Geodesic complexity of a tetrahedron

Donald M. Davis

Abstract

The topological (resp. geodesic) complexity of a topological (resp. metric) space is roughly the smallest number of continuous rules required to choose paths (resp. shortest paths) between any points of the space. We prove that the geodesic complexity of a regular tetrahedron exceeds its topological complexity by 1 or 2. The proof involves a careful analysis of shortest paths on the tetrahedron.

Keywords: Geodesic complexity, topological robotics, geodesics, tetrahedron.

MSC 2020. 53C22, 52B10, 55M30.

1   Introduction

In [3], Farber introduced the concept of the *topological complexity*, $TC(X)$, of a topological space X, which is the minimal number k such that there is a partition

$$X \times X = E_1 \sqcup \cdots \sqcup E_k$$

with each E_i being locally compact and admitting a continuous function $\phi_i : E_i \to P(X)$ such that $\phi_i(x_0, x_1)$ is a path from x_0 to x_1. Here $P(X)$ is the space of paths in X with the compact-open topology, and each ϕ_i is called a motion-planning rule. If X is the space of configurations of one or more robots, this models the number of continuous rules required to program the robots to move between any two configurations.

In [4], Recio-Mitter suggested that if X is a metric space, then we require that the paths $\phi_i(x_0, x_1)$ be minimal geodesics (shortest paths) from x_0 to x_1, and defined the *geodesic complexity*, $GC(X)$, to be the smallest number k such that there is a partition

$$X \times X = E_1 \sqcup \cdots \sqcup E_k$$

with each E_i being locally compact and admitting a continuous function $\phi_i : E_i \to P(X)$ such that $\phi_i(x_0, x_1)$ is a minimal geodesic from x_0 to x_1.[1] Each function ϕ_i is called a *geodesic motion-planning rule* (GMPR).

One example discussed by Recio-Mitter in [4] was when X is (the surface of) a cube. It is well-known that here $TC(X) = TC(S^2) = 3$, and he showed that $GC(X) \geq 4$.

In this paper, we let X be the surface of a regular tetrahedron T, and prove

Theorem 1.1. $GC(T) = 4$ *or* 5.

---

[1] Recio-Mitter's definition of $GC(X) = k$ involved partitions into sets $E_0, \ldots, E_k$, which, for technical reasons, has become the more common definition of concepts of this sort, but we prefer here to stick with Farber's more intuitive formulation.

Again, for comparison, $\mathrm{TC}(T) = \mathrm{TC}(S^2) = 3$.

In Section 2, we introduce what we call the *expanded cut locus* in order to study the geodesics on T. In Section 3, we prove $\mathrm{GC}(T) \leq 5$, and in Section 4, we prove $\mathrm{GC}(T) \geq 4$. Despite considerable effort, we have been unable to establish the precise value of $\mathrm{GC}(T)$.

2   Expanded cut locus

The *cut locus* of a point P on a convex polyhedron is the set of points Q such that there is more than one shortest path from P to Q.[2] For the regular tetrahedron T, this is conveniently sketched on a flat model, or unfolding, of T. For $P \in T$, we define the *expanded cut locus* of P to be the set of terminal points of equal shortest paths[3] from P to versions of cut-locus points Q in an unfolding of T, expanded so that the same face may appear more than once.

In Figure 2.1 we illustrate the expanded cut locus of a point P. The open segments $a\,U_0$ and $a\,U_-$ correspond to the same set of points in the tetrahedron, and the segments from P to points on each at equal distance from a depict equal shortest segments from P to a point Q in T. A similar situation holds for open segments from d to U_0 and U_+, from c to L_0 and L_+, and from b to L_0 and L_-. Also the small open segments $U_-\,L_-$ and $U_+\,L_+$ are part of the expanded cut locus of P, as they represent the same points in T, and segments from P to points at equal height on the two lines are equal minimal geodesics. The three points U_+, U_-, and U_0 represent the same point in T; the paths from P to them are equal shortest paths in T. Similarly for the three L-points. Thus the expanded cut locus of P is the entire red polygon $U_0U_+L_+L_0L_-U_-U_0$ in Figure 2.1 minus the points a, b, c, and d.

