight increase, say, in b. Everything in the diagram moves. But the point at the upper right vertex will move up and to the right at an angle of 45°because (a, b, c) is balanced—every change to one side of the square is matched by an identical change to the other. The same is true for any (sufficiently slight) change in a or c. Thus the destination in the altered system remains the same. Note that the altered path will cross the same edges and in the same order.

$\blacksquare_{\text{Claim}}$

Using the claim, we can choose rationals r', s', t' approximating r, s and t which have an equivalent unfolding, that is, $E(r', s', t') = \mathbf{000}$. But we will still have $r', s' < t' < r' + s'$ and by the remark following Proposition 4, $E(r', s', t')$ must be **101**, **011**, or **110**. This is a contradiction and the Proposition is proved.

$\blacksquare$

5 Additional Mysteries

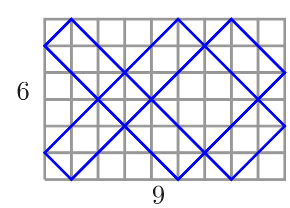
Is there a closed-form description of the destination of the path?

Is there a closed-form description of when paths are infinite?

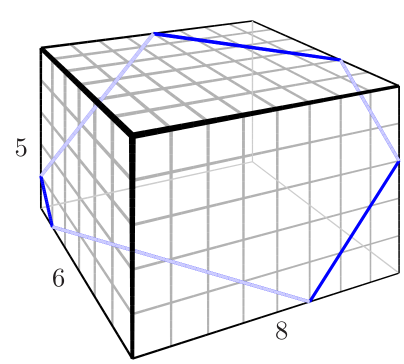
What is the length of the path?

How does this connect with work in related fields?

One particularly interesting direction involves some deviant number theory. The idea is to replicate on the box some simple features of the rectangle. A rectangle has closed loops



if and only if the dimensions have a common factor. That's not true for boxes.



On a rectangle, the number of loops is one less than the greatest common divisor of the dimensions. And the length of the path from a vertex is the least common multiple. All this suggests the possibility of wild new definitions of "relatively prime", "greatest common divisor", and "least common multiple" for triples. We may have found such definitions; you can find them at [4].

References

[1] P. Arnoux and Š. Starosta, "The Rauzy Gasket," preprint.

[2] D. B. Fuchs and E. Fuchs, "Closed geodesics on regular polyhedra," *Moscow Mathematics Journal*, **7**(2): 265-279 (2007).

[3] G. Galperin, "Convex polyhedra without simple closed geodesics," *Regular Chaotic Dynamics*, **8**(1): 45-58 (2003).

[4] F. V. Henle and J. M. Henle, "The Mystery of the Sealed Box," `www.math.smith.edu/~jhenle/Boxes`, website under construction.

[5] V. Yu Protasov, "Closed geodesics on the surface of a simplex," *Sbornik: Mathematics*, **198**(2): 103-120 (2007).

[6] Yaroslav Vorobets, "Periodic geodesics on translation surfaces," *Contemporary Mathematics*, **385**: 205-258 (2005).

[7] V. A. Zalgaller, "Symmetric geodesics on cubes: Algorithms for finding them," *St. Petersburg Mathematics Journal*, **16**(2): 423-436 (2005).

The authors would like to thank Michael Henle for discussions and fruitful ideas.