# Ovals in finite projective planes

## Talk for LSBU Maths Study Group, 3 June 2021

## Tony Forbes

### 1 Finite projective planes

A finite *projective plane* consists of a finite set of points, a finite set of lines and an incidence relation such that:

(i) for any two distinct points $a, b$, there is a unique line incident with both $a$ and $b$, which we denote by $a \vee b$;

(ii) for any two distinct lines $a, b$, there is a unique point incident with both $a$ and $b$, which we denote by $a \wedge b$;

(iii) There exists a *quadrangle*, i.e. four points, no three collinear.

As a consequence of duality there exists a *quadrilateral*, i.e. four lines, no three copunctal. There exists an integer $q \geq 2$, the *order* of the plane, such that:

(i) each line is incident with $q + 1$ points;

(ii) each point is incident with $q + 1$ lines;

(iii) the number of points is $q^2 + q + 1$;

(iv) the number of lines is $q^2 + q + 1$.

A projective plane exists whenever $q$ is a prime power. If $q \equiv 1$ or $2 \pmod 4$ and $q$ is not the sum of two integer squares, there is no projective plane of order $q$ (Bruck, Ryser, 1949). There is no projective plane of order 10 (Lam, Swiercz, Thiel, 1986). There are two major unsolved problems. Is there a plane of order $q$ that is not a prime power? Is there a non-Desarguesian plane of prime order? Contrast the situation in a projective space of dimension greater than 2. In 1943 Marshall Hall wrote the following.

> The Theorem of Desargues has played a central role in the study of the foundations of projective geometry. Its validity has been shown to be equivalent to the possibility of the introduction of homogeneous coordinates from a skew-field, and also to the possibility of imbedding a plane in a space of higher dimension. A synthetic proof of the theorem, using a third dimension, is easy, but two-dimensional proofs failed and it was eventually realized that projective planes exist in which the theorem is false.

All of the finite (3 or more)-dimensional spaces must be generated from finite fields and they necessarily have prime power order. Moreover, Desargues's theorem always holds. Indeed, the first step in the usual proof of this theorem in $\mathrm{PG}(2,q)$ is to imbed it in $\mathrm{PG}(3,q)$, an option which is not available for a finite plane created from something other than a field.

## 2 The projective plane $\mathrm{PG}(2,q)$

For the rest of this talk, we will consider only the plane $\mathrm{PG}(2,q)$, where $q$ is a prime power. The points of $\mathrm{PG}(2,q)$ are 1-dimensional subspaces of the vector space $\mathrm{GF}(q)\times\mathrm{GF}(q)\times\mathrm{GF}(q)$. A point is represented by a non-zero vector $(x,y,z)$, $x,y,z\in\mathrm{GF}(q)$, with the convention that for any non-zero $t\in\mathrm{GF}(q)$, $(tx,ty,tz)$ represents the same point. To indicate this ambiguity we use the notation $[x,y,z]$, and we will always assume that the coordinates in this representation belong to $\mathrm{GF}(q)$.

The line $[a,b,c]$, $a,b,c\in\mathrm{GF}(q)$, $a,b,c$ not all zero, is defined in terms of the points $[x,y,z]$ on it by

$$[a,b,c]\cdot[x,y,z]=ax+by+cz=0.$$

As before, if $t\in\mathrm{GF}(q)$, $t\neq 0$, then $[ta,tb,tc]=[a,b,c]$. Line $[0,0,1]$ is called *the line at infinity.*

We identify three types of point and three types of line. In each case we give the representative projective coordinates in a standard form, [1].

(i) Any point not on the line at infinity has a unique representative $[x,y,1]$, $x,y\in\mathrm{GF}(q)$. Note that $[0,0,1]\cdot[x,y,1]\neq 0$. Thus $[x,y,1]$ corresponds to a point in the affine plane $\mathrm{GF}(q)\times\mathrm{GF}(q)$ in which $(x,y)$ are the usual Cartesian coordinates. There are $q^2$ such points.

(ii) For each $m\in\mathrm{GF}(q)$ there is a point $[1,m,0]$, also denoted by $(m)$, on the line at infinity; this is the meeting point of parallel lines with gradient $m$ in the affine plane. There are $q$ such points.

(iii) There is one point $[0,1,0]$, also denoted by $(\infty)$. This is the meeting point of vertical lines in the affine plane.

