London Mathematical Society Student Texts 12

Undergraduate Algebraic Geometry

MILES REID
Mathematics Institute, University of Warwick

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge CB2 1RP, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK http://www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011–4211, USA http://www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1988
Reprinted 1990 (with corrections), 1990, 1992, 1994, 1998

Typeset in Times New Roman

A catalogue record for this book is available from the British Library

ISBN 0 521 35662 8 paperback

Transferred to digital printing 2001

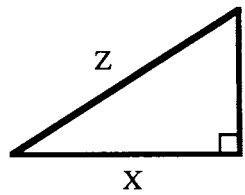
Contents

Chapter I. Playing with plane curves

§1. Plane conics

I start by studying the geometry of conics as motivation for the projective plane $\mathbb{P}^2$. Projective geometry is usually mentioned in 2nd year undergraduate geometry courses, and I recall some of the salient features, with some emphasis on homogeneous coordinates, although I completely ignore the geometry of linear subspaces and the 'cross-ratio'. The most important aim for the student should be to grasp the way in which geometric ideas (for example, the idea that 'points at infinity' correspond to asymptotic directions of curves) are expressed in terms of coordinates. The interplay between the intuitive geometric picture (which tells you what you should be expecting), and the precise formulation in terms of coordinates (which allows you to cash in on your intuition) is a fascinating aspect of algebraic geometry.

(1.1) Example of a parametrised curve. Pythagoras' Theorem says that, in the diagram



, $X^2+Y^2 = Z^2$,

so (3, 4, 5) and (5, 12, 13), as every ancient Egyptian knew. How do you find all integer solutions? The equation is homogeneous, so that $x = X/Z$, $y = Y/Z$ gives the circle $C : (x^2 + y^2 = 1) \subset \mathbb{R}^2$, which can easily be seen to be parametrised as

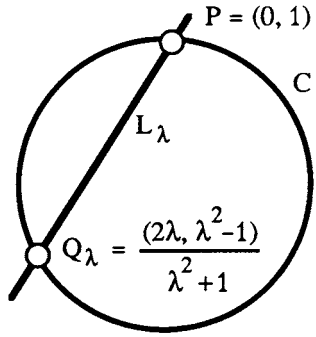
$$x = 2\lambda/(\lambda^2 + 1),\ y = (\lambda^2 - 1)/(\lambda^2 + 1), \quad \text{where} \quad \lambda = x/(1 - y);$$

so this gives all solutions:

$$X = 2\ell m,\ Y = \ell^2 - m^2,\ Z = \ell^2 + m^2 \quad \text{with} \quad \ell, m \in \mathbb{Z} \text{ coprime},$$

(or each divided by 2 if ℓ, m are both odd). Note that the equation is homogeneous, so that if (X, Y, Z) is a solution, then so is $(\lambda X, \lambda Y, \lambda Z)$.

Maybe the parametrisation was already familiar from school geometry, and in any case, it's easy to verify that it works. However, if I didn't know it already, I could have obtained it by an easy geometric argument, namely linear projection from a given point:



$P = (0, 1) \in C$, and if $\lambda \in \mathbb{Q}$ is any value, then the line L_λ through P with slope $-\lambda$ meets C in a further point Q_λ. This construction of a map by means of linear projection will appear many times in what follows.

(1.2) Similar example. C: $(2X^2 + Y^2 = 5Z^2)$. The same method leads to the parametrisation $\mathbb{R} \longrightarrow C$ given by

$$x = \frac{2\sqrt{5}\lambda}{1+2\lambda^2}, \quad y = \frac{2\lambda^2 - 1}{1+2\lambda^2}.$$

This allows us to understand all about points of C with coefficients in $\mathbb{R}$, and there's no real difference from the previous example; what about $\mathbb{Q}$?

Proposition. If $(a, b, c) \in \mathbb{Q}$ satisfies $2a^2 + b^2 = 5c^2$ then $(a, b, c) = (0, 0, 0)$.

Proof. Multiplying through by a common denominator and taking out a common factor if necessary, I can assume that a, b, c are integers, not all of which are divisible by 5; also if $5 \mid a$ and $5 \mid b$ then $25 \mid 5c^2$, so that $5 \mid c$, which contradicts what I've just said. It is now easy to get a contradiction by considering the possible values of a and b mod 5: since any square is 0, 1 or 4 mod 5, clearly $2a^2 + b^2$ is one of 0+1, 0+4, 2+0, 2+1, 2+4, 8+0, 8+1 or 8+4 mod 5, none of which can be of the form $5c^2$. Q.E.D.

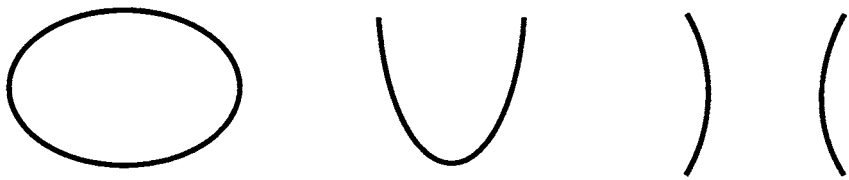
Note that this is a thoroughly arithmetic argument.

(1.3) Conics in $\mathbb{R}^2$. A conic in $\mathbb{R}^2$ is a plane curve given by a quadratic equation

$$q(x,y) = ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

Everyone has seen the classification of nondegenerate conics:

(a) ellipse (b) parabola (c) hyperbola



$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad y = mx^2 \qquad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

in addition, there are a number of peculiar cases:

(d) single point given by $x^2 + y^2 = 0$;

(e, f, g) empty set given by any of the 3 equations: (e) $x^2 + y^2 = -1$, (f) $x^2 = -1$ or (g) $0 = 1$. These three equations are different, although they define the same locus of zeros in $\mathbb{R}^2$; consider for example their complex solutions.

(h) line $x = 0$;

(i) line pair $xy = 0$;

(j) parallel lines $x(x - 1) = 0$;

(k) 'double line' $x^2 = 0$;

you can choose for yourself whether you'll allow the final case:

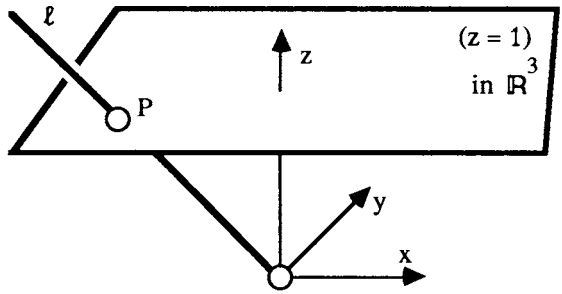
(l) whole plane given by $0 = 0$.

(1.4) Projective plane. The definition 'out of the blue':

$$\begin{aligned}\mathbb{P}^2_{\mathbb{R}} &= \{\text{lines of } \mathbb{R}^3 \text{ through origin}\} \\ &= \{\text{ratios } X : Y : Z\} \\ &= (\mathbb{R}^3 \setminus \{0\})/\sim, \quad \text{where } (X, Y, Z) \sim (\lambda X, \lambda Y, \lambda Z) \text{ if } \lambda \in \mathbb{R} \setminus \{0\}.\end{aligned}$$

(The sophisticated reader will have no difficulty in generalising from $\mathbb{R}^3$ to an arbitrary vector space over a field, and in replacing work in a chosen coordinate system with intrinsic arguments.)

