VISIT TO PAN-STARRS

Honolulu, 25 June - 30 July 2006

THREE SHORT LECTURES

ON IDENTIFICATIONS

AND ORBIT DETERMINATION

Andrea Milani

Department of Mathematics, University of Pisa

PLAN

1. **ATTRIBUTION, TOO SHORT ARCS**

2. **VIRTUAL ASTEROIDS, MULTIPLE SOLUTIONS**

3. **IDENTIFICATION MANAGEMENT, FINAL OUTPUT**

1.1 IDENTIFICATION

The **identification** problem deals with separate sets of observations, which might, and might not, belong to the same object. The identification is confirmed if all the observations can be fitted to a single least squares orbit with acceptable residuals.

The problem can be classified as **orbit identification** when the observations of both arcs are enough to solve for two least squares orbits: the input data are two sets of orbital elements, with covariance matrices. A metric in the 6-dimensional space of elements (propagated to the same epoch) has to be used to assess the proposed identifications, before the computationally intensive differential corrections.

The most difficult identification problem is **linkage**, when two arcs of observations both too short to perform orbit determination are to be joined into an arc good enough to compute an orbit. There is no way to directly compare quantities of the same nature, e.g., observations with observations: they are at different times, some interpolation function has to be used (either polynomials in time or Virtual Asteroids).

Tracklet composition is another form of identification.

1.2 ATTRIBUTION AND ATTRIBUTABLE

The identification problem can be classified as **attribution** when data insufficient to compute a usable orbit for one arc (e.g., two 2-dimensional observations) is compared to the known orbit of the other arc. Not enough information is available in the orbits space and predictions from the orbit need to be compared with the observations from the other arc. Thus it is useful to synthesize the information of the second arc into a vector observation at a single time.

A celestial body is at the heliocentric position $\mathbf{r}$ and is observed from the heliocentric position $\mathbf{q}$ on the Earth. Let (ρ, α, δ) be spherical coordinates for the topocentric position $\mathbf{r} - \mathbf{q}$. An **attributable** is a vector $A = (\alpha, \delta, \dot{\alpha}, \dot{\delta})$, representing the topocentric angular position and velocity of the body at a time $\bar{t}$.

The distance in the 4-dimensional space between the prediction A_1 issued from the orbit and the attributable A_2 computed from the observations needs to take into account the uncertainities. Then it is used as a filter to restrict the number of possible attributions between N orbits and M attributables to much less than $M \times N$ possibilities.

1.3 ATTRIBUTION METRIC

Let C_1, C_2 be the normal matrices:

$$C_1 = \Gamma_1^{-1} \ ; \ \Gamma_1 = DF \, \Gamma_{\mathbf{x}} \, DF^T$$

is obtained from the propagation of the covariance matrix $\Gamma_{\mathbf{x}}$ of the orbital elements $\mathbf{x}$ by means of the differential of the prediction $F(\mathbf{x})$. C_2 is the normal matrix of the least square fit used to obtain A_2 from the observations. The two least squares fits (for the orbit, for the attributable) have **target functions** Q_1, Q_2 (weighted sums of squares divided by number of residuals m_1, m_2)

$$Q_i(A) = Q_i(A_i) + \frac{(A - A_i)^T \, C_i \, (A - A_i)}{m_i} = Q_i(A_i) + \Delta Q_i(A) \ ,$$

neglecting higher orders. By fitting all the data together the target function becomes

$$Q(A) = Q^* + \Delta Q(A) = \frac{m_1 Q_1(A) + m_2 Q_2(A)}{m} =$$

$$= \frac{m_1}{m} \, Q_1(A_1) + \frac{m_2}{m} \, Q_2(A_2) + \frac{m_1}{m} \, \Delta Q_1(A) + \frac{m_2}{m} \, \Delta Q_2(A)$$

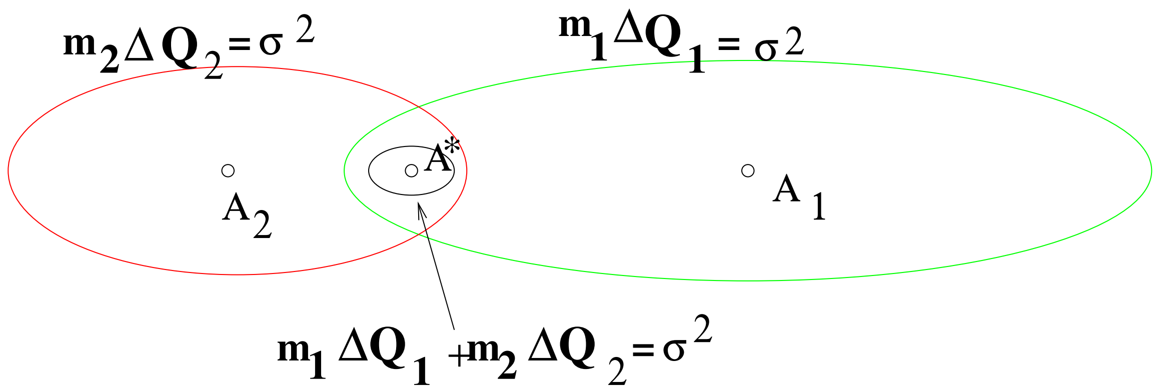
where $m = m_1 + m_2$ and the **identification penalty** $\Delta Q(A)$ is a convex combination of the two penalties, it has to be minimized, but is in > 0 unless $A_1 = A_2$.

1.4 COMPROMISE SOLUTION

The identification penalty is a function of the vector differences $(A-A_i)$ and of the normal matrices C_i

$$
\begin{aligned}
m\,\Delta Q(A) &= (A-A_1)^T\,C_1\,(A-A_1)+\\
&\quad +(A-A_2)^T\,C_2\,(A-A_2)=\\
&= (A-A_0)^T\,C_0\,(A-A_0)+K\\
C_0 &= C_1+C_2 \quad \text{assumed invertible}\\
A_0 &= C_0^{-1}\,(C_1A_1+C_2A_2)\\
K &= A_1^T\,C_1A_1+A_2^T\,C_2A_2-A_0^T\,C_0A_0=\\
&= (A_2-A_1)^T\,(C_2-C_2C_0^{-1}\,C_2)\,(A_2-A_1)
\end{aligned}
$$

thus the penalty K/m is a quadratic form in the difference A_2-A_1. The geometric interpretion is shown in the Figure (for a 2-dim case, that is with ellipses).



1.5 EFFICIENT ALGORITHM FOR ATTRIBUTION

The challenge of a high performance (=completeness + reliability + computational efficiency) algorithm can be met only by a **multi stage procedure**. Each successive step filters a smaller fraction of the $N \times M$ couples, with a larger computational load.

Filter 1 operates in the 2-dim space of (α, δ). In a simpler version it just selects the attributables within a disk (fixed radius) centered at the nominal prediction from the orbit. This can be easily implemented with $O(M\ \log M) + O(N\ \log M)$ comparisons using heap sorting and binary search in 1 variable, say α. For longer prediction time spans, a Filter 1.5 uses the 2-dim identification penalty.

Filter 2 operates in the 4-dim space of the attributables, with as control a function of the penalty, such as $\sqrt{K}$.

Filter 3 operates in the $2m$-dimensional space of residuals, looking for a nominal solution by differential corrections. As first guess, the known orbit can be used (not when it was purely hypothetic: then a first guess is computed from the compromise solution A_0 of Filter 2).

1.6 WHEN ATTRIBUTION IS USED

Known objects with *good* orbits produce tracklets which have to be attributed. This could be done *before* the identification procedure.