(i) Lines of the form $[m,-1,c]$, $m,c\in\mathrm{GF}(q)$, correspond to lines with gradient $m$ and $y$-axis intersection $c$, defined by the equation $y=mx+c$. Line $[m,-1,c]$ contains the points $[x,y,1]$ defined by $y=mx+c$ as well as the point $[1,m,0]$ on the line at infinity:

$$[m,-1,c]\cdot[x,y,1]=[m,-1,c]\cdot[1,m,0]=0.$$

(ii) Lines of the form $[-1,0,c]$, $c \in \mathrm{GF}(q)$, correspond to vertical lines with $x$-axis intersection $c$ and defined in the affine plane by $x = c$. Line $[-1,0,c]$ contains the affine points $[c,y,1]$ as well as the point $[0,1,0]$ on the line at infinity:

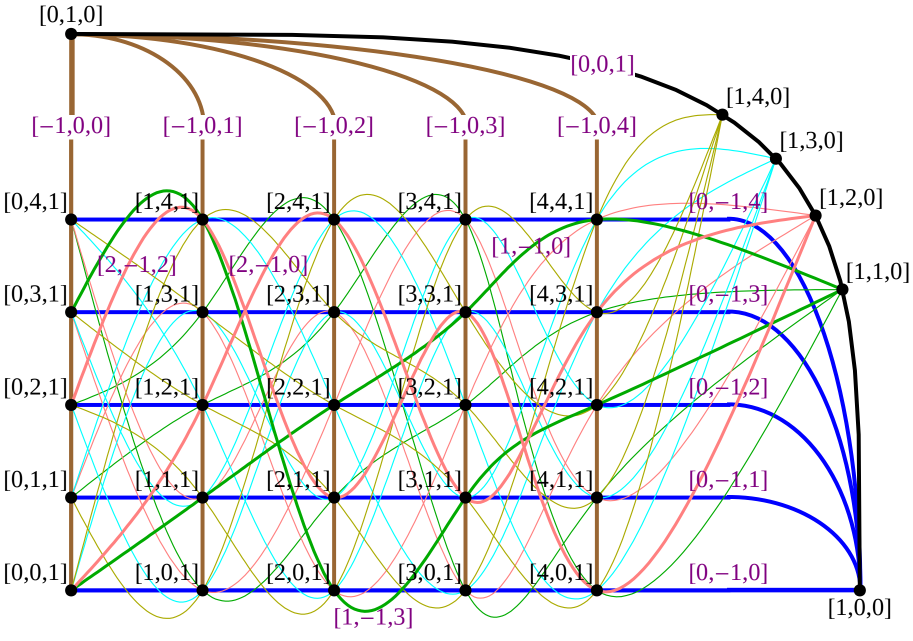
$$[-1,0,c] \cdot [c,y,1] = [-1,0,c] \cdot [0,1,0] = 0.$$

(iii) The line at infinity $[0,0,1]$ contains points $[x,y,0]$. This is the line defined by $z = 0$:

$$[0,0,1] \cdot [x,y,0] = 0.$$

(

Figure 1: Points and lines in $\mathrm{PG}(2,5)$



**Lemma 1** *The points* $[x_1,y_1,z_1]$, $[x_2,y_2,z_2]$, $[x_3,y_3,z_3]$ *are collinear iff*

$$\det \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} = 0.$$

*In particular: affine points* $[x_1, y_1, 1]$, $[x_2, y_2, 1]$, $[x_3, y_3, 1]$ *are collinear iff*

$$x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2) = 0;$$

*points* $[x_1, y_1, 1]$, $[x_2, y_2, 1]$, $[x_3, y_3, 0]$ *are collinear iff*

$$x_3(y_1 - y_2) = y_3(x_1 - x_2);$$

*distinct points* $[x_1, y_1, 1]$, $[x_2, y_2, 0]$, $[x_3, y_3, 0]$ *are not collinear; points* $[x_1, y_1, 0]$, $[x_2, y_2, 0]$, $[x_3, y_3, 0]$ *are collinear.*

**Proof** This is high-school geometry. □

**Lemma 2** *The product of the non-zero elements of* $\mathrm{GF}(q)$ *is* $-1$*.*

**Proof** If $x^2 = 1$, then $x = \pm 1$, [5]. So every element of $\mathrm{GF}(q) \setminus \{0, 1, -1\}$ can be paired with its multiplicative inverse. □

Let $P$ be a set of points. A line $\ell$ is

(i) *exterior* to $P$ if $\ell$ is incident with no point of $P$,

(ii) *tangent* to $P$ if $\ell$ is incident with precisely one point of $P$,

(iii) *secant* to $P$ if $\ell$ is incident with precisely two points of $P$,

(iv) none of the above if $\ell$ is incident with more than two points of $P$.

