Geometric Conditions for the Existence or Non-existence of a Solution to the Perspective 3-Point Problem

Michael Q. Rieck[1*]

[1*]Math and CS Dept., Drake University, 2703 Forest Street, Des Moines, 50311, IA, USA.

Corresponding author(s). E-mail(s): michael.rieck@drake.edu;

Abstract

Direct and fairly simple geometric criteria are proved to be necessary for the Perspective 3-Point (P3P) Problem to have a real solution point. This is so under the assumption that the three control points are at the vertices of an acute triangle. Collectively, these criteria appear to be sufficient as well, based on substantial experimental evidence. Proving the necessity of some of the criteria does not involve the acute triangle assumption, and so these are required for obtuse and right triangles as well. While motivated by the P3P Problem, the results are actually concerned with various constraints among six of the angles that occur in a tetrahedron. Therefore, the results likely have other applications.

Keywords: P3P: perspective: danger cylinder: deltoid: toroid: tetrahedron

1 Introduction and Main Results

Fix an acute triangle $\triangle ABC$ in ordinary three-dimensional real affine space, equipped with the standard metric. Following common practice, $\angle A$, $\angle B$ and $\angle C$ will denote the interior angles of $\triangle ABC$. All angles will be measured in radians. The side lengths of $\triangle ABC$ opposite the vertices A, B and C will be denoted a, b and c, respectively.

Consider any point P in this space that is not coplanar with A, B and C. Together, A, B, C and P form the vertices of a tetrahedron. The angles $\angle BPC$, $\angle CPA$ and $\angle APB$ will be denoted α, β and γ, respectively.

For convenience, we now equip the three-dimensional real space with a new Cartesian coordinate system (xyz) having the following properties:

- The new coordinate system respects angles in the space, but not necessarily lengths;
- The three triangle vertices A, B and C have coordinates $(\cos\phi_1, \sin\phi_1, 0)$, $(\cos\phi_2, \sin\phi_2, 0)$ and $(\cos\phi_3, \sin\phi_3, 0)$, respectively, for angles ϕ_1, ϕ_2 and ϕ_3 satisfying $\phi_1+\phi_2+\phi_3 = 0$.

It is straightforward to obtain such a coordinate system, and this imposes no restrictions on the problem being addressed in this paper. If one begins with the standard coordinate system for the space, then the desired coordinate system can be obtained through a sequence of translations, scalings and rotations, as follows.

Start by translating and rotating so as to put the triangle into the xy-plane. Then find the circumcenter of the triangle, possibly by intersecting two of the perpendicular bisector lines for the sides of the triangle. Translate so that the origin becomes the circumcenter. Compute the circumradius, *i.e.* the distance from the origin to any of the vertices A, B, C. Rescale all of space so that the circumradius becomes one.

The three vertices will now be at positions $(\cos\psi_1, \sin\psi_1, 0)$, $(\cos\psi_2, \sin\psi_2, 0)$ and $(\cos\psi_3, \sin\psi_3, 0)$, for some angles ψ_1, ψ_2, and ψ_3. For $j = 1, 2, 3$, let $\phi_j = \psi_j - (\psi_1+\psi_2+\psi_3)/3$. By rotating by an angle $(\psi_1+\psi_2+\psi_3)/3$ about the z-axis, the vertices will now have positions $(\cos\phi_1, \sin\phi_1, 0)$, $(\cos\phi_2, \sin\phi_2, 0)$ and $(\cos\phi_3, \sin\phi_3, 0)$. Note that $\phi_1+\phi_2+\phi_3 = 0$.

Also, notice that there are generally six possibilities for the vertex coordinates. To see this, notice that adding multiples of 2π to the ψ angles, produces, modulo 2π, three possibilities for the ϕ angles. Modulo 2π, these can be obtained from each other by adding the same amount $\pm 2\pi/3$ to each of the ϕ angles. Additionally, all of the ϕ_j's can be negated, resulting in the triangle being reflected about the x-axis, which yields another acceptable arrangement.

Computing the following quantity D is important, but which of the six triangle positions is used is unimportant, as can be inferred from Lemma 8 (in Section 3). Here are steps that can be used to compute D:

$$
\begin{aligned}
&x_1 = \cos\phi_1\,, x_2 = \cos\phi_2\,, x_3 = \cos\phi_3\,,\\
&y_1 = \sin\phi_1\,, y_2 = \sin\phi_2\,, y_3 = \sin\phi_3\,,\\
&x_H = x_1+x_2+x_3\,, y_H = y_1+y_2+y_3\,,\\
&C_0 = \cos\alpha\cos\beta\cos\gamma,\\
&C_1 = \cos^2\alpha\,, C_2 = \cos^2\beta\,, C_3 = \cos^2\gamma,\\
&S_0 = 1-C_0,\, S_1 = 1-C_1,\, S_2 = 1-C_2,\, S_3 = 1-C_3,\\
&H = S_1+S_2+S_3-2S_0\,,\\
&L = 2\,[\,y_H S_0 + (x_1-1)y_1 S_1 + (x_2-1)y_2 S_2 + (x_3-1)y_3 S_3\,]\,,\\
&R = 2\,[\,(1+x_H)S_0 - (x_1+1)x_1 S_1 - (x_2+1)x_2 S_2 - (x_3+1)x_3 S_3\,]\,,\\
&K = -H-R\,,\\
&D = (K^2+L^2+12HK+9H^2)^2 - 4H(2K+3H)^3\,.
\end{aligned}
\tag{1}
$$

These quantities have geometric significance that is discussed in [1] and [2]. For instance, $(x_H, y_H, 0)$ are the coordinates of the orthocenter of ΔABC.

The problem of interest in this paper is the determination of the limitations that the angles $\angle A$, $\angle B$ and $\angle C$ place on the angles α, β and γ. While this is an easily stated and geometrically compelling problem, the solution turns out to be surprisingly complicated. In the language of a well-known camera-tracking problem, the "Perspective 3-Point (P3P) Problem," we are here asking about criteria for the existence of a real-valued solution point to this problem, using A, B and C as the "control points," and prescribed values for the "viewing angles" α, β and γ.

The P3P problem was introduced quite a long time ago, in [3], and various methods for obtaining solutions are presented in [4]. For specified values for $\angle A$, $\angle B$, $\angle C$, α, β and γ, the number of real solution points can be determined using the methods in [5] and [6]. However, these methods mostly involve complicated algebraic elimination methods that only shed some light on the inherent geometric nature of the problem. Admittedly, the geometric portion of [5] does focus on the same toroid surfaces that are also seen in the present paper and in [2], but the important surface called the "companion surface to the danger cylinder (CSDC)" in these latter papers is not mentioned in [5].

The following claim lists several rather simple conditions on $\angle A$, $\angle B$, $\angle C$, α, β and γ (all assumed to be between 0 and π) that seem to be necessary for the existence of a real solution point to the P3P Problem, that is, for the existence of a point P that yields a tetrahedron $ABCP$ as described above. Moreover, experimental evidence strongly suggests that these requirements, taken together, are also sufficient for the existence of such a point P.

Conjecture 1. *Using the above setup, where triangle ΔABC is acute, the following are necessary conditions on the possible values of $\angle A$, $\angle B$, $\angle C$, α, β and γ, for the existence of a suitable tetrahedron $ABCP$, and together, these conditions are sufficient:*

1. $\alpha < \beta + \gamma$, $\beta < \gamma + \alpha$, $\gamma < \alpha + \beta$, $\alpha + \beta + \gamma < 2\pi$;
2. $\beta + \gamma - \alpha < 2(\angle B + \angle C)$, $\gamma + \alpha - \beta < 2(\angle C + \angle A)$, $\alpha + \beta - \gamma < 2(\angle A + \angle B)$;
3. $\angle A + \beta + \gamma < 2\pi$, $\alpha + \angle B + \gamma < 2\pi$, $\alpha + \beta + \angle C < 2\pi$;

4. $\alpha < \angle A \ \rightarrow\ \cos\angle C \cos\beta + \cos\angle B \cos\gamma > 0$*, and suitable permutations of this;*
5. $\alpha < \angle A \ \rightarrow\ \beta \leq \max\{\angle B, \angle C + \alpha\}$*, and suitable permutations of this;*

6. $(D > 0 \wedge \alpha < \angle A) \ \rightarrow\ (\beta \leq \angle B \vee \gamma \leq \angle C)$*, and suitable permutations of this;*
7. $D > 0 \rightarrow (\alpha < \angle A \vee \alpha > \pi - \angle A \vee \beta < \angle B \vee \beta > \pi - \angle B \vee \gamma < \angle C \vee \gamma > \pi - \angle C)$.

"Suitable permutations" means permuting the symbols "A", "B" and "C", and permuting the symbols "α", "β" and "γ", together, in the same way. The right arrows are logical implications. Of course, an implication $p \rightarrow q$ is logically equivalent to the disjunction $\neg p \vee q$. ("$\wedge$", "$\vee$" and "$\neg$" are used here for the basic logical operations of conjunction, disjunction and negation.)

There is some redundancy among the items listed in the conjecture. In particular, there is a simple connection between the first three items, as follows.

Proposition 2. *Items 1 and 2 of Conjecture 1, taken together, imply Item 3.*

Proof. Assume Items 1 and 2. If $\alpha \leq \angle A$, then $\angle A + \beta + \gamma < \angle A + \alpha + 2(\pi - \angle A) = 2\pi + \alpha - \angle A \leq 2\pi$. If, instead, $\alpha > \angle A$, then $\angle A + \beta + \gamma < \alpha + \beta + \gamma < 2\pi$. In either case, $\angle A + \beta + \gamma < 2\pi$, and by symmetry, Item 3 follows.

□

Although Proposition 2 shows that Item 3 is superfluous, it was included in the list in Conjecture 1, partly because it has an interesting geometric interpretation, as seen in the proof of Theorem 3, in Section 2.

The items in the conjecture fall into three camps, based on the nature of the analysis used to discover them. We will refer to Items 1, 2 and 3 as the "sphere-based rules;" we will refer to Items 4 and 5 as the "toroid-based rules;" and we will refer to Items 6 and 7 as the "deltoid-based rules." The necessity of Items 1, 3, 4 and 5 have already been proved in [7], though this paper is as yet unpublished. The necessity of Items 4 and 5 (the toroid-based rules) must be considered conjectural until that paper is reviewed and published, or some other published, peer-reviewed paper proves these claims. Proofs of the necessity of Items 1, 2, 3, 6 and 7 are provided in the current paper. A proof of the sufficiency claim seems far out of reach, but this claim has held up under extensive testing; see Section 5. Here now are the principal claims that will proved in this paper.