To represent a ratio $X : Y : Z$ for which $Z \neq 0$, I can set $x = X/Z$, $y = Y/Z$; this simplifies things, since the ratio corresponds to just two real numbers. In other words, the equivalence class of (X, Y, Z) under $\sim$ has a unique representative $(x, y, 1)$ with 3rd coordinate $= 1$. Unfortunately, sometimes Z might be $= 0$, so that this way of choosing a representative of the equivalence class is then no good. This discussion means that $\mathbb{P}^2_{\mathbb{R}}$ contains a copy of $\mathbb{R}^2$. A picture:



$$\mathbb{R}^2 \hookrightarrow \mathbb{R}^3 \setminus \{0\} \to \mathbb{P}^2_{\mathbb{R}} \quad \text{by} \quad (x, y) \mapsto (x, y, 1)$$

the general line in $\mathbb{R}^3$ through 0 is not contained in the plane $(Z = 0)$, so that it meets $(Z = 1)$ in exactly one point, which is a representative for that equivalence class. The lines in $(Z = 0)$ never meet $(Z = 1)$, so they correspond not to points of $\mathbb{R}^2$, but to *asymptotic directions*, or to pencils of parallel lines of $\mathbb{R}^2$; so you can think of $\mathbb{P}^2_{\mathbb{R}}$ as consisting of $\mathbb{R}^2$ together with one 'point at infinity' for every pencil of parallel lines. From this point of view, you calculate in $\mathbb{R}^2$, try to guess what's going on at infinity by some kind of 'asymptotic' argument, then (if necessary), prove it in terms of homogeneous coordinates. The definition in terms of lines in $\mathbb{R}^3$ makes this respectable, since it treats all points of $\mathbb{P}^2_{\mathbb{R}}$ on an equal footing.

Groups of transformations are of central importance throughout geometry; properties of a geometric figure must be invariant under the appropriate kind of transformations before they are significant. An *affine* change of coordinates in $\mathbb{R}^2$ is of the form $T(\mathbf{x}) = A\mathbf{x} + B$, where $\mathbf{x} = (x, y) \in \mathbb{R}^2$, and A is a 2×2 invertible matrix, B a translation vector; if A is orthogonal then the transformation T is *Euclidean*. As everyone knows, every nondegenerate conic can be reduced to one of the standard forms (a-c) above by a Euclidean transformation. It is an exercise to the reader to show that every conic can be reduced to one of the forms (a-l) by an affine transformation.

A *projectivity*, or projective transformation of $\mathbb{P}^2_{\mathbb{R}}$ is of the form $T(\mathbf{X}) = M\mathbf{X}$, where M is an invertible 3×3 matrix. It's easy to understand the effect of this transformation on the affine piece $\mathbb{R}^2 \subset \mathbb{P}^2_{\mathbb{R}}$: as a partially defined map $\mathbb{R}^2 \dashrightarrow \mathbb{R}^2$, it is the fractional-linear transformation

$$\begin{bmatrix} x \\ y \end{bmatrix} \longmapsto (A \begin{bmatrix} x \\ y \end{bmatrix} + B)/(cx + dy + e),$$

where

$$M = \left[\begin{array}{c|c} A & B \\ \hline c\ d & e \end{array}\right].$$

T is of course not defined when $cx + dy + e = 0$. Perhaps this looks rather unintuitive, but it really occurs in nature: two different photographs of the same (plane) object are obviously related by a projectivity; see for example [Berger, 4.7.4] for pictures. So a math graduate getting a job interpreting satellite photography (whether for the peaceful purposes of the Forestry Commission, or as part of the vast career prospects opened up by President Reagan's defence policy) will spend a good part of his or her time computing projectivities.

Projective transformations are implicitly in use throughout these notes, usually in the form 'by a suitable choice of coordinates, I can assume ..'.

(1.5) Equation of a conic. The inhomogeneous quadratic polynomial

$$q(x, y) = ax^2 + bxy + cy^2 + dx + ey + f$$

corresponds to the homogeneous quadratic

$$Q(X, Y, Z) = aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2;$$

the correspondence is easy to understand as a recipe, or you can think of it as the bijection $q \longleftrightarrow Q$ given by

$$q(x, y) = Q(X/Z, Y/Z, 1) \quad \text{with} \quad x = X/Z,\ y = Y/Z$$

and inversely,

$$Q = Z^2 q(X/Z, Y/Z).$$

A *conic* $C \subset \mathbb{P}^2$ is the curve given by C: $(Q(X, Y, Z) = 0)$, where Q is a homogeneous quadratic expression; note that the condition $Q(X, Y, Z) = 0$ is well defined on the equivalence class, since $Q(\lambda \mathbf{X}) = \lambda^2 Q(\mathbf{X})$ for any $\lambda \in \mathbb{R}$. As an exercise, check that the projective curve C meets the affine piece $\mathbb{R}^2$ in the affine conic given by $(q = 0)$.

'Line at infinity' and asymptotic directions. Points of $\mathbb{P}^2$ with $Z = 0$ correspond to ratios $(X : Y : 0)$. These points form the 'line at infinity', a copy of $\mathbb{P}^1_{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ (since $(X : Y) \mapsto X/Y$ defines a bijection $\mathbb{P}^1_{\mathbb{R}} \to \mathbb{R} \cup \{\infty\}$).

A line in $\mathbb{P}^2$ is by definition given by $L: (aX + bY + cZ = 0)$, and

$$L \text{ passes through } (X, Y, 0) \iff aX + bY = 0.$$

In affine coordinates the same line is given by $ax + by + c = 0$, so that all lines with the same ratio $a : b$ pass through the same point at infinity. This is called 'parallel lines meet at infinity'.

Examples. (a) The hyperbola $(x^2/a^2 - y^2/b^2 = 1)$ in $\mathbb{R}^2$ corresponds in $\mathbb{P}^2_{\mathbb{R}}$ to $C: (X^2/a^2 - Y^2/b^2 = Z^2)$; clearly this meets $(Z = 0)$ in the two points $(b, \pm a, 0) \in \mathbb{P}^2_{\mathbb{R}}$, corresponding in the obvious way to the asymptotic lines of the hyperbola.

Note that in the affine piece $(X \neq 0)$ of $\mathbb{P}^2_{\mathbb{R}}$, the affine coordinates are $u = Y/X$, $v = Z/X$, so that C becomes the ellipse $(u^2/b^2 + v^2 = 1/a^2)$. See Ex. 1.7 for an artistic interpretation.