**Lemma 3** *Let* $U$ *be a set of* $k \geq 4$ *points in* $\mathrm{PG}(2, q)$*, no three collinear, and assume* $U$ *contains the points*

$$u_0 = [1, 0, 1], \quad u_1 = [1, 0, 0], \quad u_2 = [0, 0, 1].$$

*Then*

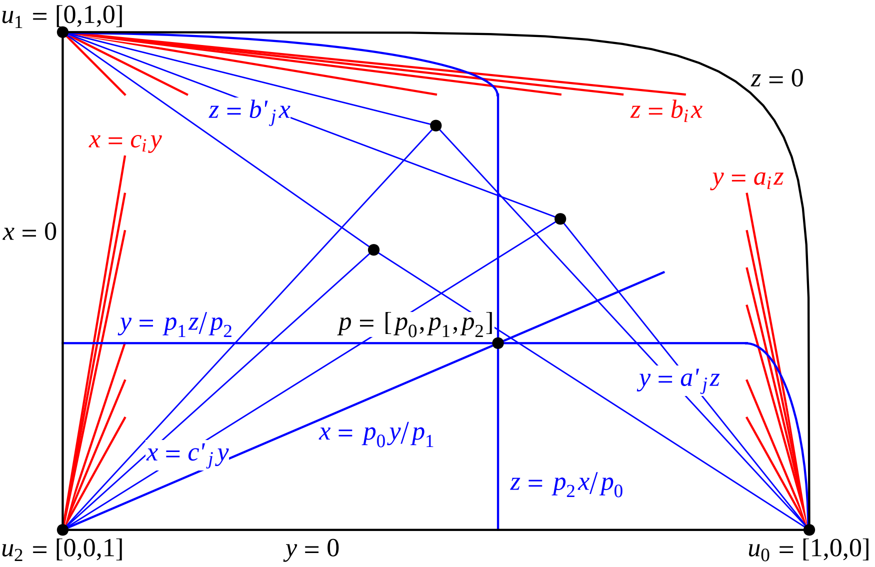
(i) *Through each point* $u_0$, $u_1$, $u_2$ *there are* $t = q - k + 2$ *tangents to* $U$ *with equations*

$$y = a_i z \text{ for } u_0, \quad z = b_i x \text{ for } u_1, \quad x = c_i y \text{ for } u_2, \qquad i = 1, 2, \ldots, t;$$

(ii) *The lines* $x = 0$, $y = 0$ *and* $z = 0$ *are secants to* $U$*.*

(iii) $a_1 b_1 c_1 \; a_2 b_2 c_2 \; \ldots \; a_t b_t c_t = -1$*.*

Figure 2: Lemma 3



**Proof** The lemma holds vacuously in PG$(2,2)$ since $k = 4$ and $t = 0$. So we may assume $q > 2$. See Figure 2, where $p$ is a typical point of $U \setminus \{u_0, u_1, u_2\}$. Tangents to $U$ are red, secants to $U$ are blue and black, points of $U$ are black dots.

(i), (ii): Straightforward.

(iii): This is based on [2, Lemma 2.5]. The $q + 1$ lines through $u_0$ consist of 2 secants (black lines), $u_0 \vee u_1$, $u_0 \vee u_2$, a further $k - 3$ secants (blue lines) with equations $y = a'_j z$, $j = 1, 2, \ldots, k-3$, say, and $t = q-k+2$ tangents (red lines) with equations $y = a_i z$, $i = 1, 2, \ldots, t$. It is clear that the $a_i$ together with the $a'_i$ are distinct and non-zero. Hence they precisely cover the non-zero elements of GF$(q)$, and therefore their product is $-1$ by Lemma 2.

Similarly for the $q + 1$ lines through $u_1$, denoting the equations of the red lines through $u_1$ by $z = b_i x$, $i = 1, 2, \ldots, t$ and the blue lines by $z = b'_j x$, $j = 1, 2, \ldots, k - 3$.

Similarly for the $q + 1$ lines through $u_2$, denoting the equations of the red lines through $u_2$ by $x = c_i y$, $i = 1, 2, \ldots, t$ and the blue lines by $x = c'_j y$,

$j = 1, 2, \ldots, k-3$.

Hence the product of all of the $a_i$, $b_i$, $c_i$, $a'_j$, $b'_j$, $c'_j$ is $-1$.