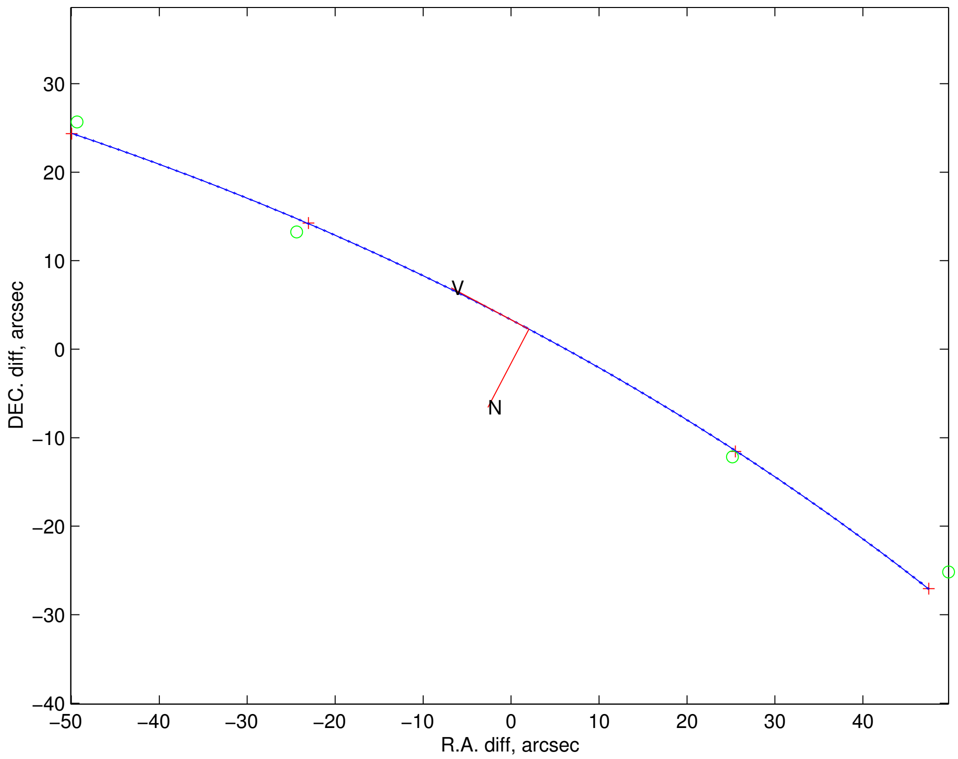
Identifications, even if just found and with few tracklets, can be augmented by adding tracklets in another night or month. (For another apparition, this is far from obvious, and Orbit Identification is more promising).

A swarm of Virtual Asteroids compatible with a tracklet can be propagated to another night and be used to seek attributions. This is a way to start the identification procedure, as a first step of recursive attribution.

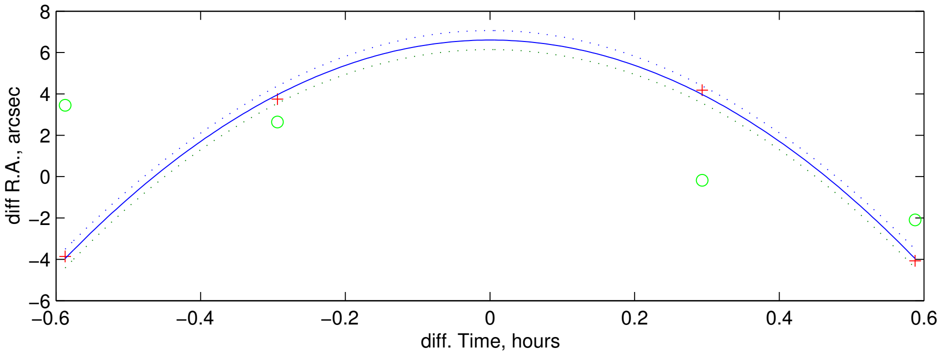
A couple of orphan detections can be attributed to a known orbit; however, if the orphans are low S/N, the orphans can only be *adopted* in pairs, because in the 2-dim space (α, δ) their number density is too large. The algorithm is somewhat different because a quadratic loop has to be avoided both in Filter 1 and in Filter 2.

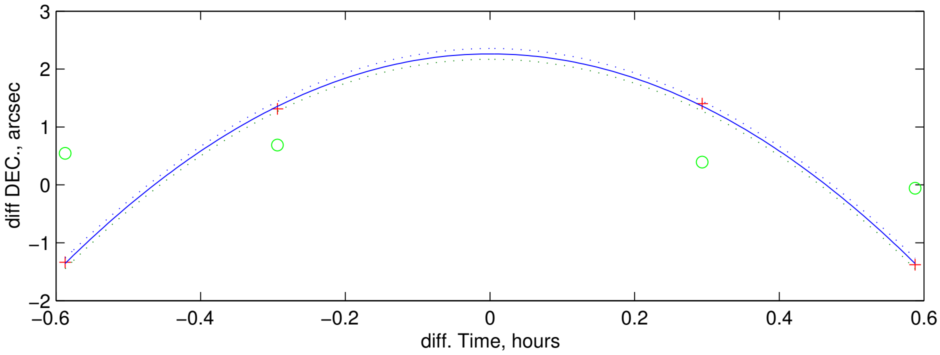
1.7 CURVATURE, TOO SHORT ARC

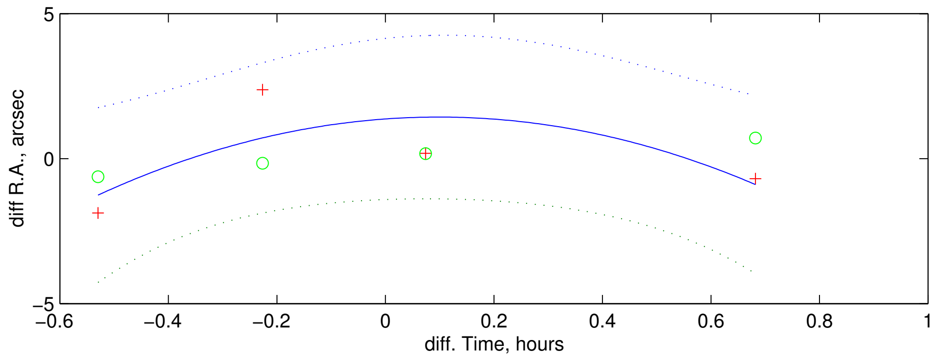
Curvature A measure of the deviation of the Observed Arc from a great circle, traced with uniform speed on the celestial sphere. The curvature is **Significant** if the deviations of the individual observations from a great circle cannot be due only to observational error (according to the Error Model).