(b) The parabola $(y = mx^2)$ in $\mathbb{R}^2$ corresponds to $C: (YZ = mX^2)$ in $\mathbb{P}^2_{\mathbb{R}}$; this now meets $(Z = 0)$ at the single point $(0, 1, 0)$. So in $\mathbb{P}^2$, the 'two branches of the parabola meet at infinity'; note that this is a statement with intuitive content (maybe you feel it's pretty implausible?), but is not a result you could arrive at just by contemplating within $\mathbb{R}^2$ – maybe it's not even meaningful.

(1.6) Classification of conics in $\mathbb{P}^2$. Let k be any field of characteristic $\neq 2$; recall two results from the linear algebra of quadratic forms:

Proposition (A). There are natural bijections

$$\left\{\begin{matrix}\text{homogeneous} \\ \text{quadratic polys}\end{matrix}\right\} = \left\{\begin{matrix}\text{quad. forms} \\ k^3 \to k\end{matrix}\right\} \overset{\text{bij}}{\longleftrightarrow} \left\{\begin{matrix}\text{symmetric bilinear} \\ \text{forms on } k^3\end{matrix}\right\},$$

given in formulas by

$$aX^2 + 2bXY + cY^2 + 2dXZ + 2eYZ + fZ^2 \longleftrightarrow \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix}.$$

A quadratic form is *nondegenerate* if the corresponding bilinear form is nondegenerate, that is, its matrix is nonsingular.

Theorem (B). Let V be a vector space over k and $Q: V \to k$ a quadratic form; then there exists a basis of V such that

$$Q = \varepsilon_1 x_1^2 + \varepsilon_2 x_2^2 + .. \varepsilon_n x_n^2,$$

with $\varepsilon_i \in k$.

(This is proved by *Gram-Schmidt orthogonalisation*, if that rings a bell.) Obviously, for $\lambda \in k \setminus \{0\}$ the substitution $x_i \mapsto \lambda x_i$ takes ε_i into $\lambda^{-2}\varepsilon_i$.

Corollary. In a suitable system of coordinates, any conic in $\mathbb{P}^2_{\mathbb{R}}$ is one of the following:

(α) nondegenerate conic, C: $(X^2 + Y^2 - Z^2 = 0)$;

(β) empty set, given by $(X^2 + Y^2 + Z^2 = 0)$;

(γ) line pair, given by $(X^2 - Y^2 = 0)$;

(δ) one point (0, 0, 1), given by $(X^2 + Y^2 = 0)$;

(ε) double line, given by $(X^2 = 0)$.

(Optionally you have the whole of $\mathbb{P}^2_{\mathbb{R}}$ given by $(0 = 0)$.)

Proof. Any real number ε is either 0, or $\pm$ a square, so that I only have to consider Q as in the theorem with $\varepsilon_i = 0$ or ± 1. In addition, since I'm only interested in the locus $(Q = 0)$, I'm allowed to multiply Q through by -1. This leads at once to the given list. Q.E.D.

There are two points to make about this corollary: firstly, the list is quite a lot shorter than that in (1.3); for example, the 3 nondegenerate cases (ellipse, parabola, hyperbola) of (1.3) all correspond to case (α), and the 2 cases of intersecting and parallel line pairs are not distinguished in the projective case. Secondly, the derivation of the list from general algebraic principles is much simpler.

(1.7) Parametrisation of a conic. Let C be a nondegenerate, nonempty conic of $\mathbb{P}^2_{\mathbb{R}}$. Then by Corollary 1.6, taking new coordinates (X+Z, Y, Z-X), C is projectively equivalent to the curve $(XZ = Y^2)$; this is the curve parametrised by

$$\Phi: \mathbb{P}^1_{\mathbb{R}} \longrightarrow C \subset \mathbb{P}^2_{\mathbb{R}},$$

$$(U:V) \longmapsto (U^2:UV:V^2).$$

Remarks 1. The inverse map $\Psi: C \to \mathbb{P}^1_{\mathbb{R}}$ is given by $(X:Y:Z) \mapsto (X:Y) = (Y:Z)$; here the left-hand ratio is defined if $X \neq 0$, and the right-hand ratio if $Z \neq 0$. In terminology to be introduced later, Φ and Ψ are inverse isomorphisms of varieties.

2. Throughout §§1-2, nonempty nondegenerate conics are tacitly assumed to be projectively equivalent to $(XZ - Y^2)$; over a field of characteristic $\neq 2$, this is justified in Ex. 1.5. (The reader interested in characteristic 2 should take this as the definition of a nondegenerate conic.)

(1.8) Homogeneous form in 2 variables. Let $F(U, V)$ be a nonzero homogeneous polynomial of degree d in U, V, with coefficients in a fixed field k; (I will follow tradition, and use the word *form* for 'homogeneous polynomial'):

$$F(U,V) = a_dU^d + a_{d-1}U^{d-1}V + .. a_iU^iV^{d-i} + .. a_0V^d.$$

F has an associated inhomogeneous polynomial in 1 variable,

$$f(u) = a_du^d + a_{d-1}u^{d-1} + .. a_iu^i + .. a_0.$$

Clearly for $\alpha \in k$,

$$f(\alpha) = 0 \iff (u - \alpha) \mid f(u) \iff (U - \alpha V) \mid F(U, V) \iff F(\alpha, 1) = 0;$$

so zeros of f correspond to zeros of F on $\mathbb{P}^1$ away from the point $(1, 0)$, the 'point $\alpha = \infty$'. What does it mean for F to have a zero at infinity?

$$F(1, 0) = 0 \iff a_d = 0 \iff \deg f < d.$$

Now define the *multiplicity* of a zero of F on $\mathbb{P}^1$ to be

(i) the multiplicity of f at the corresponding $\alpha \in k$;

or (ii) $d - \deg f$ if $(1, 0)$ is the zero.

So the multiplicity of zero of F at a point $(\alpha, 1)$ is the greatest power of $(U - \alpha V)$ dividing F, and at $(1, 0)$ it is the greatest power of V dividing F.

Proposition. Let $F(U, V)$ be a nonzero form of degree d in U, V. Then F has at most d zeros on $\mathbb{P}^1$; furthermore, if k is algebraically closed, then F has exactly d zeros on $\mathbb{P}^1$ provided these are counted with multiplicities as defined above.

Proof. Let m_∞ be the multiplicity of the zero of F at (1, 0); then by definition, $d - m_\infty$ is the degree of the inhomogeneous polynomial f, and the proposition reduces to the well-known fact that a polynomial in one variable has at most deg f roots. Q.E.D.

Note that over an algebraically closed field, F will factorise as a product $F = \Pi \lambda_i^{m_i}$ of linear forms $\lambda_i = (a_iU + b_iV)$, and treated in this way, the point (1, 0) corresponds to the form $\lambda_\infty = V$, and is on the same footing as all other points.

(1.9) Easy cases of Bézout's Theorem. Bézout's theorem says that if C and D are plane curves of degree deg C = m, deg D = n, then the number of points of intersection of C and D is mn, provided that (i) the field is algebraically closed; (ii) points of intersection are counted with the right multiplicities; (iii) we work in $\mathbb{P}^2$ to take right account of intersections 'at infinity'. See for example [Fulton, p. 112] for a self-contained proof. In this section I am going to treat the case when one of the curves is a line or conic.