Theorem 3. *The sphere-based rules, i. e. Items 1, 2 and 3 in Conjecture 1, are necessary conditions for the existence of a suitable tetrahedron ABCP.*

Theorem 4. *The deltoid-based rules, i. e. Items 6 and 7 in Conjecture 1, are necessary conditions for the existence of a suitable tetrahedron ABCP.*

A proof of Theorem 3 is provided in Section 2 of the present paper, and involves some interesting geometry. Theorem 4 is much harder to prove. The deltoid-based rules are quite obscure without the benefit of insights from [2]. However, Theorem 4 is proved in Section 4 of the current paper. Prior to this though, much of the reasoning of [2] is reproduced in Sections 3 and 4. One can already get some sense of the phrase "deltoid-based rules" by noting that when the above formula for D is regarded as a homogeneous polynomial in three variables, H, K and L, and set equal to zero, one obtains the homogeneous equation for a standard deltoid curve.[1]

Section 5 describes a couple C++ programs that have been used to test Conjecture 1, and to make a strong case for it. Some compelling visual results from Mathematica are also discussed. Web addresses for the source code are provided there.

2 The Sphere-based Rules

There are actually two very different ways to prove the necessity of the sphere-based rules, *i.e.* Theorem 3. There is what might be called an "algebraic proof" and a "geometric proof." Only the latter will be presented rigorously in this paper.[2] Neither of the two proofs makes use of the assumption that the triangle ΔABC is acute, and so the sphere-based rules are actually necessary conditions for arbitrary triangles ΔABC. Here now is the geometric proof.

Proof of Theorem 3. Begin with an arbitrary tetrahedron $ABCP$. Consider a small sphere centered at P such that the other vertices are outside this sphere. For the purposes of this proof only, rescale three-dimensional real space so as to make the sphere a unit sphere $\mathcal{S}$. The plane through P that is parallel to the plane containing ΔABC intersects $\mathcal{S}$ in a great circle $\mathcal{E}$. The

[1] This form of the equation is essentially due to Bo Wang.
[2] The former is available from the author upon request.

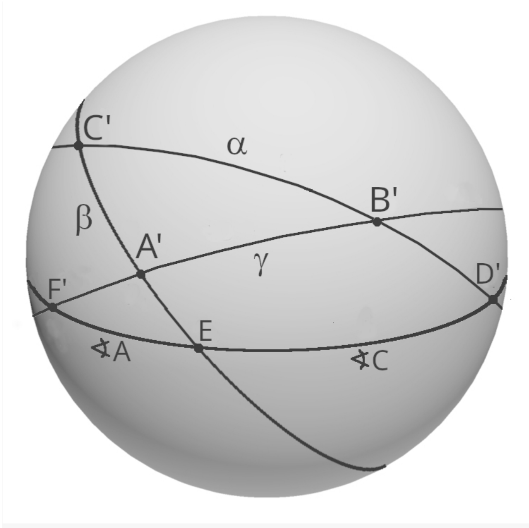

Fig. 1: Setup for the sphere-based rules

three rays from P that pass through A, B or C, each intersect $\mathcal{S}$ in a unique point. Accordingly, call these points A', B' and C'. As points on $\mathcal{S}$, they are all on the same side of $\mathcal{E}$ (a hemisphere).

Consider the great circle $\mathcal{C}_1$ that passes through B' and C', the great circle $\mathcal{C}_2$ that passes through C' and A', and the great circle $\mathcal{C}_3$ that passes through A' and B'. Notice that the great circle distance (along $\mathcal{C}_1$) between B' and C' is α, the great circle distance (along $\mathcal{C}_2$) between C' and A' is β, and the great circle distance (along $\mathcal{C}_3$) between A' and B' is γ. Also, $\mathcal{C}_1$ is on the plane containing B, C and P. Similarly for $\mathcal{C}_2$ and $\mathcal{C}_3$.

$\mathcal{C}_1$ intersects $\mathcal{E}$ in two antipodal points. Following $\mathcal{C}_1$ from B' to C', and continuing until we reach $\mathcal{E}$, call this intersection point D. Similarly, following $\mathcal{C}_1$ from C' to B', and continuing until we reach $\mathcal{E}$, call this intersection point D'. So again, D and D' are antipodal points on the sphere. Also, the directed line segment from D' to D is parallel to the directed line segment from B to C. (Both are in the plane PBC, and the plane containing $\mathcal{E}$ is parallel to the plane ABC.)

Similarly define points E and E', using $\mathcal{C}_2$, and F and F', using $\mathcal{C}_3$. It is then straightforward to check that great-circle distances between D and E', between E' and F, between F and D', between D' and E, between E and F', and between F' and D are respectively, $\angle C$, $\angle A$, $\angle B$, $\angle C$, $\angle A$ and $\angle B$. Since A', B' and C' lie in the same hemisphere, they form the vertices of a proper spherical triangle, whose side lengths are α, β and γ. Item 1 in the theorem follows immediately from this fact. See Figure 1.

Now, C'. D' and E are also the vertices of a proper spherical triangle. Moreover, the side connecting C' and D' contains B', and side connecting C' and E contains A'. Let $\alpha+\delta$ be the length of the side connecting C' and D', and let $\beta+\varepsilon$ be the length of the side connecting C' and E ($\delta, \varepsilon > 0$). The side connecting D' and E has length $\angle C$. So, $\alpha+\beta+\angle C < (\alpha+\delta)+(\beta+\varepsilon)+\angle C < 2\pi$. This proves the third inequality in Item 3, and the other two inequalities are similarly obtained. By considering two paths on the sphere connecting D' and E, we easily

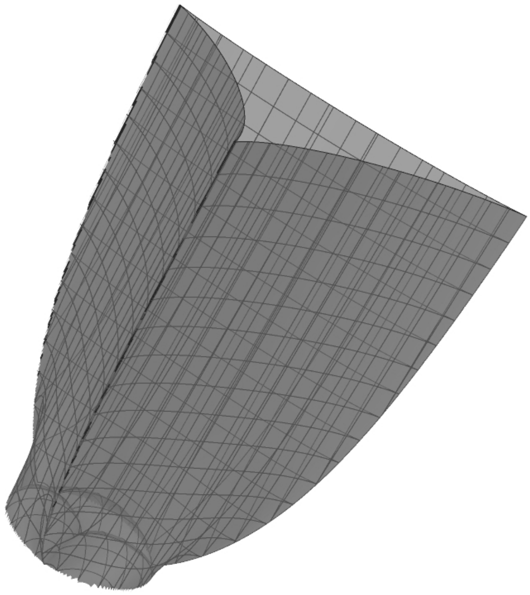

Fig. 2: Companion surface to the danger cylinder

obtain $\angle C < \gamma+\delta+\varepsilon$. Therefore, $\alpha+\beta-\gamma < \alpha+\beta-(\angle C-\delta-\varepsilon) = (\alpha+\delta)+(\beta+\varepsilon)-\angle C < 2\pi-2\angle C = 2(\angle A+\angle B)$. This and similar reasoning yields Item 2 in the conjecture.

□

Note that, in the proof, it is possible for $\angle C$ to be less than γ. Figure 1 shows an example where $\angle A < \alpha$.

3 The Danger Cylinder and its Companion Surface

Going forward, we will always assume that the triangle ΔABC is acute. It has long been understood that a P3P solution point P in space, with coordinates (x,y,z), is a repeated solution point if and only if it is on the circular cylinder that includes the circumcircle of the control points triangle. This cylinder is traditionally refered to as the "danger cylinder." Using our setup, it is the cylinder whose equation is $x^2+y^2=1$.

It is beneficial to understand the answer to a more subtle geometric question, as follows. For fixed control points, a generic point P in space, with coordinates (x,y,z), is a solution point to the P3P Problem for unique parameter values, α, β and γ (namely, $\alpha = \angle BPC$, $\beta = \angle CPA$ and $\gamma = \angle APB$). In this way, we might say that P determines an "instance" of the P3P Problem.

Assuming that P is not on the danger cylinder, we now ask for a description of the surface that P must lie on in order that its instance of the P3P Problem have a repeated solution point (which will not be P). This surface was first described in [1], and examined further in [2], where it was given the name, "the companion surface to the danger cylinder (CSDC)." Along with DC and three double toroids, this surface plays an important role in partitioning space into various regions. Points in the same region have instances of the P3P Problem that have the same number of solution points.

Moreover, the region containing a point P also determines which regions contain the other solution points for P's instance of the P3P Problem. [2] deals with this in some detail, at least for the case where the control points triangle is acute. See Theorem 4 there. This is a rather complicated, but interesting, story, the essence of which will be largely reconstructed in this and the next sections of the current paper.

If one ignores the toroids, and only uses $DC \cup CSDC$ to divide space into regions, then points in the same region have instances of the P3P Problem that have the same number of (real) "weak solution points." By this is meant points that either are themselves solution points or are "weakly related" to solution points. (See below.)

Figure 2 shows what $CSDC$ looks like for a typical acute control points triangle ΔABC, in the upper half space ($z > 0$). The lower half space ($z < 0$) portion is just the reflection of this about the xy-plane. In order to describe $CSDC$, a few results from [2] and [8] are now reproduced here, along with shorter proofs. For better motivated discussions of some of this material, consult these two documents.

Following the practice used there, the xy-plane will here be identified with the complex plane by setting $\zeta = x + iy$. The control points / triangle vertices (A, B, C) are thus points $\zeta_j = x_j + iy_j$ in the complex plane with $|\zeta_j| = 1$ $(j = 1,2,3)$. Additionally, since $\phi_1 + \phi_2 + \phi_3 = 0$, we have $\zeta_1\zeta_2\zeta_3 = 1$.

