ON THE DISCRIMINANT OF GRUNERT'S SYSTEM OF ALGEBRAIC EQUATIONS AND RELATED TOPICS

Michael Q. Rieck[1]

[1]*Mathematics and Computer Science Department, Drake University, Des Moines, IA 50310, USA*
mrieck@drake.edu

Keywords: P3P: perspective: pose: camera: resection: tracking: danger cylinder: deltoid: triangle geometry: orthocenter

Abstract: Grunert's system of equations is commonly used as a basis for mathematical investigations into the Perspective 3-Point Pose Problem, for camera resectioning and tracking. This consists of three quadratic equations involving three unknown distances. The discriminant of this system helps to determine the number of real-valued solutions, in terms of the system's parameters. In its raw form, this is a very complicated and seemingly unintelligible polynomial. However, through a series of algebraic manipulations, this article manages to bring this polynomial into a far more sensible form. In addition, by making substitutions suggested by the system of equations, the discriminant is realized as a rational function of the Cartesian coordinates of the ambient space containing the control points. Moving perpendicular to the plane containing the control points, and moving away from this plane, cross sections of the surface on which this rational function vanishes approach a deltoid curve, together with the deltoid's inscribed circle (a cross section of the danger cylinder). As long as such a cross section of the surface is not too close to the control points plane, it is homeomorphic to a union of the deltoid and its inscribed circle. The orthogonal projection of the deltoid onto the control points plane contains in its interior, the control points triangle's orthocenter.

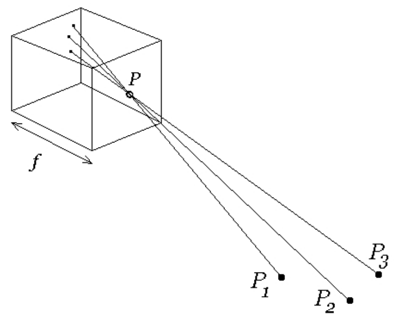

Figure 1: pinhole camera model for P3P

1 INTRODUCTION

The Perspective Three-Point Pose (P3P) problem, a well-known and basic problem in camera resectioning and tracking, was introduced by J. A. Grunert, soon after cameras were invented (Grunert, 1841). Mathematically, the difficult part of this problem is described as follows. Fix three points, P_1, P_2 and P_3, in three-dimensional real space, and also fix three angles, θ_1, θ_2 and θ_3. Is there another point P such that $\theta_1 = \measuredangle P_2PP_3$, $\theta_2 = \measuredangle P_3PP_1$ and $\theta_3 = \measuredangle P_1PP_2$? If so, determine how many such points exist, and identify these. In a practical setting, P_1, P_2 and P_3 are known points in physical space, and are often referred to as "control points." A solution point P might represent the optical center of a pinhole camera, and is mathematically a "center of perspective" corresponding to the given problem parameters P_1, P_2, P_3, θ_1, θ_2 and θ_3. When there are multiple such P, we will say that they are "related (solution) points."

In the pinhole camera model that we are considering here, it is assumed that the intrinsic characteristics of the camera are known, and in particular, the focal length f. Figure 1 shows such a pinhole camera, the three control points, and their camera images. Using the coordinates of these image points, it is straightforward to determine the cosines of the angles θ_1, θ_2 and θ_3. The details for this can be found in many introductions to the problem. By solving P3P, it can be deduced that the camera's pinhole P is among the various solution points. The Cartesian coordinates of these points can readily be determined. Some additional criterion might then be applied to guess which of these points corresponds to the actual position of the camera.

Let r_i be the distance from P to P_i, let d_i be the distance between the two control points other than P_i, and let $c_i = \cos\theta_i$ $(i = 1,2,3)$. The d_i and c_i are known a priori, but the r_i are unknown. The P3P problem amounts to solving Grunert's system of three quadratic equations:

$$\begin{cases} r_2^2 + r_3^2 - 2c_1 r_2 r_3 = d_1^2 \\ r_3^2 + r_1^2 - 2c_2 r_3 r_1 = d_2^2 \\ r_1^2 + r_2^2 - 2c_3 r_1 r_2 = d_3^2. \end{cases} \tag{1}$$

By eliminating two of the unknowns, say r_2 and r_3, it is well known that one arrives at a quartic equation in r_1^2, which will here be denoted $Q = 0$. In solving his system, Grunert actually obtained a quartic in the ratio of two of the unknowns, and other researchers have derived similar quartics (Haralick et al., 1994). The equation $Q = 0$ suggests that there can be up to four possibilities for the triple (r_1^2, r_2^2, r_3^2), but here we are only interested in triples of non-negative values, because we are interested in regarding these values as squared distances.

Each point P, not on the plane containing the control points, has a related solution point P', easily obtained by reflection through this plane. That is, since P and P' result in the same distances, r_1, r_2 and r_3, they also have the same values of c_1, c_2 and c_3, when (1) is solved for the c's. The control points plane cuts the rest of space into two halves, and for simplicity, it is usually desirable to focus only on one of these halves, and on the related points it contains. By doing so, it is well known that, given the system (1), with particular values for the c's and d's, there are a maximum of four related solution points.

Of special interest is the case when two or more solution points coalesce into a single point, a "repeated solution point." This just means that the solution point has the property that if infinitesimally small changes are made to the parameters of the system, then the original solution point will be replaced with multiple nearby solution points. This would of course require that the quartic polynomial Q have a repeated root. In the special case where a solution point is in the control points plane, Q degenerates into a quadratic polynomial, and any related solution point is also in this plane. This "coplanar case" will not be included in our discussion, so it is tacitly assumed that P_1, P_2, P_3 and P are not coplanar.

Having a repeated root for Q is necessary, but *not* sufficient, for having a repeated solution point to Grunert's system. In fact, (Wolfe et al., 1991) examines the "altitude plane" case where a non-repeated solution point corresponds to a repeated root of Q. It has been understood for a while that a solution point is a repeated solution point if and only if it is on the so-called "danger cylinder" (cf. (Zhang and Hu, 2006) and (Rieck, 2014)). This is the circular cylinder containing the circumcircle of the control points triangle, *i.e.* the triangle having the control points as vertices.

The question naturally arises as to conditions on the parameters of Grunert's system for the existence of a repeated solution point. There is actually a polynomial in the parameters of Grunert's system, d_1, d_2, d_3, c_1, c_2 and c_3, whose vanishing is necessary and sufficient for a repeated solution point to occur. This polynomial will be referred to as the "discriminant" of Grunert's system. The sign of this discriminant can be used to help determine the number of real solution points for Grunert's system. (This is analogous to the usage of discriminants of univariate polynomials in determining the number of real roots.)

Practical algorithms that employ P3P to determine the position of a camera relative to three control points typically break down due to rounding errors when the camera's pinhole is on or very near the danger cylinder. This phenomenon was recognized as early as the works (Smith, 1965) and (Thompson, 1966), where it was already understood that the singularity or near-singularity of the Jacobian matrix $(\partial c_i/\partial r_j)$ associated with the system (1) was responsible for the anomolies.

Algebraic methods, such as most of the methods in (Haralick et al., 1994), that rely solely on eliminating unknowns, and finding the roots of a resulting univariate polynomial, therefore encounter difficulties whenever the parameters of (1) are such that some solution point is on or very near the danger cylinder, even if the pinhole is actually located at a different, related solution point. However, iterative approximation methods, such as a 3-dimensional Newton-Raphson method, behave well near such related solution points, as long as they are not too near the danger cylinder, and assuming too that a reasonably accurate initial approximation is used. This is so since the Jacobian is generally well behaved in such a region.

The discriminant of Grunert's system is quite a complicated polynomial, and is a factor of the discriminant of Q, which is even more complicated. Despite its complexity, one of the achieved goals of this paper is to write the discriminant of Grunert's system in a relatively simple way. This is achieved (Theorem 2 and Corollary 2) in a manner that reveals some interesting geometry associated with Grunent's system.

In order to better study Grunert's system and P3P, it will be advantageous to use a special coordinate system for physical space. In this Cartesian coordinate system, P_i will have coordinates $(x_i, y_i, 0)$ with $x_i^2 + y_i^2 = 1$ $(i = 1,2,3)$. Letting $x_i = \cos\phi_i$ and $y_i = \sin\phi_i$, it will also be assumed that $\phi_1 + \phi_2 + \phi_3 = 0$. These requirements impose no actual restrictions on the problem and they greatly simplify the computations to be performed in this paper. If these conditions are not already satisfied in some a priori coordinate system, then it is straightforward to apply a conformal affine transformation so as to produce a new coordinate system for which these conditions are satisfied. In general, such a transformation just amounts to a combination of a translation, a rotation and a scaling.

Letting (x, y, z) denote the coordinates of a solution point P, in the special coordinate system, we thus have the additional system of equations associated with the well-known and easily solved trilateration problem:

$$\begin{cases} (x-x_1)^2 + (y-y_1)^2 + z^2 &= r_1^2 \\ (x-x_2)^2 + (y-y_2)^2 + z^2 &= r_2^2 \\ (x-x_3)^2 + (y-y_3)^2 + z^2 &= r_3^2. \end{cases} \tag{2}$$

Section 2 investigates some useful and intriguing properties of the special coordinate system. Particular attention is paid to developing successful techniques for manipulate expressions involving x_1, x_2, x_3, y_1, y_2 and y_3, so as to keep resulting expressions as simple as possible. However, some of the computations required for this study are still admittedly tedious, and symbolic manipulation software can help a great deal to reduce the labor. Nevertheless, in such cases, this article does take pains to describe how the needed calculations should be performed. In theory, this should suffice to enable an energetic reader to do all of the computations by hand. In practice, this might not really be a reasonable expectation. To assist, any interested reader can, upon request, obtain from the author, a Mathematica® notebook that includes support for the relevant computations.

Section 2 also draws some intriguing connections between the algebra and the geometry of the triangle whose vertices are the three control points. The orthocenter of this triangle plays a particularly important role. Letting (X_H, Y_H) denote its coordinates, it is found that polynomials in x_1, x_2, x_3, y_1, y_2 and y_3 that are invariant under simultaneously permuting the indices of the x's and the y's (in the same way) can be expressed instead as polynomials in X_H and Y_H. It is also shown that the orthocenter must be contained inside the standard deltoid curve that circumscribes the circumcircle of the control points triangle (*i.e.* the unit circle), and that seems to keep reappearing in this study.

Section 3 continues work that began in (Rieck, 2011) and (Rieck, 2014), but all of this is developed "from scratch" here. Attention is focused on a family of simple rational functions of the parameters of (1) (*i.e* the c's and d's) that can also be expressed as simple rational functions of the Cartesian coordinates (x, y, z) of any of its solution points. This capability is of considerable potential importance, for the theory and practice of P3P tracking. It is used in the present article to derive the formula for the discriminant, and to gain a good understanding of the surface in xyz-space upon which it vanishes. Two particularly symmetric rational functions in this family are singled out to serve as a basis for the others. Theorem 1 gives very simple formulas for these two quantities in terms of Cartesian coordinates.

Section 4 introduces the discriminant of Grunert's system, providing a careful definition for it in terms of a multi-polynomial resultant. However, the discussion then quickly takes a quite different tack, and derives a formula for the discriminant based on the results of Sections 2 and 3. Theorem 2 presents this formula, where, by means of (1) and (2), it is also observed that when the c's in the discriminant are replaced by r's, and then these are replaced by x, y and z, the result is a rational function in x, y and z^2.

The surface in xyz-space where this vanishes is then carefully examined. It is immediately seen that this surface is related to the same deltoid curve that contains the control points triangle's orthocenter in its interior. Theorem 3 then goes a long way towards describing the surface on which the discriminant vanishes, showing how well behaved and "deltoidal" it is, unless one gets too close to the control points plane.

Section 5 provides a concrete example of the preceding results.

Section 6 takes advantage of a certain birational transformation to put the formulas from Theorem 1 into a very nice, symmetrical form that better highlights the significance of the orthocenter. Theorem 4 establishes a basis for the functions from Section 3, and describes their behavior in the limit as $|z| \to \infty$. Section 7 investigates special curves obtained by intersecting contour surfaces for the functions in Section 3. These are very well behaved far from the control points plane, but exhibit interesting and somewhat erratic behavior near this plane. These curves are ultimately used to prove Theorem 3.

2 SPECIAL COORDINATES AND AN ASSOCIATED ALGEBRA

The focus of this section is on learning to take advantage of the special coordinate system in order to facilitate manipulations of expressions involving the coordinates of the control points. Recall that the control points have Cartesian coordinates $(x_i, y_i, 0) = (\cos\phi_i, \sin\phi_i, 0)$ $(i = 1,2,3)$ with $\phi_1 + \phi_2 + \phi_3 = 0$. Applications of the angle sum formulas of trigonometry immediately yield the following formulas.

Lemma 1. *Under the imposed restrictions, the Cartesian coordinates of the control points satisfy the following equations:*

(i) $x_1 = x_2x_3 - y_2y_3$,

(ii) $x_2 = x_3x_1 - y_3y_1$,

(iii) $x_3 = x_1x_2 - y_1y_2$,

(iv) $-y_1 = x_2y_3 + y_2x_3$,

(v) $-y_2 = x_3y_1 + y_3x_1$,

(vi) $-y_3 = x_1y_2 + y_1x_2$,

(vii) $x_1x_2x_3 - x_1y_2y_3 - y_1x_2y_3 - y_1y_2x_3 = 1$,

(viii) $y_1y_2y_3 - y_1x_2x_3 - x_1y_2x_3 - x_1x_2y_3 = 0$.

These, along with the basic $x_i^2 + y_i^2 = 1$, can be applied in different directions to reduce the number of factors in a term, and to reduce the number of terms in an expression. No precise prescription is offered for successfully simplifying a given expression. In fact, it is not even clear what it means for one expression to be simpler than another. Consider the following example:

$$2(1+x_1)(1+x_2)(1+x_3) = (1+x_1+x_2+x_3)^2$$

This is a true statement (see Lemma 2, item *(ix)*). However, it is not completely clear which of these expressions should be regarded as "simpler."

Let us now take up a general discussion of polynomials in x_1, x_2, x_3, y_1, y_2, and y_3 that are invariant under permutations of the indices. For instance, $x_1y_1 + x_2y_2 + x_3y_3$. We begin by taking an inventory of some of the simpler relationships between the polynomials of the x's and y's that exhibit this invariance. To do so, let us give names to two such linear polynomials, by defining $X_H = x_1 + x_2 + x_3$ and $Y_H = y_1 + y_2 + y_3$. The following relationships will now be established.

Lemma 2. *The coordinates of the control points are related to X_H and Y_H as follows:*

(i) $x_1^2 + x_2^2 + x_3^2 = 1 + [(X_H - 1)^2 - Y_H^2]/2$

(ii) $y_1^2 + y_2^2 + y_3^2 = 2 - [(X_H - 1)^2 - Y_H^2]/2$

(iii) $x_2x_3 + x_3x_1 + x_1x_2 = [(X_H + 1)^2 + Y_H^2]/4 - 1$

(iv) $y_2y_3 + y_3y_1 + y_1y_2 = [(X_H - 1)^2 + Y_H^2]/4 - 1$

(v) $x_1y_2 + x_1y_3 + x_2y_3 + x_2y_1 + x_3y_1 + x_3y_2 = -Y_H$

(vi) $x_1y_1 + x_2y_2 + x_3y_3 = (X_H + 1)Y_H$

(vii) $x_1x_2x_3 = [(X_H - 1)^2 - Y_H^2]/4$

(viii) $y_1y_2y_3 = -(X_H + 1)Y_H/2$

(ix) $(x_1 + 1)(x_2 + 1)(x_3 + 1) = (X_H + 1)^2/2$

Proof. Let $e_2 = x_2x_3 + x_3x_1 + x_1x_2$, $e_2' = y_2y_3 + y_3y_1 + y_1y_2$, $p_2 = x_1^2 + x_2^2 + x_3^2$ and $p_2' = y_1^2 + y_2^2 + y_3^2$. Then, using Lemma 1, $(X_H - 1)^2 - Y_H^2 = 1 - 2X_H + p_2 + 2e_2 - p_2' - 2e_2' = -2 + (3 - 2X_H + p_2 - p_2') + 2(e_2 - e_2') = -2 + 2(p_2 - X_H) + 2(e_2 - e_2') = -2 + 2(p_2 - X_H) + 2X_H = 2(p_2 - 1)$. This establishes *(i)*, and by adding *(i)* and *(ii)*, we immediately obtain *(ii)* as well. $(X_H + 1)^2 + Y_H^2 = 1 + 2X_H + p_2 + 2e_2 + p_2' + 2e_2' = 2(2 + X_H) + 2(e_2 + e_2') = 2(2 + X_H) + 2(2e_2 - X_H) = 4(1 + e_2)$. This establishes *(iii)*. Subtracting *(iv)* from *(iii)*, and applying Lemma 1 again, we obtain an equation that is seen to be true, and that therefore establishes the truth of *(iv)*. Part *(v)* is an immediate consequence of Lemma 1, and by adding *(v)* and *(vi)*, we immediately obtain *(vi)* as well. Applying Lemma 1 yet again, we see that $x_1x_2x_3 - x_1(x_2x_3 - x_1) - x_2(x_3x_1 - x_2) - x_3(x_1x_2 - x_3) = p_2 - 2x_1x_2x_3 = 1$, establishing *(vii)*. Likewise, $y_1y_2y_3 - y_1(y_2y_3 +$

$x_1) - y_2(y_3y_1 + x_2) - y_3(y_1y_2 + x_3) = -(x_1y_1 + x_2y_2 + x_3y_3) - 2y_1y_2y_3 = 0$, which establishes *(viii)*. To establish *(ix)*, simply expand the left side and then apply the formulas that have already been established.

□

Together, Lemmas 1 and 2 supply a good deal of flexibility in rewriting expressions involving the coordinates of the control points. They are put to work for this purpose throughout this article.

We now turn to three geometric concepts, namely, the orthocenter of the control points triangle, the area of this triangle, and a deltoid associated with this triangle. In the following, (X_H, Y_H) is identified with $(X_H, Y_H, 0)$, and more generally, (x, y) will be routinely identified with $(x, y, 0)$.

Lemma 3.

(i) The coordinates of the orthocenter of the control points triangle are (X_H, Y_H).

(ii) The following formulas all express the square of the area of the control points triangle:

$$\frac{1}{16}d_1^2d_2^2d_3^2 =$$
$$\frac{1}{2}(1 - x_2x_3 - y_2y_3)(1 - x_3x_1 - y_3y_1)\cdot (1 - x_1x_2 - y_1y_2) =$$
$$\frac{1}{16}(2d_2^2d_3^2 + 2d_3^2d_1^2 + 2d_1^2d_2^2 - d_1^4 - d_2^4 - d_3^4) =$$
$$\frac{1}{4}(x_1y_2 + x_2y_3 + x_3y_1 - x_1y_3 - x_2y_1 - x_3y_2)^2 =$$
$$[(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)/Y_H]^2 =$$
$$[(y_1 - y_2)(y_2 - y_3)(y_3 - y_1)/(X_H + 1)]^2 =$$
$$\frac{1}{16}[\, 27 + 8X_H(X_H^2 - 3Y_H^2) - (X_H^2 + Y_H^2)(X_H^2 + Y_H^2 + 18)\,].$$

(iii) The orthocenter of the control points triangle must be somewhere inside the standard deltoid curve, defined by the equation

$$27 + 8x(x^2 - 3y^2) = (x^2 + y^2)(x^2 + y^2 + 18).$$

Consequently, $X_H^2 + Y_H^2 < 9$.

Proof. To show that (X_H, Y_H) are the coordinates of the orthocenter, we simply observe that this point, together with the three control points form an orthocentric system, meaning that the line through any two of these points is perpendicular to the line through the other two points. Pairing the point whose coordinates are (X_H, Y_H) with say the control point P_1, we can check for the claimed orthogonality by computing a dot product as follows:

$$(X_H - x_1, Y_H - y_1)\cdot(x_2 - x_3, y_2 - y_3) =$$
$$(x_2 + x_3, y_2 + y_3)\cdot(x_2 - x_3, y_2 - y_3) =$$
$$x_2^2 - x_3^2 + y_2^2 - y_3^2 = 1 - 1 = 0.$$

Similarly, if we pair the points differently. This establishes the first (i).

To prove (ii), we begin by pointing out that the two formulas involving the sidelengths d_i are obtained from standard triangle area formulas. The first of these asserts that the area equals the product of the sidelengths divided by four times the circumradius, but of course the circumradius is one in the setup here. The second formula involving the d_i is easily derived from the well-known Heron's Formula.