Consider a typical point $p = [p_0, p_1, p_2]$ in $U \setminus \{u_0, u_1, u_2\}$. The equation of the secant $u_0 \vee p$ is $y = p_1 z/p_2$, the equation of the secant $u_1 \vee p$ is $z = p_2 x/p_0$, and the equation of the secant $u_2 \vee p$ is $x = p_0 y/p_1$. Moreover, $p_1/p_2 \cdot p_2/p_0 \cdot p_0/p_1 = 1$. Therefore the product of the all of the $a'_j$, $b'_j$, $c'_j$ is 1. Hence the product of all of the $a_i$, $b_i$, $c_i$ is $-1$. □

## 3 Ovals

An *oval* is a set of points $O$ such that:

(i) any line meets $O$ in at most 2 points;

(ii) for each point $p \in O$, there exists precisely one tangent to $O$ incident with $p$.

A set $P$ of points where no three are collinear cannot have size greater than $q+2$. If $|P| = q+1$, it is an oval; if $|P| = q+2$, it is called a *hyperoval.*

For any oval, there are $q+1$ tangents, $(q^2+q)/2$ secants and $(q^2-q)/2$ exterior lines. In Figure 3 we illustrate Lemma 3 when $U$ is an oval. Here we have $q = 5$, $|U| = 6$, $t = 1$, $a_1 = 3$, $b_1 = 1/3$, $c_1 = 1/4$, and $a_1 b_1 c_1 = 4 = -1$.

A good question to ask is, “How many tangents to an oval go through a point not on it?” An even better question is, “Given a number $i$, how many points not on an oval, $O$ are on $i$ tangents to $O$?” This last question can be answered completely, [3, §9.7].

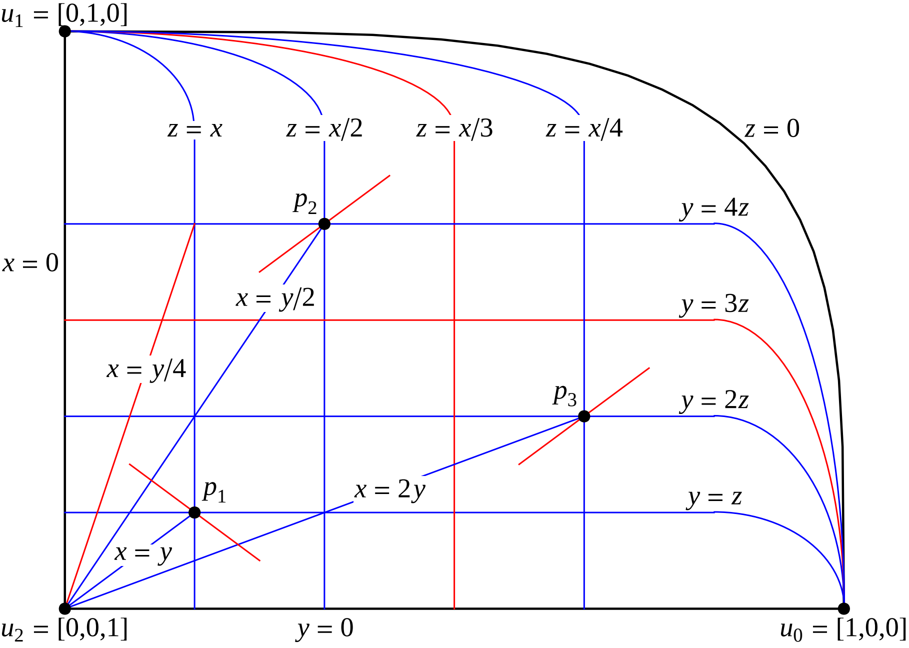
Given an oval, $O$, let $T_i$ be the number of points not on $O$ which lie on $i$ tangents to $O$.

Non-tangent lines meet $O$ in an even number of points and $|O|$ has opposite parity to $q$; therefore $T_i = 0$ if $i$ has the same parity as $q$. Also we have

$$\sum_{i=0}^{q+1} T_i = q^2, \quad \sum_{i=0}^{q+1} iT_i = (q+1)q, \quad \sum_{i=0}^{q+1} i(i-1)T_i = (q+1)q. \qquad (1)$$

The first follows because we are counting points not on $O$. For the second, we are counting ‘(tangent, point)’ combinations; there are $q+1$ tangents and each one has $q$ points not on $O$. For the third, divide by 2 to get $\sum_{i=0}^{q+1} \binom{i}{2} T_i$; then count ‘((pair of tangents), point)’ combinations. But there

Figure 3: Lemma 3 with $q = 5$ and $|U| = 6$



are $\binom{q+1}{2} = \frac{(q+1)q}{2}$ pairs of tangents to $O$ and each pair has precisely one common point, which is not in $O$.