Let $Z = z^2$. Let $c_1 = \cos\alpha$, $c_2 = \cos\beta$ and $c_3 = \cos\gamma$. Let d_1, d_2 and d_3 denote the distances between control points, specifically, $d_1 = |\zeta_3 - \zeta_2|$, $d_2 = |\zeta_1 - \zeta_3|$, $d_3 = |\zeta_2 - \zeta_1|$. It will also be useful to define $T_j = S_j - 2d_j^2 S_0 / (d_1^2 + d_2^2 + d_3^2)$ $(j = 1,2,3)$, where S_1, S_2, S_3 and S_0 were defined in Section 1. Three other quantities used in [2] and [8] will be needed here too:

$$\begin{aligned} H &= \eta^2 = T_1 + T_2 + T_3 = S_1 + S_2 + S_2 - 2S_0 \\ &= 1 - c_1^2 - c_2^2 - c_3^2 + 2c_1c_2c_3, \\ \zeta_H &= x_H + iy_H = (x_1 + x_2 + x_3) + i(y_1 + y_2 + y_3), \\ \zeta_L &= \{(\zeta_1^2 + 2\overline{\zeta_1})T_1 + (\zeta_2^2 + 2\overline{\zeta_2})T_2 + (\zeta_3^2 + 2\overline{\zeta_3})T_3\} / H. \end{aligned} \tag{2}$$

A cubic polynomial that has quite a useful role, expressed in terms of an indeterminate τ, is the following:

$$p(\tau) = \tau^3 - \zeta_L\tau^2 + \overline{\zeta_L}\tau - 1. \tag{3}$$

Lemma 5. *The discriminant of the polynomial $p(\tau)$ is*

$$\mathcal{D} \stackrel{\text{def}}{=\!=} \zeta_L^2\overline{\zeta_L}^2 - 4(\zeta_L^3 + \overline{\zeta_L}^3) + 18\zeta_L\overline{\zeta_L} - 27.$$

Moreover, this quantity vanishes if and only if the P3P problem has a repeated solution, for the specified parameters.

Proof. $\mathcal{D}$ is easily verified to be the discriminant of $p(\tau)$, using the formula for the discriminant of a general cubic polynomial. In his solution to the P3P Problem, S. Finsterwalder introduced a certain cubic polynomial; see Equation (14) of [4]. It is known that Grunert's system of equations has a repeated solution if and only if Finsterwalder's cubic polynomial has a repeated root. This, of course, is so if and only if the discriminant of this polynomial is zero.

It will now be shown that Finsterwalder's cubic polynomial can be transformed, via a Möbius transformation (with constant coefficients), into the polynomial $p(\tau)$ times a nonzero constant, and vice-versa. This will establish that the discriminant of Finsterwalder's cubic polynomial is a nonzero constant multiple of the discriminant of $p(\tau)$.

The Möbius transformation is simply this:

$$\lambda = \frac{\zeta_1(\zeta_3-\zeta_2)}{\zeta_3(\zeta_2-\zeta_1)} \cdot \frac{\zeta_3\tau-1}{\zeta_1\tau-1}, \tag{4}$$

and hence

$$\tau = \frac{1}{\zeta_1\zeta_3} \cdot \frac{\zeta_3(\zeta_1-\zeta_2)\lambda-\zeta_1(\zeta_2-\zeta_3)}{(\zeta_1-\zeta_2)\lambda-(\zeta_2-\zeta_3)}.$$

Now, Finsterwalder's cubic polynomial is

$$\mathcal{G}\lambda^3+\mathcal{H}\lambda^2+\mathcal{I}\lambda+\mathcal{J}$$

with $\mathcal{G}=d_3^2(d_3^2T_2-d_2^2T_3)$, $\mathcal{H}=d_2^2(d_2^2-d_1^2)T_3+d_3^2(d_3^2+2d_1^2)T_2$, $\mathcal{I}=d_2^2(d_2^2-d_3^2)T_1+d_1^2(d_1^2+2d_3^2)T_2$ and $\mathcal{J}=d_1^2(d_1^2T_2-d_2^2T_1)$. Upon substituting the formula for λ in terms of τ, and making other evident substitutions such as $d_1^2=(\zeta_3-\zeta_2)(\overline{\zeta_3}-\overline{\zeta_2})$, etc., $\overline{\zeta_j}=1/\zeta_j$ $(j=1,2,3)$ and $\zeta_1=1/(\zeta_2\zeta_3)$, it is straightforward (though a bit tedious) to transform Finsterwalder's polynomial to $p(\tau)$ times a nonzero constant. □

Further reasoning and motivation for the Möbius transformation is provided in [8].

Lemma 6. [3] *For a given point P in space, with coordinates (x,y,z), consider the associated instance of the P3P problem (for which P is a solution point). Then,*

$$\zeta_L = \zeta^2-2\overline{\zeta}+(\zeta\overline{\zeta}-1)(\zeta^2-\zeta_H\zeta-\overline{\zeta}+\overline{\zeta_H})/Z. \tag{5}$$

Proof. Let ζ_L' denote the right side of the equation in the lemma, and so, we need to prove that $\zeta_L'=\zeta_L$. Let $p(\tau)$ be as before, and let $q(\tau)=\tau^3-\zeta_L'\tau^2+\overline{\zeta_L'}\tau-1$. It suffices to show that the two cubic polynomials $p(\tau)$ and $q(\tau)$ are equal. To accomplish this, it will be shown that $p(\overline{\zeta_j})=q(\overline{\zeta_j})$ $(j=1,2,3)$. Clearly $p(0)=q(0)=-1$. Then, the fact that the two cubic polynomials agree for four different values of the argument τ will imply that these polynomials are actually the same. The equation $(d_3^2T_2-d_2^2T_3)/H=\zeta_1 p(\overline{\zeta_1})/(\zeta_2-\zeta_3)$ can be checked directly by expanding each side to show that they have a common expression in terms of ζ_1, ζ_2, ζ_3, T_1, T_2 and T_3.

Because its circumradius is one, the square of the area of the control points triangle ΔABC is $d_1^2d_2^2d_3^2/16$, and so the square of the volume of the tetrahedron having this triangle as a

[3] This is Theorem 1 in [2].

face, and the point P as its opposite vertex, is $d_1^2 d_2^2 d_3^2 Z/144$. The parallelepiped having the segments $\overline{PA}$, $\overline{PB}$ and $\overline{PC}$ as edges thus has squared volume $d_1^2 d_2^2 d_3^2 Z/4$. However, this must also equal the square of the Gramian determinant associated the vectors $\overrightarrow{PA}$, $\overrightarrow{PB}$ and $\overrightarrow{PC}$, which equals $H r_1^2 r_2^2 r_3^2$, where $r_1^2 = (\zeta - \zeta_1)(\overline{\zeta} - \overline{\zeta_1}) + Z$ is the squared distance between P and A, etc. Therefore, $H = d_1^2 d_2^2 d_3^2 Z/(4 r_1^2 r_2^2 r_3^2)$.

The Law of Cosines implies that $S_1 = 1 - c_1^2 = 1 - (r_2^2 + r_3^2 - d_1^2)^2/(4r_2^2 r_3^2) = (2r_2^2 r_3^2 + 2d_1^2 r_2^2 + 2d_1^2 r_3^2 - d_1^4 - r_2^4 - r_3^4)/(4r_2^2 r_3^2)$, and similarly for S_2 and S_3. Now, $(d_3^2 T_2 - d_2^2 T_3)/H = (d_3^2 S_2 - d_2^2 S_3)/H = 4r_1^2 r_2^2 r_3^2 (d_3^2 S_2 - d_2^2 S_3)/(d_1^2 d_2^2 d_3^2 Z)$, which can now be expanded and shown to equal $\zeta_1 q(\overline{\zeta_1})/(\zeta_2 - \zeta_3)$. It follows that $p(\overline{\zeta_1}) = q(\overline{\zeta_1})$. By symmetry, $p(\overline{\zeta_2}) = q(\overline{\zeta_2})$ and $p(\overline{\zeta_3}) = q(\overline{\zeta_3})$. □

Lemma 7. [4]

When the right side of (5) is substituted for ζ_L, and likewise for $\overline{\zeta_L}$, into the formula for the discriminant $\mathcal{D}$ of $p(\tau)$, the result is $Z^{-4}\,(\zeta\overline{\zeta} - 1)^2\,\mathcal{P}(\zeta, \overline{\zeta}, Z)$, where $\mathcal{P}(\zeta, \overline{\zeta}, Z)$ is a polynomial in ζ, $\overline{\zeta}$ and Z. As a polynomial in Z (with coefficients being polynomials in ζ and $\overline{\zeta}$), it has degree four.

Proof. Let $N = Z^4\,\mathcal{D} =$

$$Z_L^2\,\overline{Z_L}^2 - 4Z\,(Z_L^3 + \overline{Z_L}^3) + 18Z^2 Z_L \overline{Z_L} - 27Z^4,$$

where $Z_L = Z\zeta_L$, $\overline{Z_L} = Z\overline{\zeta_L}$, and Z is regarded as a real variable. Thus, Z_L, $\overline{Z_L}$ and N can be expressed as polynomials in ζ, $\overline{\zeta}$ and Z. As a polynomial in Z, N has degree 4.

If we momentarily set $\overline{\zeta} = 1/\zeta$, we find that $Z_L = Z(\zeta^3 - 2)/\zeta$, and $\overline{Z_L} = Z(1 - 2\zeta^3)/\zeta^2$, resulting in $N = 0$. Therefore, $\zeta\overline{\zeta} - 1$ is a factor of N (no longer using the assumption that $\overline{\zeta} = 1/\zeta$).

Now, following common practice in complex analysis, treat Z_L and $\overline{Z_L}$ as functions in independent variable ζ, $\overline{\zeta}$ and Z. (A new variable could be substituted for $\overline{\zeta}$ here if desired.) Likewise, treat N as a function in independent variables Z_L, $\overline{Z_L}$ and Z. Direct computations show that when $\overline{\zeta} = 1/\zeta$, we get

$$\left.\frac{\partial Z_L}{\partial \zeta}\right|_{\overline{\zeta}=1/\zeta} = \frac{(1+2Z)\zeta^3 - \zeta_H \zeta^2 + \overline{\zeta_H}\zeta - 1}{\zeta^2},$$

$$\left.\frac{\partial \overline{Z_L}}{\partial \zeta}\right|_{\overline{\zeta}=1/\zeta} = \frac{-(1+2Z)\zeta^3 + \zeta_H \zeta^2 - \overline{\zeta_H}\zeta + 1}{\zeta^3},$$

$$\left.\frac{\partial N}{\partial Z_L}\right|_{\overline{\zeta}=1/\zeta} = \frac{-4Z^3(1+\zeta^3)^3}{\zeta^5},$$

and

$$\left.\frac{\partial N}{\partial \overline{Z_L}}\right|_{\overline{\zeta}=1/\zeta} = \frac{-4Z^3(1+\zeta^3)^3}{\zeta^4}.$$

[4] This is part of Lemma 4 in [2]

If we now regard N as a function of ζ, $\overline{\zeta}$ and Z, we obtain, by the chain rule,

$$\left.\frac{\partial N}{\partial \zeta}\right|_{\overline{\zeta}=1/\zeta} = 0.$$

By symmetry,

$$\left.\frac{\partial N}{\partial \overline{\zeta}}\right|_{\overline{\zeta}=1/\zeta} = 0.$$

It follows that $(\zeta\overline{\zeta}-1)^2$ is a factor of N.