The formulas immediately under these two formulas are obtained from them by simply using the fact that $d_1^2 = (x_2 - x_3)^2 + (y_2 - y_3)^2 = 2(1 - x_2x_3 - y_2y_3)$, and similarly for d_2^2 and d_3^2. These formulas are invariant under permutations of the indices, and using Lemma 2, they can be rewritten in terms of X_H and Y_H, resulting in the final formula in (ii). The two remaining formulas in (ii) require further manipulations using the earlier lemmas. The details are omitted, but the process is straightforward. This establishes (ii).

Finally, the equation in (iii) is indeed the formula for a deltoid curve, and it is in fact, the same standard deltoid curve that appears in (Rieck, 2015). The quantity in the last formula in (ii) must of course be positive since it represents the square of an area. But this requirement means that the orthocenter, whose coordinates are (X_H, Y_H) must be inside the deltoid. This imposes bound $X_H^2 + Y_H^2 < 9$, since the standard deltoid is contained in the circle of radius three centered at the origin.

□

No doubt, item *(iii)* in Lemma 3 admits a purely geometric proof, and may very well have been considered elsewhere in the context of triangle geometry. Clark Kimberling introduces the deltoid in a very different way in Chapter 6, Section 1 of (Kimberling, 2003). The image there seems to suggest that any two of the lines in the pencil of lines shown there intersect in the interior of the deltoid. Moreover, the altitudes of the triangle are included among these lines, which would again seem to lead to the conclusion that the orthocenter is located inside the deltoid.

After a quick study of the algebra $\mathcal{A}$ of polynomials in x_1, x_2, x_3, y_1, y_2, y_3, this algebra can then be applied productively to a study of Grunert's system. The polynomial coefficients will always be real numbers. Though it is worthwhile to consider the complexified version of $\mathcal{A}$, this is unnecessary for our purposes here, and so will be avoided. Now, x_1, x_2, x_3, y_1, y_2, and y_3 are just required to satisfy the equation in Lemma 1, so technically, $\mathcal{A}$ can be regarded as a quotient ring of a polynomial ring in six indeterminates, by modding out by the ideal generated by the set of polynomials suggested by Lemma 1.

Let $\mathcal{S}_3$ be the symmetric group of permutations of the elements of the set $\{1,2,3\}$. Let $\mathcal{G}$ be the automorphism group of $\mathcal{A}$, isomorphic to $\mathcal{S}_3$, where each $\sigma \in \mathcal{S}_3$ induces an element $\hat{\sigma}$ of $\mathcal{G}$ that sends each x_i to $x_{\sigma(i)}$ and sends each y_i to $y_{\sigma(i)}$. Let $\mathcal{A}^{\mathcal{G}}$ be the subalgebra of $\mathcal{A}$ consisting of the fixed elements of $\mathcal{A}$ under the action of $\mathcal{G}$. Using the multiplication operation of $\mathcal{A}$, one can consider $\mathcal{A}$ to also be an $\mathcal{A}^{\mathcal{G}}$-module.

The Reynolds (average) operator π on $\mathcal{A}$ with respect to $\mathcal{G}$ is defined by the following formula:

$$\pi(f)(x_1,x_2,x_3,y_1,y_2,y_3) = \frac{1}{6}\sum_{\sigma\in\mathcal{S}_3} f(x_{\sigma(1)},x_{\sigma(2)},x_{\sigma(3)},y_{\sigma(1)},y_{\sigma(2)},y_{\sigma(3)}) \quad (3)$$

for all $f \in \mathcal{A}$. Let $\mathcal{N}$ consist of the elements of $\mathcal{A}$ that are annihilated by π (*i.e.* mapped to zero). The elements of $\mathcal{A}^{\mathcal{G}}$ are the elements of $\mathcal{A}$ left invariant (fixed) by π. $\mathcal{N}$ and $\mathcal{A}^{\mathcal{G}}$ are $\mathcal{A}^{\mathcal{G}}$-submodules of $\mathcal{A}$. Moreover, it is well known and straightforward to check that as $\mathcal{A}^{\mathcal{G}}$-modules,

$$\mathcal{A} = \mathcal{A}^{\mathcal{G}} \oplus \mathcal{N}.$$

As an algebra, $\mathcal{A}^{\mathcal{G}}$ is freely generated by two elements, namely X_H and Y_H. The fact that X_H and Y_H are algebraically independent can be seen intuitively by considering the geometry that motivated this algebra. There, (X_H,Y_H) is the orthocenter associated with the triangle whose vertices are (x_1,y_1), (x_2,y_2) and (x_3,y_3). By varying the vertices subject to the restrictions, it is possible to move (X_H,Y_H) to any point inside the deltoid. Thus there is no non-trivial polynomial equation of X_H and Y_H (alone) that is satisfied for all possible values of X_H and Y_H. It is also possible to show this fact algebraically by generating a suitable Groebner basis involving the Lemma 1 equations and the definitions for X_H and Y_H. No member of the resulting Groebner basis is a polynomial of X_H and Y_H alone.

As an $\mathcal{A}^{\mathcal{G}}$-module, $\mathcal{A}^{\mathcal{G}}$ itself is generated by the identity element. The situation for the $\mathcal{A}^{\mathcal{G}}$-module $\mathcal{N}$ is more interesting.

Lemma 4. *As an $\mathcal{A}^{\mathcal{G}}$-module, $\mathcal{N}$ is generated by the following the seven elements:*

$$\begin{aligned}
X_1 &= 2x_1 - x_2 - x_3\\
X_2 &= 2x_2 - x_3 - x_1\\
X_3 &= 2x_3 - x_1 - x_2\\
Y_1 &= 2y_1 - y_2 - y_3\\
Y_2 &= 2y_2 - y_3 - y_1\\
Y_3 &= 2y_3 - y_1 - y_2\\
\Delta &= x_1y_2 + x_2y_3 + x_3y_1 - x_1y_3 - x_2y_1 - x_3y_2.
\end{aligned}$$

In fact, since $X_1 + X_2 + X_3 = 0$ and $Y_1 + Y_2 + Y_3 = 0$, one of the X's and one of the Y's can be removed, reducing to a set of five generators. Moreover, the following product formulas hold for these elements, and exhibit decompositions into $\mathcal{A}^{\mathcal{G}}$ and $\mathcal{N}$ components:

$$\begin{aligned}
X_i^2 &= {}^1\!/\!_2\,(9 - 6X_H + X_H^2 - 3Y_H^2) +\\
&\quad {}^1\!/\!_2\,(3X_i - X_HX_i - 3Y_HY_i)\\
X_{i+1}X_{i+2} &= -{}^1\!/\!_4\,(9 - 6X_H + X_H^2 - 3Y_H^2) +\\
&\quad {}^1\!/\!_2\,(3X_i - X_HX_i - 3Y_HY_i)\\
Y_i^2 &= {}^1\!/\!_2\,(9 + 6X_H - 3X_H^2 + Y_H^2) -\\
&\quad {}^1\!/\!_2\,(3X_i + 3X_HX_i + Y_HY_i)\\
Y_{i+1}Y_{i+2} &= -{}^1\!/\!_4\,(9 + 6X_H - 3X_H^2 + Y_H^2) -\\
&\quad {}^1\!/\!_2\,(3X_i + 3X_HX_i + Y_HY_i)\\
X_iY_i &= (2X_H + 3)Y_H +\\
&\quad {}^1\!/\!_2\,(Y_HX_i + X_HY_i - 3Y_i)\\
X_iY_{i+1} &= -{}^1\!/\!_2\,(2X_H + 3)Y_H +\\
&\quad {}^1\!/\!_2\,(Y_HX_{i+2} + X_HY_{i+2} - 3Y_{i+2} + 3\Delta)\\
X_{i+1}Y_i &= -{}^1\!/\!_2\,(2X_H + 3)Y_H +\\
&\quad {}^1\!/\!_2\,(Y_HX_{i+2} + X_HY_{i+2} - 3Y_{i+2} - 3\Delta)\\
X_i\Delta &= {}^1\!/\!_3\,(3 + 2X_H)Y_H(X_{i+2} - X_{i+1}) +\\
&\quad {}^1\!/\!_6\,(9 - 6X_H + X_H^2 - 3Y_H^2)\cdot\\
&\quad (Y_{i+1} - Y_{i+2}) \in \mathcal{N}\\
Y_i\Delta &= {}^1\!/\!_3\,(3 + 2X_H)Y_H(Y_{i+1} - Y_{i+2}) +\\
&\quad {}^1\!/\!_6\,(9 + 6X_H - 3X_H^2 + Y_H^2)\cdot\\
&\quad (X_{i+2} - X_{i+1}) \in \mathcal{N}\\
\Delta^2 &= {}^1\!/\!_4\,[27 + 8X_H(X_H^2 - 3Y_H^2) -\\
&\quad (X_H^2 + Y_H^2)(X_H^2 + Y_H^2 + 18)] \in \mathcal{A}^{\mathcal{G}}.
\end{aligned}$$

(Indices here are to be read modulo 3.)

Proof. The product formulas are straightforward, though a bit tedious, to check, by replacing each generator with its definition, and by applying the relations found in Lemmas 1 and 2. The formula for Δ^2 is from Lemma 3.

Next, the elements of the group $\mathcal{G}$ permute the elements of each of the sets $\{X_1, X_2, X_3\}$, $\{Y_1, Y_2, Y_3\}$ and $\{\Delta, -\Delta\}$, and it then becomes clear that each of the seven listed elements is annihilated by π, and hence belongs to $\mathcal{N}$. Let $\mathcal{M}$ be the $\mathcal{A}^{\mathcal{G}}$-submodule of $\mathcal{N}$ generated by the listed elements. We must show that $\mathcal{M} = \mathcal{N}$.

Since $\pi(\pi(f)) = \pi(f)$ for each $f \in \mathcal{A}$, it is seen that $\mathcal{A}^{\mathcal{G}} = \{\pi(f) \mid f \in \mathcal{A}\}$ and $\mathcal{N} = \{f - \pi(f) \mid f \in \mathcal{A}\}$. Now, $x_i - \pi(x_i) = x_i - X_H/3 = X_i/3 \in \mathcal{M}$ for $i = 1,2,3$. Similarly, $y_i - \pi(y_i) \in \mathcal{M}$. Therefore, $x_i, y_i \in \mathcal{A}^{\mathcal{G}} \oplus \mathcal{M}$. We next want to establish that the $\mathcal{A}^{\mathcal{G}}$-module $\mathcal{A}^{\mathcal{G}} \oplus \mathcal{M}$ is actually an algebra, that is, it is closed under multiplication. But this is clear since the listed product formulas show that $\mathcal{M} \cdot \mathcal{M} \subseteq \mathcal{A}^{\mathcal{G}} \oplus \mathcal{M}$.

Since $1, x_1, x_2, x_3, y_1, y_2, y_3 \in \mathcal{A}^{\mathcal{G}} \oplus \mathcal{M}$, since $\mathcal{A}^{\mathcal{G}} \oplus \mathcal{M}$ is closed under multiplication, and since $1, x_1, x_2, x_3, y_1, y_2, y_3$ generate $\mathcal{A}$ as an algebra, it follows that $\mathcal{A}^{\mathcal{G}} \oplus \mathcal{M} = \mathcal{A}$. Because $\mathcal{M} \subseteq \mathcal{N}$ and $\mathcal{A}^{\mathcal{G}} \oplus \mathcal{N} = \mathcal{A}$, we can conclude that $\mathcal{M} = \mathcal{N}$.

□

3 A CLASS OF RATIONAL FUNCTIONS

Returning to Grunert's system (1), some mathematical insight into this system is gained by considering what happens if one or more of the c's are negated. All real solutions to (1) will be considered here, meaning that the r's are allowed to be negative too. This idea can be visualized by imagining that each line connecting a center of perspective to a control point is a copy of the real number line with zero at the center of perspective, arbitrarily oriented in one direction or the other. Reversing this orientation then amounts to negating the corresponding r_i.

Now, negating some c_i in (1) corresponds to replacing θ_i with its supplementary angle. Changing just one such c_i alters the arrangement significantly, leading to quite different mathematical solutions, *i.e.* very different possibilities for the r's. However, if exactly two of the c's are negated, then it is easy to see how a solution to the original setup can be converted into a solution to the altered setup. For instance, if c_1 and c_2 are negated, then by simply negating r_3 too, the system remains essentially unaffected. This merely amounts to reorienting the imagined number line connecting the center of perspective to the third control point.

Of course this change in an imagined line has absolutely no effect on the solution points in physical space (*i.e.* xyz-space). Also unaffected are these four quantities: c_1^2, c_2^2, c_2^2 and $c_1c_2c_3$. Because of these facts, it might be argued that these quantities represent more accurately the physical P3P problem than do c_1, c_2 and c_3 themselves. In any case, in Section 4, it is shown that the discriminant of Grunert's system can be written as a polynomial in c_1^2, c_2^2, c_2^2 and $c_1c_2c_3$.

Letting

$$\eta = (1 - c_1^2 - c_2^2 - c_3^2 + 2c_1c_2c_3)^{1/2},$$

the following formulas are immediately obtained from (1), (2), and Lemmas 1 and 3, and have mostly been proven already in (Rieck, 2014).

Lemma 5. *Various quantities introduced up to this point are related as follows:*

(i) $d_1^2 = (x_2 - x_3)^2 + (y_2 - y_3)^2$
$= 2(1 - x_2x_3 - y_2y_3)$
(and similarly for d_2^2 and d_3^2)

(ii) $r_i^2 = 1 - 2(x_ix + y_iy) + x^2 + y^2 + z^2$ $(i = 1,2,3)$

(iii) $(r_2^2 + r_3^2 - d_1^2)/2 =$
$(x - x_2)(x - x_3) + (y - y_2)(y - y_3) + z^2$
(and similarly with indices permuted)

(iv) $4r_1^2r_2^2r_3^2\ \eta^2 = d_1^2d_2^2d_3^2\ z^2$

(v) $4r_1^2r_2^2r_3^2\ c_1^2 = r_1^2(r_2^2 + r_3^2 - d_1^2)^2$
(and similarly for c_2^2 and c_3^2)

(vi) $8r_1^2r_2^2r_3^2\ c_1c_2c_3 =$
$(r_2^2 + r_3^2 - d_1^2)(r_3^2 + r_1^2 - d_2^2)(r_1^2 + r_2^2 - d_3^2)$

(vii) $c_1^2/\eta^2 = r_1^2(r_2^2 + r_3^2 - d_1^2)^2/(d_1^2d_2^2d_3^2z^2)$
(and similarly for c_2^2/η^2 and c_3^2/η^2)

(viii) $c_1c_2c_3/\eta^2 = (r_2^2 + r_3^2 - d_1^2)\cdot$
$(r_3^2 + r_1^2 - d_2^2)(r_1^2 + r_2^2 - d_3^2)\ /\ (2d_1^2d_2^2d_3^2z^2)$

(ix) *The quantities c_1^2, c_2^2, c_3^2, $c_1c_2c_3$ and η^2 can all be expressed as rational functions of d_1^2, d_2^2, d_3^2, x, y and z^2.*

In anticipation of the proof of the theorem in this section, the following lemma expresses significant quantities involved in the special-coordinates P3P setup in terms of the generators for $\mathcal{A}$ as a module over $\mathcal{A}^{\mathcal{G}}$. These formulas are straightforward consequences of Lemmas 1, 2, 4 and 5.

Lemma 6. *Various quantities can be expressed in terms of the generators for $\mathcal{A}$ as follows:*

$$\begin{aligned}
x_i &= {}^1\!/\!_3\,(X_H + X_i),\\
y_i &= {}^1\!/\!_3\,(Y_H + Y_i),\\
(1-x_i)y_i &= {}^1\!/\!_6\,[\,(3-X_H)Y_i\\
&\quad -Y_H X_i - 2X_H Y_H\,],\\
(1+x_i)x_i &= {}^1\!/\!_6\,[\,(3+X_H)X_i - Y_H Y_i\\
&\quad +X_H^2 - Y_H^2 + 3\,],\\
d_i^2 &= {}^1\!/\!_3\,(9 - X_H^2 - Y_H^2)\\
&\quad +{}^2\!/\!_3\,(X_H X_i + Y_H Y_i),\\
r_i^2 &= 1 + x^2 + y^2 + z^2\\
&\quad -{}^2\!/\!_3\,(X_H + X_i)x - {}^2\!/\!_3\,(Y_H + Y_i)y,\\
r_2^2 + r_3^2 - d_1^2 &= 2(x^2+y^2+z^2) - 1\\
&\quad +{}^2\!/\!_3\,[(X_1 - 2X_H)x + (Y_1 - 2Y_H)y]\\
&\quad +{}^1\!/\!_3\,[X_H^2 + Y_H^2 - 2X_H X_1 - 2Y_H Y_1]
\end{aligned}$$

(and similarly for $r_3^2 + r_1^2 - d_2^2$ and $r_1^2 + r_2^2 - d_3^2$).

Two other quantities that will be required are $r_1^2 r_2^2 r_3^2$ and $(r_2^2 + r_3^2 - d_1^2)(r_2^3 + r_1^2 - d_2^2)(r_1^2 + r_2^2 - d_3^2)$. These can be expanded as follows.

Lemma 7. *The expressions $r_1^2 r_2^2 r_3^2$ and $(r_2^2 + r_3^2 - d_1^2)(r_2^3 + r_1^2 - d_2^2)(r_1^2 + r_2^2 - d_3^2)$, when expanded in terms on x, y and z, in accord with Lemma 5, have coefficients that belong to $\mathcal{A}^{\mathcal{G}}$. As such, their expressions in terms of X_H and Y_H are as specified in the appendix.*

Proof. This is only a sketch of a proof. A complete and careful proof requires significant algebraic manipulations that cannot be detailed in the limited space. Nevertheless, a roadmap is provided here for how this can be accomplished with a determined effort.

$r_1^2 r_2^2 r_3^2$ is clearly invariant under the group $\mathcal{G}$, and hence is an element of $\mathcal{A}^{\mathcal{G}}$. For $i = 1,2,3$, write $r_i^2 = p_i + q_i$ where $p_i \in \mathcal{A}^{\mathcal{G}}$ and $q_i \in \mathcal{N}$. We have these two facts concerning the products of elements of two sets: $\mathcal{A}^{\mathcal{G}} \cdot \mathcal{A}^{\mathcal{G}} = \mathcal{A}^{\mathcal{G}}$ and $\mathcal{A}^{\mathcal{G}} \cdot \mathcal{N} = \mathcal{N}$. It follows that $r_1^2 r_2^2 r_3^2 = p_1 p_2 p_3 + p_1 q_2 q_3 + q_1 p_2 q_3 + q_1 q_2 p_3 + q_1 q_2 q_3$. According to Lemma 6, $p_i = 1 + x^2 + y^2 + z^2 - {}^2\!/\!_3\,(X_H x + Y_H y)$ and $q_i = -{}^2\!/\!_3\,(X_i x + Y_i y)$. Since $p_1 = p_2 = p_3$, we only need to compute $p_1 p_2 p_3 + p_1(q_2 q_3 + q_1 q_3 + q_1 q_2) + q_1 q_2 q_3$. Clearly $q_2 q_3 + q_1 q_3 + q_1 q_2 \in \mathcal{A}^{\mathcal{G}}$, so to compute this, it is only necessary to compute the invariant part of $q_2 q_3$ (*i.e.* the $\mathcal{A}^{\mathcal{G}}$-compenent), and multiply this by three. Since $q_1 q_2 q_3 \in \mathcal{A}^{\mathcal{G}}$, to compute this, it suffices to compute the invariant part of the product of q_1 and the annihilated part of $q_2 q_3$ (*i.e.* the $\mathcal{N}$-compenents).

Direct computation, using Lemmas 4, shows that the invariant part of $q_2 q_3$ is $\pi(q_2 q_3) = {}^4\!/\!_9\;\pi((X_2 x + Y_2 y)(X_3 x + Y_3 y)) = {}^4\!/\!_9\;[\pi(X_2 X_3)x^2 + \pi(Y_2 Y_3)y^2 + \pi(X_2 Y_3 + X_3 Y_2)xy] = {}^4\!/\!_9\,[-{}^1\!/\!_4(9 - 6X_H + X_H^2 - 3Y_H^2)x^2 - {}^1\!/\!_4(9 + 6X_H - 3X_H^2 + Y_H^2)y^2 - (2X_H + 3)Y_H xy] = \frac{1}{9}(3Y_H^2 - X_H^2 + 6X_H - 9)\,x^2 + \frac{1}{9}(3X_H^2 - Y_H^2 - 6X_H - 9)\,y^2 - \frac{4}{9}(2X_H + 3)Y_H\,xy$. Likewise, the annihilated part is $\frac{2}{9}[(3 - X_H)X_1 - 3Y_H Y_1]x^2 - \frac{2}{9}[3(1 + X_H)X_1 + Y_H Y_1]\,y^2 + \frac{4}{9}(Y_H X_1 + (X_H - 3)Y_1)\,xy$. Following the above procedure now leads directly to the claimed result concerning $r_1^2 r_2^2 r_3^2$. One can handle $(r_2^2 + r_3^2 - d_1^2)(r_2^3 + r_1^2 - d_2^2)(r_1^2 + r_2^2 - d_3^2)$ in a similar manner, though the computations are admittedly more complicated.

