

Multiple Solutions in Preliminary Orbit Determination from Three Observations

Giovanni Federico Gronchi
Dipartimento di Matematica, Università di Pisa
Largo B. Pontecorvo, 5, Pisa, Italy
gronchi@dm.unipi.it

Received: 16/9/2008 / Accepted: 24/2/2009

Abstract

Charlier's theory (1910) provides a geometric interpretation of the occurrence of multiple solutions in Laplace's method of preliminary orbit determination, assuming geocentric observations. We introduce a generalization of this theory allowing to take into account topocentric observations, that is observations made from the surface of the rotating Earth. The generalized theory works for both Laplace's and Gauss' methods. We also provide a geometric definition of a curve that generalizes Charlier's limiting curve, separating regions with a different number of solutions. The results are generically different from Charlier's: they may change according to the value of a parameter that depends on the observations.

1 Introduction

The orbit of a celestial body whose motion is dominated by the Sun can be computed from its astrometric observations by differential corrections of an initial orbit; this is an iterative method to obtain the minimum value of a target function [1]. Preliminary orbit determination gives the initial orbit to start the iterations.

Laplace [8] and Gauss [6] developed two different methods to search for a preliminary orbit from three observations of a celestial body in the sky¹, that represent the minimum amount of data necessary to completely determine a six parameter orbit.

In Laplace's method the observations are regarded as if they were made from the center of the Earth; the two-body dynamics then leads to equation (32), relating the geocentric distance of the body to its heliocentric distance at a fixed time (see Appendix 7.1). This equation, together with the simple geometric relation (33) allows to write a polynomial equation of degree eight

¹Laplace's method can easily take into account more than three observations.

(34) for the heliocentric distance of the celestial body. Actually more than one solution of (34) may exist, and this can result in multiple solutions of the orbit determination problem.

In 1910 Charlier [3], [4] gave a geometric interpretation of the occurrence of multiple solutions for a preliminary orbit with Laplace's method. He realized that (neglecting the errors in the measurements and in the model) this depends only on the position of the celestial body in a reference plane defined by the Sun, the Earth and the body at a given time, and he was able to divide this plane into four connected components by two algebraic curves, separating regions with a unique solution from regions with two solutions.

Taking into account topocentric observations, i.e. observations made from a point on the surface of the rotating Earth, is quite natural in Gauss' method (see Appendix 7.2), and also Laplace's method can be modified to consider this effect (see [9]). In both cases from the two-body dynamics we obtain an equation like (2) (see for example (47)), that has the same algebraic structure as (32) but depends on the additional parameter γ , and reduces to equation (32) for $\gamma = 1$. Thus for a generic value of γ Charlier's theory cannot be applied. This work is devoted to a generalization of Charlier's theory to give a geometric interpretation of multiple solutions also in the more realistic case of topocentric observations.

2 The intersection problem

Assume that we have three observations of a celestial body whose motion is dominated by the gravitational attraction of the Sun. Each observation consists of two angular positions (α_i, δ_i) at different subsequent times $t_i, i = 1, 2, 3$; for example α_i can be the right ascension and δ_i the declination of the body at time t_i .

Let \mathbf{q}_i be the heliocentric position vectors of the observer and let $\mathbf{r}_i, \boldsymbol{\rho}_i$ be respectively the heliocentric and topocentric position vectors of the observed body at time t_i . Moreover let $r_i = \|\mathbf{r}_i\|$, $\rho_i = \|\boldsymbol{\rho}_i\|$, $q_i = \|\mathbf{q}_i\|$ and define $\epsilon_i \in [0, \pi]$ as the co-elongations of the observed body, so that $\cos \epsilon_i = (\mathbf{q}_i \cdot \boldsymbol{\rho}_i) / (q_i \rho_i)$.

We consider an average time of observation \bar{t} : in Gauss' method $\bar{t} = t_2$, while usually in Laplace's method $\bar{t} = (t_1 + t_2 + t_3)/3$. We write r, ρ, q, ϵ for the values of the quantities corresponding to $r_i, \rho_i, q_i, \epsilon_i$ at time \bar{t} . Note that q and $\mathbf{q}_i, i = 1, 2, 3$, can be obtained from planetary ephemerides, ϵ can be computed by interpolating the values of ϵ_i (computed in turn from $\alpha_i, \delta_i, \mathbf{q}_i$), while r, ρ are unknown because r_i, ρ_i are also.

Actually the results presented in this paper do not depend on the value of q . A different value of q corresponds to different units of length, therefore we could set $q = 1$ without loss of generality. Nevertheless we prefer to leave q in all the formulae, since different units may be used in the applications of the theory to specific problems.

The geometry of the three bodies immediately gives the relation

$$r^2 = q^2 + \rho^2 + 2q\rho \cos \epsilon \quad \textit{geometric equation} . \quad (1)$$

Using the two-body dynamics we can deduce the following relation:

$$\mathcal{C} \frac{\rho}{q} = \gamma - \frac{q^3}{r^3} \quad \text{dynamical equation,} \quad (2)$$

where $\gamma, \mathcal{C} \in \mathbb{R}$ are given constants (see Appendix 7.2 for the computation of γ, \mathcal{C} in Gauss' method).

