Direct Solution of Modulus Constraints

F. Schaffalitzky
fsm@robots.ox.ac.uk
Robotics Research Group
University of Oxford
Department of Engineering Science
Parks Road, Oxford, UK

Abstract

The modulus constraint is a constraint on the position of the plane at infinity (π_∞) which applies to the problem of self-calibration in the case of constant internals. For any pair of cameras which are known to have the same internal parameters, the classical modulus constraint is the vanishing of a certain quartic polynomial whose coefficients are determined from the cameras. Given a projective three-view reconstruction, it is of practical interest to recover the plane at infinity by solving for the three parameters of π_∞. Geometrically this is the problem of intersecting three quartic surfaces in projective space, so one should expect to get 64 solutions. It is not clear how to carry out the process in practice because continuation methods are slow and non-linear optimization may produce a local minimum. This paper presents a new derivation of the classical constraints, and additionally shows how to derive novel cubic constraints which exist for any triple of views. For three views, it is shown how to use the new constraint to classify the $64 = 4 \times 4 \times 4$ classical solutions into one spurious (namely the trifocal plane), 21 feasible and 2×21 which must be rejected on physical grounds. The ambiguity is thus reduced from 64 to 21. A numerical algorithm is given to compute all 21 feasible solutions.

1 Introduction

Effective methods exist for computing projective structure of scenes from images of them [3, 6, 7]. Here, "projective" means that the coordinates of the scene tokens and the camera parameters are recovered up to a projective transformation which is *a priori* unknown. For some purposes, this ambiguity of gauge is not a great drawback (*eg* for point transfer or recognition) and in any case it can be argued that there is no natural coordinate system with which the scene should be described. For example, in a natural (*ie* not man-made) scene the choice of origin is entirely arbitrary. However, where metric properties such as angles and lengths are concerned, a reconstruction which is in the "wrong" projective frame is not convenient. For this, what is needed is to "upgrade" to a coordinate frame which is related to the scene coordinate frame by a Euclidean (similarity) transformation, and this process is called *auto-calibration* or *self-calibration* [11, 4] when it is accomplished without using any physical measurements of the camera internals or scene structure.

In this paper we consider the classical auto-calibration problem from a three-view projective reconstruction, with the assumption that the camera internals are the same for each view (the *internals* of a camera are those characteristics that do not depend on camera pose or position; *eg* focal length, the distance between the lens and the sensor array).

The contribution is a new derivation of the original quartic *modulus constraints*, a new cubic modulus constraint and a numerically feasible method of computing all the solutions. It is also shown that while the original quartic constraints admit 64 solutions, one of these is in fact the trifocal plane and that the remaining 63 solutions split up into three sets of 21 of which only one set is of practical interest.

Throughout it will be assumed that the given cameras are in sufficiently general position. There do exist critical camera motions for which auto-calibration is not possible without making further assumptions about the internal parameters (or the motion) [19, 22], but in this paper the generic case is considered.

Roadmap The paper is laid out as follows: Section 2 describes in overview how one is led to consider *horopters curves*[11] in the context of self-calibration. Section 3 is a technical digression on an algebraic formula for computing the focal point of a 3×4 camera matrix, which will be used in our analysis of horopter curves. Section 4 makes a connection between horopter curves and the modulus constraint. In section 5 we give an analysis of the properties of the classical modulus constraints derived in [14, 13] and show that it leads naturally to a classification of the classical solutions and the discovery of a new constraint, of degree three. Section 6 gives a resultant-based method for solution of these combined quartic and cubic constraints and section 7 presents some basic experimental results.

Notation The scale of quantities (vectors, matrices etc) is often significant and consequently, $=$ is used to denote equality and $\sim$ to denote equality up to scale.

Points in projective spaces $\mathbb{P}^n$ will be represented with column vectors of length $n+1$ and points in the dual space $\mathbb{P}^{n*}$ (*eg* lines in $\mathbb{P}^2$ and planes in $\mathbb{P}^3$) with row vectors of length $n+1$. Thus, a point $\mathbf{X} \in \mathbb{P}^n$ lies on a hyperplane $\boldsymbol{\alpha} \in \mathbb{P}^{n*}$ if and only if (the matrix equation) $\boldsymbol{\alpha}\mathbf{X} = 0$ holds. This convention avoids the need for transposition and helps make the meaning of symbols clear from their context.

In this convention, a plane $\boldsymbol{\alpha}$ in 3D space is the locus of points $\mathbf{X}$ which satisfy a linear equation $\alpha\mathbf{X} = 0$. Thus, if $\mathtt{G}$ is a homography which transforms $\mathbf{X}$ to $\mathtt{G}\mathbf{X}$, it is apparent that it must transform α into $\alpha\mathtt{G}^{-1}$ because the relationship of incidence $0 = \alpha\mathbf{X} = (\alpha\mathtt{G}^{-1})(\mathtt{G}\mathbf{X})$ has to be invariant under collineations.