The actual cut locus for this point P is shown in Figure 2.2, which is a flat version of part of T, but does not contain multiple versions of points.
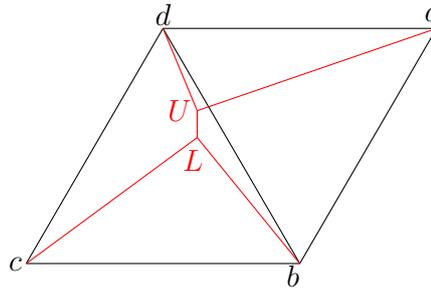
Figure 2.1. An expanded cut locus.

---
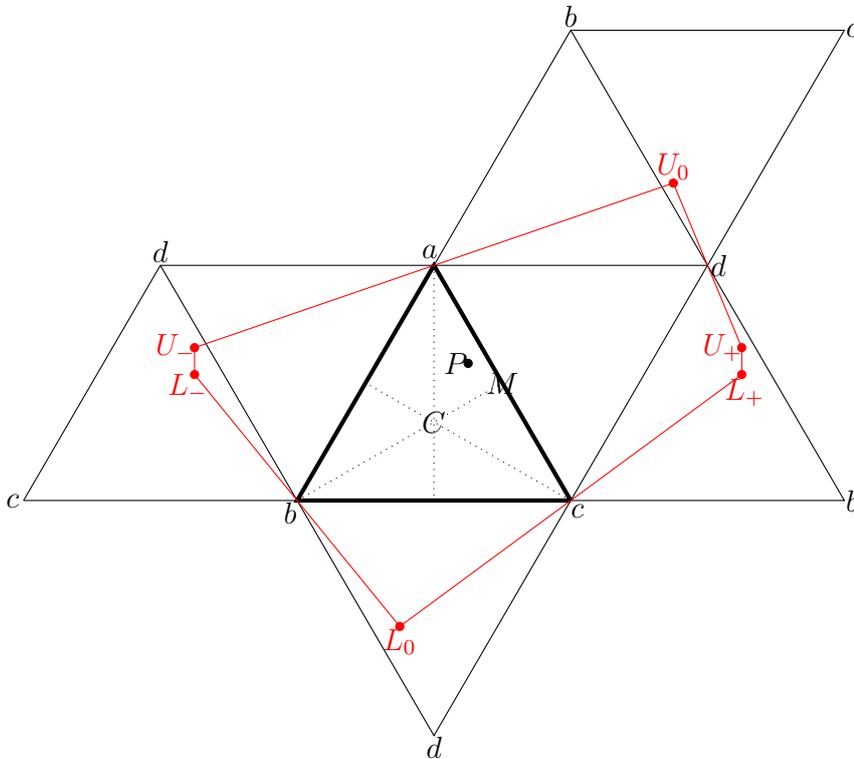
[2] The cut locus is often defined to be the closure of this set, but it is convenient here to define it as we have.

[3] We will use the phrase "equal shortest path" or "equal minimal geodesic" to refer to distinct shortest paths of equal length.

Figure 2.2. The corresponding cut locus.

The expanded cut locus of any point P in the interior of triangle aCM in Figure 2.1, where C is the centroid and M the midpoint of ac, has a form similar to the one depicted there. We make this precise in Theorem 2.3.

Theorem 2.3. *Suppose that in Figure 2.1 the coordinates of a, b, and c are, respectively, $(0, \sqrt{3})$, $(-1, 0)$, and $(1, 0)$, and $P = (x, \alpha\sqrt{3})$ with $0 < x < \frac{1}{2}$ and $\frac{1}{3} + \frac{1}{3}x < \alpha < 1 - x$. Then the expanded cut locus of P is as depicted in Figure 2.1 and described above with*

$$
\begin{aligned}
U_\pm &= \left(\pm 2 + x, \sqrt{3}\left(1 - \frac{x(2-x)}{3(1-\alpha)}\right)\right) \\
U_0 &= \left(2 - x, \sqrt{3}\left(1 + \frac{x(2-x)}{3(1-\alpha)}\right)\right) \\
L_\pm &= \left(\pm 2 + x, \sqrt{3}\,\frac{1 - x^2}{3\alpha}\right) \\
L_0 &= \left(-x, \sqrt{3}\,\frac{x^2 - 1}{3\alpha}\right).
\end{aligned}
\tag{2.4}
$$