Theorem. Let $L \subset \mathbb{P}^2_k$ be a line (respectively $C \subset \mathbb{P}^2_k$ a nondegenerate conic), and let $D \subset \mathbb{P}^2_k$ be a curve defined by $D : (G_d(X, Y, Z) = 0)$, where G is a form of degree d in X, Y, Z. Assume that $L \not\subset D$ (respectively, $C \not\subset D$); then

$$\#\{L \cap D\} \leq d \qquad \text{(respectively } \#\{C \cap D\} \leq 2d\text{)}.$$

In fact there is a natural definition of multiplicity of intersection such that the inequality still holds for 'number of points counted with multiplicities', and if k is algebraically closed then equality holds.

Proof. A line $L \subset \mathbb{P}^2_k$ is given by an equation $\lambda = 0$, with λ a linear form; for my purpose, it is convenient to give it parametrically as

$$X = a(U, V), \quad Y = b(U, V), \quad Z = c(U, V),$$

where a, b, c are linear forms in U, V. So for example, if $\lambda = \alpha X + \beta Y + \gamma Z$, and $\gamma \neq 0$, then L can be given as

$$X = U, \quad Y = V, \quad Z = -(\alpha/\gamma)U - (\beta/\gamma)V.$$

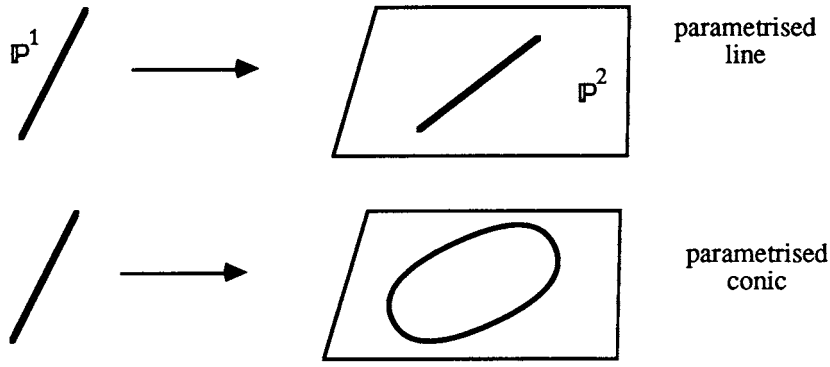
Similarly, as explained in (1.7), a nondegenerate conic can be given parametrically as

$$X = a(U, V), \quad Y = b(U, V), \quad Z = c(U, V),$$

where a, b, c are quadratic forms in U, V. This is because C is a projective transformation of $(XZ = Y^2)$, which is parametrically $(X, Y, Z) = (U^2, UV, V^2)$, so C is given by

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M \begin{bmatrix} U^2 \\ UV \\ V^2 \end{bmatrix}$$

where M is a nonsingular 3×3 matrix.



Then the intersection of L (respectively C) with D is given by finding the values of the ratios (U : V) such that

$$F(U, V) = G_d(a(U,V), b(U,V), c(U,V)) = 0.$$

But F is a form of degree d (respectively 2d) in U, V, so the result follows by (1.8).

(1.10) Corollary. If $P_1, .. P_5 \in \mathbb{P}^2_{\mathbb{R}}$ are distinct points such that no 4 are collinear, then there exists at most one conic through $P_1, .. P_5$.

Proof. Suppose by contradiction that C_1 and C_2 are conics with $C_1 \neq C_2$ such that

$$C_1 \cap C_2 \supset \{P_1, .. P_5\}.$$

C_1 is nonempty, so that if it's nondegenerate, then by (1.7), it's projectively equivalent to the parametrised curve

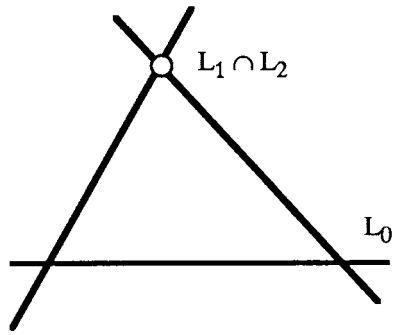
$$C_1 = \{(U^2, UV, V^2) \mid (U, V) \in \mathbb{P}^1\};$$

by (1.9), $C_1 \subset C_2$. Now if Q_2 is the equation of C_2, then $Q_2(U^2, UV, V^2) \equiv 0$ for all $(U, V) \in \mathbb{P}^1$, and an easy calculation (see Ex. 1.6) shows that Q_2 is a multiple of $(XZ - Y^2)$; this contradicts $C_1 \neq C_2$.

Now suppose C_1 is degenerate; by (1.6) again, it's either a line pair or a line, and one sees easily that

$$C_1 = L_0 \cup L_1, \quad C_2 = L_0 \cup L_2,$$

with L_1, L_2 distinct lines. Then $C_1 \cap C_2 = L_0 \cup (L_1 \cap L_2)$:



thus 4 points out of $P_1, .. P_5$ lie on L_0, a contradiction. Q.E.D.

(1.11) Space of all conics. Let

$$S_2 = \{\text{quadratic forms on } \mathbb{R}^3\} = \{3 \times 3 \text{ symmetric matrixes}\} \cong \mathbb{R}^6.$$

If $Q \in S_2$, write $Q = aX^2 + 2bXY + .. fZ^2$; then for $P_0 = (X_0, Y_0, Z_0) \in \mathbb{P}^2\mathbb{R}$, consider the relation $P_0 \in C : (Q = 0)$. This is of the form

$$Q(X_0, Y_0, Z_0) = aX_0^2 + 2bX_0Y_0 + .. fZ_0^2 = 0,$$

and for fixed P_0, this is a linear equation in (a, b,.. f). So

$$S_2(P_0) = \{Q \in S_2 \mid Q(P_0) = 0\} \cong \mathbb{R}^5 \subset S_2 = \mathbb{R}^6$$

is a 5-dimensional hyperplane. For $P_1, .. P_n \in \mathbb{P}^2\mathbb{R}$, define similarly

$$S_2(P_1, .. P_n) = \{Q \in S_2 \mid Q(P_i) = 0 \text{ for } i = 1, .. n\};$$

then there are n linear equations in the 6 coefficients (a, b,.. f) of Q. This gives the result:

Proposition. $\dim S_2(P_1, .. P_n) \geq 6 - n$.