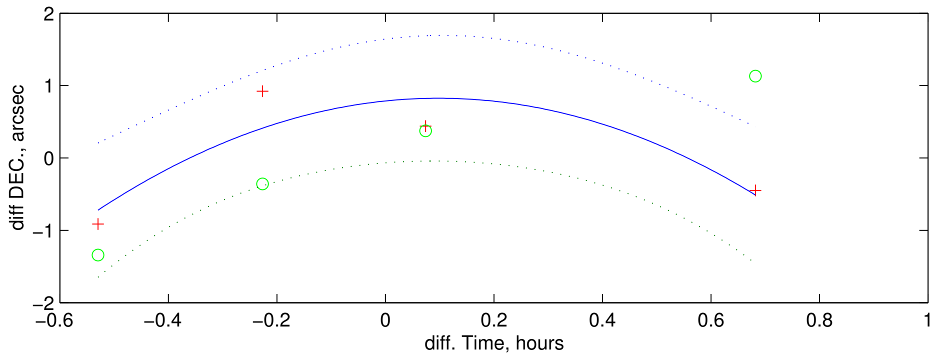


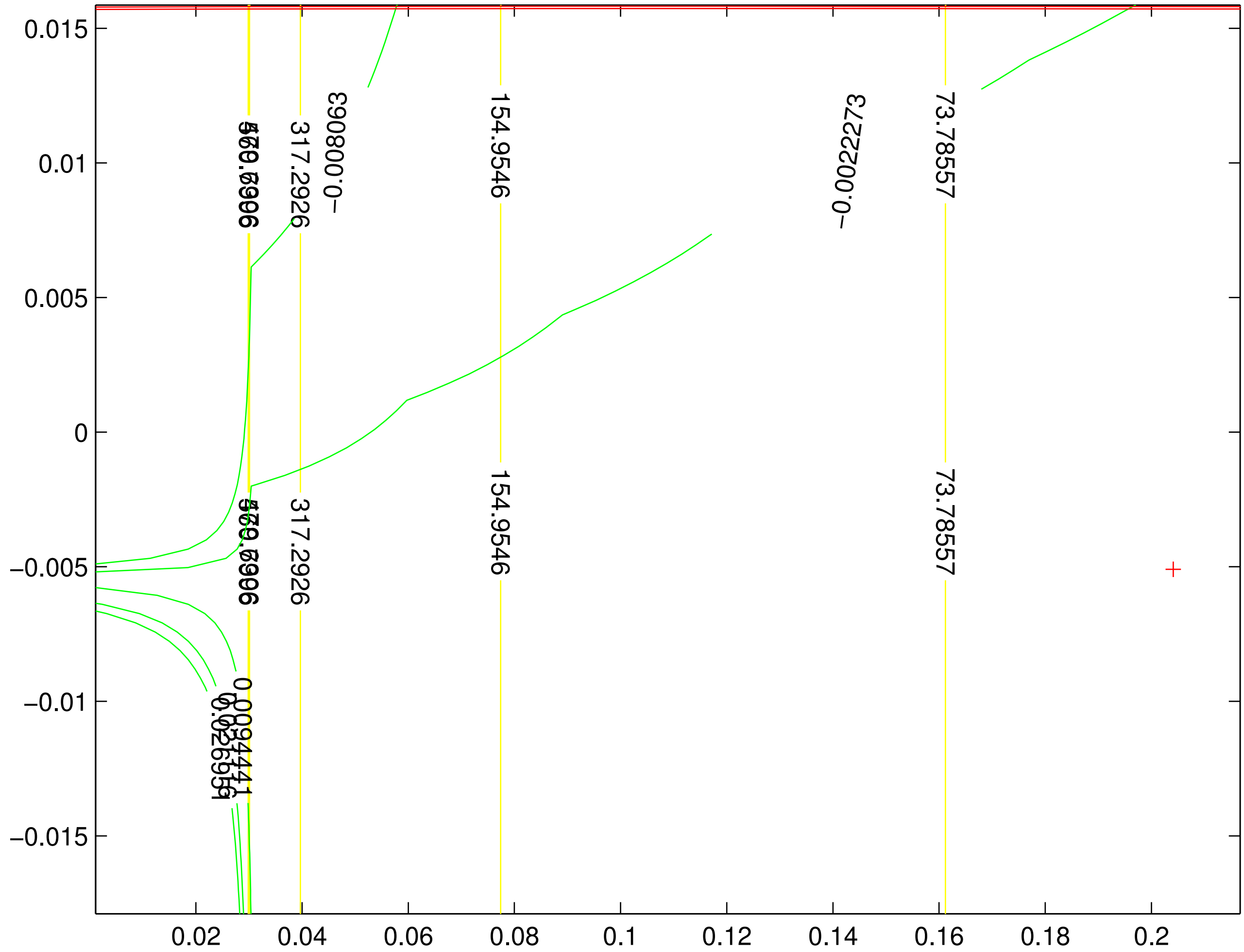
Too Short Arc (TSA) An Observed Arc too short to compute a useful Least Squares Orbit, because it has no significant curvature.


diff R.A., arcsec
8
6
4
2
0
−2
−4
−6
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
diff. Time, hours


diff DEC., arcsec
3
2
1
0
−1
−2
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
diff. Time, hours


diff R.A., arcsec
5
0
−5
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1
diff. Time, hours


diff DEC., arcsec
2
1
0
−1
−2
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1
diff. Time, hours



Plot on the $(\rho, \dot{\rho})$ plane (topocentric range in AU, range rate in AU/day) of the lines of constant geodesic curvature ($\rho = const$) and constant along track accelerationt, given the attributable of the discovery nigt of 2004 AS_1. The intersection of the nominal values would be in the lower left corner (imediate impact region), the a posteriory ground truth is the red cross. This disastrous false alarm was due to the faulty error model.

1.8 DEFINITIONS: ARC TYPE, DISCOVERY

Arc of Type N An Observed Arc which can be split into exactly N disjoint TSA in such a way that each couple of TSA consecutive in time, if joined, would show a significant curvature. This definition is meant to replace the currently used definition of *N-nighter*, an observed arc containing observations belonging to exactly N distinct nights.

Discovery A set of observations of a SSO, forming an Observed Arc of Type N with $N \geq 3$; there must be a unique full least squares orbit fitting the data with residuals compatible with the Error Model; the object needs to be a New SSO. It is also required that the data contain enough photometric information to fit an absolute magnitude. The Observations have to belong to tracklets which have been submitted to the Data Center, either at once or at different times (by one or more Observers); the Orbit, and the critical Identification (allowing a Type 3 Arc to be built) must either have been submitted to the Data Center by Orbit Computers or have been computed by Data Center itself.

Discovery of a comet A Discovery as above, complemented with enough observational data to prove that there is a directly detectable cometary activity.

2.1 VIRTUAL ASTEROIDS

When a single orbit solution is not enough to represent the possible orbits, they are replaced by a swarm of **Virtual Asteroids** (VAs). The VAs share the reality of the physical asteroid, in that only one of them is real, but we do not know which one. (Also called the multiple hypothesis method). Additional observations allow to decrease the number of VAs still compatible.

In fact, there is always an infinite number of orbits in the **confidence region** defined by some maximum allowed penalty $m\Delta Q(\mathbf{x}) \leq \sigma^2$. Thus each VA is just a representative of a patch in the confidence region. Such a finite sampling strategy can be successful only when handling the patch around a VA is matemathically much simpler than handling the entire confidence region at once.

This occur when the VAs are enough to allow the use of a linear approximation in their patch. That is, the confidence region is NOT well approximated by the ellipsoid $(\mathbf{x}-\mathbf{x}^*)^t\, C_{\mathbf{x}}(\mathbf{x}^*)\,(\mathbf{x}-\mathbf{x}^*) \leq \sigma^2$ centered at the nominal least squares solution, but the ellipsoids centered at each of the VAs are a covering of the confidence region.

2.2 LINEARIZATION

When the patches around each VAs are small enough, we can linearize the prediction function (providing the positions in some following night, etc.) at each VA and use the linear theory, for example to compute the identification penalty. That is, the VA number j can be identified with another asteroid, the others cannot, thus number j is more likely to be true than the others.

We will first discuss the simplest possible case, in which only two VAs are required, with orbits $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$. This is the *double solution* case occurring only in the sweet spot surveys. In most cases, when there are tracklets over 3 nights, each one of the two solutions is surrounded by a **small** component of the confidence region. Thus the linear approximation, expressed by the normal matrices $C(\mathbf{x}^{(1)}), C(\mathbf{x}^{(2)})$ is accurate for each of the two; on the other hand, the distance between the two solutions is large, that is the quadratic approximation of the penalty $(\mathbf{x}^{(2)} - \mathbf{x}^{(1)})^T C(\mathbf{x}^{(1)})\, (\mathbf{x}^{(2)} - \mathbf{x}^{(1)})$ is very large and completely wrong.

2.3 COMPUTATION OF THE ATTRIBUTABLE

The procedure to compute an attributable, if there are $m \geq 3$ observations, is as follows. Given the observed values $(t_i, \alpha_i, \delta_i)$ for $i = 1, m$ we can fit both angular coordinates as a function of time with a polynomial model: in the cases of interest a degree 2 model is satisfactory

$$\begin{aligned}\alpha(t) &= \alpha(\bar{t}) + \dot{\alpha}(\bar{t})\ (t-\bar{t}) + \frac{1}{2}\ddot{\alpha}(\bar{t})\ (t-\bar{t})^2 \\ \delta(t) &= \delta(\bar{t}) + \dot{\delta}(\bar{t})\ (t-\bar{t}) + \frac{1}{2}\ddot{\delta}(\bar{t})\ (t-\bar{t})^2\end{aligned}$$

with $\bar{t}$ the mean of the t_i; the solution $(\alpha, \dot{\alpha}, \ddot{\alpha}, \delta, \dot{\delta}, \ddot{\delta})$ is obtained by a linear weighted least squares fit, together with the two 3×3 covariance matrices $\Gamma_\alpha, \Gamma_\delta$.