Suppose $q$ is even. Then $T_i = 0$ when $i$ is even. Calling the sums in (1) $A$, $B$, $C$ in that order, $(q+1)(B-A)-C$ yields

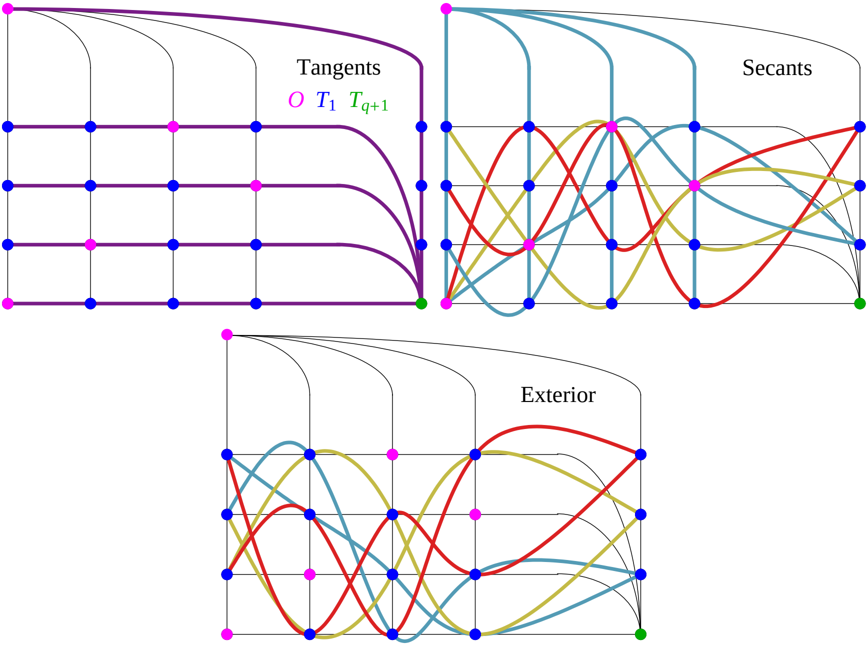
$$\sum_{i=3}^{q-1}\big((q+1)(i-1)-i(i-1)\big)T_i = (q+1)\big((q+1)q-q^2\big)-(q+1)q = 0,$$

the coefficients of $T_1$ and $T_{q+1}$ vanishing. Therefore $T_3 = T_4 = \ldots = T_q = 0$ and the only solution of (1) in non-negative integers $T_0, T_1, \ldots, T_{q+1}$ is

$$T_1 = q^2-1, \quad T_{q+1} = 1, \quad T_i = 0 \text{ for all other } i.$$

Since $T_{q+1} = 1$ there is a unique point $c$ that extends $O$ to a hyperoval. The $q+1$ tangents through $c$ each have $q-1$ points not in $O \cup \{c\}$, $q^2-1$ points altogether.

Figure 4: Oval $[x, x^2, 1] \cup [0, 1, 0]$ in PG(2, 4)



Now suppose $q$ is odd. Then $T_1 = 0$ and $C - B$ gives $\sum_{i=0}^{q+1} i(i-2)T_i = 0$. Therefore $T_i = 0$ unless $i = 0$ or $i = 2$. Hence

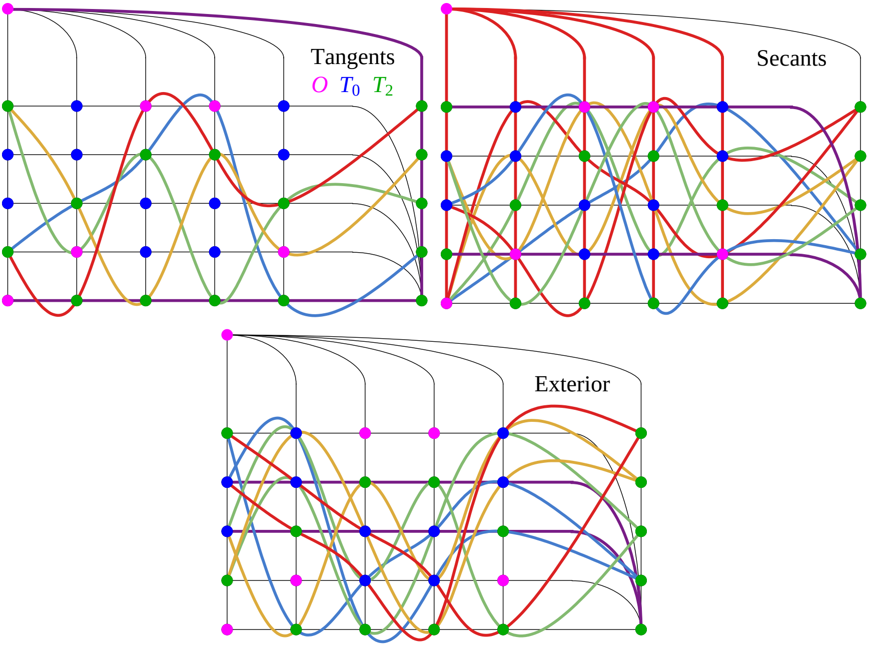
$$T_0 = \frac{q(q-1)}{2}, \quad T_2 = \frac{(q+1)q}{2}, \quad T_i = 0 \text{ for all other } i.$$

The oval cannot be extended to a hyperoval.