□

Lemma 8.

$$\mathcal{D} \;=\; \frac{N}{Z^4} \;=\; \frac{D}{H^4},$$

where N is as in the proof of Lemma 7, and D is as in (1), in Section 1. Assuming that $Z \neq 0$ (i.e. P is not in the xy-plane), the signs of $\mathcal{D}$, N and D are equal (-1, 0 or +1). Moreover, $\mathcal{D}$ is invariant under multplying each of ζ_1, ζ_2, ζ_3 and ζ by $e^{2\pi i/3}$, or multiplying each of them by $e^{-2\pi i/3}$. It is also invariant under conjugating each of ζ_1, ζ_2, ζ_3 and ζ.

Proof. The first equality is just from the definition of N. Now, since $Z \neq 0$, it follows that $H \neq 0$. This is so, since, as shown in the proof of Lemma 6, $H = d_1^2 d_2^2 d_3^2 Z \,/\, (4 r_1^2 r_2^2 r_3^2)$, and d_1, d_2, d_3, r_1, r_2 and r_3 are nonzero. It remains to show that $D = H^4 \mathcal{D}$.

From (2) we see that $H\,\zeta_L \;=\; (\zeta_1^2 + 2\overline{\zeta_1})T_1 + (\zeta_2^2 + 2\overline{\zeta_2})T_2 + (\zeta_3^2 + 2\overline{\zeta_3})T_3$. It needs to be shown that $D = \;(H\zeta_L)^2\,(H\overline{\zeta_L})^2 - 4H\,[\,(H\zeta_L)^3 + (H\overline{\zeta_L})^3\,] + 18H^2\,(H\zeta_L)\,(H\overline{\zeta_L}) - 27H^4$. Using the formulas in (1), direct computation shows that $L \,=\, 2[\,(x_1-1)y_1T_1 + (x_2-1)y_2T_2 + (x_3-1)y_3T_3\,]$, $R \,=\, -2[\,(x_1+1)x_1T_1 + (x_2+1)x_2T_2 + (x_3+1)x_3T_3\,]$, and $H\,\zeta_L \;=\; (\zeta_1^2 + 2\overline{\zeta_1})T_1 + (\zeta_2^2 + 2\overline{\zeta_2})T_2 + (\zeta_3^2 + 2\overline{\zeta_3})T_3 \;=\; K + iL$. Using the fact that T_1, T_2. T_3, K and L are real, the claim concerning D can now be checked by expanding the formulas. This amounts to establishing that $(K+iL)^2(K-iL)^2 - 4H\,[\,(K+iL)^3 + (K-iL)^3\,] + 18H^2(K+iL)(K-iL) - 27H^4 \;=\; (K^2 + L^2 + 12HK + 9H^2)^2 - 4H(2K+3H)^3$, which can be checked by expanding.

Finally, the invariance of $\mathcal{D}$ under the indicated changes to ζ_1, ζ_2, ζ_3, and ζ is actually a corollary to Lemma 6. Observe first how ζ_L and $\overline{\zeta_L}$ are affected by the substitutions, and then use the definition of $\mathcal{D}$ in terms of these.

□

CSDC is simply defined to be the surface in $(xyz-)$ space consisting of the points for which $\mathcal{P}(\zeta,\overline{\zeta},Z) = 0$. Together with the danger cylinder $(\zeta\overline{\zeta} = 1)$, we obtain all of the points for which the discriminant $\mathcal{D}$ of $p(\tau)$ vanishes. These are the points whose instance of the P3P Problem has a repeated solution point.

Lemma 9. *CSDC is unbounded, and as $Z \to \infty$, the orthogonal projection of CSDC onto the xy-plane approaches the standard deltoid curve.*

Proof. By Lemma 6, as $Z \to \infty$, we see that $\zeta_L \to \zeta^2 - 2\overline{\zeta}$, and so $\mathcal{D} \to$

$$(\zeta\overline{\zeta} - 1)^2 \, [\zeta^2\overline{\zeta}^2 - 4(\zeta^3 + \overline{\zeta}^3) + 18\zeta\overline{\zeta} - 27].$$

Thus, in the limit, $\mathcal{D}$ vanishes when ζ is on the unit circle, and also when ζ is on the standard deltoid curve (given by $\zeta^2\overline{\zeta}^2 - 4\zeta^3 - 4\overline{\zeta}^3 + 18\zeta\overline{\zeta} = 27$), and nowhere else. The unit circle is expained by the fact that $\mathcal{D}$ vanishes on the danger cylinder. The standard deltoid curve must likewise by explained by the vanishing of $\mathcal{D}$ on *CSDC*, from which the claim in the lemma now follows.

□

Lemma 10. *CSDC consists of two surfaces $CSDC_0$ and $CSDC_1$, each symmetric about the xy-plane. Their intersection is the circumcircle of the triangle ΔABC. $CSDC_0$ is unbounded, and except for the circumcircle and three vertical lines on DC, $CSDC_0$ lies outside DC. Also, except for the circumcircle, $CSDC_1$ is inside the unit sphere, and hence is bounded and inside DC.*

Proof. *CSDC* is an algebraic surface, given by the polynomial equation $\mathcal{P}(\zeta,\overline{\zeta},Z) = 0$. While $\zeta\overline{\zeta} - 1$ is not a factor of this polynomial, it is a factor (in fact, a double factor) of the polynomial that results from setting Z equal to zero. Thus, the circumcircle is part of *CSDC*. In fact, the intersection of *CSDC* and *DC* consists of this circle and three vertical lines, namely, the lines where $\zeta = -1$ or $-e^{\pm 2\pi i/3}$. This can all be seen using the formula for $\mathcal{P}(\zeta,\overline{\zeta},Z)$ in Lemma 16 of [2].

By Theorem 3 of [2], the portion of *CSDC* that is inside *DC* is also inside the unit sphere, and hence is bounded. This is $CSDC_1$. By Lemma 9 (here), we infer that $CSDC_0$ is unbounded, and that for $Z \gg 0$, it asymptotically approaches the standard deltoid curve.

□

When two points in space have the same values for α, β and γ, and hence solve the same instance of the P3P Problem, we will say that they are "strongly related." When two points are either strongly related, or else have the same value for one of α, β or γ, and supplementary values for the other two, we will say that they are "weakly related." Notice that two points are weakly related if and only if they have the same values for c_1^2, c_2^2, c_3^2 and $c_1c_2c_3$. "Strongly related" and "weakly related" are important equivalence relations. The significance of "strongly related" for the P3P Problem is self-evident. The importance of "weakly related" for the P3P Problem is made apparent in Lemma 11 (the next claim) and Lemma 12 (in the next section).

Lemma 11. *Let P be a point in the upper half-space ($z > 0$).*

1. *The reflection of P about the xy-plane, in the lower half-space, is strongly related to P.*
2. *If P satisfies $D > 0$, then there exists exactly one other point in the upper-half space that is weakly related to it.*

3. *If instead, $D < 0$, then there are exactly three other points in the upper-half space that are weakly related to P (and to each other).*
4. *If instead, P is on DC, then, generally, there are exactly two other points in the upper half-space that are weakly related to it (and to each other), and these are on $CSDC$. This is so, except when P is on the circumcircle of ΔABC or on one of the three special vertical lines mentioned in Lemma 10, i.e., when P is on $DC \cap CSDC$.*

Proof. Item 1 is immediately clear from the symmetry of the P3P Problem about the xy-plane. Items 2 and 3 are just a restatement of Theorem 2 in [2], which is proved there and will simply be assumed here. Remember that D and $\mathcal{D}$ have the same sign, by the Lemma 8 (here).

For Item 4, recall that a generic point on DC serves as a double point for its instance of the P3P system of equations. This means that infinitesimally small perturbations of the P3P parameters α, β and γ that causes $\mathcal{D}$ to be (infinitesimally) negative, would result in a P3P system with two distinct (real) solution points that are infinitesimally close to the original point. These two points are strongly related, and must be weakly related to two other points in the upper half-space. These latter two points must be infinitesimally close to two points on $CSDC$ since $\mathcal{D}$ is infinitesimally small. Therefore, by continuity, the original, unperturbed system must have two solution points on $CSDC$. However, if P is on both DC and $CSDC$, then there will be fewer than two other (distinct) points in the upper half-space that are weakly related to P.

□

The last two lemmas in this section will be used in the next section to prove Theorem 4.

4 The Deltoid-based Rules

The plan for proving Theorem 4 is to show that certain possible regions of space, designated below as $110^+, 101^+, 011^+$ and 111^+, do not actually exist. Lemmas 19 and 20 establish this claim, after which the proof of Theorem 4 quickly follows. To initiate this undertaking, it is necessary to relay some more of the terminology, notation and results in [2]. While A, B and C will remain fixed, the point P will move around in space continuously, and as such, will be called a "particle," instead of a point. P will thus trace out a continuous curve. Its movement can be regarded mathematically as a continuous map from an interval into 3-space, thereby supplying mathematical rigor. However, to gain better intuition of the analysis, thinking about P moving dynamically is quite helpful. Fixed points can be regarded as stationary particles.

As a particle P moves, its values of α, β, γ, c_1, c_2 and c_3 will change continuously, usually. However, it will sometimes be necessary to allow P (or another particle) to pass through a control point (A, B or C). At the moment when this occurs, two of the angles (α, β, γ) and two of the cosines (c_1, c_2, c_3) will not be defined. As P pass through the control point, the signs of these two cosines will instantaneously change, because the corresponding angles will instantaneously be replaced with their supplementary angles.