2 Horopters and π_∞

Assume given a projective reconstruction of the scene with cameras which have the same internal parameters. It will be useful to be quite explicit: each camera, $\mathtt{P}_i$, has a focal point, given by 3 coordinates which we stack into a 3-vector $\mathbf{v}_i$, one vector for each view index $i = 1, 2, 3$. The *pose* of the ith camera is given by a 3×3 rotation matrix $\mathtt{Q}_i$ (with determinant 1). If the internal parameters are given by the 3×3 matrix $\mathtt{K}$, the ith given camera is then

$$\mathtt{P}_i = \mathtt{K}\mathtt{Q}_i \left[\begin{array}{c|c} \mathtt{I}_{3\times 3} & -\mathbf{v}_i \end{array} \right] \mathtt{G}^{-1} \tag{1}$$

where $\mathtt{G}$ is a 4×4 matrix representing a homography of $\mathbb{P}^3$, that is, it represents the coordinate transform between the actual Euclidean coordinates in the world and the projective coordinate frame of the reconstruction. Given a projective reconstruction, we know only the $\mathtt{P}_i$, not $\mathtt{K}, \mathtt{Q}_i, \mathbf{v}_i$ or $\mathtt{G}$, the task being to compute them. Once $\mathtt{G}$ is known, the reconstruction can be put into a frame where angles and ratios of (but not absolute) lengths can be computed – this is known as *metric structure*.

Computing metric structure is very much simplified if the *affine structure* is already known. The meaning of affine structure is that the plane at infinity (in the scene) coincides with the plane at infinity in the reconstruction; *ie* the reconstruction and the scene coordinates are related by an (unknown) affine transformation. Thus, if we can determine where the plane at infinity should be, then we can determine the affine, and hence metric, structure.

From the above model it is actually very easy to determine the location of the plane at infinity. It suffices to find three points on it; for example its intersections with the three axes of rotation of the camera motions will do. The screw axis of the between-view rotation for cameras i, j intersects the plane at infinity in the point

$$\mathtt{G}\begin{pmatrix} \mathbf{n}_{ij} \\ \mathbf{0} \end{pmatrix}$$

where the 3-vector n_{ij} is the direction of the axis of rotation of the 3×3 matrix $\mathtt{Q}_{ij} = \mathtt{Q}_i^{-1}\mathtt{Q}_j$. But that axis is just the eigenvector with eigenvalue 1. From this eigenvector property it follows that point of intersection is the null-vector of

$$\begin{aligned} \mathtt{P}_i - \mathtt{P}_j &= \mathtt{K}\left[\begin{array}{c|c} \mathtt{Q}_i - \mathtt{Q}_j & -\mathtt{Q}_i\mathbf{v}_i + \mathtt{Q}_j\mathbf{v}_j \end{array} \right]\mathtt{G}^{-1} \\ &= \mathtt{K}\mathtt{Q}_i\left[\begin{array}{c|c} \mathtt{I}_{3\times 3} - \mathtt{Q}_i^{-1}\mathtt{Q}_j & -\mathbf{v}_i + \mathtt{Q}_i^{-1}\mathtt{Q}_j\mathbf{v}_j \end{array} \right]\mathtt{G}^{-1} \\ &= \mathtt{K}\mathtt{Q}_i\left[\begin{array}{c|c} \mathtt{I}_{3\times 3} - \mathtt{Q}_{ij} & -\mathbf{v}_i + \mathtt{Q}_{ij}\mathbf{v}_j \end{array} \right]\mathtt{G}^{-1} \end{aligned}$$

Since three views provide three such intersections, the plane at infinity may be determined from three views (Ignoring the case of critical motions where these three intersections are collinear or even coincident, such as described in [1] which exploits the special structure).

Unfortunately, there is a flaw in this solution. To be given a projective reconstruction as three cameras $\mathtt{P}_1, \mathtt{P}_2, \mathtt{P}_3$ means we are given three 3×4 arrays of real numbers, whose geometric meaning is unchanged by scalar factors. But the meaning of the algebraic difference $\mathtt{P}_i - \mathtt{P}_j$ *is* changed if each camera is rescaled separately. Thus, the given arrays are not really of the form in equation (1) but have an extra (unknown) scale factor, λ_i :

$$\mathtt{P}_i = \lambda_i\mathtt{K}\mathtt{Q}_i \left[\begin{array}{c|c} \mathtt{I}_{3\times 3} & -\mathbf{v}_i \end{array} \right] \mathtt{G}^{-1} \tag{2}$$

The required point of intersection is therefore the null-vector of

$$\lambda_i^{-1}\mathtt{P}_i - \lambda_j^{-1}\mathtt{P}_j$$

but since we don't know the three scale factors $\lambda_1, \lambda_2, \lambda_3$ the approach appears to be a dead end.

We have gained something, though. Namely, to find the plane at infinity it suffices to find the ratio $\lambda_1 : \lambda_2 : \lambda_3$, which is a 2D problem. As the ratios vary, three curves are traced out by the null-vectors, one for each pair of cameras. The curve for cameras i, j is given by

$$\mathbf{X} = f(\lambda_i^{-1}\mathtt{P}_i - \lambda_j^{-1}\mathtt{P}_j)$$

where $f(\mathtt{P})$ denotes the kernel of the 3×4 matrix $\mathtt{P}$. A point $\mathbf{X}$ on the curve satisfies (by definition of the kernel)

$$\mathtt{P}_i\mathbf{X} \sim \mathtt{P}_j\mathbf{X}$$

Thus, the curve traced out is the locus of $\mathbf{X} \in \mathbb{P}^3$ which project to the same point in images i and j. This locus is known as a *horopter curve* and has been used before in the context of calibration and ambiguity of reconstruction [1, 12]. It is a twisted cubic [16] which passes through both camera centres and which is asymptotic to the axis of rotation of the between-view camera motion [1].