We can also expect that 'equality holds if $P_1,.. P_n$ are general enough'. More precisely:

Corollary. If $n \leq 5$ and no 4 of $P_1,.. P_n$ are collinear, then

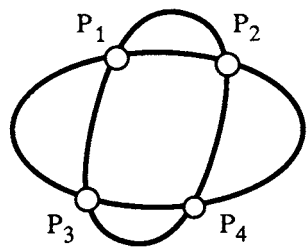
$$\dim S_2(P_1,.. P_n) = 6 - n.$$

Proof. Corollary 1.10 implies that if $n = 5$, $\dim S_2(P_1,.. P_5) \leq 1$, which gives the corollary in this case. If $n \leq 4$, then I can add in points $P_{n+1},.. P_5$ while preserving the condition that no 4 points are collinear, and since each point imposes at most one linear condition, this gives

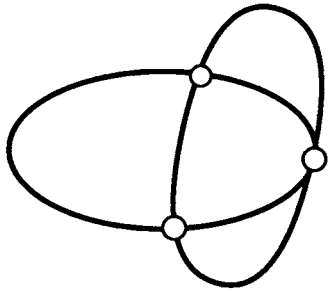
$$1 = \dim S_2(P_1,.. P_5) \geq \dim S_2(P_1,.. P_n) - (5 - n). \quad \text{Q.E.D.}$$

Note that if 6 points $P_1,.. P_6 \in \mathbb{P}^2_{\mathbb{R}}$ are given, they may or may not lie on a conic.

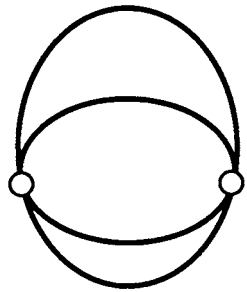
(1.12) Intersection of two conics. As we have seen above, it will often happen that two conics meet in 4 points:



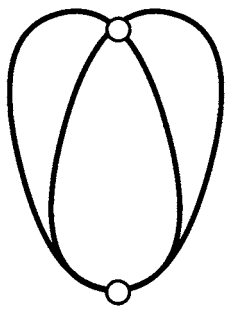
conversely according to Corollary 1.11, given 4 points $P_1,.. P_4 \in \mathbb{P}^2$, under suitable conditions $S_2(P_1,.. P_4)$ is a 2-dimensional vector space, so choosing a basis Q_1, Q_2 for $S_2(P_1,.. P_4)$ gives 2 conics C_1, C_2 such that $C_1 \cap C_2 = \{P_1,.. P_4\}$. There are lots of possibilities for multiple intersections of nonsingular conics:



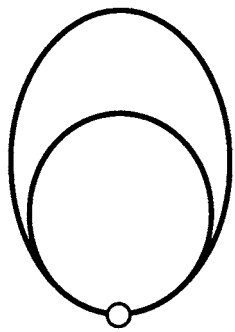
$2P_1 + P_2 + P_3$



$2P + 2Q$



$3P + Q$



$4P$

see Ex. 1.9 for suitable equations.

(1.13) Degenerate conics in a pencil.

Definition. A *pencil of conics* is a family of the form

$$C_{(\lambda,\mu)}: (\lambda Q_1 + \mu Q_2 = 0);$$

each element is a plane curve, depending in a linear way on the parameters (λ, μ); think of the ratio $(\lambda : \mu)$ as a point of $\mathbb{P}^1$.

Looking at the examples, one expects that for special values of $(\lambda : \mu)$ the conic $C_{(\lambda,\mu)}$ is degenerate. In fact, writing $\det(Q)$ for the determinant of the symmetric 3×3 matrix corresponding to the quadratic form Q, it is clear that

$$C_{(\lambda,\mu)} \text{ is degenerate} \iff \det(\lambda Q_1 + \mu Q_2) = 0.$$

Writing out Q_1 and Q_2 as symmetric matrixes expresses this condition as

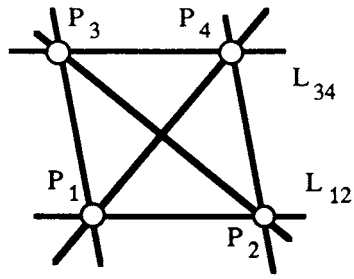
$$F(\lambda, \mu) = \det \left| \lambda \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix} + \mu \begin{bmatrix} a' & b' & d' \\ b' & c' & e' \\ d' & e' & f' \end{bmatrix} \right| = 0.$$

Now notice that $F(\lambda, \mu)$ is a homogeneous cubic form in λ, μ. In turn I can apply (1.8) to F to deduce:

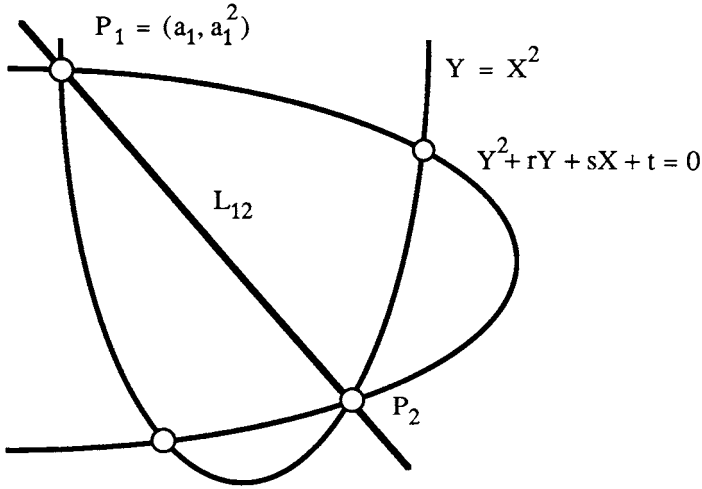
Proposition. Suppose $C_{(\lambda,\mu)}$ is a pencil of conics of $\mathbb{P}^2_k$, with at least one non-degenerate conic (so that $F(\lambda, \mu)$ is not identically zero). Then the pencil has at most 3 degenerate conics. If $k = \mathbb{R}$ then the pencil has at least one degenerate conic.

Proof. A cubic form has ≤ 3 zeros. Also over $\mathbb{R}$, it must have at least one zero.

(1.14) Worked example. Let $P_1, .. P_4$ be 4 points of $\mathbb{P}^2_{\mathbb{R}}$ such that no 3 are collinear; then the pencil of conics $C_{(\lambda,\mu)}$ through $P_1, .. P_4$ has 3 degenerate elements, namely the line pairs $L_{12} + L_{34}$, $L_{13} + L_{24}$, $L_{14} + L_{23}$, where L_{ij} is the line through P_i, P_j:



Next, suppose that I start from the pencil of conics generated by $Q_1 = Y^2 + rY + sX + t$ and $Q_2 = Y - X^2$, and try to find the points $P_1, .. P_4$ of intersection.



This can be done as follows: (1) find the 3 ratios $(\lambda : \mu)$ for which $C_{(\lambda,\mu)}$ are degenerate conics. Using what has been said above, this just means that I have to find the 3 roots of the cubic

$$F(\lambda, \mu) = \det \left| \lambda \begin{bmatrix} 0 & 0 & s/2 \\ 0 & 1 & r/2 \\ s/2 & r/2 & t \end{bmatrix} + \mu \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1/2 \\ 0 & 1/2 & 0 \end{bmatrix} \right|$$

$$= -\tfrac{1}{4}(s^2\lambda^3 + (4t - r^2)\lambda^2\mu - 2r\lambda\mu^2 - \mu^3).$$

(2) Separate out 2 of the degenerate conics into pairs of lines (this involves solving 2 quadratic equations). (3) The 4 points P_i are the points of intersection of the lines.