Certain toroids are important here. The toroid $\mathcal{T}_A$ consists of all the points in space for which $\alpha = A$. This is the surface obtained by taking the open arc on the circumcircle of ΔABC that connects B and C, and that passing through A, and rotating it about the line through B and C. Note that we are deliberately excluding the points B and C from $\mathcal{T}_A$. Using the coordinate

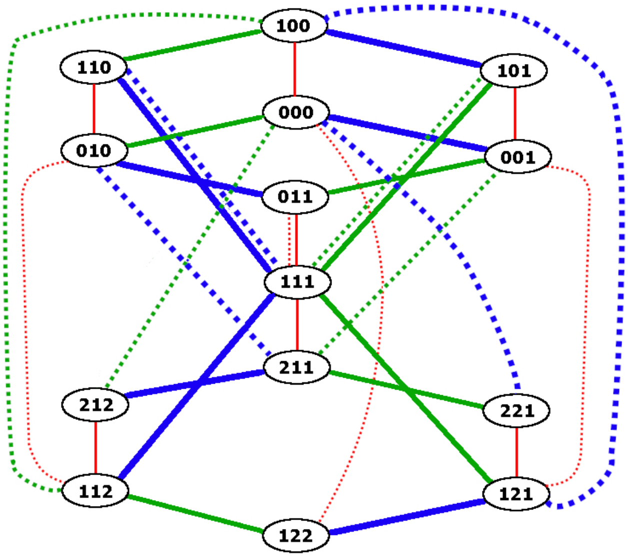

Fig. 3: Transition graph for toroidal regions

system from Section 1, this circumcircle is simply the unit circle in the xy-plane. Apart from the arc that we started with, the toroid $\mathcal{T}_A$ stays outside the unit sphere. Let $\overline{\mathcal{T}}_A = \mathcal{T}_A \cup \{B, C\}$, which will also be called a toroid.

If we instead start with the open arc on the circumcircle connecting B and C that does not pass through A, and again rotate about the line through B and C, then we obtain a different, but related, toroid, which will be denoted $\mathcal{T}_{\pi-A}$. This is the surface on which $\alpha = \pi - A$. Apart from the starting arc, this toroid stays inside the unit sphere and hence inside $\overline{\mathcal{T}}_A$. Let $\overline{\mathcal{T}}_{\pi-A} = \mathcal{T}_{\pi-A} \cup \{B, C\}$, which will also be called a toroid. The toroids $\mathcal{T}_B, \mathcal{T}_{\pi-B}, \mathcal{T}_C, \mathcal{T}_{\pi-C}, \overline{\mathcal{T}}_B, \overline{\mathcal{T}}_{\pi-B}, \overline{\mathcal{T}}_C$ and $\overline{\mathcal{T}}_{\pi-C}$ are similarly defined. All of these toroids will be refered to as "basic toroids."

Consider three-dimensional real space with these toroids removed: $\overline{\mathcal{T}}_A$, $\overline{\mathcal{T}}_{\pi-A}$, $\overline{\mathcal{T}}_B$, $\overline{\mathcal{T}}_{\pi-B}$, $\overline{\mathcal{T}}_C$ and $\overline{\mathcal{T}}_{\pi-C}$. The connected components of the resulting space will be called "toroidal regions." Each is identified by means of a three-digit code, where the first digit relates to the toroids $\overline{\mathcal{T}}_A$ and $\overline{\mathcal{T}}_{\pi-A}$, the second digit relates to the toroids $\overline{\mathcal{T}}_B$ and $\overline{\mathcal{T}}_{\pi-B}$, and the third digit relates to the toroids $\overline{\mathcal{T}}_C$ and $\overline{\mathcal{T}}_{\pi-C}$. If the first digit is 0, then the region is outside $\overline{\mathcal{T}}_A$; if it is 1, then the region is between $\overline{\mathcal{T}}_A$ and $\overline{\mathcal{T}}_{\pi-A}$; and if it is 2, then the region is inside $\overline{\mathcal{T}}_{\pi-A}$. Similarly for the second and third digits. For example, 001 refers to the region outside $\overline{\mathcal{T}}_A$ and $\overline{\mathcal{T}}_B$, but inside $\overline{\mathcal{T}}_C$.

Notice that the identification code for a toroidal region cannot involve a 0 and a 2, since that would suggest that there are points that are both outside and inside the unit sphere. Also, 222 is not a valid identification code, since no point can be inside $\overline{\mathcal{T}}_{\pi-A}$, inside $\overline{\mathcal{T}}_{\pi-B}$, and inside $\overline{\mathcal{T}}_{\pi-C}$. (These three toroids intersect in a single point, namely, the orthocenter of ΔABC.) Notice too, that when the first digit of an identification code is 0, this means that

$\alpha < \angle A$; when it is 1, this means that $\angle A < \alpha < \pi - \angle A$; when it is 2, this means that $\alpha > \pi - \angle A$. Similarly for the other two digits.

Figure 3 is reproduced from [2]. It shows a graph that indicates how it is possible for a particle to continuously transition from one toroidal region to another, either by passing through one of the basic toroids (solid edges in the graph) or by passing through one of the control points (dotted edges in the graph). The thickness/color of a solid edge indicates which toroid the particle passes through: thin/red for $\mathcal{T}_A$ or $\mathcal{T}_{\pi-A}$, medium/green for $\mathcal{T}_B$ or $\mathcal{T}_{\pi-B}$, and thick/blue for $\mathcal{T}_C$ or $\mathcal{T}_{\pi-C}$. Similarly the thickness/color of a dotted edge indicates which control point (A, B or C) the particle passes through.

For example, the solid edge connecting 121 and 122 represents moving across $\mathcal{T}_{\pi-C}$, while staying inside $\mathcal{T}_{\pi-B}$, and between $\mathcal{T}_{\pi-A}$ and $\mathcal{T}_A$. The dotted edge connecting 100 and 121 represents passing through vertex C and thereby changing toroidal regions as follows. If the particle begins between $\mathcal{T}_{\pi-A}$ and $\mathcal{T}_A$, but outside $\mathcal{T}_B$ and $\mathcal{T}_C$, it will stay between $\mathcal{T}_{\pi-A}$ and $\mathcal{T}_A$, but it will move inside $\mathcal{T}_{\pi-B}$, and be between $\mathcal{T}_{\pi-C}$ and $\mathcal{T}_C$. This can be visualized by considering one of the three double toroid ($\mathcal{T}_{\pi-A} \cup \mathcal{T}_A$, $\mathcal{T}_{\pi-B} \cup \mathcal{T}_B$, $\mathcal{T}_{\pi-C} \cup \mathcal{T}_C$) at a time. When zooming in on C, each of $\mathcal{T}_{\pi-A} \cup \mathcal{T}_A$ and $\mathcal{T}_{\pi-B} \cup \mathcal{T}_B$ can be approximated by a double cone with apex C, while $\mathcal{T}_C$ can be approximated by a plane (the tangent plane for $\mathcal{T}_C$ at C), and $\mathcal{T}_{\pi-C}$ simply vanishes, since C is outside $\overline{\mathcal{T}_{\pi-C}}$.

It will be helpful to subdivide the toroidal regions into smaller regions by cutting them with the surface $DC \cup CSDC$ (where $D = 0$). Let 000^+ and 000^- denote the portions of toroidal region 000 where $D > 0$ and where $D < 0$, respectively. And so forth. The goal now is to show that regions 011^+, 101^+, 110^+ and 111^+ do not exist. This will establish that Rules 6 and 7 in Conjecture 1 are necessary conditions, that is, it will prove Theorem 4. A series of lemmas will ultimately lead to this objective.

As we consider moving a particle P around in space, we now also consider all of the other possible particles that would stay weakly related to P. The next result gives a sense of how these will behave.

Lemma 12. *When a particle P passes through $\mathcal{T}_A$ or $\mathcal{T}_{\pi-A}$, without simultaneously passing through one of the other basic toroids, exactly one of the other particles weakly related to P in the upper half-space, and its reflection in the lower half-space (also weakly related to P), will pass through the vertex A. As they do, their values of c_2 and c_3 will instantaneously be negated because their values of the β and γ angles will instantaneously be replaced by their supplements. The particles will remain weakly related to P (and each other) after one of them passes through A. Of course, what has been said here concerning $\mathcal{T}_A$ and $\mathcal{T}_{\pi-A}$ applies in a symmetrical manner to $\mathcal{T}_B$, $\mathcal{T}_{\pi-B}$, $\mathcal{T}_C$ and $\mathcal{T}_{\pi-C}$.*

Proof. Lemma 9 in [2] immediately implies all that is said here concerning P passing through $\mathcal{T}_A$. A small adjustment in the reasoning in the proof of Lemma 9 in [2] likewise handles the case when P passes through $\mathcal{T}_{\pi-A}$ instead. In both cases, the reasoning begins as follows. Consider the three lines that pass through P and one of the control points, when P is on $\mathcal{T}_A$ or on $\mathcal{T}_{\pi-A}$. This configuration of lines can be rigidly moved so as to now intersect at A, while still passing through the control points. The angles between the two lines passing through B and C continue to equal $\angle A$ and $\pi - \angle A$. The third line now automatically passes through A. (See Lemma 9 in [2] for further details.)

□

Note that at the exact moment that when the particle passes through one of the control points, two of the angles α, β, γ will be undefined, and so the notions of "strongly related" and "weakly related" break down momentarily. However the algebraic geometric technique of "blowing up" the control points can correct this deficiency for the weak relationship, but not the strong relationship. As the particle passes through the control point, its weak relationships with other particles continues unimpeded; not so for its strong relationships with other particles. "Weakly related" is an important equivalence relation for achieving a better understanding of the nature of the P3P problem, since it is maintained under continuous motion, even when passing through a control point.

A scenario that will be of particular interest in this paper, but which was not considered in the other papers, is this: Assume that a particle P starts off on DC, high above the xy-plane ($z \gg 0$). Suppose that P moves along a vertical line on DC until it reaches the circumcircle of ΔABC. Let us say that ϕ is a fixed angle such that this line is describe by $(x,y) = (\cos\phi, \sin\phi)$.

A description of this scenario begins by asking whether or not this line intersects one of the basic toroids ($\mathcal{T}_A$, $\mathcal{T}_B$ or $\mathcal{T}_C$). It will be helpful to let ϕ_1' be either $(\phi_2+\phi_3)/2$ or $\pi + (\phi_2+\phi_3)/2$, such that the point $(\cos\phi_1', \sin\phi_1')$ lies midway between B and C along the arc of the circumcircle connecting B and C that does not contain A (the shorter arc connecting B and C). Similarly for ϕ_2' and ϕ_3'.