As the ratio $\lambda_1 : \lambda_2 : \lambda_3$ varies, the three corresponding points on the three horopters vary, spanning a varying plane, and consequently the ratio $\lambda_1 : \lambda_2 : \lambda_3$ parameterizes a 2-degrees-of-freedom family of planes which includes the plane at infinity.

In order to make effective use of this, it will first be necessary to understand better how the null-space $f(\mathtt{P})$ depends on the camera matrix $\mathtt{P}$.

3 Algebraic Nullspaces

A camera matrix is a 3×4 matrix $\mathtt{P}$ of rank 3. It has a 1D null-space which represents the centre of projection of the camera. Numerically, the null-space can be computed very satisfactorily via a singular value decomposition [5] of $\mathtt{P}$ but for theoretical algebraic purposes it is not very useful. Instead, an algebraic expression can be constructed by forming the (column) 4-vector $f(\mathtt{P})$ of 3×3 minors of the camera matrix $\mathtt{P}$. This is a well-known trick described in standard texts such as [7].

To make the choice of sign clear, and for definiteness, we note that the algebraic nullspace satifies the equation

$$\det\begin{pmatrix}\mathtt{P}\\ \pi\end{pmatrix} = \pi f(\mathtt{P}) \tag{3}$$

for all (row) vectors π, representing planes in the world. In particular, a plane π passes through the focal point if and only if the 4×4 matrix on the left is singular, *ie* iff π is a linear combination of the rows of $\mathtt{P}$. The advantage of using this algebraic form is that the scale of the null-vector is defined canonically by the scale of the matrix. We have, for example, the following properties which will be used later and which can be deduced directly from the definition or equation (3) :

Properties of f (Coordinate change). Let $\mathtt{P}$ be a 3×4 matrix. Then

- If λ is a scalar then $f(\lambda\mathtt{P}) = \lambda^3 f(\mathtt{P})$
- If $\mathtt{A}$ is a 3×3 matrix then $f(\mathtt{AP}) = \det(\mathtt{A})f(\mathtt{P})$
- If $\mathtt{G}$ is a 4×4 matrix then $f(\mathtt{PG}^{-1}) = \mathtt{G}f(\mathtt{P})/\det(\mathtt{G})$

Dual construction. By construction, $f(\mathtt{P})$ is a cubic expression for a 4-vector which is annihilated (*ie* orthogonal to) all three rows of $\mathtt{P}$. Similarly, the same method can be used to construct an explicit expression $\mathbf{span}(\mathbf{p},\mathbf{q},\mathbf{r})$ for the plane spanned by (*ie* orthogonal to) three (column) 4-vectors $\mathbf{p},\mathbf{q},\mathbf{r}$. One obvious approach is to collect p, q, r into a 4×3 matrix, transpose it, apply f and then transpose again to get a (row) 4-vector.

4 Horopters and the Quartic Constraints

We are now ready to describe the horopter curves parametrically. Recall that the horopter for cameras i, j is described by $\mathbf{X} = f(\lambda_i^{-1}\mathtt{P}_i - \lambda_j^{-1}\mathtt{P}_j)$. Now, f is a cubic function of its argument, so this formula will expand as a cubic function of λ_i^{-1} and λ_j^{-1} :

$$\mathbf{X} = \lambda_i^{-3}f(\mathtt{P}_i) - \lambda_i^{-2}\lambda_j^{-1}\sigma(\mathtt{P}_i,\mathtt{P}_j) + \lambda_i^{-1}\lambda_j^{-2}\sigma(\mathtt{P}_j,\mathtt{P}_i) - \lambda_j^{-3}f(\mathtt{P}_j)$$

for some (column) vectors $\sigma(\mathtt{P}_i,\mathtt{P}_j), \sigma(\mathtt{P}_j,\mathtt{P}_i)$ whose form could be determined explicitly. Note that this is the *definition* of $\sigma(\cdot,\cdot)$; it is defined by the above expansion and thus derives its properties from those of f. In detail, σ is defined by the requirement that

$$f(s\mathtt{A} + t\mathtt{B}) = s^3 f(\mathtt{A}) + s^2 t\sigma(\mathtt{A},\mathtt{B}) + st^2\sigma(\mathtt{B},\mathtt{A}) + t^3 f(\mathtt{B})$$

holds as an identity between $s, t, \mathtt{A}, \mathtt{B}$. It is not hard to see that σ is quadratic in its first argument and linear in its second argument.

For brevity, define $f_i = f(\mathtt{P}_i), \sigma_{ij} = \sigma(\mathtt{P}_i,\mathtt{P}_j)$ so that the parameterization of the horopter for cameras i, j is

$$\begin{aligned}(\lambda_i : \lambda_j) &\mapsto \mathbf{X}\\ &= f(\lambda_i^{-1}\mathtt{P}_i - \lambda_j^{-1}\mathtt{P}_j)\\ &= \lambda_i^{-3}f_i - \lambda_i^{-2}\lambda_j^{-1}\sigma_{ij} + \lambda_i^{-1}\lambda_j^{-2}\sigma_{ji} - \lambda_j^{-3}f_j\end{aligned} \tag{4}$$

Next, a connection between horopter curves and modulus constraints is established. Recall that any world plane induces a homography between two images of a scene, by back-projection from one image plane onto the world plane followed by re-projection into the other image plane. The homography induced from view 2 to view 1 by a plane π and two cameras $\mathtt{P}_1, \mathtt{P}_2$ is denoted $\mathtt{H}(\mathtt{P}_1,\mathtt{P}_2;\pi)$. In the case where the plane is the plane at infinity and the cameras are as in (1), the induced homography from view i to view j is $\mathtt{K}\mathtt{Q}_j\mathtt{Q}_i^{-1}\mathtt{K}^{-1}$. Since this is conjugate to a rotation, the eigenvalues must be $1, e^{-i\theta}, e^{+i\theta}$ where θ is the angle of rotation of $\mathtt{Q}_j\mathtt{Q}_i^{-1}$.