The second derivatives with respect to time are computed as an insurance against the possibility that a linear fit is not a good representation of the short arc data, but the attributable contains only the averages and rates of angular motion. The marginal covariance matrix of A, whatever the values of $(\ddot{\alpha}, \ddot{\delta})$ is obtained by extracting the relevant 4×4 submatrix.

2.4 COMPONENTS OF CURVATURE

The heliocentric position of the asteroid is the vector $\mathbf{r}$ and the topocentric position is

$$\boldsymbol{\rho} = \rho\, \hat{\boldsymbol{\rho}} = \mathbf{r} - \mathbf{q}$$

where $\mathbf{q}$ is the heliocentric position of the observer, $\hat{\boldsymbol{\rho}}$ the unit vector, ρ the distance. The angolar velocity of the the asteroid on the celestial sphere is

$$\mathbf{v} = \frac{d\hat{\boldsymbol{\rho}}}{dt} = \eta\, \hat{\mathbf{v}} \quad , \quad \hat{\mathbf{v}} \cdot \hat{\boldsymbol{\rho}} = 0$$

where $\eta = \|\mathbf{v}\|$ is the *proper motion*. By using the *arc length parameter* s, defined by $ds/dt = \eta$, we have $d\hat{\boldsymbol{\rho}}/ds = \hat{\mathbf{v}}$ and the derivative $d\hat{\mathbf{v}}/ds = \hat{\mathbf{v}}'$ has the properties

$$\begin{aligned} \hat{\mathbf{v}}' \cdot \hat{\mathbf{v}} &= \frac{1}{2}\frac{d}{ds}\|\hat{\mathbf{v}}\|^2 = 0 \\ \hat{\mathbf{v}}' \cdot \hat{\boldsymbol{\rho}} &= \frac{d}{ds}[\hat{\mathbf{v}} \cdot \hat{\boldsymbol{\rho}}] - \hat{\mathbf{v}} \cdot \hat{\boldsymbol{\rho}}' = -1 \end{aligned}$$

With the ortogonal vector $\hat{\mathbf{n}} = \hat{\boldsymbol{\rho}} \times \hat{\mathbf{v}}$ we can express $\hat{\mathbf{v}}'$ as

$$\hat{\mathbf{v}}' = -\hat{\boldsymbol{\rho}} + \kappa\, \hat{\mathbf{n}}$$

where κ is the **geodesic curvature**, measuring the deviation from a great circle (a geodesic on the sphere).

2.5 COMPUTATION OF CURVATURE

The geodesic curvature can be computed from the second derivatives $(\ddot\alpha, \ddot\delta)$ and the attributable $A = (\alpha, \delta, \dot\alpha, \dot\delta)$.

$$\kappa = \frac{1}{\eta^3}\left[(\ddot\delta\dot\alpha - \ddot\alpha\dot\delta)\cos\delta + \dot\alpha(\eta^2 + \dot\delta^2)\sin\delta\right] .$$

Another component of the path second derivative is the along track **acceleration**, that is

$$\frac{d^2\hat{\boldsymbol{\rho}}}{dt^2}\cdot\hat{\mathbf{v}} = \frac{d}{dt}(\eta\,\hat{\mathbf{v}})\cdot\hat{\mathbf{v}} = \left(\dot\eta\,\hat{\mathbf{v}} + \eta^2\,\hat{\mathbf{v}}'\right)\cdot\hat{\mathbf{v}} = \dot\eta\ ,$$

it can be computed from $(\ddot\alpha, \ddot\delta)$ by

$$\dot\eta = \frac{d^2\hat{\boldsymbol{\rho}}}{dt^2}\cdot\hat{\mathbf{v}} = \frac{\ddot\alpha\,\dot\alpha\,\cos^2\delta + \ddot\delta\,\dot\delta - \dot\alpha^2\,\dot\delta\,\cos\delta\,\sin\delta}{\eta} .$$

The third component of curvature is simply the curvature of the sphere, as shown by $\hat{\mathbf{v}}'\cdot\hat{\boldsymbol{\rho}} = -1$.

Given these formulas, it is possible to compute the covariance matrix of the quantities $(\kappa, \dot\eta)$ by propagation of the covariance matrix of the angles and their derivatives with the matrix of partial derivatives computed from the above formulae for κ and $\dot\eta$.

2.6 THE PRELIMINARY ORBITS OF LAPLACE

The time derivatives of the topocentric vector $\boldsymbol{\rho}$

$$\begin{aligned} \dot{\boldsymbol{\rho}} &= \dot{\rho}\,\hat{\boldsymbol{\rho}} + \rho\,\eta\,\hat{\mathbf{v}} = \dot{\mathbf{r}} - \dot{\mathbf{q}} \\ \ddot{\boldsymbol{\rho}} &= \ddot{\rho}\,\hat{\boldsymbol{\rho}} + (\rho\,\dot{\eta} + 2\,\dot{\rho}\,\eta)\,\hat{\mathbf{v}} + \rho\,\eta^2\,(\kappa\,\hat{\mathbf{n}} - \hat{\boldsymbol{\rho}}) \end{aligned}$$

can be decomposed along the orthonormal vectors $(\hat{\boldsymbol{\rho}}, \hat{\mathbf{v}}, \hat{\mathbf{n}})$

$$\ddot{\boldsymbol{\rho}}\cdot\hat{\mathbf{v}} = \rho\,\dot{\eta} + 2\,\dot{\rho}\,\eta \quad , \quad \ddot{\boldsymbol{\rho}}\cdot\hat{\mathbf{n}} = \rho\,\eta^2\,\kappa \quad , \quad \ddot{\boldsymbol{\rho}}\cdot\hat{\boldsymbol{\rho}} = \ddot{\rho} - \rho\,\eta^2 \,.$$

Let's assume $\mathbf{q}$ is the center of the Earth (neglecting the topocentric correction). The 2-body formula for the accelerations $\ddot{\boldsymbol{\rho}}$ and $\ddot{\mathbf{q}}$ is

$$\ddot{\boldsymbol{\rho}} = \frac{-\mu\,\mathbf{r}}{r^3} + \frac{\mu\,\mathbf{q}}{q^3}$$

where r is the heliocentric distance of the asteroid, q is the heliocentric distance of the Earth, μ the mass of the Sun times the gravitational constant. Note that the denominator of the first fraction is $r^3 = S(\rho)^{3/2}$ where the relationship between the sides ρ, r, q of the triangle is

$$r^2 = S(\rho) = \rho^2 + 2\,\cos\varepsilon\,q\,\rho + q^2$$

with $\cos\varepsilon = \hat{\mathbf{q}}\cdot\hat{\boldsymbol{\rho}}$.

2.7 MULTIPLE SOLUTIONS

The component of the relative acceleration $\ddot{\boldsymbol{\rho}}$ along the $\hat{\mathbf{n}}$ direction is

$$\ddot{\boldsymbol{\rho}}\cdot\hat{\mathbf{n}} = \frac{-\mu\,\mathbf{q}\cdot\hat{\mathbf{n}}}{r^3} + \frac{\mu\,\mathbf{q}\cdot\hat{\mathbf{n}}}{q^3} = \rho\,\eta^2\,\kappa$$

where we have used $\hat{\boldsymbol{\rho}}\cdot\hat{\mathbf{n}} = 0$ to simplify. Given κ, this is an equation for ρ. Let us define