Lemma 13. *The vertical line on DC, described by $(x,y) = (\cos\phi, \sin\phi)$ for fixed ϕ, intersects the toroid $\overline{\mathcal{T}}_A$ in the upper half-space (i.e. $z > 0$) if and only if $\phi_1' - \pi/2 < \phi < \phi_1' + \pi/2$ (mod 2π), that is, if and only if there exists an integer k such that $\phi_1' - \pi/2 < \phi - 2k\pi < \phi_1' + \pi/2$.*

Proof. The equation for the double toroid $\mathcal{T}_A \cup \mathcal{T}_{\pi-A}$ can be obtained from $c_1^2 = \cos^2\angle A$. Lemma 7 of [2] works this out to be the following:

$$\begin{aligned}&Z^2 + [2\zeta\overline{\zeta} - (\overline{\zeta_2}+\overline{\zeta_3})\zeta - (\zeta_2+\zeta_3)\overline{\zeta} - 2]Z\\&+ (\zeta\overline{\zeta}-1)[\zeta\overline{\zeta} - (\overline{\zeta_2}+\overline{\zeta_3})\zeta - (\zeta_2+\zeta_3)\overline{\zeta}\\&+ (1+\zeta_2\overline{\zeta_3}+\overline{\zeta_2}\zeta_2)] = 0.\end{aligned}$$

Again $Z = z^2$. Consider the intersection of this double toroid with the danger cylinder. DC is given by the equation $\zeta\overline{\zeta} = 1$. Substituting $1/\zeta$ for $\overline{\zeta}$ in the formula for the double toroid yields

$$Z = \frac{(\overline{\zeta_2}+\overline{\zeta_3})\zeta^2 + (\zeta_2+\zeta_3)}{\zeta},$$

provided that ζ is nonzero. The curve described by this equation, let us call it Γ, intersects the xy-plane at the two points given by

$$\zeta^2 = -\frac{\zeta_2+\zeta_3}{\overline{\zeta_2}+\overline{\zeta_3}} = -\frac{(\zeta_2+\zeta_3)^2}{|\zeta_2+\zeta_3|^2}.$$

The two points are therefore

$$\zeta = \pm i\,\frac{\zeta_2+\zeta_3}{|\zeta_2+\zeta_3|}.$$

Now, setting $\zeta_1' = (\zeta_2+\zeta_3)/|\zeta_2+\zeta_3|$, we see that .

$$\frac{(\overline{\zeta_2}+\overline{\zeta_3})\zeta_1'^2+(\zeta_2+\zeta_3)}{\zeta_1'} \;=\; 2|\zeta_2+\zeta_3| \;>\; 0,$$

We can then reason that the open semicircle portion of the unit circle in the xy-plane that connects the antipodal points $\pm i\,(\zeta_2+\zeta_3)/|\zeta_2+\zeta_3|$, and that has ζ_1' as its midpoint, is the vertical projection of Γ onto the xy-plane. Moreover, letting $\zeta = \exp(i\phi)$ range over this semicircle, we see that ϕ ranges (modulo 2π) over the values given in the lemma.

□

Note that the unit circle in the xy-plane is also part of the intersection of DC and the double toroid $\overline{\mathcal{T}}_A \cup \overline{\mathcal{T}}_{\pi-A}$. Also, no part of $\overline{\mathcal{T}}_{\pi-A}$ not on this unit circle intersects DC.

As P moves along the vertical line on DC, we consider the two other particles, P' and P'', that also start off high above the xy-plane, and that stay weakly related to P, and so stay on $CSDC$ (by Lemma 11). They will initially move downward along $CSDC$, but the situation becomes more complicated as soon as P crosses one of the three basic toroids. We need to track the movement of P, P' and P'' as carefully as possible.

Towards this goal, additional ideas and notation from [2] are needed here. Begin by observing that when P, P' and P'' are all high about the xy-plane, they are all in toroidal region 000. Let us consider how this might change as P descends along the verticle line. Notice that if P goes into toroidal region 110, either by first going into 100 or 010, then P' must either also move into toroidal region $1\underline{1}0$ or move into toroidal region 112. If could, for instance, wind up in 112 by first crossing $\overline{\mathcal{T}}_A$ to arrive in 100, and then going through B to arrive in 112. (See the discussion of Figure 3.)

However, even if we know that P' is in region 110, it is initially unclear whether P and P' have the same values for α and β, or have supplementary values for these angles. Similarly, two possibilities seemingly exist if we know that P' is in region 112 instead. To capture a sense of the possibilities, we will say that when P is in toroidal region 110, the pair of particles (P,P') will be in one of these four "configurations:" $(110,110)$, $(110,\underline{1}\,\underline{1}0)$, $(110,\underline{1}12)$, $(110,1\underline{1}2)$. In writing, for instance, the configuration $(110,\underline{1}12)$, we are indicating that P is in toroidal region 110, P' is in toroidal region 112, and that these two particles have the same value for β, but supplementary values for α, and supplementary values for γ.

The "0" at the end of "110" and the "2" at the end of "$\underline{1}$12" clearly indicate that the γ values differ. The "1" at the start of "110" and the "$\underline{1}$" at the start of "$\underline{1}$12" indicate that the α values differ. An underlined 1 suggests that the corresponding angle is the supplement of what it would be if the 1 was not underlined. The configuration $(110,\underline{1}12)$ is the same as the configuration $(\underline{1}10,112)$. In other words, which of the two leading ones we underline is irrelevant. The notation can be extended to discuss more than two particles. For instance, based on the discussion so far, it seems that (P,P',P'') might have configuration $(110,\underline{1}\,\underline{1}0,\underline{1}12)$, at some moment. Theorem 4 will now be proved via a series of lemmas.

Lemma 14. *As P descends down the line $(x,y) = (\cos\phi, \sin\phi)$, beginning with $z \gg 0$ and continuing until z is arbitrarily close to, but not equal to, zero, it will pass through one or two of the basic toroids, $\mathcal{T}_A$, $\mathcal{T}_B$, $\mathcal{T}_C$, but never all three of them.*

Proof. The control points / triangle vertices are $A\,(=\zeta_1)$, $B\,(=\zeta_2)$ and $C\,(=\zeta_3)$. With ϕ_1', ϕ_2' and ϕ_3', as defined earlier, let $A'=\zeta_1'=\exp(i\phi_1')$, $B'=\zeta_2'=\exp(i\phi_2')$ and $C'=\zeta_3'=\exp(i\phi_3')$. Restricting movement to the unit circle, denote the (arc) distance between two points ζ and ζ' by $d(\zeta,\zeta')$.

Lemma 13 gives the range of ϕ that would cause P to pass through $\mathcal{T}_A$. This corresponds to an open semicircle ($\zeta=\exp(i\phi)$) on the unit circle in the xy-plane. Call this semicircle σ_A. Similarly, let semicircles σ_B and σ_C correspond to $\mathcal{T}_B$ and $\mathcal{T}_C$, respectively. These three semicircles cannot all overlap, but must intersect pairwise. To see this, reason as follows.

First, σ_A cannot contain A because the vertical line through A clearly does not intersect $\mathcal{T}_A$ in the upper half-space. Of course, σ_A does contain the point $A'=\exp(i\phi_1')$, which is in fact the midpoint of this semicircle (see Lemma 13). Likewise, σ_B has B' as its midpoint, but does not contain B, and σ_C has C' as its midpoint, but does not contain C.

Now, $d(B,C)=2d(A',B)=2d(A',C)<\pi$. By this and symmetrical reasoning, we obtain $d(A',B)<\pi/2$, $d(A',C)<\pi/2$, $d(B',C)<\pi/2$, $d(B',A)<\pi/2$, $d(C',A)<\pi/2$ and $d(C',B)<\pi/2$. So, $d(B',C')<\pi$, $d(C',A')<\pi$ and $d(A',B')<\pi$. By the Inscribed Angle Theorem, we get $\angle C'A'B'<\pi/2$, $\angle A'B'C'<\pi/2$, $\angle C'B'A'<\pi/2$, That is, the triangle $\Delta A'B'C'$ is acute.

Letting O denote the origin, notice that O is the circumcenter of both ΔABC and $\Delta A'B'C'$. Here are additional facts that are straightforward to check: $2\angle AOC'=\angle AOB$, $2\angle AOB'=\angle AOC$, $\angle BOC=2\pi-\angle AOB-\angle AOC=2(\pi-\angle AOC'-\angle AOB')=2(\pi-\angle B'OC'$. So, $\angle B'OC'=\pi-\frac{1}{2}\angle BOC>\pi-\pi/2=\pi/2$. Likewise, $\angle C'OA'>\pi/2$ and $\angle A'OB'>\pi/2$. It follows that σ_A does not contain B' nor C', that σ_B does not contain C' nor A', and that σ_C does not contain A' nor B'.

It is now straightforward to see that any two of σ_A, σ_B and σ_C overlap in an arc of length strictly between 0 and $\pi/2$. From this, it is straightforward to argue that no point can be on all three of these semicircles. Also, every point on the unit circle (in the xy-plane) is within a (arc) distance $\pi/2$ of at least one of A', B' and C', and so is on at least one of the three semicircles. □

Lemma 15. *As P descends from $z\gg 0$, as in Lemma 14, assume that the first of the basic toroids it crosses is $\mathcal{T}_A$. As it crosses, the configuration for (P,P',P'') will change from $(000,000,000)$ to $(100,100,122)$ or $(100,122,100)$. Similarly if P crosses first through $\mathcal{T}_B$ or $\mathcal{T}_C$, instead of $\mathcal{T}_A$.*

Proof. When P crosses $\mathcal{T}_A$, it enters toroidal region 100. Because of Lemma 12, either P' or P'', but not both, must pass through the vertex A. Upon doing so, it will enter toroidal region 122 because its values of β and γ will be instantaneously replaced by their supplementary angles, and because it will also have moved inside $\overline{\mathcal{T}}_A$. The remaining particle will simply cross $\mathcal{T}_A$ as P crosses $\mathcal{T}_A$, and so also enter toroidal region 100. □

Lemma 16. *If P moves as in Lemma 15, and if P does not cross another basic toroid before reaching the xy-plane, whichever of P' and P'' passed through the vertex A will continue to move until it reaches the orthocenter H of ΔABC. The remaining particle will move until it arrives at some point on the circumcircle (unit circle). In doing so, after P passes through $\overline{\mathcal{T}}_A$, none of the three particles (P, P', P'') will cross the surface*

$\overline{\mathcal{T}}_A \cup \overline{\mathcal{T}}_{\pi-A} \cup \overline{\mathcal{T}}_B \cup \overline{\mathcal{T}}_{\pi-B} \cup \overline{\mathcal{T}}_C \cup \overline{\mathcal{T}}_{\pi-C}$.