(Strong) Modulus Constraint. If cameras $\mathtt{P}_i$, $\mathtt{P}_j$ have the same internal parameters, then a necessary condition for a plane π to be the plane at infinity is that the eigenvalues of $\mathtt{H}(\mathtt{P}_j, \mathtt{P}_i; \pi)$ all have the same modulus (absolute value).

This constraint, first given in [15], yields an easily evaluated criterion for testing whether or not a given plane might be the plane at infinity. To avoid the algebraically cumbersome absolute values, the modulus constraint is expressed in terms of the (coefficients of the) characteristic equation of the induced homography. The characteristic equation of a 3×3 matrix $\mathtt{A}$ is the polynomial $\det(T\mathtt{I}_{3\times 3} - \mathtt{A}) = aT^3 + bT^2 + cT + d$ where $a = \det(\mathtt{I}_{3\times 3}) = 1$, $-b$ is the trace of $\mathtt{A}$, c the trace of $\mathtt{A}$'s cofactor matrix and $-d = \det(\mathtt{A})$. In the case of homographies one identifies matrices $\mathtt{A} \sim \mu\mathtt{A}$ which differ by scalar multiples, so we also have to identify the characteristic equations $aT^3 + bT^2 + cT + d \sim aT^3 + b\mu T^2 + c\mu^3 T + d\mu^3$. Generalizing slightly, we identify[1] $aT^3 + bT^2 + cT + d$ with $a\lambda^3 T^3 + b\lambda^2\mu T^2 + c\lambda\mu^2 T + d\mu^3$ for all non-zero scalars λ, μ. (The second polynomial arises from multiplying the characteristic equation of $(\mu/\lambda)\mathtt{A}$ by λ^3).

(Weak) Modulus Constraint. A necessary condition for a plane π to be the plane at infinity is that the characteristic equation of $\mathtt{H}(\mathtt{P}_j, \mathtt{P}_i; \pi)$ have the form

$$T^3 - (2c+1)T^2 + (2c+1)T - 1 \qquad (5)$$

(or one equivalent to it) for some scalar c.

To see this, take the eigenvalues of the homography to be $1, e^{-i\theta}, e^{+i\theta}$ and set $c = \cos\theta$. It is a slightly weaker condition because it does not stipulate the nature of the scalar c (*eg* that it is real and bounded between -1 and $+1$).

In order to apply this constraint it is necessary to relate the coefficients of the characteristic equation to the coordinates of the plane π. The following result will allow us to express the weak modulus constraint in algebraic terms:

[1] The result of performing these identifications is a "space" which is analogous to the usual projective space, where vectors are identified if they differ by scalar factors. Getting used to this new space is not much harder than getting used to projective space.

Lemma. The induced homography $\mathtt{H}(\mathtt{P}_j, \mathtt{P}_i; \pi)$ has characteristic equation (equivalent to) :

$$\begin{pmatrix} \pi f_i & -\pi\sigma_{ij} & \pi\sigma_{ji} & -\pi f_j \end{pmatrix}$$

where the notation lists the coefficients of $T^3, T^2, T, 1$.
Proof: The general case follows from the special case where $\pi = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}$ by a change of coordinates, so it suffices to do this case. Writing

$$\mathtt{P}_i = [\mathtt{A}|\mathbf{a}], \mathtt{P}_j = [\mathtt{B}|\mathbf{b}]$$

the induced homography is $\mathtt{B}\mathtt{A}^{-1}$ so that the characteristic equation is

$$\begin{aligned} \det(T\mathtt{I}_{3\times 3} - \mathtt{B}\mathtt{A}^{-1}) &\sim \det(T\mathtt{A} - \mathtt{B}) \\ &= \det \begin{pmatrix} T\mathtt{A} - \mathtt{B} & T\mathbf{a} - \mathbf{b} \\ 0\ \ 0\ \ 0 & 1 \end{pmatrix} \\ &= \det \begin{pmatrix} T\mathtt{P}_i - \mathtt{P}_j \\ \pi \end{pmatrix} \\ &= \pi f(T\mathtt{P}_i - \mathtt{P}_j) \end{aligned}$$

as required. **QED**

Corollary - (Quartic) Modulus Constraint. A necessary condition for a plane π to be the plane at infinity is

$$\det \begin{vmatrix} \pi f_i & (\pi\sigma_{ij})^3 \\ \pi f_j & (\pi\sigma_{ji})^3 \end{vmatrix} = 0 \qquad (6)$$