$$C = \frac{\eta^2\,\kappa\,q^3}{\mu\,\hat{\mathbf{q}}\cdot\hat{\mathbf{n}}}$$

then the dynamic equation ($\hat{\mathbf{n}}$ component) takes the form

$$1 - C\,\frac{\rho}{q} = \frac{q^3}{S(\rho)^{3/2}}\ .$$

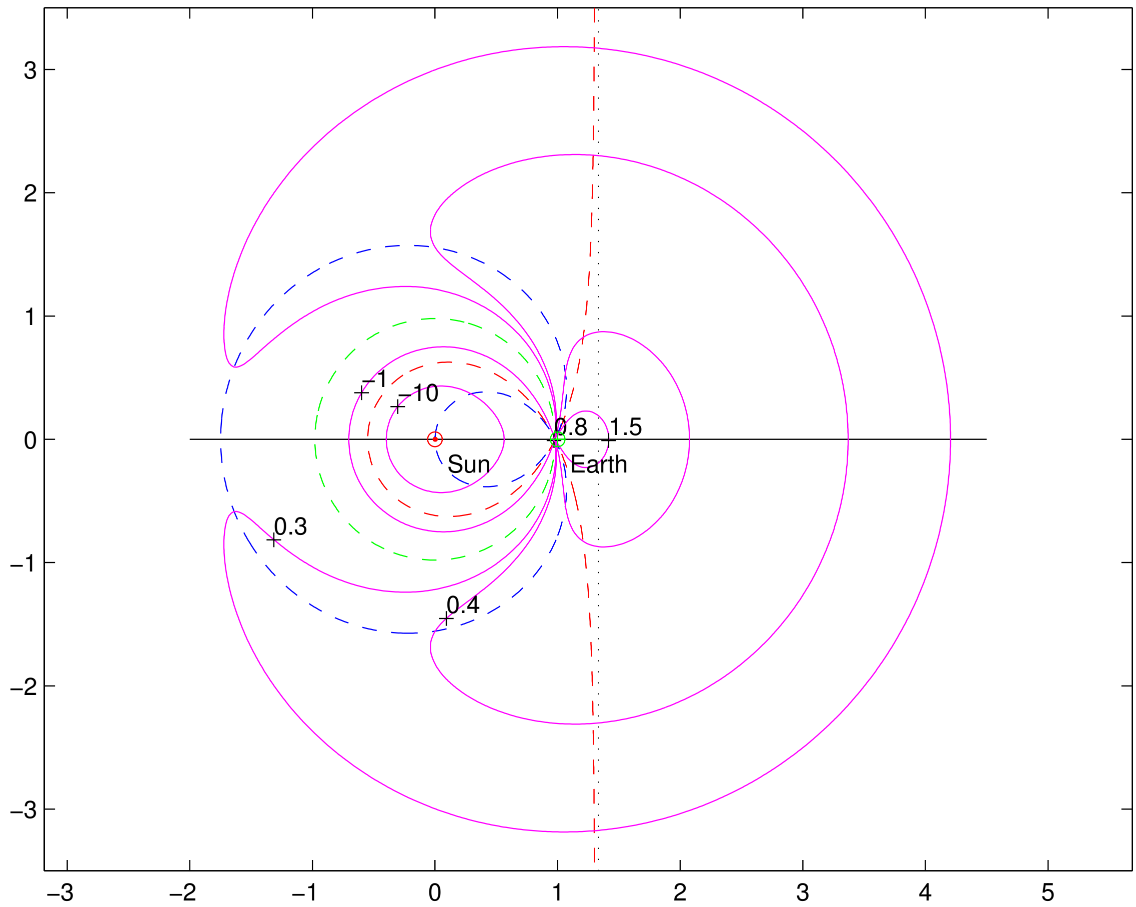
By substituting the possible values of ρ obtained by the triangle equation, after squaring, we obtain an algebraic equation of degree eight:

$$\begin{aligned} p(r) \;\; &= \;\; C^2 r^8 - q^2\left(C^2 + 2C\cos\varepsilon + 1\right) r^6 \\ &\quad + 2q^5(C\cos\varepsilon + 1)r^3 - q^8 = 0 \end{aligned}$$

which has the same solutions provided $q/\rho > C$. By the standard theory of polynomial equations, this polynomial cannot have more than three positive roots, including the “degenerate” solution $r = q$, implying $\rho = 0$.

2.8 REGIONS OF DOUBLE SOLUTIONS

As the parameter C varies, the dynamical equation $C\,\rho/q = 1 - q^3/r^3$ represents a family of curves in the plane.



The green dotted curve is $r = q$, that is $\kappa = 0$. The red dotted curve is the *limiting curve*, the double solutions can occur only in two regions: inside the loop and outside the $r = q$ curve on the left of the unlimited branch. The tangents have angles of $116^0.5$ and $63^0.5$ with the Sun.

2.9 HOW TO HANDLE DOUBLE SOLUTIONS

The main problem is that in most cases the double solutions for the preliminary orbit are one Main Belt and one NEO, often Aten ($a < 1\,AU$). Taking into account the a priori probability estimated from population models the MB solution has in fact a probability > 0.99; on the other hand the only way to discover NEOs is to make the hypothesis that each one is a NEO, otherwise a poor orbit would result in a loss of the identifications.

Thus the goal of the preliminary orbit determination (IOD) should be to get both solutions whenever they exist. This can be done in two ways:

1) Use a IOD algorithm which finds both roots of the degree 8 equation, checks which solution is admissible, computes $\rho, \dot{\rho}$, optionally improves the orbit by an iterative method, the tries differential corrections for all the preliminary orbits found.

2) Use a IOD algorithm based upon a larger swarm of VA, thus in most cases if there are two distinct solutions both will be found. Optionally, whenever a 3-night orbit is found, look for the *phantom* orbit corresponding to the other root of the deg. 8 equation.

2.10 ESTIMATES OF CURVATURE

From the dynamical equation in the $\hat{\mathbf{n}}$ direction

$$\ddot{\boldsymbol{\rho}} \cdot \hat{\mathbf{n}} = \rho\, \eta^2\, \kappa \Longrightarrow \eta^2\, \kappa = \frac{\ddot{\boldsymbol{\rho}} \cdot \hat{\mathbf{n}}}{\rho}$$

the heliocentric acceleration of a solar system body is small:

$$r \simeq 1 \Rightarrow \ddot{\boldsymbol{\rho}} \cdot \hat{\mathbf{n}} \leq 0.6\, cm\, s^{-2} = 3 \times 10^{-4}\, AU\, d^{-2}\ .$$

The geocentric distance ρ cannot be too small, else the object is too small to be aprimary goal; simple example: $H \leq h \Rightarrow \rho > 1\, AU$ at opposition (phase 0). Thus

$$\eta^2\, \kappa \leq 3 \times 10^{-4}\, d^{-2}$$

From the Taylor formula for $\hat{\rho}$ as a function of the arc length

$$\hat{\boldsymbol{\rho}}(\Delta s) = \hat{\boldsymbol{\rho}}(0) + \hat{\mathbf{v}}\, \Delta s + (\kappa\, \hat{\mathbf{n}} - \hat{\boldsymbol{\rho}})\, \frac{\Delta s^2}{2} + O(\Delta s^3)\ ;$$

neglecting $O(\Delta t^3)$ we have $\Delta s = \eta\, \Delta t + \dot{\eta}\, \Delta t^2/2$ thus

$$|(\hat{\boldsymbol{\rho}}(\Delta t) - \hat{\boldsymbol{\rho}}(0)) \cdot \hat{\mathbf{n}}| = \eta^2\, \kappa\, \frac{\Delta t^2}{2} \leq$$
$$\leq 1.5 \times 10^{-4}\, \Delta t^2 \simeq 0^{\circ}.01\, \Delta t^2$$

The uncertainty $\Delta\hat{\mathbf{v}}$ in the direction of angular motion contributes less: $\eta\, \Delta\hat{\mathbf{v}} \simeq 0^{\circ}.004\, d^{-1}$ for PS.