Proof. The points on the circumcircle, other than the vertices A, B and C, are weakly related to each other. By the Inscribed Angle Theorem, while moving a particle around on the circumcircle, its α, β and γ values remain constant until one of the three vertices is crossed. Upon crossing, two of these angles get replaced by their supplements, while the third is unaffected. At each of the non-vertex points on the circumcircle, (α, β, γ) equals one of the following: $(\pi - \angle A, \angle B, \angle C)$, $(\angle A, \pi - \angle B, \angle C)$, $(\angle A, \angle B, \pi - \angle C)$.

Now, the orthocenter H of ΔABC is weakly related to all of these non-vertex points on the circumcircle. Here $\alpha = \pi - A$, $\beta = \pi - B$, $\gamma = \pi - C$. Reasoning using continuity, it can be seen that as a particle moves in three dimensions, and as it approaches a non-vertex point on the circumcircle, staying above the xy-plane, one and only one of its weakly related particles above the xy-plane must approach H.

Returning to the setup in the lemma, one of P' and P'' must therefore approach H as P approaches a non-vertex point on the circumcircle. The other has no place to go to other than a different non-vertex point on the circumcircle. Moreover, after P passes through $\overline{\mathcal{T}}_A$, since it never crosses $\overline{\mathcal{T}}_A \cup \overline{\mathcal{T}}_{\pi-A} \cup \overline{\mathcal{T}}_B \cup \overline{\mathcal{T}}_{\pi-B} \cup \overline{\mathcal{T}}_C \cup \overline{\mathcal{T}}_{\pi-C}$ again, its weakly related particles P' and P'' cannot (again) pass through this surface. Whichever of P' and P'' went through A is the only one that can reach the point H, which is inside ΔABC, since this triangle is acute.

□

Lemma 17. *Suppose that P, P' and P'' move as in Lemma 15, but that after P crosses $\overline{\mathcal{T}}_A$, it next crosses $\overline{\mathcal{T}}_B$ before arriving on the circumcircle of ΔABC. Assuming that P'' goes through the vertex A, later, P' will go through the vertex B. The configuration for (P, P', P'') will change like so:*

$$(000,000,000) \to (100,100,122) \to (110,\underline{1}12,1\underline{1}2).$$

Similarly, if we permute the vertices and/or switch the roles of P' and P''.

Proof. The intermediate configuration is clear from Lemma 15 and its proof. Now, either P' or P'' (but not both) must then go through B. However, P'' cannot go through B, because, at the moment when it would need to do so, it will be inside $\overline{\mathcal{T}}_{\pi-B}$, and so cannot reach B. So P' must go through B. It is straightforward to then check that the final configuration must be as claimed.

□

Lemma 18. *Consider two particles P and P' in the upper half-space which maintain a weak relationship with each other, and which keep $D > 0$. Then, P and P' cannot both be in toroidal region 110. Similarly for 101 and 011.*

Proof. By Lemma 11, no other particle in the upper half-space can be weakly related to P and P'. Now, the region 110^+ is outside the unit sphere since it is outside $\overline{\mathcal{T}}_C$. By Lemma 10, it is outside $CSDC_0$ too since $D > 0$ throughout the region. It is thus outside DC too. It is also a bounded region since toroidal region 110 is bounded. Assume that P are P' are both in toroidal region 110^+. Then, (P,P') must be in one of these two configuration: $(110,110)$ or $(110,\underline{11}0)$.

First suppose the configuration is $(110,110)$. Suppose the particles move so as to stay inside 110^+. P cannot approach $\mathcal{T}_A$ because P' cannot approach the vertex A. If P' could, then, upon passing through A, the configuration would need to change to $(010,0\underline{1}2)$, but there is no toroidal region 012. Similarly, P cannot approach $\mathcal{T}_B$. The boundary of region 110^+ must therefore be a subset of $\overline{\mathcal{T}}_C \cup CSDC_0$. However, the region 110^+ is outside $\overline{\mathcal{T}}_C$ and outside $CSDC_0$. Since 110^+ is bounded, this gives a contradiction. So the configuration $(110,110)$ is not possible.

Now consider the $(110,\underline{11}0)$ possibility. From here, P cannot reach $\mathcal{T}_A$ since P' cannot reach A simultaneously. This is because when P arrives at $\mathcal{T}_A$, its α will equal $\angle A$, so P' will need to have its α equal $\pi - \angle A$. However, P' has $\alpha = \angle A$ when it passes through A. Likewise, P cannot reach $\mathcal{T}_B$. By the same reasoning as before, this leads to a contradiction. We must conclude that $(110,\underline{11}0)$ is impossible too. Consequently, P and P' cannot both be in toroidal region 110. Likewise, still assuming $D > 0$, they cannot both be in toroidal region 101, nor toroidal region 011.

□

Lemma 19. *Consider again two particles P and P' in the upper half-space which maintain a weak relationship with each other, and which keeps $D > 0$. P cannot be in toroidal region 110 while P' is in toroidal region 112. Therefore, $D \leq 0$ throughout toroidal region 110, and $D = 0$ can only happen for a point on DC. Ditto for toroidal regions 101 and 011.*

Proof. Assume that P is in toroidal region 110 and that P' is in toroidal region 112. Then, (P,P') must be in one of these two configuration: $(110,1\underline{1}2)$ or $(110,\underline{1}12)$. By symmetry, we may assume $(110,1\underline{1}2)$. Now, P' cannot reach vertex C, so P cannot reach $\mathcal{T}_C$. Also, P' cannot reach vertex B while P arrives at $\mathcal{T}_B$, since P and P' have supplementary values for β, instead of equal values.

If P goes through $\mathcal{T}_A$, and P' goes through A, then the configuration will become $(010,010)$. P could not now pass through $\mathcal{T}_B$ because that would require P' to pass through B, resulting in the configuration $(000,\underline{2}02)$, which is impossible. Region 010^+ is bounded, being inside $\overline{\mathcal{T}}_B$. However, it is outside $\overline{\mathcal{T}}_A$ and outside $\overline{\mathcal{T}}_C$. Reasoning as before, it is outside the unit sphere, and so outside $CSDC_0$. So $\mathcal{T}_B$ must be part of the boundary of 010^+, approachable by P. This produces a contradiction, and hence, P cannot be in toroidal region 110 while P' is in toroidal region 112.

Since P and P' cannot both be in toroidal region 110 (by Lemma 18), we must conclude that P cannot be in toroidal region 110. That is, D must be less than or equal to zero throughout toroidal region 110. However, Lemmas 15 and 17 make equality impossible unless the particle P is on DC. Those lemmas capture a complete sense of where a particle for which $D = 0$ can be. Such a particle is on the surface $DC \cup CSDC$, and we have seen that if the particle is on $CSDC$, but not on DC, then it must be in one of these toroidal regions: 000, 100, 010, 001, 122, 212, 221, 112, 121, 211. It cannot be in toroidal region 110, 101 nor 011.

□

Lemma 20. *$D < 0$ throughout the toroidal region 111.*

Proof. A particle for which $D = 0$ must be on DC or else weakly related to a particle on DC. Lemmas 14 and 16 make it clear that D cannot be zero inside toroidal region 111, because these lemmas describe which toroidal regions a particle can be in when $D = 0$. The only possibilities are 000, 100, 010, 001, 110, 101, 011, 112, 121, 211, 122, 212 and 221.

The boundary of the subregion 111^+ of toroidal region 111 must thus be contained in the set $\overline{\mathcal{T}_A} \cup \overline{\mathcal{T}_{\pi-A}} \cup \overline{\mathcal{T}_B} \cup \overline{\mathcal{T}_{\pi-B}} \cup \overline{\mathcal{T}_C} \cup \overline{\mathcal{T}_{\pi-C}}$. However, because this region is bounded, part of its boundary must come from $\overline{\mathcal{T}_A} \cup \overline{\mathcal{T}_B} \cup \overline{\mathcal{T}_C}$. If a particle is in region 111^+, then it must be able to reach this part of the boundary (staying in this region), cross the boundary, and so arrive inside 110^+, 101^+ or 001^+. But, these regions do not exist, by Lemma 19. Therefore, region 111^+ does not exist either.

□

We are now prepared to rapidly prove that the deltoid-based rules are necessary conditions, when ΔABC is an acute triangle.

Proof of Theorem 4. Using the notation developed in this section, $(D > 0 \wedge \alpha < \angle A) \rightarrow (\beta \leq \angle B \vee \gamma \leq \angle C)$ means $\neg(D > 0 \wedge \alpha < \angle A \wedge \beta > \angle B \wedge \gamma > \angle C)$, which just means that the region 011^+ does not exist. Item 6 in Conjecture 1 thus means that none of the regions 110^+, 101^+, 011^+ exists. Lemma 19 confirms this fact. Similarly, Item 7 in Conjecture 1 means that the region 111^+ does not exist, which Lemma 20 confirms.

□

5 Experimental Results

Conjecture 1 has been extensively tested over a two year period, and has always held up. In fact, tests even suggest that Items 1, 2, 4 and 6 might suffice, meaning that Items 3, 5 and 7 follow from them. For testing purposes, two C++ programs, `tetrahedron_test.cpp` and `dynamic_tetrahedron_test.cpp` are available at the time of this publication[5], and the reader is encourages to experiment with these if possible.

In these programs, values for $\angle A$, $\angle B$ and $\angle C$ are selected by the user, and fixed. The cube $[0,\pi]^3$ is used for possible values of (α,β,γ). It is divided into small cubic cells, the number of which can be set via a preprocessor constant. Each cell is then determined to be "allowable" or "unallowable," by testing one or more points in the cell to see whether or not its coordinates, *i.e.* its values of α, β and γ, satisfy the conditions listed in Conjecture 1. The number of points tested inside a given cell is determined by another preprocessor constant. If any of these points satisfies all of the conditions, then this cell is declared to be "allowable;" otherwise, it is declared to be "unallowable." This is an imperfect way to proceed, but as long as each cell is sufficiently small and enough points are tested, it does a reasonably good job of distinguishing cells that contain some point satisfying the conditions from cells that do not contain such a point.