This is equivalent to the quartic constraint constructed by Pollefeys and Van Gool and used in [13].
Proof: Note that for the characteristic equation $(\pi f_i)T^3 - (\pi\sigma_{ij})T^2 + (\pi\sigma_{ji})T - (\pi f_j)$ to be equivalent to $T^3 - (2c+1)T^3 + (2c+1)T - 1$ there must be scalars λ, μ such that $\pi f_i = \lambda^3$, $\pi\sigma_{ij} = \lambda^2\mu(2c+1)$, $\pi\sigma_{ji} = \lambda\mu^2(2c+1)$ and. $\pi f_j = \mu^3$. Eliminating the scalars leads at once to the given condition. **QED**

Given three views, there will be three such constraints and by Bézout's theorem [18] they will have a total of $4 \times 4 \times 4 = 64$ solutions counting both real and complex, with appropriate multiplicities

To spell out the form of these constraints on the three parameters of π_∞, one can arrange the vectors which enter them in a triangular diagram (7) in which each side of the triangle gives a constraint. The constraint for one side of the triangle is that the ratio arising from the last and second elements (the fs) is the cube of the ratio arising from the two middle elements (the σs).

$$\begin{array}{ccccccc} & & & f_1 & & & \\ & & \sigma_{12} & & \sigma_{13} & & \\ & \sigma_{21} & & ? & & \sigma_{31} & \\ f_2 & & \sigma_{23} & & \sigma_{32} & & f_3 \end{array} \qquad (7)$$

5 Cubic Constraints

Since only one of the 64 solutions to the quartic constraints is right, it is of practical interest to see where the spurious ones come from. One spurious solution is easily exhibited – it is the trifocal plane. The trifocal plane $\pi_{\mathbf{TF}}$ is the plane spanned by the three camera centres f_1, f_2, f_3 and satisfies $\pi_{\mathbf{TF}} f_1 = \pi_{\mathbf{TF}} f_2 = \pi_{\mathbf{TF}} f_3 = 0$, so that equation (6) makes it transparent that $\pi_{\mathbf{TF}}$ satisfies the quartic modulus constaints.

This leaves $64 - 1 = 63$ solutions to account for and it will now be shown that these can be divided naturally into three categories. The reader may want to refer to diagram (7) again in what follows.

Firstly, suppose that π is a solution passing through both f_1 and f_2 but not f_3. Then, using the two quartic constraints for view 3, it follows (refer to equation (6)) that π must also pass through σ_{23} and σ_{13}. However, these four points are not in general coplanar, which rules out the possibility of a solution through f_1, f_2 but not f_3.

Secondly, suppose that π is a solution passing through f_1 but not f_2 or f_3. Then, using the two quartic constraints for view 1, it follows (refer to equation (6) again) that π must also pass through σ_{12} and σ_{13}. Again, it can be verified that the plane spanned by these three points does not in general satisfy the quartic modulus constraint for views $2, 3$.

Thirdly, suppose that π is a solution which does not pass through any of the three focal points. Being a solution means that the ratio $(\pi f_i)/(\pi f_j)$ is the cube of the ratio $(\pi\sigma_{ij})/(\pi\sigma_{ji})$, for each pair of views i, j. Thus,

$$\left(\frac{(\pi\sigma_{12})}{(\pi\sigma_{21})}\frac{(\pi\sigma_{23})}{(\pi\sigma_{32})}\frac{(\pi\sigma_{31})}{(\pi\sigma_{13})}\right)^3 = \frac{(\pi f_1)}{(\pi f_2)}\frac{(\pi f_2)}{(\pi f_3)}\frac{(\pi f_3)}{(\pi f_1)} = 1$$

which shows that

$$\frac{(\pi\sigma_{12})}{(\pi\sigma_{21})}\frac{(\pi\sigma_{23})}{(\pi\sigma_{32})}\frac{(\pi\sigma_{31})}{(\pi\sigma_{13})} = 1, \omega \text{ or } \bar{\omega} \tag{8}$$

where $\omega, \bar{\omega}$ are the complex cube roots of 1. The 63 solutions can thus be divided into three categories according to the value of (8). While it is not *a priori* clear that these three categories are non-empty, it is certainly true that any physically plausible solution must give a value that is real, *ie* the value of (8) must necessarily equal 1. This proves

Theorem - (Cubic) Modulus Constraint. A necessary condition for a plane π to be the plane at infinity is

$$(\pi\sigma_{12})(\pi\sigma_{23})(\pi\sigma_{31}) = (\pi\sigma_{21})(\pi\sigma_{32})(\pi\sigma_{13}) \tag{9}$$

This is a new constraint which has not been exploited before. It has the advantage of being of degree three, less than four, and of incorporating information from all three views into a single constraint (refer to diagram (7) again).

By Bézout's theorem, taking the cubic constraint and the two quartic constraints for view 1 (say) gives exactly $3 \times 4 \times 4 = 48$ solutions, counting as before both real and complex, with suitable multiplicities. In particular, there must be some solutions to the three quartic constraints which do not satisfy the cubic constraint and this shows that all three categories (8) are non-empty.

In addition to the quartic and cubic constraints, there are higher-order constraints (of degree six), related to the central element of diagram (7), and these are described in a longer technical report which is available on request from the author.