In the case of a point $p$ counted by $T_2$ the line joining the two tangent points is called the *polar line* of $p$. Observe also that this relation defines a 1–1 correspondence between $T_2$ points and secants.

Observe also that $T_0$ the same as the number of exterior lines. Moreover, each exterior line contains $(q+1)/2$ $T_0$ points and $(q+1)/2$ $T_1$ points. To see this, let $e$ be an exterior line to $O$, let $t$ be a tangent, and let $q = t \wedge e$. Then $q$ is a $T_2$ point on $e$ an so there must be a second tangent, $u$ say, such that $q = u \wedge e$. Continue to pair off the tangents.

Figure 5: Oval $[x, x^2, 1] \cup [0, 1, 0]$ in PG$(2, 5)$



## 4 Conics

A *conic*, $C$ is the set of points defined by

$$C = \{[x, y, z] : ax^2 + by^2 + cz^2 + fyz + gzx + hxy = 0\}, \qquad (2)$$

provided that the quadratic is not degenerate. A homogeneous ternary quadratic form $Q(x, y, z)$ is degenerate, if there exists a non-singular linear substitution $(x, y, z) \mapsto P(x, y, z)$ that transforms $Q(x, y, z)$ into a form with fewer than three variables, [3, §9.7].

As an illustration of a degenerate quadratic, which does not generate a conic, consider $q = 7$ and

$$D = \{[x, y, z] : x^2 + y^2 + hxy = 0\} \qquad (3)$$

for various $h$. Observe that $D$ has fewer than three variables. Completing the square, multiplying by 4 and square-rooting, we obtain

$$2y + hx = \pm\sqrt{h^2 - 4}\, x. \qquad (4)$$

If $h \in \{0,3,4\}$, the square root does not exist and the only element of $D$ is $[0,0,1]$. If $h=1$, (4) reduces to $y=(3\pm 1)x$ and $D$ consist of two lines that meet at $[0,0,1]$,

$$\{[x,2x,1] : 0 \le x < 7\} \cup \{[1,2,0]\}, \quad \{[x,4x,1] : 0 \le x < 7\} \cup \{[1,4,0]\}.$$

Similarly, when $h=6$ we get $y=(4\pm 1)x$, two lines that meet at $[0,0,1]$,

$$\{[x,3x,1] : 0 \le x < 7\} \cup \{[1,3,0]\}, \quad \{[x,5x,1] : 0 \le x < 7\} \cup \{[1,5,0]\}.$$

If $h=5=-2$, (4) reduces to $(y-x)^2=0$ and now $D$ consist of a single line,

$$\{[x,x,1] : 0 \le x < 7\} \cup \{[1,1,0]\}.$$

Finally, when $h=2$, (4) reduces to $(y+x)^2=0$ and $D$ consist of the line

$$\{[x,6x,1] : 0 \le x < 7\} \cup \{[1,6,0]\}.$$

For an alternative definition, a quadratic form $Q(x,y,z)$ in $\mathrm{PG}(2,q)$ is *degenerate* if $\{[x,y,z] : Q(x,y,z)=0\}$ either consists of a single point or contains a line.

Let $Q(x,y,z)$ be a not necessarily non-degenerate homogeneous ternary quadratic form and let

$$R = \{[x,y,z] : Q(x,y,z) = 0\}.$$

Then we have the following possibilities.

(i) $|R|=1$ and $Q(x,y,z)$ is degenerate.

(ii) $|R|=2q+1$, $Q(x,y,z)$ is degenerate and $R$ consists of two lines.

(iii) $|R|=q+1$, $Q(x,y,z)$ is degenerate and $R$ consists of a line. This happens when the quadratic is something squared.

(iv) $|R|=q+1$, $Q(x,y,z)$ is not degenerate and $R$ is a conic.