Proof. Since

$$
\langle x, \sqrt{3}(\alpha - 1)\rangle \cdot \langle 2 - x, \sqrt{3}\,\frac{x(2-x)}{3(1-\alpha)}\rangle = 0,
$$

$\overrightarrow{aP} \perp \overrightarrow{aU_0}$. This implies that if Q_0 and Q_- are points on aU_0 and aU_-, respectively, at equal distances from a, then the segments PQ_0 and PQ_- have equal length. Since Q_0 and Q_- represent the same point in face abd of the tetrahedron, we deduce that this point is in the cut locus of P.

Similarly the red lines through b, c, and d are perpendicular to the segments from P to those points. Another easy verification is that $\frac{1}{2}(U_0 + U_-) = a$, and so Pa is the perpendicular bisector of U_0U_-, and similarly for b, c, and d. That $\frac{1}{2}(U_0 + U_+) = d$ shows that U_0 and U_+ lie in the same relative position in triangle bcd. One readily sees that the region inside the red polygon in Figure 2.1 exactly covers the four triangles that comprise the tetrahedron. ∎

This slick verification hides the way in which the formulas (2.4) were obtained. We initially used the method of star unfolding and Voronoi diagrams developed in [1], and applied to the cube in [2] using perpendicular bisectors.

The triangle abc in Figure 2.1 is divided into six congruent subtriangles. The formulas (2.4) only apply to points P in the interior of the upper right subtriangle aCM, but the expanded cut locus of points in the other five subtriangles can be obtained by obvious rotations and reflections. We now consider the form of the expanded cut locus for points on the boundary of triangle aCM.

As P approaches the edge aM, $L_\pm$ approaches $U_\pm$. When P is on the edge, they coincide, and the two multiplicity-3 points U and L in the cut locus become a single multiplicity-4 point, which we will later call B, for "both." In Figure 2.5, we depict the two extreme cases, $P = a$ and $P = M$. The

continuum between them should be clear. We label the left one $P \approx a$, because when $P = a$, the line passing through a is not part of the expanded cut locus, since the line connecting P with points on the lines at equal distance from a in each direction are actually the same line in T. But for points P arbitrarily close to a, the lines from P to points on the line are not the same line in T.

Figure 2.5. P on an edge.
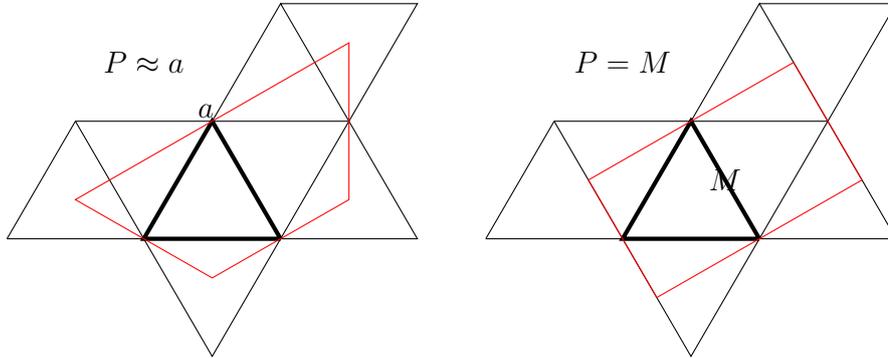
As P approaches the line $x = 0$, U_+ and U_0 approach d_+ (the version of d on the positive side in Figure 2.1), and U_- approaches d_-. The diagram when $x = 0$ is in Figure 2.6.