6 Direct Solution Using Resultants.

This section describes a direct method of solution of the quartic and cubic constraints on π_∞. As a by-product it will be demonstrated analytically that there are 21 solutions. Geometrically, the idea is to parameterize the variety defined by the cubic constraint, substitute the parameterization into the quartic constraints and then simplify and solve the resulting system. Algebraically, there is a more direct route which will be given here.

6.1 Parameterization

Section 2 described how the use of parameterized horopters leads to a 2-degrees-of-freedom parameterization of the position of the plane at infinity. The idea is that if the ratio $\lambda_1 : \lambda_2 : \lambda_3$ is known, then so are the points

$$\mathbf{N}_{ij} = f(\lambda_i^{-1}\mathtt{P}_i - \lambda_j^{-1}\mathtt{P}_j) \sim f(\lambda_j \mathtt{P}_i - \lambda_i \mathtt{P}_j)$$

and these lie on the plane at infinity (they are the vanishing points of the axes of rotation of the camera motions). Thus, taking all three pairs of cameras, we obtain $\mathbf{N}_{12}, \mathbf{N}_{23}, \mathbf{N}_{31}$ as explicit homogenenous cubic expressions in $\lambda_1, \lambda_2, \lambda_3$. Taking the plane spanned the $\mathbf{N}_{ij}$ gives a parameterization of π_∞, of total degree $3 + 3 + 3 = 9$ (because each point is of degree 3 – the reader may want to refer back to the construction of "**span**" dual to "f" in section 3).

This is rather high, but we can do much better, by parameterizing instead the locus defined by the cubic modulus constraint (9). By inspecting equations (8) and (9) it can be seen that any solution to the cubic constraint is characterised by the three ratios that make up the product on the left in (8), in the sense that if these ratios are known the plane π can be recovered by solving three linear equations.

To explicitly construct the parameterization, note that if the three ratios in equation (8) have product equal to 1, there must exist non-zero scalars λ_i such that the ratios, in order, are $\lambda_1/\lambda_2, \lambda_2/\lambda_3, \lambda_3/\lambda_1$ (it turns out these are the same scalars λ_i mentioned above but this fact is not needed for what follows). Put differently, the points $\lambda_j \sigma_{ij} - \lambda_i \sigma_{ji}$ must lie on π, that is, they span π. For a given $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ we thus define $\phi(\lambda)$ to be the plane spanned by these three points.

Since each of these is linear in λ, we get a parameterization of degree $1+1+1=3$:

$$\phi(\lambda) = \mathbf{span}(\sigma_{12}\lambda_2 - \sigma_{21}\lambda_1, \sigma_{23}\lambda_3 - \sigma_{32}\lambda_2, \sigma_{31}\lambda_1 - \sigma_{13}\lambda_3) \tag{10}$$

Substitution of this parameterization into the three quartic modulus constraints $(\pi_\infty f_i : \pi_\infty f_j) = (\pi_\infty \sigma_{ij} : \pi_\infty \sigma_{ji})^3$ shows that these can now be written jointly as

$$\mathbf{rank} \begin{pmatrix} \lambda_1^3 & \phi(\lambda) f_1 \\ \lambda_2^3 & \phi(\lambda) f_2 \\ \lambda_3^3 & \phi(\lambda) f_3 \end{pmatrix} \leq 1$$

because by construction of ϕ, the ratio $\phi(\lambda)\sigma_{ij}$: $\phi(\lambda)\sigma_{ji}$ equals the ratio $\lambda_i : \lambda_j$.

The rank condition means that there is a scalar κ_0 such that $\phi(\lambda) f_i = \kappa_0 \lambda_i^3$ for all $i = 1, 2, 3$. We can consider the equations

$$\phi(\lambda) f_i = \kappa \lambda_i^3$$

not as a system of three plane cubics with an unknown parameter κ, but as a system of three equations on $\mathbb{P}^1 \times \mathbb{P}^2$, with κ providing the coordinate on $\mathbb{P}^1$ and λ the coordinates on $\mathbb{P}^2$. In this form, iterative methods such as Newton's method can be used to "polish" an initial estimate to give accuracy down to machine precision.

For a general choice of κ, the system will have no solution, but for $\kappa = 0$ there are three obvious solutions, namely

$$\lambda = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{and} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

These are solutions because, in each case, one of the factors in the triple vector product (10) making up $\phi(\lambda)$ is zero. If each λ_i is non-zero, the only way the triple vector product can vanish is for the three factors to be collinear points, *ie* linearly dependent vectors. Writing out this condition shows that there are in general three more values of λ for which $\phi(\lambda)$ is zero. In any case, a classical result [17, 8] states that a rational parameterization of any (smooth) cubic surface has six "basepoints" which is what the solutions to $\phi(\lambda) = 0$ are.

The method of solution now focuses on finding the possible values of κ_0, using the Macaulay multi-polynomial resultant [2, 9, 10] (see also [20] for an introduction in the context of camera calibration). A plane cubic is described by a homogeneous cubic polynomial in three variables. Suppose C_1, C_2, C_3 are three plane cubics (such polynomials have 10 coefficients). In general, any two of them will meet in a finite number of points (namely $3 \times 3 = 9$ by Bézout's theorem), so that the three plane cubics have no common points at all, in general. The condition for the three cubics to have a common point is given by the *resultant*, which is a polynomial function $\mathbf{Res}_{3,3,3}(C_1, C_2, C_3)$ of the coefficients of the cubics. This function vanishes if and only if the three cubics have a common point. The resultant $\mathbf{Res}_{3,3,3}$ has degree 9 in each C_i and a total degree of 27.