□

Next, consider moving a point P with coordinates (x,y,z) around in space, and allowing the r's and c's to change to match P in the systems of equations (1) and (2). As $|z|$ tends to infinity, so too do $|r_1|$, $|r_2|$ and $|r_3|$, while $|c_1|$, $|c_2|$ and $|c_3|$ tend to one. Nevertheless, it is possible to find useful rational function in c_1^2, c_2^2, c_2^2 and $c_1 c_2 c_3$ that converge and become non-trivial as $|z|$ tends to infinity. In fact, such functions were employed in some of the author's previous work. Based in part on the results in (Rieck, 2015), it becomes possible to identify all such rational functions whose numerator and denominator are both linear, and to understand clearly what happens as $|z|$ goes to infinity. We can write such a rational function as the ratio of two other rational functions, each of the form

$$\frac{\alpha + \beta\, c_1 c_2 c_3 + \gamma_1\, c_1^2 + \gamma_2\, c_2^2 + \gamma_3\, c_3^2}{\eta^2}. \qquad (4)$$

Lemma 8. *The quantity (4) can be rewritten as a rational function in x, y, z^2, having $d_1^2 d_2^2 d_3^2\, z^2$ as its denominator, and having a numerator that is a polynomial in z^2 of degree three or less, with polynomials in x and y as coefficients. When $\alpha + \beta + \gamma_1 + \gamma_2 + \gamma_3 \neq 0$,* *this rational function diverges and is asymptotic to* $4(\alpha+\beta+\gamma_1+\gamma_2+\gamma_3)\, z^4 \,/\, d_1^2 d_2^2 d_3^2$, *as* $|z| \to \infty$.

Proof. Using Lemma 5, the quantity (4) can be rewritten as a rational function of x, y and z^2, with the denominator being $d_1^2 d_2^2 d_3^2 z^2$. The numerator is clearly a cubic polynomial in z^2 whose coefficients are polynomials in x and y, and the leading coefficient can be seen to be $4(\alpha+\beta+\gamma_1+\gamma_2+\gamma_3)$. The last claim in the lemma thus follows. □

The next lemma is proved by just continuing the analysis begun in the proof of the Lemma 8, but arranging for the z^6 term in the numerator to vanish, and then computing the coefficient of z^4. This too is straightforward using Lemma 5. Here it is helpful to introduce a new quantity, namely, set $\iota = 3 - x_2x_3 - x_3x_1 - x_1x_2 - y_2y_3 - y_3y_1 - y_1y_2 = (9 - X_H^2 - Y_H^2)/2 = (d_1^2+d_2^2+d_3^2)/2$.

Lemma 9. *Setting* $\alpha = -\beta-\gamma_1-\gamma_2-\gamma_3$ *in (4), this quantity can now be expressed as a rational function in x, y, z, having* $d_1^2 d_2^2 d_3^2 z^2$ *as its denominator, and having a numerator that is a polynomial in* z^2 *of degree two or less. When* $\iota\beta + d_1^2\gamma_1 + d_2^2\gamma_2 + d_3^2\gamma_3 \neq 0$, *this rational function diverges and is asymptotic to* $-4(\iota\beta + d_1^2\gamma_1 + d_2^2\gamma_2 + d_3^2\gamma_3)\, z^2 \,/\, d_1^2 d_2^2 d_3^2$, *as* $|z| \to \infty$.

As a consequence of Lemmas 8 and 9, the following becomes immediately clear.

Lemma 10. *A quantity of the form (4) diverges as* $|z| \to \infty$, *except possible when the parameters satisfy both*

$$\begin{aligned} &\alpha+\beta+\gamma_1+\gamma_2+\gamma_3 = 0 \quad \text{and} \\ &\iota\beta + d_1^2\gamma_1 + d_2^2\gamma_2 + d_3^2\gamma_3 = 0. \end{aligned} \tag{5}$$

Moreover, the ratio of two functions of the form (4) must either diverge or else approach a constant value, independent of x and y, as $|z| \to \infty$, *except possibly when both functions satisfy (5).*

A goal now will be to try to gain a solid understanding of functions of the form (4) that satisfy (5). Two particular functions of this sort are central to this investigation. These will be called $\mathcal{L}$ and $\mathcal{R}$, for "left" and "right." Here are the definitions:

$$\begin{aligned} \mathcal{L} = \frac{2}{\eta^2}\big[&(1-x_1)y_1(c_1^2-1) + (1-x_2)y_2(c_2^2-1) + \\ &(1-x_3)y_3(c_3^2-1) + Y_H(1-c_1c_2c_3)\big], \end{aligned} \tag{6}$$

$$\begin{aligned} \mathcal{R} = \frac{2}{\eta^2}\big[&(1+x_1)x_1(c_1^2-1) + (1+x_2)x_2(c_2^2-1) + \\ &(1+x_3)x_3(c_3^2-1) + (1+X_H)(1-c_1c_2c_3)\big]. \end{aligned}$$

The reason for choosing these two functions will become clearer. This begins with the following lemma.

Lemma 11. $\mathcal{L}$, $\mathcal{R}$ *and the constant function 1 are all* $\mathcal{S}_3$*-invariant, where* $\mathcal{S}_3$ *here acts by permuting the subscripts on the c's, x's and y's. Additionally, they are all of the form (4) and satisfy conditions (5). Linear combinations of these functions (with constant coefficients) are also of the form (4) and satisfy (5). In fact, all functions of the form (4) that satisfy (5) are uniquely expressible in this way.*

Proof. Treating the constant function 1 as η^2/η^2, it is easy to check that $\mathcal{L}$, $\mathcal{R}$ and 1 are all of the form (4), and they are clearly $\mathcal{S}_3$-invariant. Let us now establish condition (5) for each of these three functions. In the case of the constant function 1, this can quickly be checked directly using only part *(i)* of Lemma 5.

In the case of $\mathcal{L}$, $\beta = -Y_H$, $\gamma_1 = (1-x_1)y_1$, $\gamma_2 = (1-x_2)y_2$ and $\gamma_3 = (1-x_3)y_3$. The first of the two required conditions is clearly satisfied, that is, $\alpha = -\beta-\gamma_1-\gamma_2-\gamma_3$. The other condition requires some effort to establish. One way to proceed uses the following observation about symmetric polynomials: $y_1^2(y_2+y_3) + y_2^2(y_3+y_1) + y_3^2(y_1+y_2) = Y_H^3/3 - Y_H(y_1^2+y_2^2+y_3^2)/3 + Y_H(y_2y_3+y_3y_1+y_1y_2)/3 - 3y_1y_2y_3$.

By applying Lemma 2, this reduces to $Y_H(3 + 4X_H + X_H^2 + Y_H^2)/4$. When this amount is subtracted from $\iota\beta + d_1^2\gamma_1 + d_2^2\gamma_2 + d_3^2\gamma_3$, the difference is seen to equal $-Y_H - 2(x_1y_1 + x_2y_2 + x_3y_3) + (x_1x_2 + x_1x_3 + x_2x_3)(y_1+y_2+y_3) + 2x_1x_2x_3(y_1+y_2+y_3) - 2(y_1x_2x_3 + x_1y_2x_3 + x_1x_2y_3) + 2y_1y_2y_3X_H - 3y_1y_2y_3$. Applying Lemma 2 again, reduces this to $-Y_H(3 + 4X_H + X_H^2 + Y_H^2)/4$. Therefore $\iota\beta + d_1^2\gamma_1 + d_2^2\gamma_2 + d_3^2\gamma_3$ is zero.

In the case of $\mathcal{R}$, $\beta = -1 - X_H$, $\gamma_1 = (1+x_1)x_1$, $\gamma_2 = (1+x_2)x_2$ and $\gamma_3 = (1+x_3)x_3$. Again, the first condition is immediate, but the second condition requires a few manipulations. A possible approach here we makes use of the fact that $x_1^2y_2y_3 + x_2^2y_3y_1 + x_3^2y_1y_2 = (1-y_1^2)y_2y_3 + (1-y_2^2)y_3y_1 + (1-$

$y_3^2)y_1y_2 = y_2y_3 + y_3y_1 + y_1y_2 - y_1y_2y_3Y_H = (-3 - 2X_H + X_H^2 + 3Y_H^2 + 2X_HY_H^2)/4$. When -2 times this amount is subtracted from $\iota\beta + d_1^2\gamma_1 + d_2^2\gamma_2 + d_3^2\gamma_3$, the difference is seen to equal $(1+X_H)(y_2y_3+y_3y_1+y_1y_2) - 2(x_1y_2y_3+x_2y_3y_1+x_3y_1y_2) - 2x_1x_2x_3X_H + (1+X_H)(x_2x_3+x_3x_1+x_1x_2) + 2(x_1^2+x_2^2+x_3^2) - 6x_1x_2x_3 - X_H - 3 = (-3 - 2X_H + X_H^2 + 3Y_H^2 + 2X_HY_H^2)/2$. Therefore $\iota\beta + d_1^2\gamma_1 + d_2^2\gamma_2 + d_3^2\gamma_3$ is zero.

The constant function 1 satisfies (5) because $2\iota = d_1^2+d_2^2+d_3^2$. Properties (4) and (5) having been established for all three functions, the linear nature of these properties makes it clear that they also hold for linear combinations of the functions. Since these functions are easily seen to be linearly independent, together they span a three-dimensional space of functions. But the space of all functions of the form (4) is five-dimensional, and condition (5) imposes two linearly independent restrictions, so the space of all functions of the form (4) that satisfy (5) is three-dimensional. Therefore, this space must be the space having $\mathcal{L}$, $\mathcal{R}$ and 1 as a basis.

□

Using (1) and (2), we will often regard that c_1, c_2 and c_3 are expressed as functions of x, y and z. In this way, any quantity expressible as a function of c_1, c_2 and c_3 is also expressible as a function of x, y and z (by composing functions). This is particularly true of $\mathcal{L}$ and $\mathcal{R}$. This thinking is behind Theorem 1, below. It indicates precisely how the quantities $\mathcal{L}$ and $\mathcal{R}$, which were defined in terms of the parameters of Grunert's system, can also be expressed simply in terms of the coordinates of any of its solution points. In Section 4, the knowledge gained here concerning $\mathcal{L}$ and $\mathcal{R}$ is applied to obtain a better understanding of the discriminant polynomial associated with Grunert's system.

The first proof of Theorem 1 offered here is purely algebraic, with the algebra $\mathcal{A}$ used to simplify the computations involved. Nevertheless, the computations are still tedious and greatly benefit from the use of symbolic manipulation software. There is however another known way to prove the theorem. Corollary 1 is shown to follow from the theorem by simply applying a certain linear transformation. However, this corollary is really just a restatement of Theorem 1 in (Rieck, 2014). Since the latter was previously proven via a geometric argument, one can simply invert the linear transformation to obtain a rapid proof of Theorem 1 here, based on Theorem 1 there. This alternative proof of Theorem 1 is outlined following Corollary 1.

Theorem 1. *The quantities $\mathcal{L}$ and $\mathcal{R}$ defined in (6) in terms of c_1, c_2 and c_3, are also equal to the following expressions involving x, y and z instead:*

$$\mathcal{L} = \big[2(1+x)y - (1+X_H)y - Y_H(1+x)\big]\cdot (x^2+y^2-1)\,z^{-2} \;+\; 2(1+x)y$$

and

$$\mathcal{R} = \big[y^2-(1-x)^2+(1-X_H)(1-x)-Y_Hy\big]\cdot (x^2+y^2-1)\,z^{-2} \;+\; y^2-(x-1)^2.$$

In the special case of an equilateral triangle, where $X_H = Y_H = 0$, these simplify as follows:

$$\mathcal{L} = 2(1+x)y + (1+2x)\,y\,(x^2+y^2-1)\,z^{-2}$$

and $\mathcal{R}$ =

$$y^2-(1-x)^2 + (y^2-x^2+x)\,(x^2+y^2-1)\,z^{-2}.$$

First proof of Theorem 1. Similar to Lemma 7, only a sketch of a proof is provided here. This should supply adequate guidance to allow the motivated reader to fill in the details. The polynomial manipulations are difficult (but not impossible) to do by hand. They are much easier to perform with the aid of a symbolic manipulation software system.

By Lemma 5, we can express r_1^2, r_2^2, r_3^2, $r_2^2+r_3^2-d_1^2$, $r_3^2+r_1^2-d_2^2$ and $r_1^2+r_2^2-d_3^2$ as polynomials in x, y and z, with the coefficients being polynomials in x_1, x_2, x_3, y_1, y_2 and x_3. Define

$$\begin{aligned}
N_0 &= 8r_1^2r_2^2r_3^2\\
N_1 &= 2r_1^2(r_2^2+r_3^2-d_1^2)^2\\
N_2 &= 2r_2^2(r_3^2+r_1^2-d_2^2)^2\\
N_3 &= 2r_3^2(r_1^2+r_2^2-d_3^2)^2\\
N_4 &= (r_2^2+r_3^2-d_1^2)(r_3^2+r_1^2-d_2^2)(r_1^2+r_2^2-d_3^2).
\end{aligned}$$

By Lemma 5 again, for $i = 1,2,3$, we have $c_i^2\,/\,\eta^2 = N_i\,/\,(2d_1^2d_2^2d_3^2z^2)$. Likewise, $c_1c_2c_3\,/\,\eta^2 = N_4\,/\,(2d_1^2d_2^2d_3^2z^2)$. Letting $\mathcal{L}_0 = d_1^2d_2^2d_3^2\,z^2\mathcal{L}$, we see from (1), (6), and Lemma 1 that

$$\mathcal{L}_0 = (x_1y_1+x_2y_2+x_3y_3)\,N_0 + (1-x_1)y_1\,N_1 + (1-x_2)y_2\,N_2 + (1-x_3)y_3\,N_3 - Y_H\,N_4.$$

Likewise, letting $\mathcal{R}_0 = d_1^2 d_2^2 d_3^2 z^2 \mathcal{R}$, we see that

$$\mathcal{R}_0 = (1 - x_1^2 - x_2^2 - x_3^2) N_0 + (1 + x_1) x_1 N_1 \\ + (1 + x_2) x_2 N_2 + (1 + x_3) x_3 N_3 - (1 + X_H) N_4.$$

Note that $N_0, N_1, N_2, N_3, N_4, \mathcal{L}_0$ and $\mathcal{R}_0$ can also be expressed as polynomials in x, y and z, with the coefficients being polynomials in x_1, x_2, x_3, y_1, y_2 and x_3. So these equations can be expressed as polynomial equations, and it suffices to compare corresponding coefficients in order to prove the theorem. However, the manipulations involved in doing this directly, including repeated applications of Lemma 1, are cumbersome. The algebra $\mathcal{A}$ introduced in the previous section, and particularly Lemma 5, can be employed to reduce the labor involved.

Lemma 7 expresses N_0 and N_4 (apart from a constant factor for N_0) in terms of X_H and Y_H. Because of symmetry, expressing N_1 in terms of $X_H, Y_H, X_1, X_2, X_3, Y_1, Y_2, Y_3$ and Δ, automatically provides similar expressions for N_2 and N_3. Now, $N_1 = 2r_1^2(r_2^2 + r_3^2 - d_1^2)^2$, and Lemma 6 provides expansions for the irreducible factors of this, namely r_1^2 and $r_2^2 + r_3^2 - d_1^2$. If we again let p_1 and q_1 be the invariant part and the annihilated part of r_1^2, respectively, and now let p_1' and q_1' be the similar parts of $r_2^2 + r_3^2 - d_1^2$, then $N_1 = 2(p_1 + q_1)(p_1' + q_1')^2$. Reasoning along the lines of the proof of Lemma 7 will then provide an expansion for N_1.

In fact, proceeding in this way, it is seen that the invariant part (*i.e.* $\mathcal{A}^{\mathcal{G}}$-component) and the annihilated part (*i.e.* $\mathcal{N}$-component) of $3(r_2^2 + r_3^2 - d_1^2)^2$ are as specified in the appendix. Moving forward with this information, one next discovers that the invariant and annihilated parts of $3r_1^2(r_2^2 + r_3^2 - d_1^2)^2$ are as given in the appendix.

Now, using Lemmas 2 and 6, we have $\mathcal{L}_0 = (X_H + 1) Y_H N_0 + {}^1\!/_6[(3 - X_H) Y_1 - Y_H X_1 - 2X_H Y_H] N_1 + {}^1\!/_6[(3 - X_H) Y_2 - Y_H X_2 - 2X_H Y_H] N_2 + {}^1\!/_6[(3 - X_H) Y_3 - Y_H X_3 - 2X_H Y_H]\, N_3 - Y_H N_4$. Likewise, $\mathcal{R}_0 = {}^1\!/_2(Y_H^2 - (X_H - 1)^2) N_0 + {}^1\!/_6[(3 + X_H) X_1 - Y_H Y_1 + X_H^2 - Y_H^2 + 3] N_1 + {}^1\!/_6[(3 + X_H) X_2 - Y_H Y_2 + X_H^2 - Y_H^2 + 3] N_2 + {}^1\!/_6[(3 + X_H) X_3 - Y_H Y_3 + X_H^2 - Y_H^2 + 3]\, N_3 - (1 + X_H) N_4$.

But, ${}^1\!/_6[(3 - X_H) Y_1 - Y_H X_1 - 2X_H Y_H] N_1 + {}^1\!/_6[(3 - X_H) Y_2 - Y_H X_2 - 2X_H Y_H] N_2 + {}^1\!/_6[(3 - X_H) Y_3 - Y_H X_3 - 2X_H Y_H] N_3$ is clearly invariant, and indeed is just three times the invariant part of ${}^1\!/_6[(3 - X_H) Y_1 - Y_H X_1 - 2X_H Y_H] N_1$. This helps to simplify the computations by using reasoning similar to the proof of Lemma 7. Similarly for ${}^1\!/_6[(3 + X_H) X_1 - Y_H Y_1 + X_H^2 - Y_H^2 + 3]\, N_1 + {}^1\!/_6[(3 + X_H) X_2 - Y_H Y_2 + X_H^2 - Y_H^2 + 3]\, N_2 + {}^1\!/_6[(3 + X_H) X_3 - Y_H Y_3 + X_H^2 - Y_H^2 + 3]\, N_3$.

For both $\mathcal{L}_0$ and $\mathcal{R}_0$, the N_0 and N_4 parts are simply a matter of multiplying the expressions already achieved for N_0 and N_4 by a couple of simple elements of $\mathcal{A}^{\mathcal{G}}$, and so these present no further difficulties. After obtaining suitable expressions for the parts of both $\mathcal{L}_0$ and $\mathcal{R}_0$, and using the expansion for d_i^2 in Lemma 6, the equations in the theorem can immediately be confirmed.

□

Theorem 1 from (Rieck, 2014) will now be essentially restated, as Corollary 1 here, and proved. However, the proof here does not rely on the somewhat intricate geometric details found in (Rieck, 2014). Instead, if follows rapidly from the present article's Theorem 1, by means of an invertible linear transformation.

Corollary 1. *A certain rational expression of d_1, d_2, c_1, c_2 and c_3 can be expressed in terms of x, y, z and the control points coordinates as follows:*

$$[d_1^2(1 - c_2^2) - d_2^2(1 - c_1^2)]\,/\,\eta^2 = \frac{2(1+x_3)y_3}{y_1 - y_2} + \frac{y_3}{y_1 - y_2}\left[y^2 - (x-1)^2\right] + \frac{1+x_3}{y_1 - y_2}\left[2(x+1)y\right] + \left\{ \frac{(1+2x_3)y_3}{y_1 - y_2} - \frac{y_3}{y_1 - y_2}\left(y^2 - x^2\right) - \frac{1+x_3}{y_1 - y_2}(2xy) + \frac{2(x_1+x_2)x_3}{y_1 - y_2} y - \frac{2(x_1+x_2)y_3}{y_1 - y_2} x \right\} \cdot \frac{1 - x^2 - y^2}{z^2}.$$

(Similar expressions result by permuting the indices.)