Figure 2.6. P on the line $x = 0$.
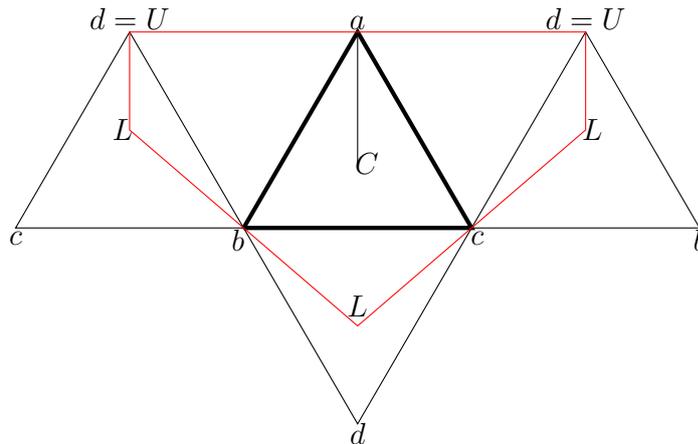
As P moves from a to C along the line $x = 0$, the point L in Figure 2.6 moves from the centroid of bcd to d. The limiting case $P = a$ has already been discussed. However, if the L in Figure 2.6 is moved to the centroid of bcd, we obtain a picture which looks quite different from the left side of Figure 2.5, which also depicts the case $P = a$. Even accounting for the fact that when $P = a$, the line emanating from a is not part of the expanded cut locus, the diagrams still differ in that Figure 2.6 has a vertical line on the left side, whereas Figure 2.5 has a vertical line in the upper right. The explanation is that paths from a to corresponding points on those lines are exactly the same path on T.
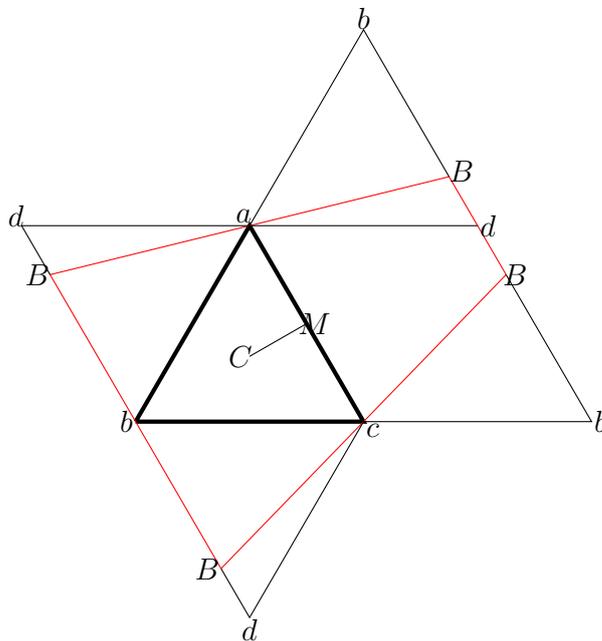
In Figure 2.7 we show the expanded cut locus when P is at the centroid C of abc, which is the case $L = d$ in Figure 2.6.

Figure 2.7. P **at the centroid.**
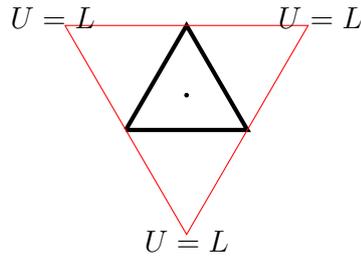
Finally, if P is on the segment CM, $U_\pm = L_\pm (= B)$, and they lie on edge bd. This is depicted in Figure 2.8. As P moves from C to M, B moves from d to the midpoint of bd.

Figure 2.8. P **on the segment** CM**.**

3   Upper bound

Theorem 3.1. *There is a partition*

$$T \times T = E_1 \sqcup E_2 \sqcup E_3 \sqcup E_4 \sqcup E_5$$

with E_i locally compact and a GMPR ϕ_i on E_i.

Proof. Let G_P denote the polygon associated to the point P sketched in red in any of the figures of Section 2. More precisely, one must, of course, use the formulas (2.4) to determine the vertices of the polygon, and if P is reflected across the line $x = 0$ in Figure 2.1, then one must modify the formulas to give the reflection of the polygon. If P is at a vertex, there are two choices for G_P, either as in Figure 2.5 or 2.6. It doesn't matter, but let's choose 2.6.

The set E_1 is the complement of the total cut locus of T. It consists of pairs (P, Q) such that Q is interior to the polygon G_P, together with those for which Q is a vertex of T, except for cases such as (P, d) in Figure 2.6. (The only cases when a vertex V is in the cut locus of a point P is when P lies on a half-open segment $[C, V')$ connecting the centroid C of the face opposite V with one of the other vertices V'.) Here $\phi_1(P, Q)$ is the straight line from P to Q in our expanded cut locus diagram.