The point of all this is that, treating $\lambda_1, \lambda_1, \lambda_3$ as the variables of plane cubics, the expression

$$\begin{aligned} p(\kappa) = \mathbf{Res}_{3,3,3}(\quad & \kappa\lambda_1^3 - \phi(\lambda) f_1, \\ & \kappa\lambda_2^3 - \phi(\lambda) f_2, \\ & \kappa\lambda_3^3 - \phi(\lambda) f_3 \quad) \\ = \kappa^{27} + {} & p_{25}\kappa^{25} + \ldots + p_6 \kappa^6 \end{aligned} \tag{11}$$

is a univariate polynomial in κ of degree 27 which has a root at $\kappa = \kappa_0$. It is a property of the resultant that the leading term of $p(\kappa)$ has coefficient 1. The fact that the term of degree 26 is zero is a property of the particular cubics used. The discussion above showed that p has a zero of order 6 at $\kappa = 0$.

Having solved for the $21 = 27 - 6$ non-zero values of κ, the λ_i can be recovered linearly. One way to recover them is to use the description in [2] of the resultant as the determinant of a 15×15 matrix whose null-vector(s) is the vector of quartic monomials in λ. In practice, finding the roots of a univariate polynomial of degree 21 is a notoriously ill-conditioned problem [2, 21] and some care has to be taken. The approach adopted for the work described here was to avoid explicit polynomial arithmetic by representing $p(\kappa)$ functionally by the formula (11). This approach allows incremental reduction of the order of p by "peeling off" known solutions as they become available from the Newton-Raphson polishing step.

6.2 Counting: $4 \times 4 \times 4 = 1 + 3 \times 21$

It has been shown that the quartic and cubic modulus constraints together admit exactly 21 solutions. Using the analogue of Bézout's theorem for intersections in $\mathbb{P}^1 \times \mathbb{P}^2$ gives the same result [18]. This completes the classification of the 64 solutions to the three quartic modulus constraints. They are:

- The trifocal plane, $\pi_{\mathbf{TF}}$, spanned by the camera centres.
- The 21 solutions which additionally satisfy the cubic modulus constraint (9).
- The two complex conjugate sets of 21 solutions which give values $\omega, \bar{\omega}$ in (8).

Of these, only the second set of 21 solutions is of practical interest.

7 Experimental Results.

The algorithm described can be run on well-conditioned real data (but the same conditioning must be used in each image, or the assumption of constant internals will be violated).

Figures 1 and 2 show three views from a sequence of 9 images of a turntable. A projective reconstruction was obtained by manually selecting feature points and running a standard projective reconstrunction algorithm, culminating in full projective bundle adjustment.

The zeroth, third and seventh cameras were then used as input to the modulus constraint solver. Of the 21 solutions found, 13 were complex and so discarded. Of the remaining 8, only two gave values between -1 and 3 for the middle two coefficients in (5) and so the other 6 were discarded on the grounds that this is the possible range of values taken by $1 + 2\cos\theta$ for all angles θ.

For the remaining two solutions (see figure 2), a metric upgrade was attempted by the standard method [7] of first upgrading to affine and then estimating the image of the absolute conic using the eigenvectors of the infinite homographies. One was obviously wrong (by visual inspection) and the other looks good. As a test of correctness, the angles at the corners of the base of the object are 89.760, 89.808, 89.152 and 89.200 degress (the base of the real object is planar, but the base of the reconstructed object is not planar, so these angles do not add up to 360 degrees).

Unfortunately, the modulus constraint is only a partial constraint (it is necessary but not sufficient) and having a solution does not gaurantee that the recovered between-view camera motions are true Euclidean motions. To be specific, the modulus constraint imposes three conditions on the plane at infinity and thus determines finitely many solutions for it. However, the “infinite homographies” recovered from such solutions may not fix a common conic.

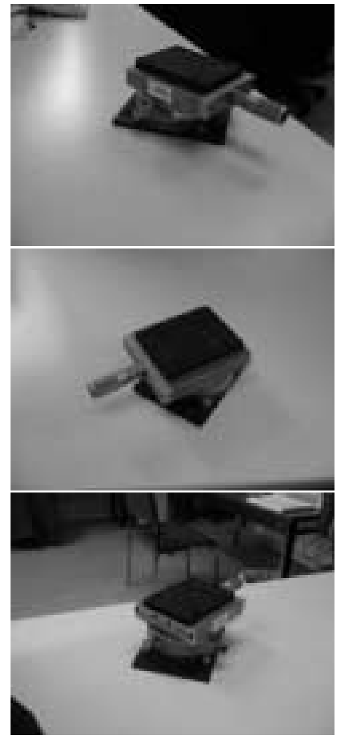

Figure 1: Three images from a sequence of 9 used in the experiments.

For ideal, synthetic data it *is* the case that there is a common fixed conic, but in the presence of measurement noise the methods described in [7] can fail due to ill-conditioning in the linear least squares problem used. We stress that this is a weakness of the modulus constraint and not of the numerical algorithm we have presented. For example, the example in figure 2 computed using views zero, three and seven works fine, but using views zero, three and six from the same sequence does not produce a good result because this triplet is close to being a single-axis motion. In practice this shows up as a near-ambiguity in the solution of the least squares problem for the absolute conic. It is thus easy to detect by inspecting the two smallest singular values of the linear system being solved.