**Theorem 4** *A conic is an oval.*

**Proof** [3, §9.7] Let $C$ be the conic

$$C = \{[x,y,z] : ax^2+by^2+cz^2+fyz+gzx+hxy=0\},$$

Take a line, $\ell$, which w.l.o.g. we may assume is the line at infinity, $z=0$. Then

$$C \cap \ell = \{[x,y,0] : ax^2+by^2+hxy=0\}. \tag{5}$$

Suppose $a = b = h = 0$. Then the conic would be defined by $z(cz + fy + gx) = 0$ and a non-singular linear substitution would reduce it to less than three variables. To see this, observe that if $g = 0$, the quadratic has at most two variables. Otherwise the substitution $x = x' - fy' - cz'$, $y = gy'$, $z = gz'$ has determinant $g^2$ and transforms the quadratic to $g^2x'z'$, which involves only two variables. Hence $a$, $b$, $h$ are not all zero.

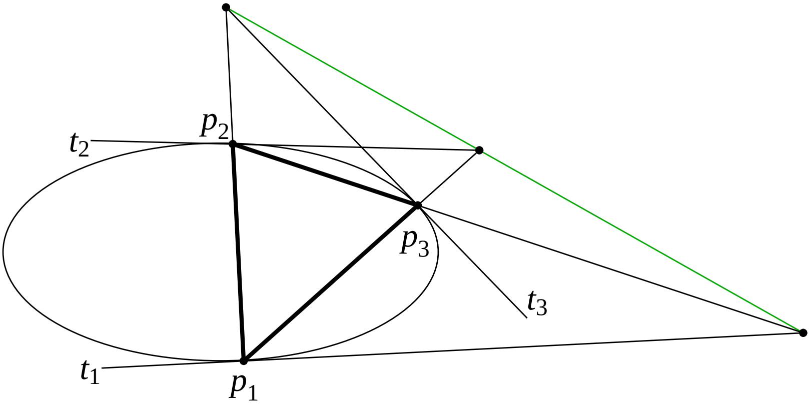
Suppose $a \neq 0$. The point $[1, 0, 0]$ does not lie on the conic (5). Any other point of $\ell$ has representative $[x, 1, 0]$, which lies on (5) iff $ax^2+hx+b = 0$, and this quadratic has at most two solutions.

If $b \neq 0$, the argument is similar.

If $a = b = 0$, then (5) becomes $hxy = 0$ and the only two points of $\ell$ that satisfy it are $[1, 0, 0]$ and $[0, 1, 0]$.

Hence at most two points of $\ell$ are on the conic. Therefore, since any conic has $q + 1$ points, $C$ must be an oval. □

Figure 6: Pascal's 3-point theorem in $\mathbb{R}^2$



**Theorem 5 (Pascal)** *Let $C$ be a conic in* $\mathrm{PG}(2, q)$*. Let $p_1$, $p_2$ and $p_3$ be distinct points on $C$. Let $t_i$ be the tangent to $C$ at $p_i$, $i = 1, 2, 3$. Then the points $t_1 \wedge (p_2 \vee p_3)$, $t_2 \wedge (p_3 \vee p_1)$ and $t_3 \wedge (p_1 \vee p_2)$ are collinear.*

**Proof** To illustrate what the theorem is saying see Figure 6, the picture corresponding to $\mathbb{R}^2$. This is actually a degenerate form of Pascal's hexagon theorem. □

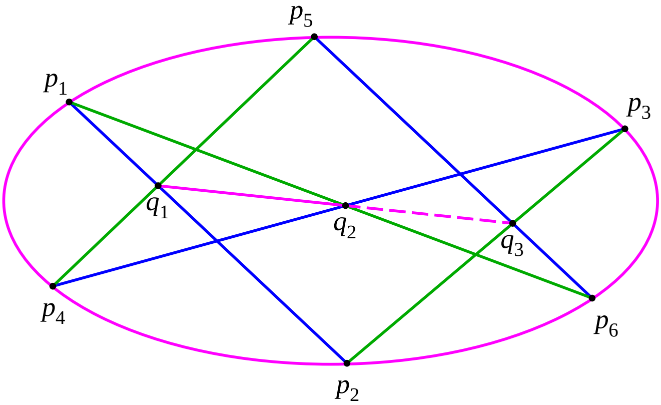
**Theorem 6 (Pascal's hexagon theorem)** *Let $C$ be a conic in* $\mathrm{PG}(2, q)$,

*$q \geq 5$, let $(p_1, p_2, p_3, p_4, p_5, p_6)$ be a hexagon with distinct points on $C$ and let*

$$q_1 = (p_1 \vee p_2) \wedge (p_4 \vee p_5), \quad q_2 = (p_3 \vee p_4) \wedge (p_6 \vee p_6), \quad q_3 = (p_5 \vee p_6) \wedge (p_2 \vee p_3).$$