This procedure gives a geometric interpretation of the reduction of the general quartic in Galois theory (see for example [van der Waerden, Algebra, Ch. 8, §64]): let k be a field, and $f(X) = X^4 + rX^2 + sX + t \in k[X]$ a quartic polynomial. Then the two parabolas C_1 and C_2 meet in the 4 points $P_i = (a_i, a_i^2)$ for $i = 1,.. 4$, where the a_i are the 4 roots of f.

Then the line $L_{ij} = P_iP_j$ is given by

$$L_{ij}: (Y = (a_i + a_j)X - a_ia_j),$$

and the reducible conic $L_{12} + L_{34}$ is given by

$$Y^2 + (a_1a_2+a_3a_4)Y + (a_1+a_2)(a_3+a_4)X^2 + sX + t = 0,$$

that is, by $Q_1 - (a_1+a_2)(a_3+a_4)Q_2 = 0$. Hence the 3 values of μ/λ for which the conic $\lambda Q_1 + \mu Q_2$ breaks up as a line pair are

$$-(a_1+a_2)(a_3+a_4), \quad -(a_1+a_3)(a_2+a_4), \quad -(a_1+a_4)(a_2+a_3).$$

The cubic equation whose roots are these 3 quantities is called the *auxilliary cubic* associated with the quartic; it can be calculated using the theory of elementary symmetric functions; this is a fairly laborious procedure. On the other hand, the geometric method sketched above gives an elegant derivation of the auxilliary cubic which only involves evaluating a 3×3 determinant.

The above treatment is taken from [M.Berger, 16.4.10 and 16.4.11.1].

Exercises to §1.

1.1. Parametrise the conic C: $(x^2 + y^2 = 5)$ by considering a variable line through (2, 1) and hence find all rational solutions of $x^2 + y^2 = 5$.

1.2. Let p be a prime; by experimenting with various p, guess a necessary and sufficient condition for $x^2 + y^2 = p$ to have rational solutions; prove your guess (a hint is given after Ex. 1.9 below - bet you can't do it for yourself!).

1.3. Prove the statement in (1.3), that an affine transformation can be used to put any conic of $\mathbb{R}^2$ into one of the standard forms (a-l). (Hint: use a linear transformation $x \mapsto Ax$ to take the leading term $ax^2 + bxy + cy^2$ into one of $\pm x^2 \pm y^2$ or $\pm x^2$ or 0; then complete the square in x and y to get rid of as much of the linear part as possible.)

1.4. Make a detailed comparison of the affine conics in (1.3) with the projective conics in (1.6).

1.5. Let k be any field of characteristic $\neq 2$, and V a 3-dimensional k-vector space; let $Q: V \to k$ be a nondegenerate quadratic form on V. Show that if $0 \neq e_1 \in V$ satisfies $Q(e_1) = 0$ then V has a basis e_1, e_2, e_3 such that $Q(x_1e_1 + x_2e_2 + x_3e_3) = x_1x_3 + ax_2^2$. (Hint: work with the symmetric bilinear form φ associated to Q; since φ is nondegenerate, there is a vector e_3 such that $\varphi(e_1, e_3) = 1$. Now find a suitable e_2.)

Deduce that a nonempty, nondegenerate conic $C \subset \mathbb{P}^2_k$ is projectively equivalent to $(XZ = Y^2)$.

1.6. Let k be a field with at least 4 elements, and C: $(XZ = Y^2) \subset \mathbb{P}^2_k$; prove that if Q(X, Y, Z) is a quadratic form which vanishes on C then $Q = \lambda(XZ - Y^2)$. (Hint: if you really can't do this for yourself, compare with the argument in the proof of Lemma

2.5.)

1.7. In $\mathbb{R}^3$, consider the two planes A: (Z = 1) and B: (X = 1); a line through 0 meeting A in (x, y, 1) meets B in (1, y/x, 1/x). Consider the map φ: A $\dashrightarrow$ B defined by $(x, y) \mapsto (y' = y/x, z' = 1/x)$; what is the image under φ of

(i) the line ax = y + b; the pencil of parallel lines ax = y + b (fixed a and variable b);

(ii) circles $(x-1)^2 + y^2 = c$ for variable c (distinguish the 3 cases $c > 1$, $c = 1$ and $c < 1$).

Try to imagine the above as a perspective drawing by an artist sitting at $0 \in \mathbb{R}^3$, on a plane (X = 1), of figures from the plane (Z = 1). Explain what happens to the points of the two planes where φ and φ^{-1} are undefined.

1.8. Let $P_1, .. P_4$ be distinct points of $\mathbb{P}^2$ with no 3 collinear. Prove that there is a unique coordinate system in which the 4 points are (1, 0, 0), (0, 1, 0), (0, 0, 1) and (1, 1, 1). Find all conics passing through $P_1, .. P_5$, where P_5 = (a, b, c) is some other point, and use this to give another proof of Corollary 1.10 and Proposition 1.11.

1.9. In (1.12) there is a list of possible ways in which two conics can intersect. Write down equations showing that each possibility really occurs. Find all the singular conics in the corresponding pencils. (Hint: you will save yourself a lot of trouble by using symmetry and a well-chosen coordinate system.)

Hint for 1.2: it is known from elementary number theory that −1 is a quadratic residue modulo p if and only if p = 2 or $p \equiv 1 \bmod 4$.

1.10. (Sylvester's determinant). Let k be an algebraically closed field, and suppose given a quadratic and cubic form in U, V as in (1.8):

$$q(U, V) = a_0U^2 + a_1UV + a_2V^2,$$
$$c(U, V) = b_0U^3 + b_1U^2V + b_2UV^2 + b_3V^3.$$

Prove that q and c have a common zero $(\eta : \tau) \in \mathbb{P}^1$ if and only if

$$\det \begin{vmatrix} a_0 & a_1 & a_2 & & \\ & a_0 & a_1 & a_2 & \\ & & a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 & b_3 & \\ & b_0 & b_1 & b_2 & b_3 \end{vmatrix} = 0$$

(Hint: Show that if q and c have a common root then the 5 elements

$$U^2q,\ UVq,\ V^2q,\ Uc \text{ and } Vc$$

do not span the 5-dimensional vector space of forms of degree 4, and are therefore linearly dependent. Conversely, use unique factorisation in the polynomial ring $k[U, V]$ to say something about relations of the form $Aq = Bc$ with A and B forms in U, V, $\deg A = 2$, $\deg B = 1$.)