Proof. The function on the left side of the equation is another example of a function of the form (4) that satisfies (5). The parameters for this are $\alpha = d_1^2 - d_2^2$, $\beta = 0$, $\gamma_1 = d_2^2$, $\gamma_2 = -d_1^2$ and $\gamma_3 = 0$. It is possible to write it as a linear combination of $\mathcal{L}$, $\mathcal{R}$ and 1, specifically, as $[(1 + x_3)\mathcal{L} + y_3 \mathcal{R} + 2(1 + x_3) y_3]/(y_1 - y_2)$. To see this, note the following equalities of vectors:

$$2(1 + x_3)\,\big((1 - x_1) y_1,\ (1 - x_2) y_2,\ (1 - x_3) y_3,\ -(y_1 + y_2 + y_3)\big) + 2y_3\,\big((1 + x_1) x_1,\ (1 + x_2) x_2,\ (1 + x_3) x_3, -(1 + x_1 + x_2 + x_3)\big) + 2(1 + x_3) y_3\,\big(-1,\ -1,\ -1,\ 2\big) =$$

$$2\,\big(y_1 - y_3 - x_1 y_1 - x_3 y_3 + x_1 y_3 + x_3 y_1 - x_1 x_3 y_1 + x_1^2 y_3,\ y_2 - y_3 - x_2 y_2 - x_3 y_3 + x_2 y_3 + x_3 y_2 - x_2 x_3 y_2 + x_2^2 y_3,\ 0,\ -y_1 - y_2 - x_1 y_3 - x_2 y_3 - x_3 y_1 - x_3 y_2\big)$$

$= 2(y_1 - y_2)\,\big(1 - x_1x_3 - y_1y_3,\ -1 + x_2x_3 + y_2y_3,\ 0,\ 0\big) \;=\; (y_1 - y_2)\,\big(d_2^2,\ -d_1^2,\ 0,\ 0\big).$

The middle equality follows from Lemma 1. It then follows from (6) that $[(1+x_3)\,\mathcal{L} + y_3\,\mathcal{R} + 2(1+x_3)y_3)]\,/\,(y_1 - y_2) \;=\; [d_1^2(1-c_2^2) - d_2^2(1-c_1^2)]\,/\,\eta^2$. By Theorem 1, this also equals

$$\begin{aligned}&\big[\,(1+x_3)\,\{[(X_H-1)y + Y_H(1+x) - 2xy]\cdot\\&(1-x^2-y^2)z^{-2} + 2(1+x)y\} \;+\; y_3\cdot\\&\{\,[\,(x-(X_H+1)/2)^2 - (y - Y_H/2)^2 +\\&(Y_H^2 - (X_H-1)^2)/4\,]\,(1-x^2-y^2)z^{-2} +\\&(y^2 - (x-1)^2)\,\} + 2(1+x_3)y_3\,\big]\,/\,(y_1-y_2).\end{aligned}$$

The right-hand side of the corollary can be produced from this by applying Lemma 1 a couple times.

□

Unlike the functions $\mathcal{L}$ and $\mathcal{R}$ which are invariant under the group action of $\mathcal{S}_3$, *i.e.* under permutations of the subscripts, the function in Corollary 1 does not have this symmetry, and in fact is antisymmetric when interchanging the indices 1 and 2. By cycling the indices, two similar functions are produced. While the sum of all three of these functions is zero, any two of them are linearly independent. Let us refer to these functions as "the triplet." Since each of these is obtained as a linear combination of $\mathcal{L}$, $\mathcal{R}$ and 1, it is reasonable to ask about inverting this process. This is indeed possible, and allows for a way to prove Theorem 1 in this paper based on Theorem 1 in (Rieck, 2014).

Second proof of Theorem 1. Theorem 1 in (Rieck, 2014) can be rewritten in the form presented in Corollary 1 here. So this corollary is valid. Examining the proof that was presented for this corollary, and considering the symmetry involved in permuting indices here, it is easily discovered that there is a certain two-by-two constant matrix and a certain constant vector, such that two of the triplet functions can be obtained by multiplying the constant matrix by the vector $[\mathcal{L}\ \mathcal{R}]^T$ and then adding the constant vector. Using a particular pair from the triplet, it is seen this matrix is

$$\begin{bmatrix} \frac{1+x_2}{y_3-y_1} & \frac{y_2}{y_3-y_1} \\ \frac{1+x_3}{y_1-y_2} & \frac{y_3}{y_1-y_2} \end{bmatrix}.$$

The determinant of this is generically non-zero, and so the process can be reversed to obtain $\mathcal{L}$ and $\mathcal{R}$ from two of the functions in the triplet. Using the fact that Corollary 1 has already been established, the reasoning in the above proof can now simply be "run in reverse" to establish the current paper's Theorem 1.

□

4 THE DISCRIMINANT AND A RELATED RATIONAL FUNCTION

All of the many algebraic approaches to truly solving Grunert's system, including Grunert's original approach (Grunert, 1841), invariably involve a quartic polynomial with quite a complicated discriminant. This discriminant is *not* what we shall mean when speaking about the discriminant of the system (1), though the latter is necessarily a factor of the former. Both are polynomials in the parameters of the system, but the discriminant of the system has these important properties:

(i) it vanishes for specific values of the parameters if and only if these parameters produce a repeated solution to the system,

(ii) its sign provides useful information as to the number of real-valued solutions to the system,

(iii) it is square-free, meaning that no non-constant irreducible factor occurs with a multiplicity greater than one.

In the case of Grunert's system, it is known that a solution point (*i.e.* a point (x,y,z) that solves the combined system (1) and (2)) is a repeated solution point if and only if it is on the danger cylinder, as already mentioned. Expressed in terms of r_1, r_2 and r_3, this condition can be stated algebraically (Rieck, 2011) as follows:

$$\begin{aligned}&d_1^2d_2^2d_3^2 + (d_2^2+d_3^2-r_1^2)r_2r_3 + (d_3^2+d_1^2-r_2^2)r_3r_1\\&+ (d_1^2+d_2^2-r_3^2)r_1r_2 - d_1^2r_1^4 - d_2^2r_2^4 - d_3^2r_3^4 \;=\; 0.\end{aligned} \tag{7}$$

Expressed in Cartesian coordinates, this becomes $4\Delta\cdot(x^2+y^2-1) \;=\; 0$, but since the control points triangle is non-degenerate, this just means $x^2+y^2 = 1$, which is the equation of the danger cylinder. The discriminant of Grunert's system can intuitively be understood as starting with the Grunert system (1), supplementing this system with equation (7), and then eliminating r_1, r_2 and r_3 from the resulting system of four equations to obtain a polynomial in the parameters of (1), *i.e.* the c's and d's.

However, to avoid certain technical complications, it has long been recognized that it is better to work with homogeneous equations when performing such eliminations. For this purpose, let us introduce a new variable s, and regard that the original r_i are to be replaced with r_i/s $(i = 1,2,3)$. The augmented and homogeneous system of equations is thus

$$\begin{cases} r_2^2 + r_3^2 - 2c_1 r_2 r_3 - d_1^2 s^2 = 0 \\ r_3^2 + r_1^2 - 2c_2 r_3 r_1 - d_2^2 s^2 = 0 \\ r_1^2 + r_2^2 - 2c_3 r_1 r_2 - d_3^2 s^2 = 0 \\ d_1^2 d_2^2 d_3^2 s^4 + (d_2^2 s^2 + d_3^2 s^2 - r_1^2) r_2 r_3 \\ + (d_3^2 s^2 + d_1^2 s^2 - r_2^2) r_3 r_1 \\ + (d_1^2 s^2 + d_2^2 s^2 - r_3^2) r_1 r_2 \\ - d_1^2 r_1^4 - d_2^2 r_2^4 - d_3^2 r_3^4 = 0. \end{cases} \tag{8}$$

As a system of four homogeneous equations in four variables (r_1, r_2, r_3 and s), this system is well suited to the elimination of these variables using a multi-polynomial resultant (cf. Section 3.2 of (Cox et al., 2004)). This is an polynomial in the c's and the squares of the d's, but we will treat these differently, usually regarding the d's as fixed parameters having to do with the fixed positions of the control points.

On the other hand, we will want to think about a camera pose as changing dynamically. When we allow such movement, the c's will of course change. In this way, the multi-polynomial resultant will be regarded as a polynomial in c_1, c_2 and c_3, having coefficients that are polynomials in d_1^2, d_2^2 and d_3^2.

Apart from a nonzero constant factor (of no concern), there is a unique square-free polynomial of the same sort, whose non-constant irreducible factors are the same as those of the multi-polynomial resultant. It is this square-free polynomial that will here be referred to as the "discriminant" of Grunert's system (1). It vanishes if and only if (1) has a repeated solution, *i.e.* a solution point on the danger cylinder.

The coefficients of the c-monomials depend on the parameters for the control points. These coefficients could be expressed in terms of d_1^2, d_2^2 and d_3^2, but in the new approach for finding the discriminant, developed below, the parameters x_1, x_2, x_3, y_1, y_2 and y_3 will be used instead. Recall that $x_H = x_1 + x_2 + x_3$, $y_H = y_1 + y_2 + y_3$, and that Lemmas 4 and 5 shows how to express Δ, d_1^2, d_2^2 and d_3^2 in terms of x_1, x_2, x_3, y_1, y_2 and y_3. The new approach is quite different than the elimination approach just outlined. It is based on and expedited by Theorem 1, and also requires the following four lemmas.

Lemma 12. *For indeterminates L and R, let $D =$*

$$[L^2 + (R+1)^2]^2 + 8(R+1)[(R+1)^2 - 3L^2] + 18[L^2 + (R+1)^2] - 27.$$

(i) If $L = 2(x+1)y$ and $R = y^2 - (x-1)^2$, then $D = (x^2+y^2-1)^2\,[\,(x^2+y^2)^2 + 8x(3y^2 - x^2) + 18(x^2+y^2) - 27\,]$.

(ii) $y[\mathcal{L} - 2(x+1)y] - (x+1)[\mathcal{R} - y^2 + (x-1)^2] = (x - X_H)(x^2+y^2-1)^2/z^2$.

(iii) If $L = \mathcal{L}$ and $R = \mathcal{R}$, then the quantity D equals $p(X_H, Y_H; x, y, z^2)\,(x^2+y^2-1)^2\,z^{-8}$ for a polynomial $p(X_H, Y_H; x, y, z^2)$.

Proof. Items *(i)* and *(ii)* can be checked directly, using Theorem 1 for item *(ii)*. We turn now to *(iii)*. Using Theorem 1 again, treat $\mathcal{L}$ and $\mathcal{R}$ as polynomials in X_H, Y_H, x, y and z^{-2}. To simplify the presentation here, let L_0, L_1, R_0, R_1 and Γ respectively denote $2(x+1)y$, $2xy + (1 - X_H)y - Y_H(1+x)$, $y^2 - x^2 + 2x$, $y^2 - x^2 + (1 + X_H)x - Y_H y - X_H$ and $x^2 + y^2 - 1$. So $\mathcal{L} = L_0 + L_1 \Gamma z^{-2}$ and $\mathcal{R} + 1 = R_0 + R_1 \Gamma z^{-2}$.

We start by computing D modulo Γ^2, observing first that $\mathcal{L}^2 \equiv L_0^2 + 2L_0 L_1 \Gamma z^{-2}$ and $(\mathcal{R}+1)^2 \equiv R_0^2 + 2R_0 R_1 \Gamma z^{-2}$. From these we also get $(\mathcal{R}+1)^3 \equiv R_0^3 + 3R_0^2 R_1 \Gamma z^{-2}$ and $(\mathcal{R}+1)\mathcal{L}^2 \equiv R_0 L_0^2 + 2R_0 L_0 L_1 \Gamma z^{-2} + L_0^2 R_1 \Gamma z^{-2}$. Additionally we get, $\mathcal{L}^4 \equiv L_0^4 + 4L_0^3 L_1 \Gamma z^{-2}$, $(\mathcal{R}+1)^4 \equiv R_0^4 + 4R_0^3 R_1 \Gamma z^{-2}$, and $\mathcal{L}^2(\mathcal{R}+1)^2 \equiv L_0^2 R_0^2 + 2L_0^2 R_0 R_1 \Gamma z^{-2} + 2R_0^2 L_0 L_1 \Gamma z^{-2}$.

Thus, $D \equiv [(L_0^2 + R_0^2)^2 + 8R_0(R_0^2 - 3L_0^2) + 18(L_0^2 + R_0^2) - 27] + [4L_0^3 L_1 + 4R_0^3 R_1 + 4L_0^2 R_0 R_1 + 4R_0^2 L_0 L_1 + 24R_0^2 R_1 - 48R_0 L_0 L_1 - 24L_0^2 R_1 + 36L_0 L_1 + 36R_0 R_1] \cdot \Gamma z^{-2}$. The first part of this is congruent to zero, *i.e.* divisible by Γ^2, by part *(i)*. To establish that D is divisible by Γ^2, it suffices now to show that the above bracketed coefficient of Γz^{-2} is divisible by Γ. Now, item *(ii)* means that $yL_1 \equiv (x+1)R_1$, modulo Γ. So in computing the coefficient of Γz^{-2} modulo Γ, we can identify yL_1 with $(x+1)R_1$. Expanding L_0 and R_0 then leads quickly to zero (modulo Γ). Thus D equals Γ^2 times a polynomial in x, y and z^{-2}. A quick check reveals that the most negative power of z resulting from the computations is z^{-8}, and so part *(iii)* is now established.

□

Lemma 13. *The polynomial $p(X_H, Y_H\,; x, y, z^2)$ in Lemma 12, part (iii), is irreducible, as a polynomial in all five quantities, X_H, Y_H, x, y and z^2.*

Proof. The general $p(X_H, Y_H\,; x, y, z^2)$ is very complicated, so let us start with the version of this for the equilateral triangle case: $p(0, 0; x, y, z^2)$. By direct computation, this is seen to equal $[(x^2+y^2)^2+8x(3y^2-x^2)+18(x^2+y^2)-27]\,z^8 + 4[(x^2+y^2)^3+7x(3y^2-x^2)(x^2+y^2)+13(x^2+y^2)^2+3x(x^2-3y^2)-18(x^2+y^2)]\,z^6 + 2[3(x^2+y^2)^4+18x(3y^2-x^2)(x^2+y^2)^2+(3x^2+y^2)(9x^4+18x^2y^2+25y^4)+12x(x^2-3y^2)(x^2+y^2)-39(x^2+y^2)^2+6x(x^2-3y^2)+9(x^2+y^2)]\,z^4 + 4(x^2+y^2-1)[(x^2+y^2)^4-5x(x^2-3y^2)(x^2+y^2)^2+(7x^6+3x^4y^2+33x^2y^4+5y^6)-x(x^2-3y^2)(x^2+y^2)-6xy^2-4(x^2+y^2)^2+2x^3]\,z^2 + (x^2+y^2-1)^2[(x^2+y^2)^2+2x(3y^2-x^2)+(x^2+y^2)]^2$.

Now, z^8 is the highest power of z, and its coefficient is $(x^2+y^2)^2+8x(3y^2-x^2)+18(x^2+y^2)-27$. This is the polynomial for the standard deltoid, which is known to be irreducible. So if $p(0, 0; x, y, z^2)$ has a non-trivial factorization into a product of two non-constant factors, then, up to rescaling by constant factors, one of the non-constant factors must be either *(i)* $(x^2+y^2)^2+8x(3y^2-x^2)+18(x^2+y^2)-27$, or *(ii)* a monic polynomial in z^2 whose highest z-degree term is z^2, z^4 or z^6.

Since $(x^2+y^2)^2+8x(3y^2-x^2)+18(x^2+y^2)-27$ does not divide each of the coefficients powers of z in $p(0, 0; x, y, z^2)$, this rules out case *(i)*. To rule out case *(ii)*, set $x = 1$ and $y = 1$, to obtain the polynomial $p(0, 0; 1, 1, z) = 29z^8+184z^6+404z^4+352z^2+100$. This polynomial is easily seen to be irreducible over the field of rational numbers, which would be impossible in case *(ii)*. Therefore, $p(0, 0; x, y, z^2)$ is irreducible.

Similar reasoning is used now to move from $p(0, 0; x, y, z^2)$ to the general $p(X_H, Y_H\,; x, y, z^2)$. If $p(X_H, Y_H\,; x, y, z^2)$ had a non-constant factor that was just a polynomial of X_H and Y_H, then this would divide the coefficients for all of the monomials $x^i y^j z^{2k}$. But, as with $p(0, 0; x, y, z^2)$, direct inspection reveals that the coefficient of z^8 in $p(X_H, Y_H\,; x, y, z^2)$ is again $(x^2+y^2)^2+8x(3y^2-x^2)+18(x^2+y^2)-27$. Any other non-trivial factorization of $p(X_H, Y_H\,; x, y, z^2)$ would result in a non-trivial factorization of $p(0, 0; x, y, z^2)$. Since $p(0, 0; x, y, z^2)$ is irreducible, we must conclude that $p(X_H, Y_H\,; x, y, z^2)$ is also irreducible.

□

Lemma 14. *Let $P(u_0, u_1, u_2, u_3)$ be a non-constant polynomial in four indeterminates. Consider taking $P(c_1c_2c_3, c_1^2, c_2^2, c_3^2)$ and applying substitutions from Lemma 5 to convert this to a rational function in x, y and z, whose denominator is a positive power of the expansion of $8r_1^2r_2^2r_3^2$. Then the resulting numerator cannot be a polynomial that is independent of z.*

Proof. For brevity, let us here write R_1, R_2, R_3, $\bar{R_1}$, $\bar{R_2}$, and $\bar{R_3}$ to denote r_1^2, r_2^2, r_3^2, $r_2^2+r_3^2-d_1^2$, $r_3^2+r_1^2-d_2^2$ and $r_1^2+r_2^2-d_3^2$, respectively. So, c_1^2, c_2^2, c_3^2 and $c_1c_2c_3$ equal $2R_1\bar{R_1}^2$, $2R_2\bar{R_2}^2$, $2R_3\bar{R_3}^2$ and $\bar{R_1}\,\bar{R_2}\,\bar{R_3}$, respectively, divided by $8R_1R_2R_3$. Assume the claim is false and let P be a counterexample. By applying the Reynolds operator (3) to P, we obtain an $\mathcal{S}_3$-invariant counterexample. Without loss of generality, assume this equals P. So P can be written as an $\mathcal{S}_3$-invariant, homogeneous polynomial in $R_1, R_2, R_3, \bar{R_1}$, $\bar{R_2}$ and $\bar{R_3}$, divided by a positive power of $8R_1R_2R_3$. We are assuming that when this is expanded in terms of x, y and z, and simplified, the resulting numerator is independent of z.

We now show that there are no non-constant, $\mathcal{S}_3$-invariant, homogeneous polynomials in R_1, R_2, R_3, $\bar{R_1}$, $\bar{R_2}$ and $\bar{R_3}$ that are independent of z, upon making the substitutions. To do so, assume that $p(R_1, R_2, R_3, \bar{R_1}, \bar{R_2}, \bar{R_3})$ is a counterexample, and expand it to obtain a polynomial $\hat{p}(x, y, Z)$, with $Z = z^2$. Now look at $\hat{p}(0, 0, Z)$; that is, set $x = 0$ and $y = 0$. We will show that this is not a constant, that is, it must depend on Z.

By Lemma 6, upon setting $x = 0$ and $y = 0$, the quantities R_1, R_2, R_3, $\bar{R_1}$, $\bar{R_2}$ and $\bar{R_3}$ become respectively $Z+1$, $Z+1$, $Z+1$, $2Z-1+{}^1\!/\!_3(X_H^2+Y_H^2-2X_HX_1-2Y_HY_1)$, $2Z-1+{}^1\!/\!_3(X_H^2+Y_H^2-2X_HX_2-2Y_HY_2)$ and $2Z-1+{}^1\!/\!_3(X_H^2+Y_H^2-2X_HX_3-2Y_HY_3)$. Call the last three of these $\lambda_1(Z)$, $\lambda_2(Z)$ and $\lambda_3(Z)$. $\hat{p}(0, 0, Z)$ is expressible as a homogeneous polynomial in $Z+1$, $\lambda_1(Z)$, $\lambda_2(Z)$ and $\lambda_3(Z)$.

Now, being $\mathcal{S}_3$-invariant, $\hat{p}(0, 0, Z)$ is also expressible as a polynomial in $Z+1$, $\lambda_1(Z)+\lambda_2(Z)+\lambda_3(Z)$, $\lambda_2(Z)\lambda_3(Z)+\lambda_3(Z)\lambda_1(Z)+\lambda_1(Z)\lambda_2(Z)$ and $\lambda_1(Z)\lambda_2(Z)\lambda_3(Z)$. These equal $Z+1$, $6Z+X_H^2+$

Y_H^2-3, $12Z^2+4(X_H^2+Y_H^2-3)Z+(2X_H^3-6X_HY_H^2-5X_H^2-5Y_H^2+3)$ and $8Z^3+4(X_H^2+Y_H^2-3)Z^2+2(2X_H^3-6X_HY_H^2-5X_H^2-5Y_H^2+3)Z+(X_H^4+Y_H^4+2X_H^2Y_H^2-4X_H^3+12X_HY_H^2+4X_H^2+4Y_H^2-1)$.