2.11 ESTIMATES OF ACCELERATION

A similar argument applies to $\dot{\eta}$

$$\ddot{\boldsymbol{\rho}} \cdot \hat{\mathbf{v}} = \rho\, \dot{\eta} + 2\, \dot{\rho}\, \eta \Longrightarrow \dot{\eta} = \frac{\ddot{\boldsymbol{\rho}} \cdot \hat{\mathbf{v}}}{\rho} - 2\, \eta\, \frac{\dot{\rho}}{\rho}$$

$\dot{\rho}$ and $\eta\ \rho$ are $\leq 1/60\, AU\ d^{-2}$ thus

$$\dot{\eta} \leq 9 \times 10^{-4}\, d^{-2}\ .$$

However, this is a very generous estimate, if fact the equal sign can apply only for very strange orbits; for a realistic sample of orbits, the values should be significantly less.

$$|(\hat{\boldsymbol{\rho}}(\Delta t) - \hat{\boldsymbol{\rho}}(0)) \cdot \hat{\mathbf{v}} - \eta\ \Delta t| = |\dot{\eta}\ \frac{\Delta t^2}{2}| \leq$$
$$\leq 4.5 \times 10^{-4}\, \Delta t^2 \simeq 0^\circ.03\, \Delta t^2$$

The uncertainty in η does not matter. Thus the area to be scanned to find the same object after Δt days is

$$\pi\, 0^\circ.01\ 0^\circ.03\, \Delta t^2 \simeq 10^{-3}\, \Delta t^2\, deg^2$$

Even with number density $N \simeq 400$ per deg^2, after one day there is on average less than one detection, and the number of false identifications can be controllable even after a few days. This is the basis of the method of Kubica. In the future could be the base for an even more aggressive strategy, combining some tracklets and some single observations.

3.1 THE QUALITY OF THE PAN-STARRS OUPUT

Pan-STARRS will be the first Solar System survey to take full responsibility for processing the observations to obtain the discoveries, that is for identification and orbit determination. Thus the strengths of PS are not just the gigapixel firepower, but also MOPS. However, the question is to make sure that the output of MOPS is something to be proud of.

Quality control of the data output has many aspects, we shall discuss QC for orbits and for identifications. The main problem is to find the right balance between quantity and quality, that is in the simulations we want to measure both completeness and reliability.

To push completeness even in the probabilistic detection range, we wish to include in the output identifications with few tracklets (maybe even 2) and orbits even when the conventional differential corrections fail. This is acceptable provided each identification and its orbit are documented with an exact specification of its quality. Then the output can be partitioned in high quality, reliable full "discoveries" and in low quality, dubious, incomplete discoveries (or even more than two categories if needed).

3.2 THE ORBITS ARE NOT ALL EQUAL

The orbits can be **preliminary** (also called initial, hence IOD) and **least squares**, in the latter case with different number of **fit parameters**: either 4, or 5, or 6.

Nominal solutions; given a set of $m = 2M$ scalar observations, and some first guess $\mathbf{x}$ for the orbital elements, the observation residuals $\boldsymbol{\xi}$ (an m-dim vector) with their weight matrix W (an $m \times m$ symmetric matrix) and the normal matrix C at $\mathbf{x}$

$$B(\mathbf{x}) = \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{x}}(\mathbf{x}) \quad , \quad C(\mathbf{x}) = B(\mathbf{x})^T \, W \, B(\mathbf{x}) \; ,$$

differential corrections is a modified Newton's method, an iterative procedure with, at each step, a correction $\Delta\mathbf{x}$ such that

$$C(\mathbf{x}) \, \Delta\mathbf{x} = D(\mathbf{x}) = -B^T \, W \, \boldsymbol{\xi}$$

where the right hand side D is proportional to the gradient of the target function $Q(\mathbf{x}) = \boldsymbol{\xi}^T \, W \, \boldsymbol{\xi}/m$. At convergence $\mathbf{x} \to \mathbf{x}^*$ with $D(\mathbf{x}^*) = \underline{0}$, then $\mathbf{x}^*$ is a (local) minimum of $Q(\mathbf{x})$; this is the 6-parameter fit. Starting from a different first guess can result in a different **local minimum** (as in the case of double preliminary orbit near quadrature).

3.3 CONSTRAINED ORBITS

Let $\lambda_j(\mathbf{x}), j = 1,\ldots,6$ be the eigenvalues of $C(\mathbf{x})$, with $\lambda_1(\mathbf{x})$ the smallest one; let $V_1(\mathbf{x})$ be an eigenvector with eigenvalue $\lambda_1(\mathbf{x})$, that is

$$C(\mathbf{x})\, V_1(\mathbf{x}) = \lambda_1(\mathbf{x})\; V_1(\mathbf{x})$$

that is the **weak direction** of least information, also the eigenspace of the largest eigenvalue of $\Gamma(\mathbf{x})$, that is the direction of greatest uncertainty. Let $\mathcal{H}(\mathbf{x})$ be the 5-dim hyperplane orthogonal to $V_1(\mathbf{x})$

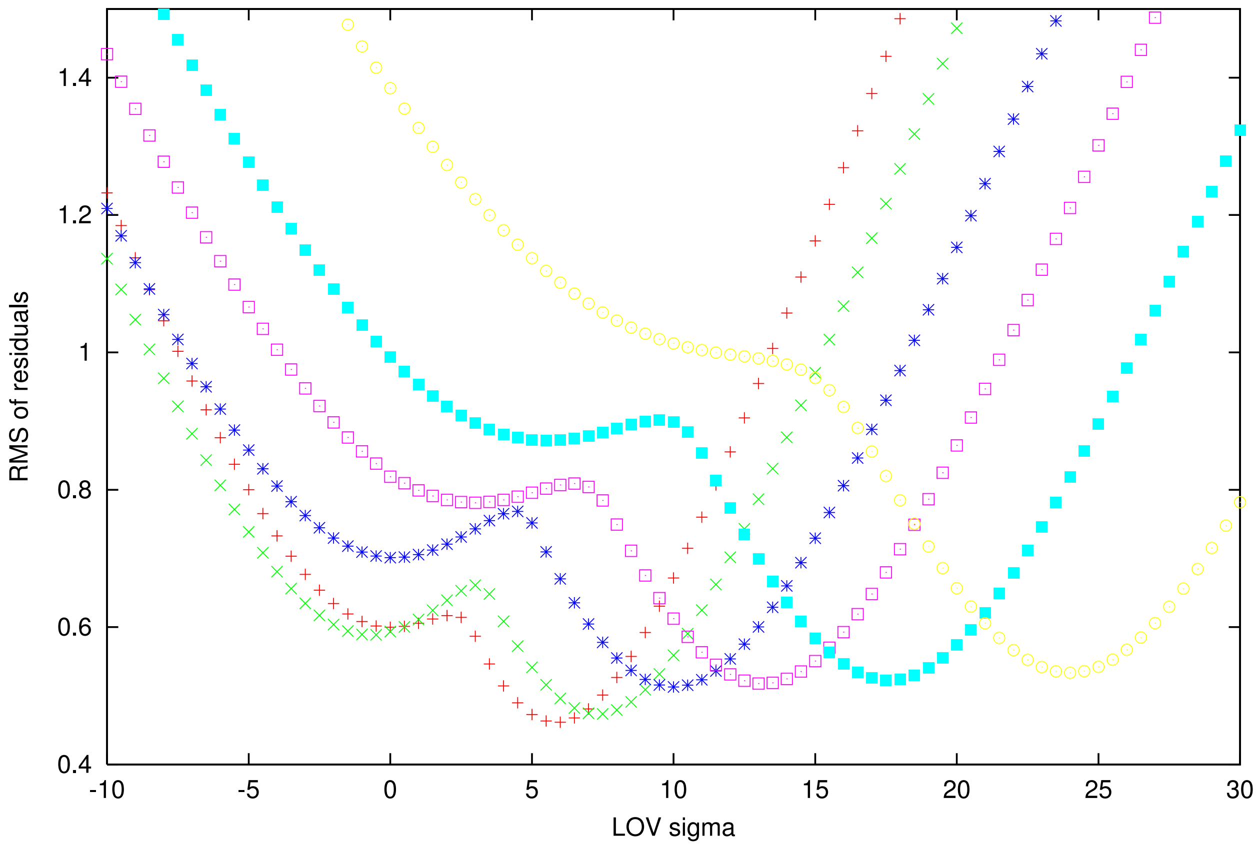
$$\mathcal{H}(\mathbf{x}) = \{\mathbf{y} | (\mathbf{y} - \mathbf{x}) \cdot V_1(\mathbf{x}) = 0\} \; .$$

Constrained differential corrections is the iterative procedure in which the equation for each step is

$$C(\mathbf{x})\; \Delta\mathbf{x} = \pi_{\mathcal{H}(\mathbf{x})} D(\mathbf{x})$$

with the right hand side projected onto the hyperplane $\mathcal{H}(\mathbf{x})$, that is refusing to correct along the weak direction. At convergence $\mathbf{x} \to \mathbf{x}^*$ the gradient of the target function is parallel to the weak direction. Such points form a curve in 6-dim space, called **Line Of Variation (LOV)**.