*Then $q_1$, $q_2$ and $q_3$ are collinear.*

Figure 7: Pascal's hexagon theorem in $\mathbb{R}^2$



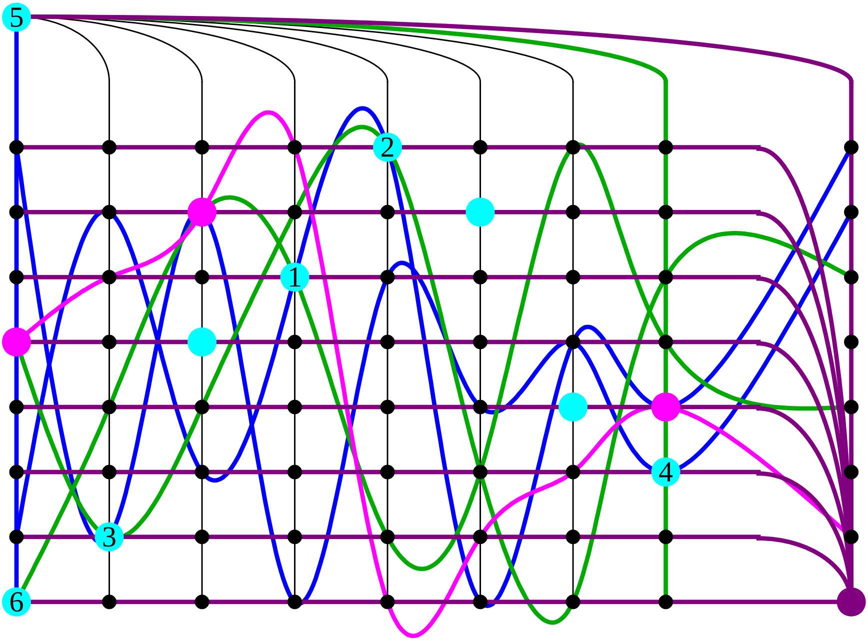
**Proof** As a diversion from ovals in finite planes, we prove this theorem in $\mathbb{R}^2$; see Figure 7 and [4]. We use the Cayley–Bacharach theorem, a corollary of which states that if two general cubics intersect in 9 points, then any other cubic through 8 of the points must go through the 9th. See Theorem 4 in `https://www.theoremoftheday.org/MathsStudyGroup/ADF34C.pdf`. In our case we consider the two cubics $(p_1 \vee p_2) \cup (p_3 \vee p_4) \cup (p_5 \vee p_6)$ (blue lines) and $(p_1 \vee p_6) \cup (p_2 \vee p_3) \cup (p_4 \vee p_5)$ (green lines). These intersect in the 9 points $p_1$, $p_2$, $p_3$, $p_4$, $p_5$, $p_6$, $q_1$, $q_2$, $q_3$. But the cubic $C \cup (q_1 \vee q_2)$ (magenta) goes through 8 of these points, and therefore it must go through the 9th, $q_3$. So either $q_3$ is on the conic, which it obviously isn't, or $q_3$ is on the line $(q_1 \vee q_2)$. □

**Theorem 7 (Segre, 1955)** *If $q$ is odd, an oval in* $\mathrm{PG}(2, q)$ *is a conic.*

**Proof** See [3, §9.7]. □

The oddness of $q$ is relevant in Segre's theorem. Let $C$ be a conic, and hence an oval, generated by a quadratic $Q$ in $\mathrm{PG}(2, q)$, $q$ even, $q \geq 8$. As

Figure 8: Pascal's hexagon theorem for oval $[x, x^2, 1] \cup [0, 1, 0]$ in $\mathrm{PG}(2, 8)$



we have seen, there is a unique point $c$ that extends $C$ to a hyperoval. Now exchange $c$ for a point $d$ on $C$ to obtain another oval. Since a quadratic is defined by 5 points, the set $C \setminus \{d\}$ can only be generated by $Q$, which excludes $c$. Therefore the oval $(C \setminus \{d\}) \cup \{c\}$ cannot be generated by a quadratic.

## References

[1] A. A. Albert and R. Sandler, *An Introduction to Finite Projective Planes*, Holt, Reinhart & Winston, 1968, Dover, 2015.

[2] M. Brown, (Hyper)ovals and ovoids in projective spaces, *Finite Geometry and its Applications*, Socrates Intensive Course, 2000.

[3] P. J. Cameron, *Combinatorics: Topics, Techniques, Algorithms*, Cambridge Univ. Press, 1994.

[4] J. Conway and A. Ryba, The Pascal Mysticum Demystified, *Math. Intelligencer*, September, 2012.

[5] L. E. Dickson, *Linear Groups with an Exposition of the Galois Field Theory*, Teubner, 1901, Dover, 1958.