Calling these polynomials $q_0(Z)$, $q_1(Z)$, $q_2(Z)$ and $q_3(Z)$, respectively, a Groebner basis with slack variables can be generated to establish that there is only one syzygy relating these polynomials: $12q_0(Z)^2-8q_0(Z)^3-4q_0(Z)q_1(Z)+4q_0(Z)^2q_1(Z)-q_1(Z)^2+4q_2(Z)-2q_0(Z)q_2(Z)+q_3(Z)=0$. That is, the polynomial $12Q_0^2-8Q_0^3-4Q_0Q_1+4Q_0^2Q_1-Q_1^2+4Q_2-2Q_0Q_2+Q_3$ in indeterminates Q_0, Q_1, Q_2 and Q_3 generates a principal ideal containing all of the polynomials in these for which the substitutions $Q_i \to q_i(Z)$ $(i=0,1,2,3)$ result in zero. By adding constants to these, we obtain the set S containing all of the polynomials in Q_0, Q_1, Q_2 and Q_3 that have the property that they become constant upon making the substitutions.

We are seeking a nonconstant polynomial in $Z+1$, $\lambda_1(Z)+\lambda_2(Z)+\lambda_3(Z)$, $\lambda_2(Z)\lambda_3(Z)+\lambda_3(Z)\lambda_1(Z)+\lambda_1(Z)\lambda_2(Z)$ and $\lambda_1(Z)\lambda_2(Z)\lambda_3(Z)$ such that the corresponding polynomial in Q_0, Q_1, Q_2 and Q_3 is a member of S. However, since $\hat{p}(0,0,Z)$ is homogeneous as a polynomial in $Z+1$, $\lambda_1(Z)$, $\lambda_2(Z)$ and $\lambda_3(Z)$, the polynomial in Q_0, Q_1, Q_2 and Q_3 must be weighted homogeneous with respect to assigning weights 1, 1, 2 and 3 to Q_0, Q_1, Q_2 and Q_3, respectively. Since each polynomial in S is of the form $12Q_0^2-8Q_0^3-4Q_0Q_1+4Q_0^2Q_1-Q_1^2+4Q_2-2Q_0Q_2+Q_3$ times another polynomial in Q_0, Q_1, Q_2 and Q_3, plus a constant, this quickly yields a contradiction.

□

Lemma 15. *Let R be a unique factorization domain. Let R' denote the polynomial ring $R[c_1,c_2,c_3]$ (with c_1, c_2 and c_3 being treated as indeterminates here). Let R'' be the subring $R[c_1^2,c_2^2,c_3^2,c_1c_2c_3]$, consisting of polynomials generated by c_1^2, c_2^2, c_3^2 and $c_1c_2c_3$. For $q \in R'$, the following are equivalent:*

- *(i) $q \in R''$*
- *(ii) q is left invariant whenever any two of c_1, c_2 and c_3 are negated*
- *(iii) each monomial in q involves powers of c_1, c_2 and c_3 that are all even or all odd.*

Moreover, if $q \in R''$, irreducible as an element of R'', but reducible as an element of R', then, writing $q=q_1\cdots q_k$ $(k>1)$ with non-constant irreducible $q_1,\ldots,q_k \in R'$, either $k=2$ or $k=4$, and up to scaling by constant factors, each q_i can be converted to any other q_j by negating a pair of c's. (A "constant" here means any element of R.)

In the $k=2$ case, q_1 and q_2 are each invariant under the negation of some pair of c's, but are interchanged under the other two negations of pairs of c's, up to constant factors. In the $k=4$ case, no q_i is fixed by any of the negations of a pair of c's, and in fact, these negations map q_i to the three other q_j, up to constant factors.

Proof. The first claim is easily established by checking that (i) implies (ii), (ii) implies (iii) and (iii) implies (i). Now assume the hypotheses of the second claim. Since R is a UFD, so too is R'. The three transformations that each negate a pair of c's, along with the identity transformation, form a Klein 4-group acting on R'. It also acts on the set $\{1,2,\ldots,k\}$, where a transformation that negates two of the c's takes i to j, if the polynomial q_i is transformed into the polynomial q_j, up to a constant factor.

Consider the orbit of $1 \in \{1,2,\ldots,k\}$ under this group action. The product q' of all of the q_i as i ranges over this orbit is a non-constant factor of q, and is invariant under the group action, and so is an element of R''. Up to a constant factor, q/q' is also an element of R''. By the irreducibility assumption concerning q in R'', we must have $q'=q$, up to a constant factor. The group action is therefore transitive on $\{1,2,\ldots,k\}$, which is only possible if $k=2$ or $k=4$. The rest of the lemma follows quickly by considering the possible actions of the group.

□

We are now ready to state and prove one of the principal results of this work.

Theorem 2. *Up to a constant factor, the discriminant of Grunert's system is the following $\mathcal{S}_3$-invariant polynomial in c_1, c_2 and c_3, whose coefficients are polynomials in x_1, x_2, x_3, y_1, y_2, y_3:*

$$\{\, [\mathcal{L}^2+(\mathcal{R}+1)^2]^2+18[\mathcal{L}^2+(\mathcal{R}+1)^2] + 8(\mathcal{R}+1)[(\mathcal{R}+1)^2-3\mathcal{L}^2]-27 \,\}\, \eta^8.$$

It has degree twelve as a polynomial in c_1, c_2 and c_3, and is irreducible. It can be regarded as a polynomial *in c_1^2, c_2^2, c_3^2 and $c_1c_2c_3$, and can also be expressed as a rational function of x, y and z^2.*

Proof. Using (6), and upon cancelling, the quantity in the theorem becomes a polynomial in $x_1, y_1, x_2, y_2, x_3, y_3, c_1^2, c_2^2, c_3^2$ and $c_1c_2c_3$, because each term inside the curly braces is a rational function whose denominator is a power of η^2, not exceeding $(\eta^2)^4$. However, we will mostly need to regard this as a polynomial $q = q(c_1, c_2, c_3)$ in c_1, c_2 and c_3, whose coefficients are polynomials in x_1, y_1, x_2, y_2, x_3 and y_3.

Upon making the substitutions suggested by part *(ix)* of Lemma 5, q can also be expressed as a rational function of x, y and z^2, whose coefficients are polynomials in x_1, y_1, x_2, y_2, x_3 and y_3. By Lemmas 12 and 13, and part *(iv)* of Lemma 5, we see that this equals

$$\frac{p(X_H, Y_H\,; x, y, z^2)\,(x^2+y^2-1)^2\,d_1^8 d_2^8 d_3^8}{256\,r_1^8 r_2^8 r_3^8},$$

where $p(X_H, Y_H\,; x, y, z^2)$ is the irreducible polynomial in Lemmas 12 and 13, and where r_1^8 is written here in place of $[(x-x_1)^2+(y-y_1)^2+z^2]^4$, and similarly for r_2^8 and r_3^8.

Since the numerator is divisible by x^2+y^2-1, it follows that q vanishes whenever values for the c's, used as parameters for the combined system (1) and (2), result in a solution point (x,y,z) that satisfies $x^2+y^2=1$. This of course means that (x,y,z) is a multiple solution point. But the discriminant of Grunert's system (1) vanishes if and only if this is so. We see that when this discriminant vanishes, so too does q. Since the discriminant is square-free, it must divide q, up to "constant" factors, *i.e.* factors that are polynomials in the parameters x_1, x_2, x_3, y_1, y_2 and y_3.

Notice next that q is irreducible as a polynomial of c_1^2, c_2^2, c_3^2, and $c_1c_2c_3$. If it could be factored nontrivially as such polynomials, then one of its factors, upon making the (x,y,z)-substitutions would become a rational function whose numerator would be divisible by $p(X_H, Y_H\,; x, y, z^2)$, and another factor that would become a rational function whose numerator would be divisible by x^2+y^2-1, and be independent of z. But Lemma 14 ensures that this cannot happen.

As a polynomial in c_1, c_2 and c_3, the quantity q can be seen to have total degree 12. To see this, we just need to look at the contributions of the various $c_1c_2c_3$ terms to $[\mathcal{L}^2\eta^4 + (\mathcal{R}\eta^2+\eta^2)^2]^2 + 18[\mathcal{L}^2\eta^4 + (\mathcal{R}\eta^2+\eta^2)^2]\eta^4 + 8(\mathcal{R}\eta^2+\eta^2)[(\mathcal{R}\eta^2+\eta^2)^2 - 3\mathcal{L}^2\eta^4]\eta^2 - 27\eta^8$, so as to produce a $c_1^4c_2^4c_3^4$ term that does not vanish. In fact, this term is $\{[(-2Y_H)^2 + ((-2(1+X_H))+(2))^2]^2 + 18[(-2Y_H)^2 + ((-2(1+X_H))+(2))^2](2)^2 + 8((-2(1+X_H))+(2))[(-2(1+X_H))+(2))^2 - 3(-2Y_H)^2](2) - 27(2)^4\}\,c_1^4c_2^4c_3^4 = -64\Delta^2\,c_1^4c_2^4c_3^4 \neq 0$.

A further consequence of this unique highest degree term concerns the possible reducibility of q as a polynomial in c_1, c_2 and c_3 whose coefficients are polynomials in x_1, x_2, x_3, y_1, y_2 and y_3. Suppose that q is reducible, and let $q_1\cdots q_k$, with $k>1$, be the factorization of q into irreducible polynomials, which is unique up to scaling by constants. By Lemma 15, $k=2$ or $k=4$, and the q_i are permuted transitively, up to constant factors, by the Klein 4-group. Since q is $\mathcal{S}_3$-invariant, *i.e.* invariant under simultaneously permuting the indices of the c's, x's and y's (in the same way), the case $k=2$ is not possible, basically because the symmetric nature of the indices would break down in such a factorization.

We are left with the possibility that $q = q_1q_2q_3q_4$, and that the Klein 4-group permutes the factors transitively, up to scaling by constant factors. Begin here by noting that Δ is a polynomial in x_1, x_2, x_3, y_1, y_2 and y_3 (see Lemma 4), but this does not divide q, as polynomials in x_1, x_2, x_3, y_1, y_2, y_3, c_1, c_2 and c_3. Now, since the only highest c-degree term of q is $-64\Delta^2\,c_1^4c_2^4c_3^4$, it is readily checked that we may assume, after rescaling, that the highest c-degree term of each q_i is $\sqrt{\Delta}\,c_1c_2c_3$, and that $q = -64\,q_1q_2q_3q_4$. But this is impossible since $\sqrt{\Delta}$ is not a polynomial in x_1, x_2, x_3, y_1, y_2 and y_3. Therefore, q does not factor as $q_1q_2q_3q_4$ either. So q must be irreducible as a polynomial in c_1, c_2, c_3, with coefficients being polynomials in x_1, x_2, x_3, y_1, y_2 and y_3. Therefore, up to scaling by a constant, q is the discriminant of the system (1). □

Henceforth, let $\hat{\mathcal{D}}$ denote the discriminant of the Grunert system of equations (1), and let $\mathcal{D}$ denote the rational function $\hat{\mathcal{D}}\,/\,\eta^8$. Now, once again, consider continuously varying the point P, with coordinates (x,y,z), and its associated values of c_1, c_2 and c_3. As we do this, $\hat{\mathcal{D}}$ will vary as well. It is of practical concern to wonder when this vanishes, since the number of solutions to the system can only change when crossing the surface of such points. If P is on the danger cylinder, then of course $\hat{\mathcal{D}}$ vanishes, since P itself

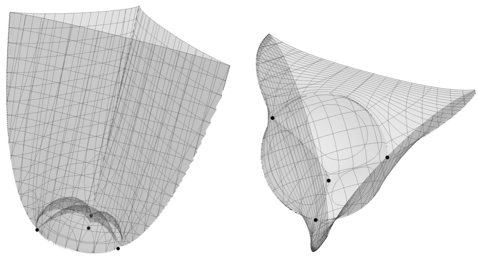

Figure 2: two views of a deltoidal surface

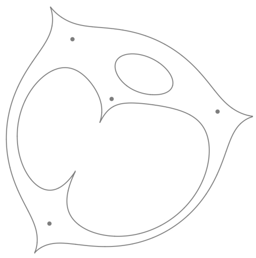

Figure 3: a z-cross section ($z = 0.3$)

is a repeated solution, but there are other points that cause $\hat{\mathcal{D}}$ to vanishes. Such points have a related solution point on the danger cylinder.

The surface on which $\hat{\mathcal{D}}$ vanishes is comprised of the danger cylinder together with another surface that is mostly wrapped around the danger cylinder. Let us refer to this other surface as the "deltoidal surface." Figure 2 shows two views of half of such a surface (say $z \geq 0$). Here the control points are $(^{63}/_{65}, {}^{16}/_{65}), (-^{5}/_{13}, {}^{12}/_{13}), (-^{3}/_{5}, -^{4}/_{5})$. The control points and the orthocenter of their triangle are indicated by dots. Figure 3 shows the $z = {}^{3}/_{10}$ slice (cross section) of the same deltoidal surface.

As z tends to infinite, the z-cross section (*i.e.* intersection with a constant-z plane) of the surface tends to approach a deltoid curve, a behavior that was described in (*Rieck*, 2015). As discussed there, this deltoid curve, which is symmetric with respect to 120-degree rotations, is approached even when the control points triangle in not equilateral; it happens for any triangle. Though the deltoidal surface is very well behaved far away from the control points plane, it is somewhat erratic near this plane (as suggested by the figures). A result concerning the z-cross sections will now be stated, and later proved in Section 7.

Theorem 3. *For a fixed value of z with $|z| \geq 3$, the z-cross section of the deltoidal surface is a simple closed curve that wraps around the outside of a circle on the danger cylinder. Together these comprise the set of points on the constant-z plane where the discriminant vanishes. The closed curve intersects the circle in three points, which are positioned symmetrically with respect to 120-degree rotations of the constant-z plane about the z-axis. Moreover, as z is allowed to vary, these three points move along three straight lines on the danger cylinder, and as $|z| \to \infty$, the simple closed curve tends to a deltoid curve.*

While it certainly does not provide proof, Figure 2 suggests that this theorem might be true, though the danger cylinder in not shown in this figure. Actually, preliminary evidence suggests that the lower bound 3 could be replaced with 2, but that the theorem would certainly be false if a value smaller than 2 was substituted instead.

Before proving this theorem, we will need to consider a family of surfaces that include as a member, the surface on which the discriminant vanishes, *i.e.* the union of the danger cylinder and the deltoidal surface. The surfaces in question are just the contour surfaces (*i.e.* constant-value surfaces) of the quantity $\mathcal{D}$. Assuming that $0 < |z| < \infty$, note that $\eta \neq 0$, and so $\mathcal{D}$ vanishes if and only if $\hat{\mathcal{D}}$ vanishes.

To understand these contour surfaces, it is useful and interesting to also explore the contour surfaces for $\mathcal{L}$ and $\mathcal{R}$. Of course, Theorem 2 shows how to express $\mathcal{D}$ in terms of $\mathcal{L}$ and $\mathcal{R}$. Consequently, a contour surface for $\mathcal{L}$ intersects a contour surface $\mathcal{R}$ in a curve that is also on a contour surface for $\mathcal{D}$.

5 AN EXAMPLE

Let us now use the example from the previous section, the one with the control points $(^{63}/_{65}, {}^{16}/_{65}), (-^{5}/_{13}, {}^{12}/_{13}), (-^{3}/_{5}, -^{4}/_{5})$, to provide evidence for the correctness of Theorems 1 and 2. Many of the computations here still require symbolic manipulation software to work out, but these computations are very direct and relatively short. Observe that $X_H = -^{1}/_{65}$, $Y_H = {}^{24}/_{65}$, $d_1^2 = {}^{196}/_{65}$, $d_2^2 = {}^{1156}/_{325}$, $d_3^2 = {}^{1936}/_{845}$, $r_1^2 = (65 - 126x + 65x^2 - 32y + 65y^2 + 65z^2)/65$, $r_2^2 = (13 + 10x + 13x^2 - 24y + 13y^2 + 13z^2)/13$ and $r_3^2 = (5 + 6x + 5x^2 + 8y + 5y^2 + 5z^2)/5$.

Here is a quick check for Theorem 1. Let's consider the point $(x, y, z) = (1, -1, 2)$. Plugging into the formulas for $\mathcal{L}$ and $\mathcal{R}$ in Theorem 1, we obtain the numbers $^{-321}/_{65}$ and $^{349}/_{260}$, respectively. We need to check these numbers against the definitions for $\mathcal{L}$ and $\mathcal{R}$ in (6). Using (1) and (2), we find that $r_1^2 = {}^{361}/_{65}$, $r_2^2 = {}^{125}/_{13}$, $r_1^2 = {}^{33}/_{5}$, $c_1^2 = {}^{429}/_{625}$, $c_2^2 = {}^{177419}/_{351975}$, $c_3^2 = {}^{29604481}/_{38130625}$ and $c_1c_2c_3 = {}^{7601077}/_{14665625}$. Plugging into (6), gives $\mathcal{L} = {}^{-321}/_{65}$ and $\mathcal{R} = {}^{349}/_{260}$, in agreement with the Theorem 1 formulas.

Let's switch now to Theorem 2, using the same control points, but starting with a general point (x,y,z). When the formulas in Theorem 1 are substituted for $\mathcal{L}$ and $\mathcal{R}$, and $d_1^8 d_2^8 d_3^8 z^8 \,/\, 256\, r_1^8 r_2^8 r_3^8$ is substituted for η^8, in the formula in Theorem 2, the result is a constant times $(x^2+y^2-1)^2\, p(x,y,z^2) \,/\, r_1^8 r_2^8 r_3^8$, where $p(x,y,z^2)$ is the polynomial given in Part 5 of the Appendix. Up to a constant factor, this is the same as $p(X_H, Y_H, x, y, z^2)$ in Lemma 13, using $X_H = -{}^1/_{65}$, $Y_H = {}^{24}/_{65}$.

The surface $p(x,y,z^2) = 0$ is the deltoidal surface, and Figure 3 suggests that there are two points on this surface for which $y = 1$ and $z = 0.3$. Now, $160000\, p(x, 1, 0.09) =$

$$61148105928 - 26548665624x - 1364362770420x^2 - 6152060366712x^3 - 15092981650881x^4 - 30624594308640x^5 - 47135678216400x^6 - 37034034576000x^7 + 7442527460000x^8 + 12070177600000x^9 - 25450724000000x^{10} + 11248640000000x^{11} - 2856100000000x^{12},$$

and this has two real roots, which are approximately -0.58612 and 0.145677.

Let's focus on the point $(x,y,z) = (-0.58612, 1, 0.3)$. This is a point on the deltoidal surface, that is, $p(-0.58612, 1, 0.09) = 0$. For this point, we see from (1) and (2) that $r_1 = 1.75425$, $r_2 = 0.369488$, $r_3 = 1.82488$, $c_1 = 0.334679$, $c_2 = 0.445236$ and $c_3 = 0.711843$.

Now, solving (1) using these c values leads to one other solution for which all of the r's are non-negative, namely, $r_1' = 2.10316$, $r_2' = 1.82759$ and $r_3' = 0.83397$. (Symbolic manipulation software reveals that this is in fact a double solution to (1).) Then solving (2), using these r values, yields the solution point $(x', y', z') = (-0.737641, -0.675193, 0.813012)$. This satisfies $(x')^2 + (y')^2 = 1$. That is, this is a point on the danger cylinder. The point $(-0.58612, 1, 0.3)$ therefore has a related solution point on the danger cylinder, and hence the discriminant of the Grunert system must vanish at $(-0.58612, 1, 0.3)$. (Technically, since this discriminant is a polynomial function of the c's, we are here speaking about the rational function of x, y and z that results by using (1) and (2) to perform substitutions.)

Since $(-0.58612, 1, 0.3)$ was a rather arbitrary point on the deltoidal surface $p(x,y,z^2) = 0$, this supplies evidence for the claim that points (x,y,z) where $p(x,y,z^2)$ vanish are also points where the discriminant of the Grunert system (1) vanish.