The $\mathbf{x}^*$ always depend upon the initial guess, and they are **LOV solutions** with 5 free parameters. (An alternative, and older, method is to fix one variable, such as eccentricity, and fit the other 5.)



The asteroid 1998 XB, discovered at an elongation of 93°, had a double IOD. The Figure shows the $RMS(\xi)$ (arcsec) vs. the LOV parameter σ. The lines are marked with plusses (arc time span 9 days), crosses (10 days), stars (11), boxes (13), full boxes (14) and circles (16).

The first orbit published by the MPC, with 10 days of obs., had $a = 1.021$ AU, corresponding to the higher local minimum. In the following days the orbit had a decreasing until 0.989 (at 13 days). With 16 days, the semimajor axis jumped to 0.906.

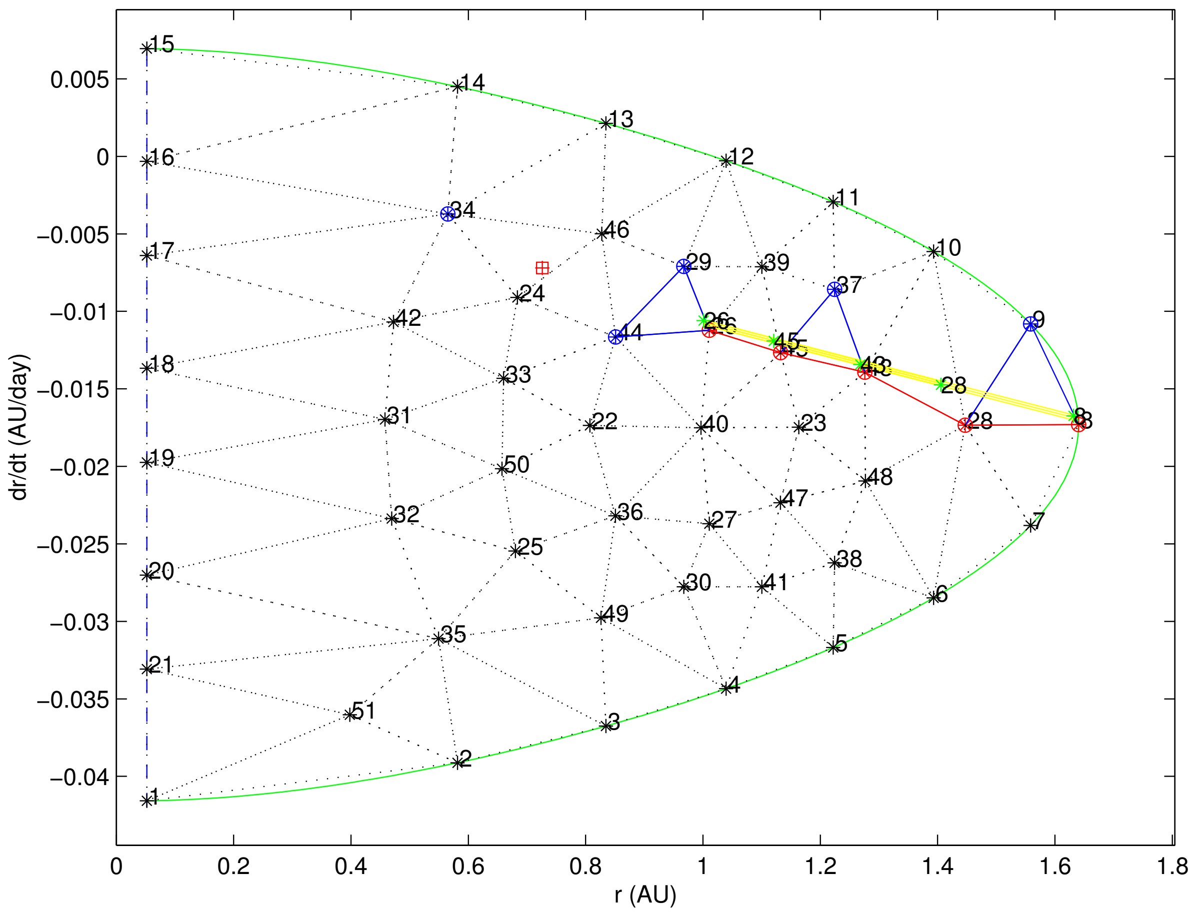
3.4 TOO SHORT ARCS

In some cases, mostly TNO, even the constrained differential corrections may not converge. This because the curvature is not significant, it is a **Type 1 Arc**, that is, we can only compute an attributable for all the tracklets together. For TNO 2-ids, this happens in 94% of the cases.

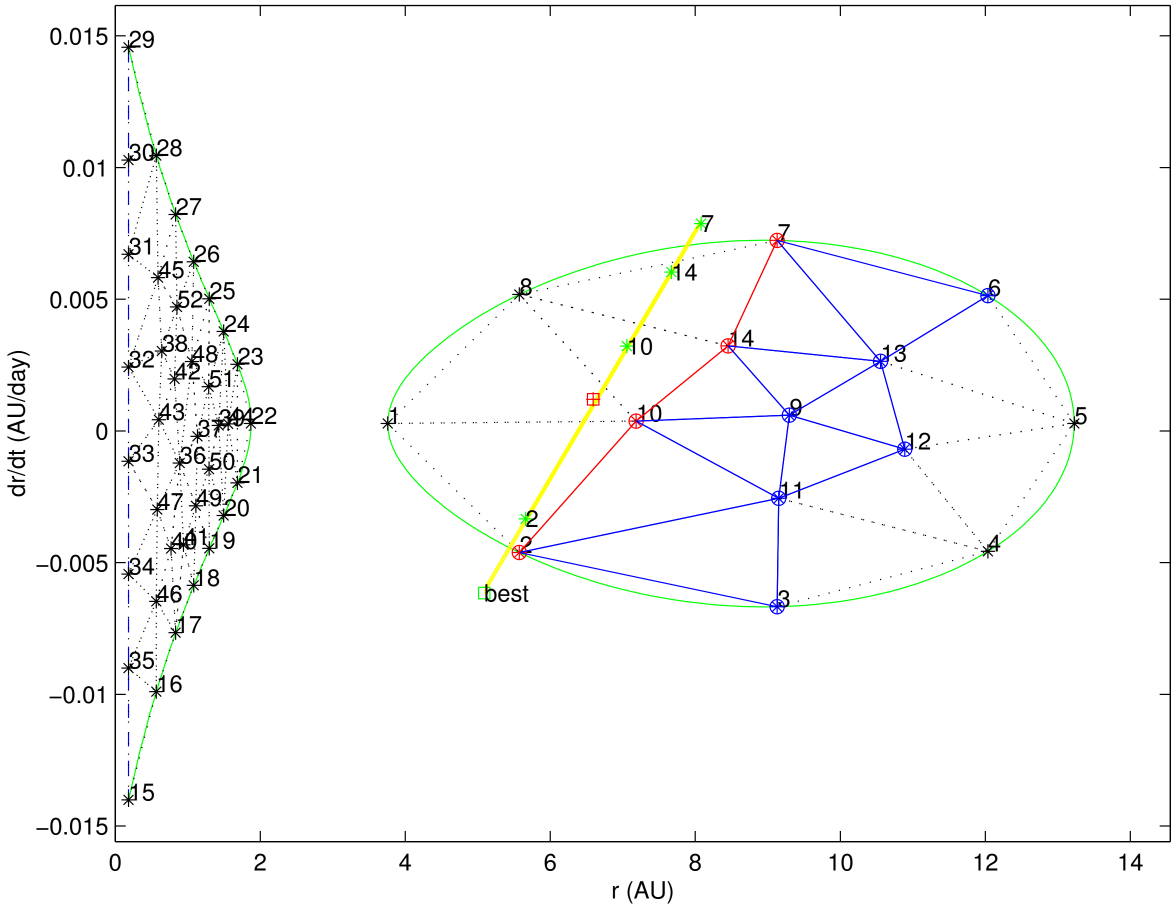
To join two tracklets in a single attributable, we use **attributable elements** $\mathbf{x} = (\alpha, \delta, \dot{\alpha}, \dot{\delta}, \rho, \dot{\rho})$ where ρ is the topocentric distance. The epoch time is $t_0 = \bar{t} - \rho/c$ ($\bar{t}$ the average observation time). Then it is a good approximation to assume that $(\rho, \dot{\rho})$ are undetermined and just leave them at the first guess value, correcting the first four.