The set E_2 consists of pairs (P, Q) where P is not a vertex and Q lies in the interior of a cut-locus segment from a vertex V to a U or L point, excluding cases in which P lies on a segment from a

vertex of face *abc* to the centroid C of *abc*, and $V = d$. We choose $\phi_2(P, Q)$ to be the path from P to the appropriate point on the right side of the vector from P to V. For example, in Figures 2.1 and 2.2, E_2 contains (P, Q) for all Q in the open segments aU, bL, cL, and dU in 2.2, and in 2.1 we choose the segments connecting P with points on aU_0, bL_-, cL_0, and dU_+. To maintain continuity of ϕ_2, we had to exclude points (P, Q) with P on the segment aC in Figure 2.1 and Q on dU because shortest paths from the point P in Figure 2.1 to dU must pass through side ac, whereas for points P on the left side of aC the diagram is reflected and the shortest paths from P to dU will pass through side ab.

This requires some care because, for example, if P is in face *abc*, the cut-locus line out from vertex d plays a different role than the others. Because we have excluded points with P on segments from a vertex to a centroid, we can consider the domain of points P for which Q is on a cut-locus line from vertex d as three topologically disjoint[4] sets, the interiors of $aCbd$, $adcC$, and $bCcd$ in Figure 3.2.

Figure 3.2. *P*-domains for lines through d.
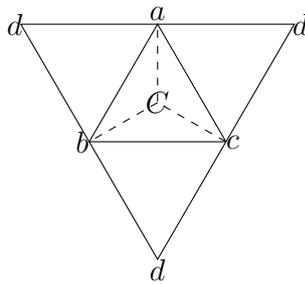
The continuity of ϕ_2 on each of these domains should be fairly clear, but because of the different roles played by points in face *abc* and the other points, Figure 3.3 should make it clearer. What is pictured here is a breakdown of the region $aCcd$ in Figure 3.2 into subregions together with, for each subregion, the endpoints of the cut-locus segments out of vertex d corresponding to points P in the subregion. For example, output region 2 is points U_+ in Figure 2.1 corresponding to points in input region 2, and output region 6 is points U_0 in a rotated version of Figure 2.1 corresponding to points in input region 6.[5] The entire segment between input regions 5 and 6 maps to output point b. The dashed boundary of output regions 5 and 6 are not in the image. We call the points $Q_{\max}$ in Figure 3.3 because they are the Q farthest from d for a point P.

Figure 3.3. Largest Q for varying P.
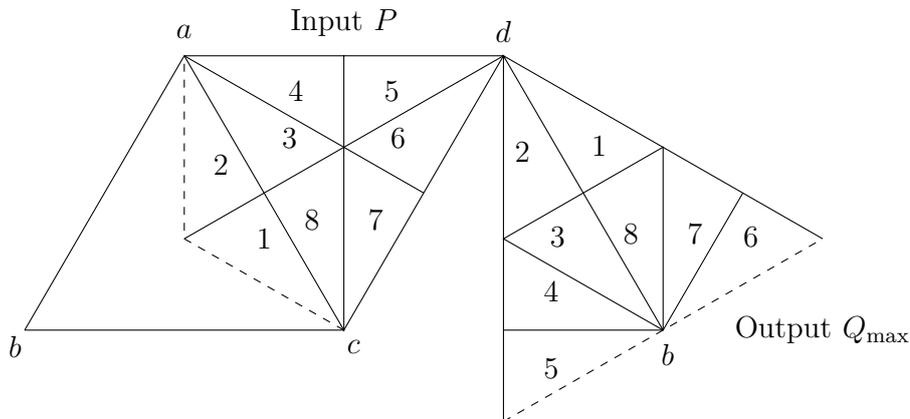
The set E_3 consists of points (P, Q) of two types. Type (1) has P in sets $\mathcal{I}$ defined as the interior of the set of points in a face which are closer to a vertex V than to the other vertices. For example, in Figure 2.1, one such region, with $V = a$, would be the interior of the quadrilateral in the upper

---

[4] Sets are *topologically disjoint* if the closure of each is disjoint from the others; then continuous functions on each combine to a continuous function on their union.