6 A BIRATIONAL TRANSFORMATION

This section is concerned with a surprising birational transformation that winds up simplifying the statement of Theorem 1, putting it into a form that facilitates discussing limit points as $|z| \to \infty$. The transformation is as follows. Define:

$$\begin{cases} \mu = [\, 2x\,(x^2+y^2+z^2-1) + X_H(1-x^2-y^2) \,] \,/\, (\, x^2+y^2+2z^2-1 \,) \\ \nu = [\, 2y\,(x^2+y^2+z^2-1) + Y_H(1-x^2-y^2) \,] \,/\, (x^2+y^2+2z^2-1) \\ \xi = (x^2+y^2-1) \,/\, (x^2+y^2+2z^2-1) \end{cases}$$

Notice that as $|z| \to \infty$, μ is asymptotic to x, ν is asymptotic to y, and ξ approaches zero. Substituting the above formulas into the following formulas provides a direct proof of the following claim.

Lemma 16.

$$\begin{cases} x = (\mu + X_H\,\xi) \,/\, (1+\xi) \\ y = (\nu + Y_H\,\xi) \,/\, (1+\xi) \\ z^2 = (\xi - 1)(1-x^2-y^2) \,/\, 2\xi \end{cases}$$

In this way, there is a birational transformation between (x,y,z^2) and (μ,ν,ξ), meaning that each member of each of these triples is obtained as a rational function of the members of the other triple.

Now, using the new coordinate system, Theorem 1 can immediately be rewritten in a more symmetric form that highlights the significance of the orthocenter of the control points triangle.

Corollary 2. *The quantities $\mathcal{L}$ and $\mathcal{R}$ defined in (6) can also be expressed as follows:*

$$\mathcal{L} \;=\; 2\,\frac{(1+\mu)\nu \;-\; (1+X_H)Y_H\,\xi^2}{1\;-\;\xi^2}$$

and

$$\mathcal{R} \;=\; \frac{[\nu^2-(\mu-1)^2] \;-\; [Y_H^2-(X_H-1)^2]\,\xi^2}{1\;-\;\xi^2}\;.$$

A lemma concerning the new coordinates, that will be required in the next section, is as follows.

Lemma 17. *Provided that* $|z| \geq 2$, *we have* $\mu^2+\nu^2=1$ *if and only if* $x^2+y^2=1$.

Proof. Let $Z=z^2$. Then $\mu^2+\nu^2-1$, expressed in terms of x, y and Z, factors as follows:

$$(x^2+y^2-1)\,[\,4(x^2+y^2+Z)^2+4(1+X_Hx+Y_Hy)\cdot(1-x^2-y^2-Z)+(1-X_H^2-Y_H^2)(1-x^2-y^2) - 4\,]\;/\;(x^2+y^2+2Z-1\,)^2.$$

The long factor in the numerator is seen to be nonzero, as follows. Its partial derivative with respect to Z is

$$8Z \;+\; 8(x-X_H/4)^2+8(y-Y_H/4)^2 -(X_H^2+Y_H^2+8)/2.$$

Recall that $X_H^2+Y_H^2<9$ (Lemma 3, part *(iii)*). The partial derivative is seen to be positive if $Z>17/16$, which we are assuming to be the case. So the long factor is increasing as a function of Z. If we substitute $Z=4$ into this factor, we get $\Phi^2-\Psi^2-4(X_H^2+Y_H^2-1)$, where $\Phi \;=2[\,(x-X_H/4)^2+(y-Y_H/4)^2\,]+(X_H^2+Y_H^2+54)/8$ and $\Psi \;=\; X_Hx+Y_Hy-(X_H^2+Y_H^2+3)/4$. Now, $\Phi+\Psi \;=\; 2(x^2+y^2+3)\geq 6$ and $\Phi-\Psi \;=\; 2[\,(x-X_H/2)^2+(y-Y_H)^2+15/2\geq 15/2$. So $\Phi^2-\Psi^2 \;=\; (\Phi+\Psi)(\Phi-\Psi)\geq 45$. Since (X_H,Y_H) is the orthocenter of the triangle, we know by Lemma 3 that $X_H^2+Y_H^2<9$, and so $-4(X_H^2+Y_H^2-1)>-32$. Thus, $\Phi^2-\Psi^2-4(X_H^2+Y_H^2-1)>45-32=13>0$. Since the long factor is strictly positive, the claim in the lemma is confirmed.

□

Going forward, it will generally be understood, when considering a point P with Cartesian coordinates (x,y,z), that P is a point in real Euclidean space, and so $z\neq\infty$. Of course, when switching to the (μ,ν,ξ)-coordinates, we can naturally discuss "points at infinity," by which we will always mean points with $|z|=\infty$, since the (μ,ν,ξ)-coordinates of such points are just characterized by the condition that $\xi=0$. These points nicely belong when discussing facts in terms of (μ,ν,ξ)-coordinates. However, there will be no need to allow $|x|$ or $|y|$ to go to infinity.

We will now revisit the issue of trying to understanding all functions of the form (4) that satisfy (5). Recall from Lemma 11 that $\mathcal{L}$, $\mathcal{R}$ and the constant function 1 form a basis for the linear space of all such functions. To assist in understanding such functions better, we will apply the following lemma, which is basically just an exercise in trigonometry.

Lemma 18. *The following equality holds:*

$$\begin{aligned}\cos 2\phi\cdot 2(1+\mu)\nu \;+\; \sin 2\phi\cdot[\nu^2-(\mu-1)^2] \;=\;\\ 2[\,\cos\phi\cdot(\mu+\cos 4\phi)+\sin\phi\cdot(\nu+\sin 4\phi)\,]\cdot\\ [\,-\sin\phi\cdot(\mu+\cos 4\phi)+\cos\phi\cdot(\nu+\sin 4\phi)\,]\\ -4\cos^2 2\phi\sin 2\phi.\end{aligned}$$

Using Corollary 2, and applying the previous lemma leads immediately to a proof of the following lemma concerning linear combinations of $\mathcal{L}$ and $\mathcal{R}$.

Lemma 19. *The following equality holds:*

$$\begin{aligned}\cos 2\phi\cdot\mathcal{L} \;+\; \sin 2\phi\cdot\mathcal{R} \;=\;\\ \{\;\big[\cos\phi\cdot(\mu+\cos 4\phi)+\sin\phi\cdot(\nu+\sin 4\phi)\big]\\ \cdot\big[-\sin\phi\cdot(\mu+\cos 4\phi)+\cos\phi\cdot(\nu+\sin 4\phi)\big]\\ -\;\big[\cos\phi\cdot(X_H+\cos 4\phi)+\sin\phi\cdot(Y_H+\sin 4\phi)\big]\\ \cdot\big[-\sin\phi\cdot(X_H+\cos 4\phi)+\cos\phi\cdot(Y_H+\sin 4\phi)\big]\,\xi^2\\ \}\;\cdot\;[\,2\,/\,(1-\xi^2)\,] \quad-\quad 4\cos^2 2\phi\sin 2\phi.\end{aligned}$$

In the limit as $|z|$ *goes to infinity, and so that* $\xi\to 0$, $\mu\to x$ *and* $\nu\to y$, *this tends towards*

$$\begin{aligned}2\,\big[\,\cos\phi\cdot(x+\cos 4\phi)+\sin\phi\cdot(y+\sin 4\phi)\,\big]\cdot\\ \big[-\sin\phi\cdot(x+\cos 4\phi)+\cos\phi\cdot(y+\sin 4\phi)\big]\\ -\quad 4\cos^2 2\phi\sin 2\phi.\end{aligned}$$

The contour curves for this function of x and y are rectangular hyperbolas, centered at a point on the unit circle, specifically, $(-\cos 4\phi,-\sin 4\phi)$.

The phenomenon of generating such a family of rectangular hyperbolas whose centers form a circle, by taking linear combinations of two rectangular hyperbola, has long been understood. For example, see (Carver, 1956) and (Alperin, 2010). This is also related to "Feuerbach's Conic Theorem." The following claim is now easily established.

Theorem 4. *Every rational function $f(c_1,c_2,c_3)$, of the form (4) that satisfies (5) is uniquely expressible as a linear combination of the functions $\mathcal{L}$, $\mathcal{R}$ and the constant function 1, and has the following property. Given a varying point P with coordinates x,y and z, and setting the values of c_1, c_2 and c_3 to match these coordinates (by means of (1) and (2)), yields a function g of x, y and z^2, defined by $g(x,y,z^2) = f(c_1,c_2,c_3)$. As $|z| \to \infty$, $g(x,y,z^2)$ tends either to a constant function, or to a quadratic function of x and y whose contours are rectangular hyperbolas, centered at some point on the unit circle.*

Proof. The first part of this is basically just a restatement of Lemma 11. Now, Lemma 19 considers a linear combination of $\mathcal{L}$, $\mathcal{R}$ having coefficients whose squares sum to one. The formula on the right side is essentially $g(x,y,z^2)$ but expressed in terms of the coordinates (μ,ν,ξ). Of course, $|z| \to \infty$ implies that $\mu \to x$, $\nu \to y$ and $\xi \to 0$. The limiting behavior is certainly as described in the theorem, for this case. But scaling such a function, *i.e.* multiplying it by a nonzero constant, does not affect this property. Since the zero function also satisfies the described limiting behavior, we now see that any linear combination of $\mathcal{L}$, $\mathcal{R}$ exhibits this behavior. Adding a constant function to such a function does not affect this limiting behavior claim. So it is true of all linear combination of the functions $\mathcal{L}$, $\mathcal{R}$ and the constant function 1.

□

7 A CLASS OF CURVES

In this section, we will explore a class of interesting curves that are directly associated with the contour surfaces for $\mathcal{D}$, including the surface for $\mathcal{D} = 0$. This latter surface is the surface in xyz-space consisting of points where the Grunert system discriminant vanishes when it is expanded as a rational function of x, y and z. This is the union of the danger cylinder and the deltoidal surface. Each of the curves studied in this section is the intersection of a contour surface for $\mathcal{L}$ and a contour surface for $\mathcal{R}$.

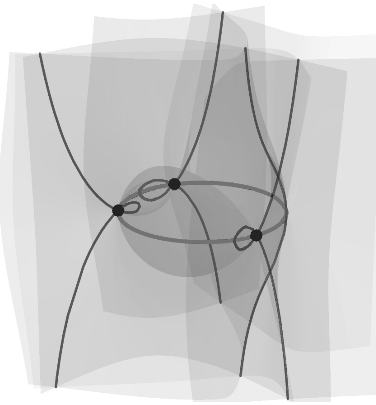

Figure 4: an example of the curves

The following basic fact will by required a couple times in this section.

Lemma 20. *Given any real numbers λ and ρ, the system of equations*

$$\begin{cases} 2(x+1)y & = & \lambda \\ y^2-(x-1)^2 & = & \rho \end{cases}$$

has a real-valued solution.

Proof. [1]

By the first equation, $y = \lambda/[2(x+1)]$. Substituting into the second equation yields $-4+8x^2-4x^4+\lambda^2-4\rho-8\rho x-4\rho x^2 = 0$. This is so provided that $x \neq -1$, which is so as long as $\lambda \neq 0$, which will be supposed for the moment. Consider the polynomial $-4+8x^2-4x^4+\lambda^2-4\rho-8\rho x-4\rho x^2$ as a continuous function of x. As $x \to \infty$, this tends to $-\infty$. But when $x = -1$, it equals λ^2, which is positive. By continuity, there is a value of $x > -1$ that makes this polynomial equal zero. This together with the corresponding value of y provide a real-valued solution to the system. This leaves the case where $\lambda = 0$. If $\rho \geq -4$, then $x = -1$ and $y = \sqrt{4+\rho}$ provides a real-valued solution. Otherwise, take $y = 0$ and $x = 1+\sqrt{-\rho}$ to obtain a real-valued solution.

□

By taking a contour surface for $\mathcal{L}$, and intersecting it with a contour surface for $\mathcal{R}$, we obtain a one-dimensional subspace that has some interesting and useful properties. Though this subspace will generally be disconnected, we will here refer to the whole

[1] this short proof was essentially provided by a reviewer

one-dimensional subspace as a "curve." By definition, $\mathcal{L}$ and $\mathcal{R}$ are constant along such a curve. The curve actually lies on a unique contour surface for every non-constant function of the form (4) that satisfies (5).

For instance, Figure 4 shows such a curve and three such contour surfaces containing the curve. (It also shows the control points and the circle containing them.) The contour surfaces shown in the figure happen to be for the function considered in Corollary 1 and the two functions obtained from it by cycling the indices. But again, for this curve, every non-trivial function of the form (4) satisfying (5) has a contour containing it. However, there is considerably redundancy in these contour surfaces, as explained in the following lemma.

Lemma 21. *Fix real values, λ and ρ. The two surfaces corresponding to the equations $\mathcal{L} = \lambda$ and $\mathcal{R} = \rho$ intersect in a curve γ with the following properties. Every non-constant function f of the form (4) and satisfying (5) has a unique contour surface containing γ. Multiplying f by a non-zero constant does not alter the corresponding contour surface. Neither does adding a constant to f. Therefore, allowing f to vary now, we obtain (only) a one-dimensional family of surfaces containing γ*

Proof. The equations $\mathcal{L} = \lambda$ and $\mathcal{R} = \rho$ have common real solutions that constitute a one-dimensional subspace, *i.e.* a curve. To see this, work with the (μ, ν, ξ) coordinates, and fix for the moment, a value for ξ. The equations in μ and ν are thus $2(1+\mu)\nu = \lambda(1-\xi^2)+2(1+X_H)Y_H\xi^2$ and $\nu^2-(\mu-1)^2 = \rho(1-\xi^2)+Y_H^2-(X_H-1)^2$. By Lemma 20, these intersect in at least one real point. By now varying ξ, we see that the common real solutions for $\mathcal{L} = \lambda$ and $\mathcal{R} = \rho$ do indeed form a one-dimensional subspace γ.

Now, let f be any function of the form (4) satisfying (5). Express it as $f = a\mathcal{L} + b\mathcal{R} + c$ for real constants a, b and c (a unique expression). Consider the equation $f = a\lambda + b\rho + c$. Each point on γ satisfies this equation since $\mathcal{L} = \lambda$ and $\mathcal{R} = \rho$ for such points. So γ is contained in a contour of f. Now, if d and e are also real numbers, with $d \neq 0$, then the function $g = df + e$ is uniquely expressed as a linear combination of $\mathcal{L}$, $\mathcal{R}$ and 1 as $g = ad\mathcal{L} + bd\mathcal{R} + (cd+e)$. The unique contour surface of g containing γ is thus given by the equation $g = ad\lambda + bd\rho + (cd+e)$. But since $g = df + e$, this is equivalent to $f = a\lambda + b\rho + c$. Therefore the contour surfaces for f and g are the same.

Since the space of functions of the form (4) that satisfy (5) is three-dimensional, and since we have seen that each of these functions is a member of a two-dimensional family of functions that produce the same contour surface containing γ, it follows that the family of contour surfaces containing γ is one-dimensional at the most. But Lemma 19 makes it clear that there is at least a one-dimensional family of contour surfaces containing γ. Therefore, the family of contour surfaces containing γ is one-dimensional.

$\square$

By Theorem 2, $\mathcal{D}$ (= $\hat{\mathcal{D}}/\eta^2$) is also constant along any of the special curves being considered here. In fact, the following is immediately clear.

Lemma 22. *Each contour surface for $\mathcal{D}$ is a disjoint union of the special curves. In particular the union of the danger cylinder and the deltoidal surface is a disjoint union of such curves, since this surface is the surface corresponding to $\mathcal{D} = 0$ (which is the same as $\hat{\mathcal{D}} = 0$).*

Letting $\Xi = \xi^2$, it is useful to think about the images of the curves in (μ, ν, Ξ) - space. It turns out that the image of such a curve has a rational parameterization.

Lemma 23. *For fixed real numbers μ_0, ν_0, κ and ϕ, consider the rational curve in (μ, ν, Ξ) - space, parameterized by t as follows:*

$$\begin{aligned} 2(1+t)\,\mu \quad = \quad & 2\mu_0 \,+\, 2\,[\,-\sin 2\phi \\ & +\mu_0\,(1-\sin 2\phi)+\nu_0\cos 2\phi\,]\,t \\ & +[\,-\cos 4\phi - \sin 2\phi \\ & +\mu_0\,(1-\sin 2\phi)+\nu_0\cos 2\phi\,]\,t^2 \end{aligned}$$

$$\begin{aligned} 2(1+t)\,\nu \quad = \quad & 2\nu_0 \,+\, 2\,[\,-\cos 2\phi \\ & +\mu_0\cos 2\phi+\nu_0\,(1+\sin 2\phi)\,]\,t \\ & +[\,-\sin 4\phi - \cos 2\phi \\ & +\mu_0\cos 2\phi+\nu_0\,(1+\sin 2\phi)\,]\,t^2 \end{aligned}$$

$$
\begin{aligned}
2\kappa(1+t)^2\,\Xi \quad = \quad & 4(\mu_0^2+\nu_0^2-1)\,t \\
& +2\,[\,-3-2\mu_0\sin 2\phi+\mu_0^2\,(3-2\sin 2\phi) \\
& -2\nu_0\cos 2\phi+4\mu_0\nu_0\cos 2\phi \\
& +\nu_0^2\,(3+2\sin 2\phi)\,]\,t^2 \\
& +4\,[\,\mu_0\,(\sin 2\phi-1)(2\sin 2\phi+1) \\
& +\mu_0^2\,(1-\sin 2\phi)-\nu_0\cos 2\phi(2\sin 2\phi+1) \\
& +2\mu_0\nu_0\cos 2\phi+\nu_0^2\,(1+\sin 2\phi)\,]\,t^3 \\
& +[\,(1-\sin 2\phi)(1+2\sin 2\phi)^2 \\
& -2\mu_0\,(1-\sin 2\phi)(1+2\sin 2\phi) \\
& +\mu_0^2\,(1-\sin 2\phi) \\
& -2\nu_0\cos 2\phi(1+2\sin 2\phi) \\
& +2\mu_0\nu_0\cos 2\phi+\nu_0^2\,(1+\sin 2\phi)\,]\,t^4
\end{aligned}
$$

This curve satisfies the following system of equations:

$$
\left\{
\begin{array}{lcll}
2(\mu+1)\nu & = & 2(\mu_0+1)\nu_0 & +\ \kappa\cos 2\phi\cdot\Xi \\
\nu^2-(\mu-1)^2 & = & \nu_0^2-(\mu_0-1)^2 & +\ \kappa\sin 2\phi\cdot\Xi.
\end{array}
\right.
\tag{9}
$$

Proof. Using Newton's dot notation, differentiating the system (9) with respect to t yields

$$
\left\{
\begin{array}{lcl}
2\nu\,\dot\mu+2(1+\mu)\,\dot\nu & = & \kappa\cos 2\phi\cdot\dot\Xi \\
2(1-\mu)\,\dot\mu+2\nu\,\dot\nu & = & \kappa\sin 2\phi\cdot\dot\Xi.
\end{array}
\right.
$$

Solving this for $\dot\mu$ and $\dot\nu$ yields

$$
\left\{
\begin{array}{lcl}
2(\mu^2+\nu^2-1)\,\dot\mu & = & [\,\cos 2\phi\,\nu-\sin 2\phi\,(1+\mu)\,]\,\kappa\dot\Xi \\
2(\mu^2+\nu^2-1)\,\dot\nu & = & [\,\cos 2\phi\,(\mu-1)+\sin 2\phi\,\nu\,]\,\kappa\dot\Xi
\end{array}
\right.
\tag{10}
$$

One can multiplying both sides of these equations by $4(1+t)^3$, and rewrite the first one as $\big([2(1+t)\mu]^2+[2(1+t)\nu]^2-[2(1+t)]^2\big)\,[2(1+t)\,\dot\mu]=\big(\cos 2\phi\,[2(1+t)\nu]-\sin 2\phi\,[2(1+t)+2(1+t)\mu]\big)\,[2\kappa(1+t)^2\,\dot\Xi]$, and similarly for the other equation. This makes it comparatively easy to "plug in" and check that the formulas for the curve satisfy the differential equations. A few applications of the double angle formula are involved in doing so. Returning to the system of algebraic equations (9), and setting $t=0$, we see that in this case, $\mu=\mu_0$ and $\nu=\nu_0$ and that the algebraic equations are valid. This together with the validity of the differential equations establishes that the algebraic equations are valid in general.

□

From the Cartesian coordinates (x,y,z) of any point P in physical space, apart from the control points, equations (1) and (2) provide corresponding values for c_1, c_2 and c_3. Assuming that P is not on the control points plane, *i.e.* $z\neq 0$, by part (iv) of Lemma 5, it is seen that $\eta^2\neq 0$, and so there are corresponding values for $\mathcal{L}$ and $\mathcal{R}$, as defined by (6). The totality of points that share the same values of $\mathcal{L}$ and $\mathcal{R}$ as the point P comprise one of the special curves being discussed in the section. Of course, this curve includes the point P itself. Now, as long as $x^2+y^2+2z^2\neq 1$, then P also has (μ,ν,ξ)- coordinates, and so we can also think about its corresponding curve in (μ,ν,ξ)-space, or in (μ,ν,Ξ)-space. These include points for which $\xi=0$ ($\Xi=0$). These will be referred to as "points at infinity corresponding to P."