The values of $(\rho, \dot{\rho})$ cannot take an arbitrary value in the half plane $\rho > 0$ because, for a given attributable, some values in that plane correspond to either hyperbolic orbits or too close to be of significant size. Thus $(\rho, \dot{\rho})$ must belong to a compact **admissible region**.

Thus the constrained fit with 4 parameters gives an **admissible region solution**. (This method was originally proposed by Tholen and implemented in KNOBS, then used by Chesley in his analysis of the 2004 AS$_1$ case).



For 1998 XB, attempting to identify the attributable from the night of November 25 with another attributable, based only upon the data of December 26. The plot shows the admissible region (parabolic boundary green, shooting star limit dotted) in the plane of the variables $(\rho, \dot{\rho})$ at the time of the first tracklet. The LOV (yellow) is close to the nodes with low penalty (red), but the true solution (red square) is far along the same direction. Anyway identification with a third night is possible.



The Centaur *(31824) Elatus* was discovered on October 29, 1999 by the Catalina Survey. The admissible region has two connected components, as it is usual for TNO at opposition (Warning: double solutions occur at quadrature, double components occur at opposition!).

For the precovery by the LONEOS survey on October 17, 1998 we have propagated $\simeq 40$ VAs and computed the identification penalties: we were able to obtain 4 LOV solutions (one hyperbolic). The nominal solution also corresponds to a hyperbolic orbit.

3.5 ORBIT COMPUTATION: AN EFFECTIVE PROCEDURE

The procedure for orbit determination can be organized as a **sequence of steps**.

1) IOD: for 3-nighters (with significant curvature) Gauss' method takes into account the topocentric corrections, thus it is more effective than Laplace's (unless the latter is corrected, but the topocentric correction in acceleration is large!). For 2-nighters a different IOD is based upon the compromise attributable A_0 (see 1.4). In both cases there can be multiple preliminary orbits.

2) If there is no significant curvature, a 4-fit is performed to find the **attributable** combining the information from the ≥ 2tracklets. If the curvature is good this is skipped.

3) The constrained differential corrections are used to 5-fit a **LOV solution**.

4) The full differential corrections are used to 6-fit a **nominal solution**.

5) Quality control is applied to the residuals at convergence, the orbit is accepted if the normalized $RMS(\xi)$ is small, plus additonal checks on systematic trends.

Each stage is used as **first guess** for the next one, provided it passes some (looser) quality control.

3.6 THE QUALITY OF AN IDENTIFICATION

In a large identification procedure (several $100,000$ ids in a full size PS simulation), how to assess quality? Let us define an **order** among them: an id with more tracklets (should be nights?) is better, among the ones with the same number the lower $RMS(\xi)$ is better. The list of identification is sorted by this order. The identifications on the top of are the most reliable; however, we must take into account all the other ids having tracklets in common: they can be **discordant** or **compatible**.

As a very small example, let us assume the sorted list of identifications is as follows (here capital letters stand for tracklet unique identifiers, OIDs in DB jargon). We proceed to the **normalization** of the ids DB, removing all the discordancies (and possibly the multiple solutions).

1. $A = B = C = D$ in normalized DB
2. $A = B = C$ compatible and inferior to 1.
3. $E = F = C$ discordant and inferior to 1.
4. $A = B$ compatible and inferior to 1.
5. $F = C$ discordant and inferior to 1.
6. $E = F_1$ in normalized DB
7. $E = F_2$ compatible and inferior to 6.

3.7 THE DISCOVERY CLAIMS

As discussed in 1.8, the requirements for discovery are:

1) astrometry with enough information, or **Arc of Type 3**, For PS, NEO and MBA with 3 tracklets in 3 distinct nights are Type 3 Arcs in 90% of the cases.

2) unique nominal solution (with 6 parameters): both double nominal solutions and multiple LOV solutions (also 1 LOV 1 nominal) are to be discarded (problem with example $E = F_1$ above).

3) previously unknown. All the tracklets should be tested for attribution to previously known objects before identification, then the corresponding ids should be added to the newly found one and normalization repeated.

What with identifications not surviving to normalization? They are to be considered possible ids. Thus the **possible identification database** contains identifications

1) with Arc Type 2 (possibly 1 for TNO)

2) with multiple orbits: double nominal (near quadrature) and LOV solutions

3) discordant with others of same quality (same number of tracklets/nights and comparable $RMS(\boldsymbol{\xi})$).

About 1/4 of TNOs observed at opposition over 3 nights will end up in this category.

REFERENCES I

M.E. Sansaturio, A. Milani & L. Cattaneo: *Nonlinear optimisation and the asteroid identification problem*, in *Dynamics, Ephemerides and Astrometry of the Solar System*, Ferraz-Mello, S. ed., Kluwer, 193–198, 1996.

A. Milani: *The asteroid identification problem I: recovery of lost asteroids*, Icarus, 137, 269-292, 1999.

A. Milani, A. La Spina, M.E. Sansaturio & S.R. Chesley: *The asteroid identification problem III: proposing identifications*, Icarus, 144, 39–53, 2000.

A. Milani, M.E. Sansaturio & S.R. Chesley: *The asteroid identification problem IV: attributions*, Icarus, 151, 150–159, 2001.

M. Carpino, A. Milani & S.R. Chesley *Error statistics of asteroid optical astrometry observations*, Icarus, 166, 248–270, 2003.

A. Milani, G.-F. Gronchi, M. de' Michieli Vitturi & Z. Knežević: *Orbit Determination with Very Short Arcs. I Admissible Regions*, Celestial Mechanics, 90, 57–85, 2004.

A. Milani, M.E. Sansaturio, G. Tommei, O. Arratia & S.R. Chesley: *Multiple solutions for asteroid orbits: computational procedure and applications*, Astronomy and Astrophysics, 431, 729–746, 2005.

REFERENCES II

A. Milani: *Virtual Asteroids and Virtual Impactors*, in *Dynamics of Populations of Planetary Systems*, Z. Knežević and A. Milani, eds., Cambridge University Press, pp. 219–228, 2005.

A. Milani & Z. Knežević: *From Astrometry to Celestial Mechanics: Orbit Determination with Very Short Arcs*, Celestial Mechanics, 92, 1–18, 2005.

A. Milani, G.F. Gronchi, Z. Knežević, M.E. Sansaturio & O. Arratia: *Orbit determination with very short arcs. II Identifications*, Icarus, 79, 350–374, 2005.

A. Milani et al.: *Unbiased orbit determination for the next generation asteroid/comet surveys*, in *Asteroids Comets Meteors 2005*, D. Lazzaro et al., eds., Cambridge University Press, pp.367–380, 2006. (Common paper with MOPS!)

A. Milani, G.F. Gronchi & Z. Knežević: *New Definition of Discovery for Solar System Objects*, submitted for publication, 2006. (To be discussed in a joint session of Division III and Commision 20 at the IAU General Assembly in Prague, August 2006).

A. Milani & G.F. Gronchi: *The Theory of Orbit Determination*, Cambridge University Press, 2008. (This you cannot yet read: thank you for helping me in writing this textbook, for which we already have a contract).