8 Conclusion and further work

Traditionally the modulus constraint has been solved using an initial guess followed by non-linear minimization. Our method is comparable in speed and has the added advantage of providing all 21 solutions, which avoids the risk of computing a local minimum. The extra complexity is a disadvantage but the exposition itself has thrown new light on the problem.

The experiments show that the results obtained from the algorithm can be used as the starting point

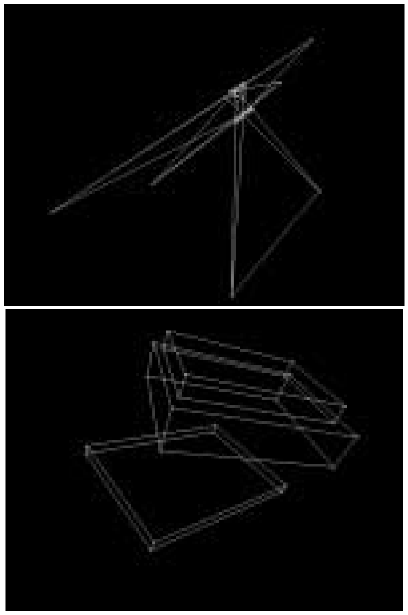

Figure 2: Result of metric rectification from the images shown. Of the 21 solutions, only two were admissible (for reasons described in the main text). The first solution gives the implausible upper-most reconstruction. The second solution gives the bottom-most reconstruction.

for non-linear optimization if some care is taken to detect degenerate configurations.

The problem of deciding which of the 21 solutions is correct remains. Apart from rejecting complex solutions and solutions for which the constant c from equation (5) does not lie between -1 and $+1$, the obvious approach is to try to perform a metric rectification and test the fit of the computed model to the original image data. An interesting possibility is that the configuration of 21 planes has a special structure which might single out certain planes. For example, a simple counting argument (a homogenous cubic in three variables has 20 coefficients) shows that given 19 planes in general position, there exists a unique cubic which they all satisfy and so, since the 21 solutions to the modulus constraint satisfy a cubic constraint, they cannot be in general position (To appreciate the significance of this, consider that conics have 6 coefficients and so 6 points which lie on a conic cannot be in general position because any 5 of them determine the conic).

Acknowledgements

This work was supported by an EPSRC studentship and EC Esprit Project Improofs. I would like to thank Andrew Zisserman for originally suggesting the problem to me, for encouraging me to publish it, for proofreading the later versions and for helping me make them readable. Also thanks to David Liebowitz who volunteered to suffer the first version.

References

[1] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. In *Proc. ECCV*, LNCS 1064/5, pages 3–16. Springer-Verlag, 1996.

[2] D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry*. Springer, 1998.

[3] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proc. ECCV*, LNCS 588, pages 563–578. Springer-Verlag, 1992.

[4] O. D. Faugeras, Q.-T. Luong, and S. Maybank. Camera self-calibration: Theory and experiments. In *Proc. ECCV*, LNCS 588, pages 321–334. Springer-Verlag, 1992.

[5] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore, MD, second edition, 1989.

[6] R. I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. CVPR*, 1992.

[7] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[8] R. Hartshorne. *Algebraic Geometry*. Springer, 1997.

[9] F. MacAulay. On some formulae in elimination. In *Proceedings of the London Mathematical Society (May 1902, 3-27)*, May 1902.

[10] F. MacAulay. Note on the resultant of a number of polynomials of the same degree (june 1921, 14-21). In *Proceedings of the London Mathematical Society*, 1921.

[11] S. Maybank and O. D. Faugeras. A theory of self-calibration of a moving camera. *IJCV*, 8(2):123–151, 1992.

[12] S. J. Maybank. *Theory of reconstruction from image motion*. Springer-Verlag, Berlin, 1993.

[13] M. Pollefeys. *Self-calibration and metric 3D reconstruction from uncalibrated image sequences*. PhD thesis, ESAT-PSI, K.U.Leuven, 1999.

[14] M. Pollefeys and L. Van Gool. A stratified approach to metric self calibration. In *Proc. CVPR*, 1997.

[15] M. Pollefeys, L. Van Gool, and A. Oosterlinck. The modulus constraint: a new constraint for self-calibration. In *Proc. ICPR*, pages 31–42, 1996.

[16] J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1979.

[17] J. G. Semple and L. Roth. *Introduction to Algebraic Geometry*. Oxford University Press, 1949.

[18] I. Shafarevich. *Basic Algebraic Geometry*. Springer, 1994.

[19] P. Sturm. *Vision 3D non calibrée: Contributions à la reconstruction projective et étude des mouvements critiques pour l'auto calibrage*. PhD thesis, INRIA Rhône-Alpes, 1997.

[20] W. Triggs. Camera pose and calibration from 4 or 5 known 3D points. In *Proc. ICCV*, pages 278–284, 1999.

[21] J.. Wilkinson. The evaluations of the zeroes of ill-conditioned polynomials, part 1. *Numerische Matematik*, (1):150–166, 1959.

[22] A. Zisserman, D. Liebowitz, and M. Armstrong. Resolving ambiguities in auto-calibration. *Phil. Trans. R. Soc. Lond. A*, 356(1740):1193–1211, 1998.