Theorem 5. *Let $\hat P$ be a point in physical space, with coordinates $(\hat x,\hat y,\hat z)$ for which $\hat z\neq 0$ and $\hat x^2+\hat y^2+2\hat z^2\neq 1$. $\hat P$ lies on a unique curve γ consisting of all points P sharing the same values of $\mathcal{L}$ and $\mathcal{R}$ as $\hat P$.*

There exists at least one point at infinity for γ. Let P_0 be one of these, and let (μ_0,ν_0) denote the (μ,ν)-coordinates of P_0. γ contains, as a sub-curve, one of the curves described in Lemma 23, using these particular values for μ_0 and ν_0.

Moreover, with $(\hat\mu,\hat\nu,\hat\Xi)$ being the values of (μ,ν,Ξ) for $\hat P$, and with (μ,ν,Ξ) denoting these coordinates for an arbitrary point P on γ, the constants κ and ϕ satisfy the requirements that

$$
\begin{array}{lcl}
\kappa\cos 2\phi\cdot\hat\Xi & = & 2(\hat\mu+1)\hat\nu-2(\mu_0+1)\nu_0 \quad \textit{and} \\
\kappa\sin 2\phi\ \cdot\hat\Xi & = & \hat\nu^2-(\hat\mu-1)^2-\nu_0^2+(\mu_0-1)^2.
\end{array}
\tag{11}
$$

κ and ϕ also satisfy the equations

$$
\begin{array}{lcl}
\kappa\cos 2\phi\cdot(1-\hat\Xi) & = & \\
 & & 2(X_H+1)Y_H-2(\hat\mu+1)\hat\nu \quad \textit{and} \\
\kappa\sin 2\phi\ \cdot(1-\hat\Xi) & = & \\
 & & Y_H^2-(X_H-1)^2-\hat\nu^2+(\hat\mu-1)^2,
\end{array}
\tag{12}
$$

as well as the equations

$$
\begin{array}{lcl}
\kappa\cos 2\phi & = & 2(X_H+1)Y_H-2(\mu_0+1)\nu_0 \quad \textit{and} \\
\kappa\sin 2\phi & = & Y_H^2-(X_H-1)^2-\nu_0^2+(\mu_0-1)^2,
\end{array}
\tag{13}
$$

which implies that

$$\begin{cases} 2(\mu+1)\nu & = & 2(X_H+1)Y_H\,\Xi \\ & + & 2(\mu_0+1)\nu_0\,(1-\Xi) \\ \nu^2-(\mu-1)^2 & = & Y_H^2-(X_H-1)^2\,\Xi \\ & + & [\nu_0^2-(\mu_0-1)^2]\,(1-\Xi). \end{cases} \tag{14}$$

Proof. γ contains P, and by definition, $\mathcal{L}$ and $\mathcal{R}$ are constant when restricted to γ. So, by Corollary 2, we obtain the following for points (μ,ν,Ξ) along γ:

$$\begin{aligned}&(1-\hat{\Xi})\,[\,(1+\mu)\nu-(1+X_H)Y_H\,\Xi\,]\;=\\&(1-\Xi)\,[\,(1+\hat{\mu})\hat{\nu}-(1+X_H)Y_H\,\hat{\Xi}\,]\quad\text{and}\\&(1-\hat{\Xi})\,[\,\nu^2-(\mu-1)^2)-[Y_H^2-(1-X_H)^2]\,\Xi\,]\;=\\&(1-\Xi)\,[\,\hat{\nu}^2-(\hat{\mu}-1)^2)-[Y_H^2-(1-X_H)^2]\,\hat{\Xi}\,].\end{aligned}$$

These yield

$$\begin{aligned}&(1-\hat{\Xi})\cdot 2(1+\mu)\nu\;=\\&\qquad[\,2(1+X_H)Y_H-2(1+\hat{\mu})\hat{\nu}\,]\,\Xi\\&\quad+\;[\,2(1+\hat{\mu})\hat{\nu}-2(1+X_H)Y_H\,\hat{\Xi}\,]\end{aligned}$$

and

$$\begin{aligned}&(1-\hat{\Xi})\cdot[\,\nu^2-(\mu-1)^2)\,]\;=\\&\qquad[\,Y_H^2-(X_H-1)^2-\hat{\nu}^2+(\hat{\mu}-1)^2)\,]\,\Xi\\&\quad+[\,\hat{\nu}^2-(\hat{\mu}-1)^2-[Y_H^2-(X_H-1)^2]\,\hat{\Xi}\,].\end{aligned}$$

These are equations of the form (9), using the values of κ and λ suggested by (12). The values of μ_0 and ν_0 must satisfy

$$2(\mu_0+1)\nu_0 \quad = \quad \frac{2(1+\hat{\mu})\hat{\nu}-2(1+X_H)Y_H\hat{\Xi}}{1-\hat{\Xi}}$$

and

$$\nu_0^2-(\mu_0-1)^2 \quad = \quad \frac{\hat{\nu}^2-(\hat{\mu}-1)^2-[X_H^2-(X_H-1)^2]\hat{\Xi}}{1-\hat{\Xi}}$$

Such values for μ_0 and ν_0 exist by Lemma 20. So points on γ satisfy (12) for some μ_0 and ν_0. But for these particular parameter values, the reasoning can be run in reverse. That is, points satisfying (9) agree concerning their values of $\mathcal{L}$ and $\mathcal{R}$, and so are on γ. The point $(\mu_0,\nu_0,0)$ is one such point, and of course this is a "point at infinity" when considering the original (x,y,z)-space. Of course, (11) now holds by (9), since $(\hat{\mu},\hat{\nu},\hat{\Xi})$ is on γ. Adding (11) and (12) yields (13), and substituting this into (9) yields (14). Now, the particular parameterized curve in Lemma 23 also satisfies (9), using the same parameters, and so this curve must at least be a sub-curve of γ.

□

It turns out that the special curves considered in this section are actually the integral curves of a certain vector field.

Lemma 24. *In (μ,ν,ξ)-space, except at points on the danger cylinder and possibly some other points with $|z|<2$, the special curves are integral curves for the following non-vanishing vector field:*

$$\Big(\ \{\,[(X_H-1)^2-Y_H^2]\,(\mu+1)+2(X_H+1)Y_H\,\nu \\ -(\mu+1)[(\mu-1)^2+\nu^2]\,\}\,\xi\,,$$

$$\{\,2(X_H+1)Y_H\,(\mu-1)+[Y_H^2-(X_H-1)^2]\,\nu \\ -\nu\,[\nu^2+(\mu-1)(\mu+3)]\,\}\,\xi\,,$$

$$(\mu^2+\nu^2-1)(1-\xi^2)\ \ \Big)$$

Proof. Starting with an arbitrary point $(\hat{\mu},\hat{\nu},\hat{\xi})$, let γ be the special curve containing it. Now use (12) to substitute into (10), and evaluate this at $(\mu,\nu,\xi)=(\hat{\mu},\hat{\nu},\hat{\xi})$. Since $(\hat{\mu},\hat{\nu},\hat{\xi})$ is arbitrary (and we no longer need to consider any other point), let us rename it (μ,ν,ξ) for simplicity. Then,

$$2(\mu^2+\nu^2-1)\begin{bmatrix}\dot{\mu}\\ \dot{\nu}\end{bmatrix} = \begin{bmatrix}\nu & -\mu-1\\ \mu-1 & \nu\end{bmatrix}\cdot$$

$$\begin{bmatrix}2(X_H+1)Y_H-2(\mu+1)\nu\\ Y_H^2-(X_H-1)^2-\nu^2+(\mu-1)^2\end{bmatrix}\dot{\Xi}\,/\,(1-\Xi)\ .$$

Now, replace Ξ with ξ^2, and $\dot{\Xi}$ with $2\xi\dot{\xi}$. We find that $(\dot{\mu}\ \dot{\nu}\ \dot{\xi})^T$ is proportional to the following:

$$\begin{bmatrix}\nu & -\mu-1 & 0\\ \mu-1 & \nu & 0\\ 0 & 0 & 1\end{bmatrix}\cdot$$

$$\begin{bmatrix}(2(X_H+1)Y_H-2(\mu+1)\nu)\,\xi\\ (Y_H^2-(X_H-1)^2-\nu^2+(\mu-1)^2)\,\xi\\ (\mu^2+\nu^2-1)(1-\xi^2)\end{bmatrix}$$

Multiplying this out gives the stated vector field.

Assume henceforth that $|z|\geq 2$. By Lemma 17, satisfying $\mu^2+\nu^2=1$ is equivalent to satisfying $x^2+y^2=1$, *i.e.* being on the danger cylinder. At a point on the danger cylinder, $\xi=0$, and by Lemma 16, $\mu=x$ and $\nu=y$. Thus the vector field vanishes on the cylinder. Yet, by Corollary 2, points on the cylinder that have the same values of x and y also have the same values of $\mathcal{L}$ and $\mathcal{R}$, and therefore are on the same special curve. These curves are not integral curves for the vector field.

Finally, off of the danger cylinder, the vector field is non-vanishing. In fact, the third component is not zero, because $\mu^2+\nu^2 \neq 1$ and $\xi^2 \neq 1$, since $\xi = 1$ means that $z = 0$, and $\xi = -1$ means that $x^2+y^2+z^2 = 1$, which are not true here.

□

Attention will now be turned to using the curves to prove Theorem 3. This is the final objective of this paper.

Lemma 25. *Regarding the special curves in (x,y,z)-space, a tangent line for such a curve at a point with coordinates (x,y,z) (with $z \neq 0$) will be parallel to the xy-plane when and only when*

$$\begin{aligned}&1-X_H^2-Y_H^2+2(Y_H^2-X_H^2+3X_H)x\\&+(3X_H^2+3Y_H^2+4X_H-7)x^2-2(1+5X_H)x^3+8x^4\\&+2(3+2X_H)Y_Hy-8Y_Hxy-10Y_Hx^2y\\&+(3X_H^2+3Y_H^2-4X_H-7)y^2+2(3-5X_H)xy^2\\&+16x^2y^2-10Y_Hy^3+8y^4-4z^2-8X_Hxz^2+12x^2z^2\\&-8Y_Hyz^2+12y^2z^2+4z^4 \;=\; 0.\end{aligned}$$

This condition does not hold when $|z| \geq 3$, and so the tangent lines cannot be parallel to the xy-plane when $|z| \geq 3$.

Proof. Define $Z = z^2$, and let us refer to the long polynomial in the lemma as $p(x,y,Z)$. Using Lemma 16, first express Z as a rational function of μ, ν and ξ. Then compute the partial derivatives of Z with respect to these, and combine them to form a gradient vector for Z, as a function of μ, ν and ξ. From Lemma 24, we have a vector that is proportional to $(\dot{\mu}\ \ \dot{\nu}\ \ \dot{\xi})$. The dot product of these two vectors, when expressed in terms of x, y and Z equals $2Z(x^2+y^2+Z-1)p(x,y,Z)\,/\,(x^2+y^2+2Z-1)$. This vanishes when $\dot{Z}$ does.

Since we are assuming that $Z \neq 0$, this expression vanishing only if either $p(x,y,Z) = 0$ or $x^2+y^2+Z-1 = 0$. However, it can be checked that x^2+y^2+Z-1 is also a factor of $\dot{x}$ and $\dot{y}$, and by factoring it out, we can deduce that a tangent vector to the curve, at such a point, is not parallel to the xy-plane. In contrast, the $p(x,y,Z)$ is not a factor of $\dot{x}$ nor $\dot{y}$, and so its vanishing does imply that a tangent vector, at such a point, is parallel to the xy-plane. This is the only way in which this can occur. Thus the first part of the lemma is established.

Now, $\partial p/\partial Z = 8Z+4(3x^2+3y^2-2X_Hx-2Y_Hy-1) = 8Z+12[\,(x-X_H/3)^2+(y-Y_H/3)^2-(3+X_H^2+Y_H^2)/9\,]$. Recall that $X_H^2+Y_H^2 < 9$ (Lemma 3). So $\partial p/\partial Z > 8(Z-2)$. Henceforth, Z is assumed to be at least 2, ensuring that $\partial p/\partial Z$ is positive, and hence p is increasing as a function of Z.

We next decompose $p(x,y,Z)$ as follows. $p(x,y,Z) = q(x,y) + r(x,y,Z)$, where $q(x,y) = 8[(x-(1+5X_H)/16)^2+(y-5Y_H/16)^2]^2+(1+5X_H)/16)(y-5Y_H/16)^2$ and $r(x,y,Z) = 4Z^2+4(3x^2+3y^2-2X_Hx-2Y_Hy-1)Z+s(x,y)$, for a certain quadratic polynomial $s(x,y)$.

We see that $q(x,y)$ has the form $8[(u^2+v^2)^2+uv^2]$, using $u = x-(1+5X_H)/16$ and $v = y-5Y_H/16$. It is interesting that this has a minimum value of -1/54. The unique local extreme is found by computing the two partial derivatives with respect to u and v, and then computing the resultant of these polynomials with respect to v, which equals a constant times $u^5(1+6u)^2$. Continuing with this, it is discovered that the minimum value of -1/54 occurs when $u = -1/6$ and $v = 1/\sqrt{18}$.

Henceforth, we will limit attention to the situation for $z = 3$, so $Z = 9$, since this is all that is required. In order to guarantee that $p(x,y,9)$ is always positive, it will be argued that $r(x,y,9) > 1/54$ for all x and y. This will then ensure that $p(x,y,Z) > 0$ for all x, y and Z with $Z \geq 9$. Applying calculus again, it is discovered that the quadratic polynomial $r(x,y,9)$ has one local extremum, where it takes on a global minimum value. This minimum value can be expressed as a rational function of X_H and Y_H. Switching to polar coordinates by setting $(X_H,Y_H) = (\rho\cos\omega, \rho\sin\omega)$, and with $\chi = \cos\omega$, it is seen that the minimum value of $r(x,y,9)$ equals N/D with $N = 6197841846567 - 270144022668\rho^2 + 585260050\rho^4 + 55321476\rho^6 - 330625\rho^8 + 807334704\rho\chi + 6949989152\rho^3\chi + 439735024\rho^5\chi - 7176000\rho^7\chi - 1186409352\rho^2\chi^2 + 50495184\rho^4\chi^2 - 37399400\rho^6\chi^2 - 9310310016\rho^3\chi^3 - 589707072\rho^5\chi^3 + 9568000\rho^7\chi^3 - 3292080\rho^4\chi^4 + 99026400\rho^6\chi^4 + 4840000\rho^5\chi^5 - 66017600\rho^6\chi^6$, and $D = 21445828608 - 60440576\rho^2 - 5087232\rho^4 + 2785280\rho\chi - 41779200\rho^3\chi - 4096000\rho^2\chi^2 + 55705600\rho^3\chi^3$. Since $-1 \leq \chi \leq 1$ and $0 \leq \rho < 3$, it becomes a simple matter to bound D between 17,812,488,192 and 24,086,274,048. Likewise,

N is bounded below by 2,948,641,317,306. So $N/D \geq 122.42$, and this establishes that $r(x,y,9) \geq 122.42$, for all x and y, which is far bigger than required.

□

The algebraic manipulations involved in the proof of the lemma are cumbersome without symbolic manipulation software. However, using such software, it is straightforward to alter the argument to prove that the lower bound can be reduced from 3 down to a number just slightly bigger than 2. Moreover, graphical evidence suggests that the lemma is still true for a bound of 2, but no lower number. We are now prepared to prove Theorem 3.

Proof of Theorem 3. Consider the region of (x,y,z)-space for which $3 \leq z < \infty$. By Lemma 25, the tangent lines of the special curves are never parallel to the xy-plane. The claim now is that when restricting to this region, each connected component of each special curve in the region contains a unique point for each value of z with $3 \leq z < \infty$. Because of the smoothness of these curves and the fact just mentioned about the tangent lines, none of these connected components can contain two points with the same z coordinate.

In addition, each of these connected components must include a point for each value of z with $3 \leq z < \infty$, because the vector field in Lemma 24 never vanishes in this region. This means that for each point (x,y,z) in this region, following the connected components of the curve on which it lies, out towards $z = \infty$, leads to a unique limit point (x_0, y_0, ∞). Conversely, each point at infinity (x_0, y_0, ∞) is the limit for a particular connected component of one of the special curves.

Consider now, for each value of t with $0 < t \leq 1/3$, the following homeomorphism δ_t of the Cartesian plane. For a general point (x_0, y_0) in the plane, and for a value of t, start at the point at infinity (x_0, y_0, ∞). Consider the connected component of a special curve that has this as its limit point. Let $(x_t, y_t, 1/t)$ be the unique point on this for which $z = 1/t$.

Now define $\delta_t((x_0, y_0)) = (x_t, y_t)$. Extend the definition of δ by defining δ_0 to be the identity function. For each $t \in [0, 1/3]$, δ_t is a homeomorphism, that is, a continuous invertible mapping whose inverse is also continuous. The mapping from the closed interval $[0, 1/3]$ that maps t to the mapping δ_t is then a continuous deformation of the plane. Moreover, for each $t \in [0, 1/3]$, δ_t takes the union of the unit circle and the standard deltoid to the intersection of the $z = 1/t$ plane and the surface $\mathcal{D} = 0$. But since $\mathcal{D} = 0$ if and only if $\hat{\mathcal{D}} = 0$, this is a z-cross section of the surface upon which the discriminant vanishes.

δ_t maps points on the unit circle to corresponding points on the danger cylinder, with only the z coordinate changing as t varies. In particular, the three symmetrically arranged points (symmetric under 120-degree rotations) where the standard deltoid intersects the unit circle are mapped to three similarly arranged points on the danger cylinder.

So the claims made in Theorem 3 are true for the region $3 \leq z < \infty$. By symmetry, the claims are also true for the region $-\infty < z \leq -3$.

□

The curves considered in this section are well behaved and easy to visualize far from the control points plane. Looking briefly at the other extreme, graphical evidence, along with some calculus reasoning, suggests that these curves loop around quite a bit near the control points plane, and can only intersect the control points plane in restricted ways. Seemingly, they can pass through a control point from any direction, and are allowed to intersect other points on the circumcircle but only along trajectories tangent to the danger cylinder. In addition, it appears that each curve must pass through the orthocenter of the control points triangle, and in doing so, must be parallel to the z-axis. The curves are distinguished by their curvature at the orthocenter. These appear to be the only ways in which these curves can intersect this plane. These claims are somewhat exhibited in Figure 4.

8 CONCLUSION

This work was largely concerned with producing a manageable formula for the discriminant of Grunert's system of equations, and trying to understand the surface where this quantity vanishes. This was largely successful, and along the way, a useful vector space of functions, a useful algebra of functions, and some useful surfaces and curves were developed and employed. However, exploring the Perspective 3-point Pose Problem to expose its hidden secrets remains challenging. At the moment, it is necessary to combine various aspects of geometry with substantial algebraic manipulations.

Mathematical "tricks" like special Cartesian coordinate systems, birational transformations and groups of symmetries have proven quite helpful, but, so far at least, have by no means succeeded in truly taming this subject. Yet they have managed to reveal a number of treasures, leaving one with a sense that one is more or less on the right path. There is certainly reason to believe that an even better combination of geometric and algebraic reasoning could lead to a breakthrough, opening doors and revealing interesting and useful aspects of P3P, while reducing the need to handle unwieldy algebraic formulas.