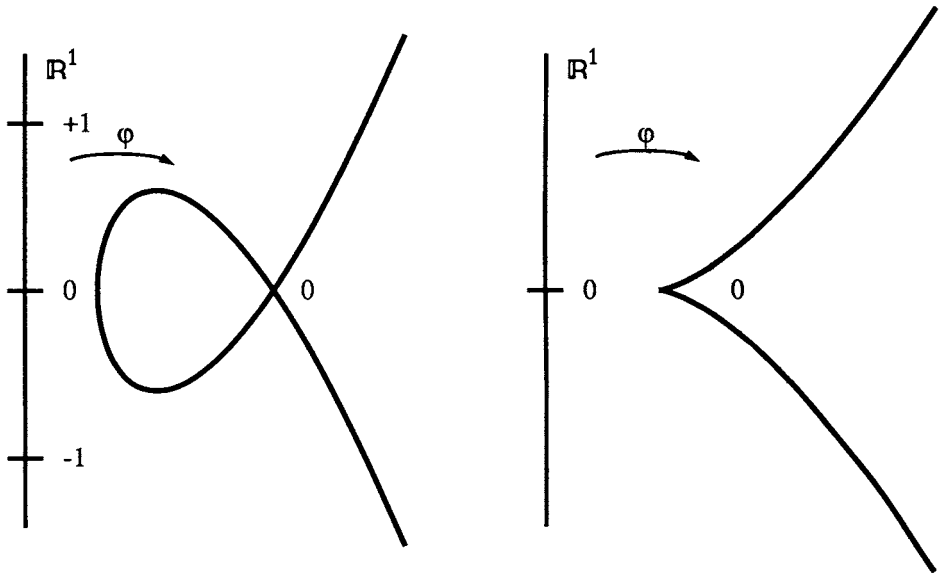
1.11. Generalise the result of Ex. 1.10 to two forms in U, V of any degrees n and m.

§2. Cubics and the group law

(2.1) Examples of parametrised cubics. Some plane cubic curves can be parametrised, just as the conics:

Nodal cubic. $C : (y^2 = x^3 + x^2) \subset \mathbb{R}^2$ is the image of the map $\varphi: \mathbb{R}^1 \to \mathbb{R}^2$ given by $t \mapsto (t^2 - 1, t^3 - t)$ (check it and see);

Cuspidal cubic. $C : (y^2 = x^3) \subset \mathbb{R}^2$ is the image of $\varphi: \mathbb{R}^1 \to \mathbb{R}^2$ given by $t \mapsto (t^2, t^3)$:



nodal cubic

$y^2 = x^3 + x^2$

cuspidal cubic

$y^2 = x^3$

Parametrised cubic curves

Think about the singularities of the image curve, and of the map φ. These examples will occur throughout the course, so spend some time playing with the equations; see Ex. 2.1–2.

(2.2) The curve $(y^2 = x(x-1)(x-\lambda))$ has no rational parametrisation.

Parametrised curves are nice; for example, if you're interested in Diophantine problems, you could hope for a rule giving all $\mathbb{Q}$-valued points, as in (1.1). The parametrisation of (1.1) was of the form $x = f(t)$, $y = g(t)$, where f and g were *rational functions*, that is, quotients of two polynomials.

Theorem. Let k be a field of characteristic $\neq 2$, and let $\lambda \in k$ with $\lambda \neq 0, 1$; let $f, g \in k(t)$ be rational functions such that

$$f^2 = g(g-1)(g-\lambda). \qquad (*)$$

Then $f, g \in k$.

This is equivalent to saying that there does not exist any nonconstant map $\mathbb{P}^1 \dashrightarrow C : (y^2 = x(x-1)(x-\lambda))$ given by rational functions. This reflects a very strong 'rigidity' property of varieties.

The proof of the theorem is arithmetic in the field $k(t)$ using the fact that $k(t)$ is the field of fractions of the UFD $k[t]$. It's quite a long proof, so either be prepared to study it in detail, or skip it for now (GOTO 2.4). In Ex. 2.12, there is a very similar example of a nonexistence proof by arithmetic in $\mathbb{Q}$.

Proof. Using the fact that $k[t]$ is a UFD, I write

$$f = r/s \text{ with } r, s \in k[t] \text{ and coprime,}$$

$$g = p/q \text{ with } p, q \in k[t] \text{ and coprime.}$$

Clearing denominators, $(*)$ becomes

$$r^2q^3 = s^2p(p-q)(p-\lambda q).$$

Then since r and s are coprime, the factor s^2 on the right-hand side must divide q^3, and in the same way, since p and q are coprime, the left-hand factor q^3 must divide s^2. Therefore,

$$s^2 \mid q^3 \text{ and } q^3 \mid s^2, \quad \text{so that} \quad s^2 = aq^3 \text{ with } a \in k$$

(a is a unit of $k[t]$, therefore in k).

Then

$$aq = (s/q)^2 \text{ is a square in } k[t].$$

Also,

$$r^2 = ap(p - q)(p - \lambda q),$$

so that by considering factorisation into primes, there exist nonzero constants $b, c, d \in k$ such that

$$bp,\ c(p - q),\ d(p - \lambda q)$$

are all squares in $k[t]$. If I can prove that p, q are constants, then it follows from what's already been said that r, s are also, proving the theorem. To prove that p, q are constants, set K for the algebraic closure of k; then $p, q \in K[t]$ satisfy the conditions of the next lemma.

(2.3) Lemma. Let K be an algebraically closed field, $p, q \in K[t]$ coprime elements, and assume that 4 distinct linear combinations (that is, $\lambda p + \mu q$ for 4 distinct ratios $(\lambda : \mu) \in \mathbb{P}^1_K$) are squares in $K[t]$; then $p, q \in K$.

Proof (Fermat's method of 'infinite descent'). Both the hypotheses and conclusion of the lemma are not affected by replacing p, q by

$$p' = ap + bq,\ q' = cp + dq,$$

with $a, b, c, d \in K$ and $ad - bc \neq 0$. Hence I can assume that the 4 given squares are

$$p,\ p - q,\ p - \lambda q,\ q.$$

Then $p = u^2$, $q = v^2$, and $u, v \in K[t]$ are coprime, with

$$\max \{\deg u, \deg v\} < \max \{\deg p, \deg q\}.$$

Now by contradiction, suppose that $\max \{\deg p, \deg q\} > 0$ and is minimal among all p, q satisfying the condition of the lemma. Then both of

$$p - q = u^2 - v^2 = (u - v)(u + v)$$

and

$$p - \lambda q = u^2 - \lambda v^2 = (u - \mu v)(u + \mu v)$$

(where $\mu = \sqrt{\lambda}$) are squares in $K[t]$, so that by coprimeness of u,v, I conclude that each of $u - v$, $u + v$, $u - \mu v$, $u + \mu v$ are squares. This contradicts the minimality of $\max \{\deg p, \deg q\}$. Q.E.D.

(2.4) Linear systems. Write $S_d = \{\text{forms of degree } d \text{ in } (X, Y, Z)\}$; (recall that a form is just a homogeneous polynomial). Any element $F \in S_d$ can be written in a unique way as

$$F = \sum a_{ijk} X^i Y^j Z^k$$

with $a_{ijk} \in k$, and the sum taken over all $i, j, k \geq 0$ with $i + j + k = d$; this means of course that S_d is a k-vector space with basis

$$\begin{array}{ccccc} & & Z^d & & \\ & Z^{d-1}X & & Z^{d-1}Y & \\ .. & & & & .. \\ X^{d-1}Z & & X^{d-2}YZ & .. & Y^{d-1}Z \\ X^d & X^{d-1}Y & & X^{d-2}Y^2 & .. \quad Y^d \end{array}$$

and in particular, $\dim S_d = \binom{d+2}{2}$. For $P_1, .. P_n \in \mathbb{P}^2$, let

$$S_d(P_1, .. P_n) = \{ F \in S_d \mid F(P_i) = 0 \text{ for } i = 1, .. n \} \subset S_d.$$

Each of the conditions $F(P_i) = 0$, (or more precisely, $F(X_i, Y_i, Z_i) = 0$, where $P_i = (X_i : Y_i : Z_i)$) is one linear condition on F, so that $S_d(P_1, .. P_n)$ is a vector space of dimension $\geq \binom{d+2}{2} - n$.