At a different stage in the programs, many possible triples (α,β,γ) are generated that correspond to actual tetrahedra $ABCP$. The number of such is also controlled by a preprocessor constant. If such a triple occurs in a cell, then that cell is called "occupied;" otherwise it is

[5] Available upon request from the author, or at
`https://github.com/mqrieck/tetrahedron_test.cpp`

$\angle A : \angle B : \angle C$	unocc., all. %	occ. unall. %
1:1:1	0.0000 %	0.0017 %
3:3:5	0.0000 %	0.0067 %
3:4:5	0.0008 %	0.0083 %
3:5:7	0.0000 %	0.4217 %
5:7:8	0.0000 %	0.5683 %

Table 1: Results from a C++ test

called "unoccupied." In an ideal situation, using a sufficient number of data points, and testing each cell sufficiently, a cell should always be either occupied and allowable, or else, unoccupied and unallowable. While in actuality, the programs generally produce a few unoccupied and allowable cells, and a few occupied and unallowable cells, these only occur near the surface in $[0,\pi]^3$ that is the boundary between the region of triples (α,β,γ) that satisfy the list of conditions, and the region of triples that do not. Unoccupied allowable cells seem to arise simply due to an inadequate number of generated data points, i.e. triples (α,β,γ) that actually correspond to tetrahedra *ABCP*. Occupied unallowable cells seem to arise due to testing too few points in the cell for allowability.

The manner in which the data points are generated should be explained. Rather than moving the point *P* around in space in a uniform way, and computing its values of (α,β,γ), it is far better to work with three "tilt planes," vary their "tilt angles" uniformly, and considering the point of intersection of these planes, which then serves as *P*. Each of these planes is simply a plane through one of the sidelines of the triangle ΔABC, and the tilt angle for such a plane is just the dihedral angle that it makes with the plane containing the triangle (*i.e.* the *xy*-plane). This method produces many points close to the triangle vertices, which is helpful since two of the quantities α, β and γ vary rapidly near a vertex.

The program `tetrahedron_test.cpp` is useful in running experiments on a given triangle ΔABC, collecting and analyzing data. For visualization purposes, `dynamic_tetrahedron_test.cpp` is quite handy. It presents an image of a layer of cells, cells having the same first coordinate, and makes it easy to scroll through different layers. The nature of each cell is represented by a different character, as explained during the execution of the program. Both programs use a number of preprocessor constants, besides those already mentioned. Using these, the various conditions in Conjecture 1 can be activated or deactivated.

Consider an example of the sort of results that `tetrahedron_test.cpp` produces. Working with an equilateral triangle ΔABC, using all of the conditions in Conjecture 1, and setting specific values for certain preprocessor constants[6], the program produced 24,652 occupied allowable cells, 0 unoccupied allowable cells, 2 occupied unallowable cells, and 95,346 unoccupied unallowable cells. If we regard unoccupied allowable cells and occupied unallowable cells as errors, then the percentage of errors in this case was 0.0017%.

Similar results, using the same preprocessor constant values, were obtained for other acute triangles, as seen in Table 1, though the results are not quite as impressive. The first column shows the ratios of the angles $\angle A$, $\angle B$ and $\angle C$ for the triangle that was tested. The second

[6] M=1000, N=50, REF_NUM=10

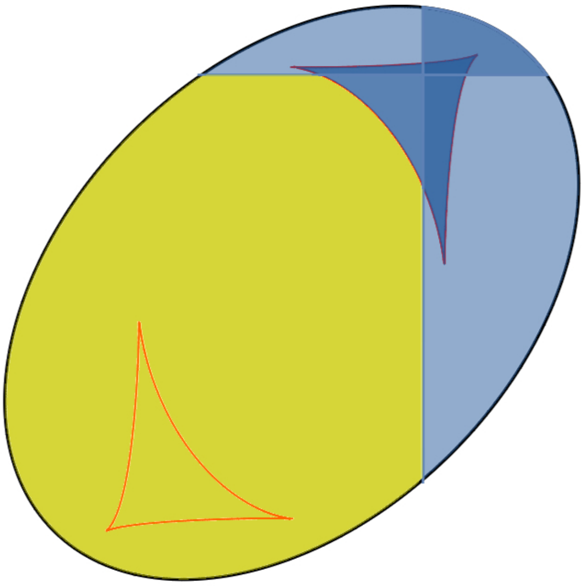

Fig. 4: Mathematica regions test

column gives the percentage of cells that were unoccupied and allowable. The third column gives the percentage of cells that were occupied and unallowable. With the settings used for the preprocessor constants, unoccupied allowable cells almost never happen, but occupied unallowable cells do occur to a significant extent. Using `dynamic_tetrahedron_test.cpp`, one can see that such anomalies occur near the boundary between the allowable and the unallowable regions of $[0,\pi]^3$, as discussed earlier.

Figure 4 shows a typical image created by Mathematica code[7] for particular values of the parameters $\angle A$, $\angle B$, $\angle C$ and α. It plots $c_2\,(=\cos\beta)$ against $c_3\,(=\cos\gamma)$. The interior of the large ellipse is where Rule 1 is satisfied (equivalently, where $H>0$), and the focus here is on this region. The blue subregion (all except the lightest subregion, in a grayscale rendering) consists of points whose coordinates (c_2,c_3) are possible. These are determined by carefully generating possible P3P setups. This is done using the extension of "Sullivan's method" discussed in [9], and this procedure does *not* involve the claims made in Conjecture 1. Darker shades of blue reflect multiple ways to obtain the allowable (c_2,c_3) pair, and this is of no importance here.

The yellowish subregion (the lightest subregion, in a grayscale rendering) shows the points (c_2,c_3) for which Item 1 in Conjecture 1 is satisfied, but at least one of the other item is not satisfied. Together, the blue subregion and the yellowish subregion partition the region inside of an ellipse. It was carefully checked that these two subregions do not overlap. This suggests that Conjecture 1 holds for the prescribed values of $\angle A$, $\angle B$, $\angle C$ and α.

[7] Available upon request from the author, or at
`https://www.mqrieck.com/P3P/Rieck_Solid_Geometry_Problems.nb`

The figure also shows a (red) curve, which is where $D = 0$ inside the ellipse. Notice that part of this curve serves as a boundary between the blue and yellow subregions, highlighting the importance of the quantity D. Many images similar to Figure 4 have been produced, all supporting Conjecture 1.

6 Conclusion

Conjecture 1 provides fairly simple criteria that appear to be required in order for a tetrahedron $ABCP$ to exist with prescribed values for $\angle A$ (= $\angle CAB$), $\angle B$ (= $\angle ABC$), $\angle C$ (= $\angle BCA$), α (= $\angle BPC$), β (= $\angle CPA$) and γ (= $\angle APB$). Most of these claims concerning necessity have been proved in this paper. Conjecture 1 can be regarded as criteria for the Perspective 3-Point Problem to have a real solution point. Experimental evidence from a couple C++ programs and several Mathematica notebooks lends substantial support to the conjecture. This is based on the assumption that the triangle ΔABC is acute.

Work is underway on extending the results to the case where ΔABC is an obtuse, rather than acute, triangle. Already it is clear from the evidence that this is reasonable, and that probably all that is required is some sort of tweaking of the formulas in Conjecture 1. However, no claim concerning the obtuse triangle case is being advanced in this paper, beyond the fact that the “sphere-based rules” (Items 1, 2 and 3 in Conjecture 1) are also required for this case too. Using C++ and Mathematica programs, one can plainly observe that some sort of alteration of the deltoid-based rules (Items 6 and 7 in Conjecture 1) is needed. The importance of the quantity D here too is unmistakable.

References

Rieck, M.Q.: On the discriminant of Grunert’s system of algebraic equations and related topics. J. Math. Imaging and Vision **60**(5), 737–762 (2018)

Rieck, M.Q., Wang, B.: Locating perspective three-point problem solutions in spatial regions. J. Math. Imaging and Vision **63**(8), 953–973 (2021)

Grunert, J.A.: Das pothenotische problem in erweiterter gestalt nebst über seine anwendungen in der geodäsie. Grunerts Archiv für Mathematik und Physik **1**, 238–248 (1841)

Haralick, R.M., Lee, C.-N., Ottenberg, K., Nölle, N.: Review and analysis of solutions of the three point perspective pose estimation problem. J. Computer Vision **13**(3), 331–356 (1994)

Gao, X.-S., Hou, X.-R., Tang, J., Cheng, H.-F.: Complete solution classification for the perspective-three-point problem. IEEE Trans. Pattern Analysis and Machine Intelligence **25**(8), 930–943 (2003)

Faugère, J.-C., Moroz, G., Rouillier, F., El-Din, M.S.: Classification of the perspective-three-point problem, discriminant variety and real solving polynomial systems of inequalities. 21st Int. Symp. Symbolic and Algebraic Computation (ISSAC ’08), Hagenberg, Austria (2008)

Rieck, M.Q.: Angular properties of a tetrahedron with an acute triangular base. Annali di Matematica Pura et Applicata (under review)

Rieck, M.Q.: Understanding the deltoid phenomenon in the perspective 3-point problem (2023). Preprint at https://www.techrxiv.org (search for ’Michael Rieck’)

Rieck, M.Q.: Related problems in spherical and solid geometry. Int. J. Geometry **12**(2), 117–124 (2023)

Declarations

- Funding: No funding
- Conflict of interest/Competing interests (check journal-specific guidelines for which heading to use): None
- Ethics approval: Not applicable; no IRB involvement
- Consent to participate: Not applicable; no IRB involvement
- Consent for publication: Not applicable; no IRB involvement
- Availability of data and materials: (See next item)
- Code availability: The author is willing to share all of his C++ and Mathematica programs related to this work, along with detailed instructions on how to use these to reproduce the results (numeric data and images), with anybody requesting them. Much of this has been posted already (see the footnotes).
- Authors' contributions: Michael Rieck is the sole author of this manuscript, as well as the C++ and Mathematica programs that support the results.