APPENDIX

1. $r_1^2 r_2^2 r_3^2 \;=$

$$
\begin{array}{lll}
1 & -\,2X_H\,x & +\,(2X_H+X_H^2+Y_H^2)\,x^2 \\
+\,2(Y_H^2-X_H^2-1)\,x^3 & +\,(2X_H+X_H^2+Y_H^2)\,x^4 & -\,2X_H\,x^5 \\
+\,x^6 & -\,2Y_H\,y & -\,4Y_H\,xy \\
+\,4X_HY_H\,x^2y & -\,4Y_H\,x^3y & -\,2Y_H\,x^4y \\
+\,(Y_H^2+X_H^2-2X_H)\,y^2 & +\,2(Y_H^2-X_H^2+3)\,xy^2 & +\,2(X_H^2+Y_H^2)\,x^2y^2 \\
-\,4X_H\,x^3y^2 & +\,3\,x^4y^2 & +\,4X_HY_H\,y^3 \\
-\,4Y_H\,xy^3 & -\,4Y_H\,x^2y^3 & +\,(Y_H^2+X_H^2-2X_H)\,y^4 \\
-\,2X_H\,xy^4 & +\,3\,x^2y^4 & -\,2Y_H\,y^5 \\
+\,y^6 & +\,3\,z^2 & -\,2X_H\,xz^2 \\
+\,3\,x^2z^2 & -\,2Y_H\,yz^2 & +\,3\,y^2z^2 \\
+\,z^4. & &
\end{array}
$$

2. $(r_2^2+r_3^2-d_1^2)(r_2^3+r_1^2-d_2^2)(r_1^2+r_2^2-d_3^2) \;=$

$$
\begin{aligned}
&-1 + 4X_H^2 - 4X_H^3 + X_H^4 + 4Y_H^2 + 12X_HY_H^2 + 2X_H^2Y_H^2 + Y_H^4 + \\
&+2X_H - X_H^2 + Y_H^2)(X_H^2 + Y_H^2 - 5)\,x \\
&+(9 - 8X_H - 18X_H^2 + 8X_H^3 + X_H^4 - 10Y_H^2 - 8X_HY_H^2 + 2X_H^2Y_H^2 + Y_H^4)\,x^2 \\
&+2(1 + 15X_H - 5X_H^2 - 3X_H^3 + 3Y_H^2 - 3X_HY_H^2)\,x^3 \\
&+2(-9 + 2X_H + 7X_H^2 + 3Y_H^2)\,x^4 - 16X_H\,x^5 + 8\,x^6 \\
&+2(1 + 2X_H)(-5 + X_H^2 + Y_H^2)\,Y_H\,y + 8(2 - 2X_H - X_H^2 - Y_H^2)\,Y_H\,xy \\
&+2(15 + 8X_H - 3X_H^2 - 3Y_H^2)Y_H\,x^2y + 8(-1 + 2X_H)Y_H\,x^3y - 16Y_H\,x^4y \\
&+(9 + 8X_H - 10X_H^2 + X_H^4 - 18Y_H^2 - 16X_HY_H^2 + 2X_H^2Y_H^2 + Y_H^4)y^2 \\
&+2(-3 + 15X_H - X_H^2 - 3X_H^3 + 7Y_H^2 - 3X_HY_H^2)\,xy^2 \\
&+4(-9 + 5X_H^2 + 5Y_H^2)\,x^2y^2 - 32X_H\,x^3y^2 + 24\,x^4y^2 \\
&+2Y_H(15 + 8X_H^2 - 3X_H^2 - 3Y_H^2)y^3 + 8(-1+2X_H)Y_Hxy^3 - 32Y_Hx^2y^3 \\
&+2(-9-2X_H+3X_H^2+7Y_H^2)\,y^4 - 16X_H\,xy^4 + 24x^2y^4 - 16Y_H\,y^5 + 8\,y^6 \\
&+2(3-5X_H^2+2X_H^3-5Y_H^2-6X_HY_H^2)\,z^2 \\
&+4(7X_H-2X_H^2-X_H^3+2Y_H^2-X_HY_H^2)\,xz^2 \\
&+2(-15+2X_H+9X_H^2+5Y_H^2)\,x^2z^2 - 32X_H\,x^3z^2 + 24\,x^4z^2 \\
&+4Y_H(7+4X_H-X_H^2-Y_H^2)\,yz^2 + 8(-1+2X_H)Y_Hxyz^2 - 32Y_Hx^2yz^2 \\
&+2(-15-2X_H+5X_H^2+9Y_H^2)\,y^2z^2 - 32X_Hxy^2z^2 + 48x^2y^2z^2 - 32Y_Hy^3z^2 \\
&+24\,y^4z^2 + 4(-3+X_H^2+Y_H^2))\,z^4 - 16X_H\,xz^4 + 24\,x^2z^4 - 16Y_Hyz^4 \\
&+24\,y^2z^4 + 8\,z^6.
\end{aligned}
$$

3. The invariant part of $3(r_2^2+r_3^2-d_1^2)^2$ equals

$$
\begin{aligned}
&3 + 4X_H^2 - 4X_H^2 + X_H^4 + 4Y_H^2 + 12X_HY_H^2 + 2X_H^2Y_H^2 + Y_H^4 \\
&-4(X_H - 2X_H^2 + X_H^3 + 2Y_H^2 + X_HY_H^2)\,x \\
&+2(-3-2X_H+5X_H^2+Y_H^2)\,x^2 - 16X_H\,x^3 + 12\,x^4 \\
&-4Y_H(1+4X_H+X_H^2+Y_H^2)\,y + 8(1+2X_H)Y_H\,xy - 16Y_Hx^2y \\
&+2(-3+2X_2+X_H^2+5Y_H^2)\,y^2 - 16X_H\,xy^2 + 24\,x^2y^2 - 16Y_Hy^3 + 12y^4 \\
&+4(-3+X_H^2+Y_H^2)\,z^2 - 16X_H\,xz^2 + 24\,x^2z^2 - 16Y_H\,yz^2 + 24\,y^2z^2 + 12\,z^4,
\end{aligned}
$$

and the annihilated part equals

$$
\begin{aligned}
&2(1+X_H)(2X_H-X_H^2-Y_H^2)X_1 + 2Y_H(2-2X_H-X_H^2-Y_H^2)Y_1 \\
&+4(1+2X_H)[(X_H-1)X_1+Y_HY_1]\,x \\
&+2[(1-7X_H)X_1-5Y_HY_1]\,x^2 + 8X_1\,x^3 \\
&+4[(1+2X_H)Y_HX_1+(-1+X_H+2Y_H^2)]\,y - 4[Y_HX_1+(1+X_H)Y_1]\,xy \\
&+8Y_1\,x^2y - 2[(1+5X_H)X_1+7Y_H]\,y^2 + 8X_1\,xy^2 + 8Y_1\,y^3 \\
&-8(X_HX_1+Y_HY_1)\,z^2 + 8X_H\,xz^2 + 8Y_H\,yz^2.
\end{aligned}
$$

4. The invariant part of $3r_1^2(r_2^2+r_3^2-d_1^2)^2$ equals

$$
\begin{aligned}
&3 + 4X_H^2 - 4X_H^3 + X_H^4 + 4Y_H^2 + 12X_HY_H^2 + 2X_H^2Y_H^2 + Y_H^4 \\
&+2(-9X_H+5X_H^2+X_H^3-X_H^4-5Y_H^2+X_HY_H^2+Y_H^4)\,x \\
&+(9-14X_H^2+8X_H^3+X_H^4-6Y_H^2-8X_HY_H^2+2X_H^2Y_H^2+Y_H^4)\,x^2 \\
&+2(-3+15X_H-9X_H^2-3X_H^3+7Y_H^2-3X_HY_H^2)\,x^3 \\
&+2(-9+6X_H+9X_H^2+5Y_H^2)\,x^4 - 24X_H\,x^5 + 12\,x^6 \\
&+2Y_H(-9-10X_H+X_H^2+2X_H^3+Y_H^2+2X_HY_H^2)\,y - 8Y_H(2X_H+X_H^2+Y_H^2)\,xy \\
&+2Y_H(15+16X_H-3X_H^2-3Y_H^2)\,x^2y + 8(-3+2X_H)Y_H\,x^3y - 24Y_Hx^4y \\
&+(9-6X_H^2+X_H^4-14Y_H^2-16X_HY_H^2+2X_2^2Y_H^2+Y_H^4)\,y^2 \\
&+2(9+15X_H-5X_H^2-3X_H^3+11Y_H^2-3X_HY_H^2)\,xy^2 + 4(-9+7x_H^2+7y_H^2)\,x^2y^2 \\
&-48X_H\,x^3y^2 + 36\,x^4y^2 + 2Y_H(15+16X_H-3X_H^2-3Y_H^2)\,y^3 \\
&+8Y_H(-3+2X_H)\,xy^3 - 48Y_H\,x^2y^3 + 2(-9-6X_H+5X_H^2+9Y_H^2)\,y^4 \\
&-24X_H\,xy^4 + 36\,x^2y^4 - 24Y_Hy^5 + 12\,y^6 \\
&+(-9+8X_H^2-4X_H^3+X_H^4+8Y_H^2+12X_HY_H^2+2X_H^2Y_H^2+Y_H^4)\,z^2 \\
&+4(3X_H-2X_H^2-X_H^3+2Y_H^2-X_HY_H^2)\,xz^2 \\
&+2(-9+6X_H+11X_H^2+7Y_H^2)\,x^2z^2 - 48X_H\,x^3z^2 + 48\,x^4z^2 \\
&+4Y_H(3+4X_H-X_H^2-Y_H^2)\,yz^2 + 8(-3+2X_H)Y_H\,xyz^2 - 48Y_H\,x^2yz^2 \\
&+2(-9-6X_H+7X_H^2+11Y_H^2)\,y^2z^2 - 48X_H\,xy^2z^2 + 72\,x^2y^2z^2 \\
&-48Y_H\,y^3z^2 + 36y^4z^2 + 4(X_H^2+Y_H^2)\,z^4 - 24X_H\,xz^4 + 36\,x^2z^4 \\
&-24Y_H\,yz^4 + 36\,y^2z^4 + 12z^6,
\end{aligned}
$$

and the annihilated part equals

$$
\begin{aligned}
&2[(1+X_H)(2X_H-X_H^2-Y_H^2)X_1+Y_H(2-2X_H-X_H^2-Y_H^2)Y_1] + \\
&2[(-3-4X_H+X_H^2+2X_H^3-Y_H^2-2X_HY_H^2)X_1+2Y_H(2+X_H+X_H^2-Y_H^2)Y_1]\,x \\
&+2[(3-X_H-5X_H^2-X_H^3+Y_H^2-X_HY_H^2)X_1-Y_H(3+2X_H+X_H^2+Y_H^2)Y_1]\,x^2 \\
&+2[(3+6X_H+3X_H^2+Y_H^2)X_1+2Y_H(X_H-1)Y_1]\,x^3 - 2[(1+X_H)X_1+Y_HY_1]\,x^4 \\
&+2[2Y_H(2+X_H-2Y_H)X_1+(-3+4X_H-X_H^2+Y_H^2-4X_HY_H^2)Y_1]\,y \\
&+4[Y_H(1+4X_H)X_1+(-3+X_H+4Y_H^2)Y_1]\,xy \\
&+2[2Y_H(X_H-2)X_1+(3-4X_H+X_H^2+3Y_H^2)Y_1]\,x^2y \\
&+4[-Y_HX_1+(X_H-3)Y_1]\,x^3y + 2[(-3-3X_H+3X_H^2-X_H^3+ \\
&\quad Y_H^2-X_HY_H^2)X_1+Y_H(-1+6X_H-X_H^2-Y_H^2)Y_1]\,y^2 \\
&+2[(3-2X_H+3X_H^2+Y_H^2)X_1+2Y_H(X_H-5)Y_1]\,xy^2
\end{aligned}
$$

$$
\begin{aligned}
&-8(X_H X_1+Y_H Y_1)x^2y^2 \;+\; 2[(2Y_H(X_H-2)X_1+(3-4X_H+X_H^2+3Y_H^2)Y_1]\,y^3 \\
&-4[Y_H X_1+(X_H-3)Y_1]\,xy^3 \;-\; 2[(X_H-3)X_1+Y_H Y_1]\,y^4 \\
&-2[(2X_H-X_H^2+X_H^3+Y_H^2+X_H Y_H^2)X_1+Y_H(2+2X_H+X_H^2+Y_H^2)Y_1]\,z^2 \\
&+4[(3+X_H+2X_H^2)X_1+Y_H(2X_H-1)Y_1]\,xz^2-2[(3+7X_H)X_1+5Y_H Y_1]\,x^2z^2 \\
&+4[(Y_H(2X_1-1)X_1+(3-X_1+2Y_1^2)Y_1]\,yz^2-4[Y_H X_1+(X_H-3)Y_1]\,xyz^2 \\
&-2[(5X_H-3)X_1+7Y_H Y_1]y^2z^2 \;-\; 8(X_H X_1+Y_H Y_1)\,z^4.
\end{aligned}
$$

5. $p(x,y,z^2)$, for the example in Section 5, equals

$$
\begin{aligned}
&332929 + 1477120x + 6214010x^2 - 928000x^3 - 9220593x^4 - 67338240x^5 + \\
&106069740x^6 + 61248000x^7 - 193147761x^8 + 75845120x^9 + 71901050x^{1}0 - \\
&70304000x^{1}1 + 17850625x^{1}2 - 3711264y - 22634880xy - 57340896x^2y - \\
&22571520x^3y + 277965504x^4y + 59508480x^5y - 395427264x^6y + 39237120x^7y + \\
&204877920x^8y - 53539200x^9y - 26364000x^{1}0y + 15518346y^2 + 119679760xy^2 + \\
&251319262x^2y^2 + 369600x^3y^2 - 866750684x^4y^2 - 162561440x^5y^2 + 750830076x^6y^2 + \\
&108422080x^7y^2 - 258020750x^8y^2 - 65910000x^9y^2 + 107103750x^{1}0y^2 - 28736736y^3 - \\
&252952320xy^3 - 555837312x^2y^3 + 368866560x^3y^3 + 951442368x^4y^3 + \\
&309154560x^5y^3 - 235048320x^6y^3 - 425068800x^7y^3 - 131820000x^8y^3 + \\
&18531855y^4 + 102859840xy^4 + 137177412x^2y^4 - 579170880x^3y^4 - 240375206x^4y^4 - \\
&94652480x^5y^4 - 46221500x^6y^4 + 439400000x^7y^4 + 267759375x^8y^4 + 9845184y^5 + \\
&309358080xy^5 + 1401870528x^2y^5 + 500597760x^3y^5 - 1090764480x^4y^5 - \\
&953971200x^5y^5 - 263640000x^6y^5 - 32442164y^6 - 355361440xy^6 - 1220951684x^2y^6 - \\
&211190720x^3y^6 + 986030500x^4y^6 + 1010620000x^5y^6 + 357012500x^6y^6 + \\
&55000896y^7 + 230680320xy^7 - 656872320x^2y^7 - 846892800x^3y^7 - 263640000x^4y^7 - \\
&36598641y^8 - 83961280xy^8 + 719137250x^2y^8 + 790920000x^3y^8 + 267759375x^4y^8 - \\
&6034080y^9 - 264451200xy^9 - 131820000x^2y^9 + 16807050y^{1}0 + 215306000xy^{1}0 + \\
&107103750x^2y^{1}0 - 26364000y^{1}1 + 17850625y^{1}2 + 898040z^2 + 58905080xz^2 + \\
&30603300x^2z^2 - 257069280x^3z^2 + 227997120x^4z^2 + 256336080x^5z^2 - \\
&752623560x^6z^2 + 295545120x^7z^2 + 421722600x^8z^2 - 353717000x^9z^2 + \\
&71402500x^{1}0z^2 - 19269120yz^2 - 330058560xyz^2 - 260713440x^2yz^2 + \\
&675230400x^3yz^2 + 368353440x^4yz^2 - 185365440x^5yz^2 - 9278880x^6yz^2 - \\
&159806400x^7yz^2 - 79092000x^8yz^2 + 121825340y^2z^2 + 390365040xy^2z^2 + \\
&13742040x^2y^2z^2 - 692867760x^3y^2z^2 - 700349520x^4y^2z^2 + 16690960x^5y^2z^2 + \\
&480399400x^6y^2z^2 + 13182000x^7y^2z^2 + 357012500x^8y^2z^2 - 275015520y^3z^2 + \\
&483787200xy^3z^2 + 2332848960x^2y^3z^2 + 664360320x^3y^3z^2 - 1609676640x^4y^3z^2 - \\
&1323067200x^5y^3z^2 - 316368000x^6y^3z^2 + 10176920y^4z^2 - 1001931840xy^4z^2 - \\
&2759428360x^2y^4z^2 - 800525440x^3y^4z^2 + 2303030600x^4y^4z^2 + 2161848000x^5y^4z^2 + \\
&714025000x^6y^4z^2 + 699023520y^5z^2 + 849725760xy^5z^2 - 1926044640x^2y^5z^2 - \\
&2166715200x^3y^5z^2 - 474552000x^4y^5z^2 - 526822400y^6z^2 - 521671280xy^6z^2 + \\
&2566873400x^2y^6z^2 + 2869282000x^3y^6z^2 + 714025000x^4y^6z^2 - 325646880y^7z^2 - \\
&1003454400xy^7z^2 - 316368000x^2y^7z^2 + 322519600y^8z^2 + 1074333000xy^8z^2 + \\
&357012500x^2y^8z^2 - 79092000y^9z^2 + 71402500y^{1}0z^2 + 43880850z^4 - 19266000xz^4 + \\
&215627100x^2z^4 + 190936200x^3z^4 - 1335133800x^4z^4 + 448188000x^5z^4 + \\
&949053300x^6z^4 - 639327000x^7z^4 + 107103750x^8z^4 - 234842400yz^4 + \\
&322857600xyz^4 + 1029818400x^2yz^4 - 4867200x^3yz^4 - 556888800x^4yz^4 - \\
&158995200x^5yz^4 - 79092000x^6yz^4 + 124705100y^2z^4 - 821914600xy^2z^4 - \\
&2606385600x^2y^2z^4 - 635440000x^3y^2z^4 + 2286755900x^4y^2z^4 + 652509000x^5y^2z^4 + \\
&428415000x^6y^2z^4 + 1029818400y^3z^4 + 835536000xy^3z^4 - 1641057600x^2y^3z^4 - \\
&1372550400x^3y^3z^4 - 237276000x^4y^3z^4 - 1274969800y^4z^4 - 1066052000xy^4z^4 + \\
&3396071900x^2y^4z^4 + 3222999000x^3y^4z^4 + 642622500x^4y^4z^4 - 662344800y^5z^4 - \\
&1213555200xy^5z^4 - 237276000x^2y^5z^4 + 915929300y^6z^4 + 1931163000xy^6z^4 + \\
&428415000x^2y^6z^4 - 79092000y^7z^4 + 107103750y^8z^4 - 19773000xz^6 - \\
&1281949500x^2z^6 + 228488000x^3z^6 + 920543000x^4z^6 - 498719000x^5z^6 + \\
&71402500x^6z^6 + 474552000yz^6 + 158184000xyz^6 - 342732000x^2yz^6 - \\
&52728000x^3yz^6 - 26364000x^4yz^6 - 1288540500y^2z^6 - 628342000xy^2z^6 + \\
&1869647000x^2y^2z^6 + 1001832000x^3y^2z^6 + 214207500x^4y^2z^6 - 342732000y^3z^6 - \\
&474552000xy^3z^6 - 52728000x^2y^3z^6 + 931528000y^4z^6 + 1500551000xy^4z^6 + \\
&214207500x^2y^4z^6 - 26364000y^5z^6 + 71402500y^6z^6 - 481966875z^8 + 321311250x^2z^8 - \\
&142805000x^3z^8 + 17850625x^4z^8 + 321311250y^2z^8 + 428415000xy^2z^8 + \\
&35701250x^2y^2z^8 + 17850625y^4z^8.
\end{aligned}
$$

REFERENCES

Alperin, R. C. (2010). The Poncelet pencil of rectangular hyperbolas. *Forum Geometricorum*, 10: 15-20.

Carver, W. B. (1956). The rectangular hyperbola. *The American Mathematical Monthly*, 63(9): 77-84.

Cox, D., Little, J., O'Shea, D. (1996) *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, 2nd ed., Springer.

Cox, D., Little, J., O'Shea, D. (2004) *Using Algebraic Geometry*, 2nd ed., Springer.

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

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

Kimberling, C. (2003) *Geometry in Action: A Discovery Approach using The Geometer's Sketchpad*, Key College Publishing.

Rieck, M. Q. (2011). An algorithm for finding repeated solutions to the general perspective three-point pose problem. *J. Mathematical Imaging and Vision*, 42(1): 92-100.

Rieck, M. Q. (2014). A fundamentally new view of the perspective three-point pose problem. *J. Mathematical Imaging and Vision*, 48(3):499-516.

Rieck, M. Q. (2015). Related solutions to the perspective three-point pose problem. *J. Mathematical Imaging and Vision*, 53(2):225-232.

Smith, A. D. N. (1965). The explicit solution of the single picture resolution problem, with a least squares adjustment to redundant control. *Photogrammetric Record*, 5(26): 113–122.

Thompson, E. H. (1966). Space resection: failure cases. *Photogrammetric Record*, 5(27): 201–204.

Wolfe, W. J., Mathis, D., Sklair, C. W., and Magee, M. (1991). The perspective view of three points. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(1): 66–73.

Zhang, C.-X. and Hu, Z.-Y. (2006). Why is the danger cylinder dangerous in the P3P problem? *Acta Automatica Sinica*, 32(4): 504–511.