[5] In this case, the segment ad in the right side of Figure 2.1 corresponds to segment db in Figure 3.3.

third of triangle *abc*. The points Q associated to P are the closed interval UL. Type (2) has P all points on segments connecting a vertex V of a face *abc* with its centroid C, including V but not C, and Q in the closed segment connecting the other vertex d with the point L associated with P as in Figure 2.6. Note that this can be considered as a UL segment, too.

For $P \in \mathcal{I}$ and Q in the closed interval UL, we can choose $\phi_3(P, Q)$ to be the appropriate point in U_+L_+ in Figure 2.1. Note that Figure 2.1 applies to any $\mathcal{I}$ by appropriate choice of a, b, and c. For type (2), in Figure 2.6 we would choose as $\phi_3(P, Q)$ the path that goes to the right from a point P on aC to the appropriate Q on the segment dL.

The rest is easy. Let E_4 consist of pairs (P, Q) such that P is a vertex and Q the centroid of the opposite face, or P is a centroid and Q the opposite vertex. Since this is a discrete set, ϕ_4 can be chosen arbitrarily.

Let E_5 be the set of (P, Q) such that P lies in one of six topologically disjoint sets $X(M)$, each of which is the union of lines from the midpoint M of an edge of T to the adjacent vertices and centroids, including M but not the vertices or centroids. A unique point $Q = B$ is associated to each point P. Recall that when $U = L$, we call it B. These are points of multiplicity 4, as in Figures 2.5 and 2.8. See Figure 3.4, which varies continuously with $P \in X(M)$. For $P \in X(M)$, we define $\phi_5(P, B)$ to be the path from P to B_1.

Figure 3.4. Typical set for E_5.
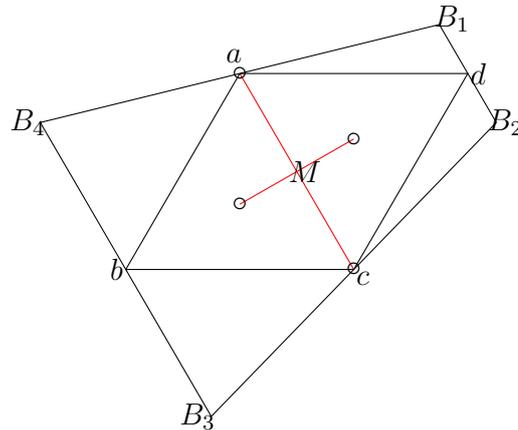
4   Lower bound

In this section we prove $\mathrm{GC}(T) \geq 4$, by a method similar to that used by Recio-Mitter for the cube in [4].

Theorem 4.1. *The space $T \times T$ cannot be partitioned as $E_1 \sqcup E_2 \sqcup E_3$ with a GMPR on each E_i.*

Proof. Let M be the midpoint of ac in Figure 2.1, and P' a point on the segment connecting M and P in that figure. The expanded cut locus for P' is of the same form as that in the figure, and as P' approaches M, $L_\pm$ approaches $U_\pm$, and they and U_0 and L_0 approach the midpoint of bd, which we call B.

Suppose $(M, B) \in E_1$, and $\phi_1(M, B)$ is the path which goes down (toward the midpoint of bd close to L_0 in Figure 2.1). (Going up is handled similarly, reversing the roles of U and L. We will consider later how to handle it when $\phi_1(M, B)$ goes left or right.) We cannot have a sequence of P' as in the figure with $P' \to M$ and $(P', U_{P'}) \in E_1$ because that would imply $\phi_1(P', U_{P'}) \to \phi_1(M, B)$, which is impossible since $\phi(P', U_{P'})$ must go either left, right, or up. There is a sequence of such P'_n all in the same E_i, which we call E_2, and, restricting more, all $\phi_2(P'_n, U_{P'_n})$ going in the same direction, which

we will suppose is left; i.e., toward U_-. We will consider later the minor modifications required if $\phi_2(P'_n, U_{P'_n})$ goes right or up.