(2.5) Lemma. Suppose that k is an infinite field, and let $F \in S_d$.

(i) Let $L \subset \mathbb{P}^2_k$ be a line; if $F \equiv 0$ on L, then F is divisible in $k[X, Y, Z]$ by the equation of L. That is, $F = H \cdot F'$ where H is the equation of L and $F' \in S_{d-1}$.

(ii) Let $C \subset \mathbb{P}^2_k$ be a nonempty nondegenerate conic; if $F \equiv 0$ on C, then F is divisible in $k[X, Y, Z]$ by the equation of C. That is, $F = Q \cdot F'$ where Q is the equation of C and $F' \in S_{d-2}$.

If you think this statement is obvious, congratulations on your intuition: you have just guessed a particular case of the Nullstellensatz. Now find your own proof (GOTO 2.6).

Proof. (i) By a change of coordinates, I can assume $H = X$. Then for any $F \in S_d$, there exists a unique expression $F = X \cdot F'_{d-1} + G(Y, Z)$: just gather together all the monomials involving X into the first summand, and what's left must be a polynomial in Y, Z only. Now $F \equiv 0$ on $L \iff G \equiv 0$ on $L \iff G(Y, Z) = 0$. The last step holds because of (1.8): if $G(Y, Z) \neq 0$ then it has at most d zeros on $\mathbb{P}^1_k$, whereas if k is infinite, then so is $\mathbb{P}^1_k$.

(ii) By a change of coordinates, $Q = XZ - Y^2$. Now let me prove that for any $F \in S_d$, there exists a unique expression $F = Q \cdot F'_{d-2} + A(X, Z) + YB(X, Z)$: if I just substitute $(XZ - Q)$ for Y^2 wherever it occurs in F, what's left has degree ≤ 1 in Y, and is therefore of the form $A(X, Z) + YB(X, Z)$. Now as in (1.7), C is the parametrised conic given by $X = U^2$, $Y = UV$, $Z = V^2$, so that

$$F \equiv 0 \text{ on } C \iff A(U^2, V^2) + UVB(U^2, V^2) \equiv 0 \text{ on } C$$

$$\iff A(U^2, V^2) + UVB(U^2, V^2) = 0 \in k[U, V] \iff A(X, Z) = B(X, Z) = 0.$$

Here the last equality comes by considering separately the terms of even and odd degrees in the form $A(U^2, V^2) + UVB(U^2, V^2)$. Q.E.D.

Ex. 2.2 gives similar cases of 'explicit' Nullstellensatz.

Corollary. Let $L: (H = 0) \subset \mathbb{P}^2_k$ be a line (respectively $C: (Q = 0) \subset \mathbb{P}^2_k$ a nondegenerate conic); suppose that points $P_1, \ldots P_n \in \mathbb{P}^2_k$ are given, and consider $S_d(P_1, \ldots P_n)$ for some fixed d. Then

(i) If $P_1, \ldots P_a \in L$, $P_{a+1}, \ldots P_n \notin L$ and $a > d$, then

$$S_d(P_1, \ldots P_n) = H \cdot S_{d-1}(P_{a+1}, \ldots P_n).$$

(ii) If $P_1, \ldots P_a \in C$, $P_{a+1}, \ldots P_n \notin C$ and $a > 2d$, then

$$S_d(P_1, \ldots P_n) = Q \cdot S_{d-2}(P_{a+1}, \ldots P_n).$$

Proof. (i) If F is homogeneous of degree d, and the curve $D: (F = 0)$ meets L in points $P_1, \ldots P_a$ with $a > d$, then by (1.9), I must have $L \subset D$, so that by the lemma, $F = H \cdot F'$; now since $P_{a+1}, \ldots P_n \notin L$, obviously $F' \in S_{d-1}(P_{a+1}, \ldots P_n)$. (ii) is exactly the same. Q.E.D.

(2.6) Proposition. Let k be an infinite field, and $P_1,.. P_8 \in \mathbb{P}^2_k$ distinct points; suppose that no 4 of $P_1,.. P_8$ are collinear, and no 7 of them lie on a non-degenerate conic; then

$$\dim S_3(P_1,.. P_8) = 2.$$

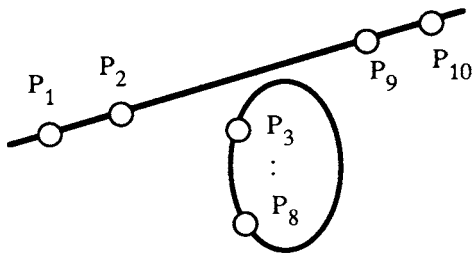
Proof. For brevity, let me say that a set of points are *conconic* if they all lie on a nondegenerate conic. The proof of (2.6) breaks up into several cases.

Main case. No 3 points are collinear, no 6 conconic. This is the 'general position' case.

Suppose for a contradiction that $\dim S_3(P_1,.. P_8) \geq 3$, and let P_9, P_{10} be distinct points on the line $L = P_1P_2$. Then

$$\dim S_3(P_1,.. P_{10}) \geq \dim S_3(P_1,.. P_8) - 2 \geq 1,$$

so that there exists $0 \neq F \in S_3(P_1,.. P_{10})$. By Corollary 2.5, $F = H \cdot Q$, with $Q \in S_2(P_3,.. P_8)$. Now I have a contradiction to the case assumption: if Q is non-degenerate then the 6 points $P_3,.. P_8$ are conconic, whereas if Q is a line pair or a double line, then at least 3 of them are collinear.



First degenerate case. Suppose $P_1, P_2, P_3 \in L$ are collinear, and let $L: (H = 0)$. Let P_9 be a 4th point on the line L. Then by Corollary 2.5,

$$S_3(P_1,.. P_9) = H \cdot S_2(P_4,.. P_8).$$

Also, since no 4 of $P_4,.. P_8$ are collinear, by Corollary 1.11, $\dim S_2(P_4,.. P_8) = 1$, and then $\dim S_3(P_1,.. P_9) = 1$, which implies $\dim S_3(P_1,.. P_8) \leq 2$.

Second degenerate case. Suppose $P_1,.. P_6 \in C$ are conconic, with $C: (Q = 0)$ a nondegenerate conic. Then choose $P_9 \in C$ distinct from $P_1,.. P_6$. By Corollary 2.5 again,

$$S_3(P_1,.. P_9) = Q \cdot S_1(P_7, P_8);$$