For each such P'_n, there is an interval of Q's in the cut locus of P'_n abutting $U_{P'_n}$ along the segment from d to $U_{P'_n}$. (The corresponding points in Figure 2.1 are close to U_0 and U_+.) There cannot be a sequence of these converging to U with $(P'_n, Q) \in E_2$ since $\phi(P'_n, Q)$ must go right or up, but $\phi_2(P'_n, U_{P'_n})$ goes left. If there were, for infinitely many n, a sequence $Q_{n,m}$ approaching $U_{P'_n}$ with $(P_n, Q_{n,m}) \in E_1$, then the sequence $(P_n, Q_{n,n})$ would approach (M, B), but $\phi_1(P_n, Q_{n,n})$ cannot approach $\phi_1(M, B)$, since the possible directions differ. Thus there exist sequences $Q_{n,m} \to U_{P'_n}$ with $(P'_n, Q_{n,m})$ in a new set E_3, and we may assume that $\phi_3(P'_n, Q_{n,m})$ all have the same direction, which we may assume to be "up," i.e., toward the vicinity of U_0.

For each (n, m), there exists a sequence $Q_{n,m,\ell} \to Q_{n,m}$ such that the unique minimal geodesic from P'_n to $Q_{n,m,\ell}$ goes to the right, i.e., in the vicinity of U_+. These points $Q_{n,m,\ell}$ are not in the cut locus of P'_n. For each (n, m), there cannot be infinitely many ℓ with $(P'_n, Q_{n,m,\ell}) \in E_3$, since $\phi_3(P'_n, Q_{n,m})$ and $\phi(P'_n, Q_{n,m,\ell})$ have different directions. We restrict now to, for each (n, m), an infinite sequence of ℓ such that $(P'_n, Q_{n,m,\ell}) \notin E_3$. Taking a diagonal limit on m and ℓ, $(P'_n, Q_{n,m,\ell}) \to (P'_n, U_{P'_n})$; since $\phi_2(P'_n, U_{P'_n})$ and $\phi(P'_n, Q_{n,m,\ell})$ have opposite directions, $(P'_n, Q_{n,m,\ell}) \notin E_2$ for an infinite sequence of m's and all $\ell \geq L_m$ for an increasing sequence of integers L_m. Now taking a diagonal limit over n, m, and ℓ, we approach (M, B). Since the directions of $\phi_1(M, B)$ and $\phi(P'_n, Q_{n,m,\ell})$ differ, there must be an infinite sequence of $(P'_n, Q_{n,m,\ell})$ not in E_1. So it requires a fourth set E_4.

Now we discuss the minor changes for other cases to which we alluded above. If $\phi_2(P'_n, U_{P'_n})$ went right, instead of left, then the Q's will be chosen on the segment from vertex a to $U_{P'_n}$, close to U, with corresponding points in Figure 2.1 close to U_0 and U_-, and the rest of the argument proceeds similarly. If $\phi_2(P'_n, U_{P'_n})$ went up, then the Q's will be chosen on the segment connecting U and L, converging to U, and the argument proceeds as before.

If instead of going down or up, $\phi_1(M, B)$ goes left, then we consider P' on a little segment going sharply down and left from M in Figure 2.1. The expanded cut locus will be similar to that in Figure 2.1, but with U_+L_+ and U_0 interchanged (and moved slightly to the other side of line bdb), and similarly for U_-L_- and L_0. These P' have $\phi(P', U_{P'})$ going up, down, or right, and an argument like the one above works. ∎

References

[1] P. Agarwal, B. Arnov, J. O'Rourke, and C. Schevon, *Star unfolding of a polytope with applications*, SIAM J. Comput. **26** (1997), 1689–1713.

[2] D.M.Davis and M.Guo, *Isomorphism classes of cut loci on a cube*, arXiv:2301.11366.

[3] M.Farber, Topological complexity of motion planning, Discr. Comp. Geom. **29** (2003), 211–221.

[4] D. Recio-Mitter, *Geodesic complexity of motion planning*, J. Appl. Comput. Topol. **5** (2021), 141–178.

Donald M. Davis

Department of Mathematics, Lehigh University, Bethlehem, PA 18015, USA

E-mail address: dmd1@lehigh.edu