Equations (1) and (2) define surfaces of revolution around the axis $\mathbb{R}\mathbf{q}$ passing through the center of the Sun and the observer. If the center of the Sun, the observer and the observed body are not collinear at time \bar{t} , the observation line (also called line of sight: a half-line from the observer's position defined by ϵ) and the axis $\mathbb{R}\mathbf{q}$ define univocally a *reference plane*, that we shall use to study the intersection of these surfaces². Otherwise we can choose as reference plane any of the planes containing the observer's position and the center of the Sun.

We introduce the *intersection problem*

$$\begin{cases} D(r, \rho) = (q\gamma - \mathcal{C}\rho)r^3 - q^4 = 0 \\ G(r, \rho) = r^2 - q^2 - \rho^2 - 2q\rho \cos \epsilon = 0 \\ r, \rho > 0 \end{cases} \quad , \quad (3)$$

that is, given $(\gamma, \mathcal{C}, \epsilon) \in \mathbb{R}^2 \times [0, \pi]$ we search for pairs (r, ρ) of strictly positive real numbers, solutions of (2) and (1). For given values of $(\gamma, \mathcal{C}, \epsilon)$ the solutions of (3) correspond to the intersections of the observation line with the planar algebraic curve defined by (2) in the reference plane.

We can perform elimination of the variable ρ by means of resultant theory (see [5]), thus from (3) we obtain the *reduced problem*

$$\begin{cases} P(r) = \mathbf{res}(D, G, \rho) = 0 \\ r > 0 \end{cases} \quad (4)$$

where $\mathbf{res}(D, G, \rho)$ stands for the resultant of the polynomials $D(r, \rho)$ and $G(r, \rho)$ with respect to the variable ρ . The resultant computation gives

$$P(r) = \mathcal{C}^2 r^8 - q^2(\mathcal{C}^2 + 2\mathcal{C}\gamma \cos \epsilon + \gamma^2)r^6 + 2q^5(\mathcal{C} \cos \epsilon + \gamma)r^3 - q^8. \quad (5)$$

The reduced formulation (4) is suitable to obtain an upper bound for the maximum number of solutions, in fact $P(r)$ has only four monomials, thus by Descartes' sign rule there are at most three positive roots of $P(r)$, counted with multiplicity.

Note that, if $r = \bar{r}$ is a component of a solution of (3), then from (2) we obtain a unique value $\bar{\rho}$ for the other component and, conversely, from a value $\bar{\rho}$ of ρ we obtain a unique \bar{r} . In particular there are no more than three values of ρ that are components of the solutions of (3).

Definition 1. We define as spurious solution of (4) a positive root \bar{r} of $P(r)$ that is not a component of a solution $(\bar{r}, \bar{\rho})$ of (3) for any $\bar{\rho} > 0$.

²the observed body is here regarded as a single point.

Thus a positive root of $P(r)$ is a spurious solution of (4) iff it gives a non-positive ρ through the dynamical equation (2).

In the following we shall deal with possible multiple solutions of the intersection problem: from each solution of (3) a full set of orbital elements can be determined, in fact the knowledge of the topocentric distance ρ allows to compute the corresponding value of $\dot{\rho}$.

3 Charlier's theory

Charlier's theory describes the occurrence of multiple solutions in the problem defined by equations (32), (33), with geocentric observations. Nevertheless, if we interpret ρ and q as the geocentric distance of the observed body and the heliocentric distance of the center of the Earth, then equation (2) with $\gamma = 1$ corresponds to (32) and equation (1) corresponds to (33). Therefore we shall discuss Charlier's theory by studying the multiple solutions of (3) with $\gamma = 1$, and we shall see that in this case the solutions of (3) can be at most two. The discussion presented in this section is based on [10].

Charlier has been the first to realize that *'the condition for the appearance of another solution simply depends on the position of the observed body'*. We shall explain later the meaning of this statement, but we stress that it assumes that the two-body model for the orbit of the observed body is exact and neglects the observation and interpolation errors in the parameters \mathcal{C}, ϵ . The previous hypotheses imply the following assumption:

$$\text{the parameters } \mathcal{C}, \epsilon \text{ are such that the corresponding intersection problem with } \gamma = 1 \text{ admits at least one solution.} \quad (6)$$

In the real astronomical applications this assumption may not be fulfilled and the intersection problem may have no solution; a reason for that is just the presence of errors in the observations, that affect mostly the computation of \mathcal{C} . However we observe that condition (6) may hold also taking into account these errors, therefore it is more interesting for the applications.

For each choice of \mathcal{C}, ϵ the polynomial $P(r)$ in (4) has three changes of sign in the sequence of its coefficients, in fact the coefficient of r^3 (the only that needs a comment) is positive because from (2) and (1) we have

$$\mathcal{C} \cos \epsilon + 1 = \frac{1}{2\rho^2 r^3} [(r^3 - q^3)(r^2 - q^2) + \rho^2(r^3 + q^3)] > 0,$$

thus the positive roots of $P(r)$ can indeed be three.

As $P(q) = 0$, there is always the solution corresponding to the center of the Earth, in fact, from the dynamical equation, $r = q$ corresponds to $\rho = 0$. This solution must be discarded for physical reasons. Using (6), Descartes' sign rule and the relations

$$P(0) = -q^8 < 0; \quad \lim_{r \rightarrow +\infty} P(r) = +\infty,$$

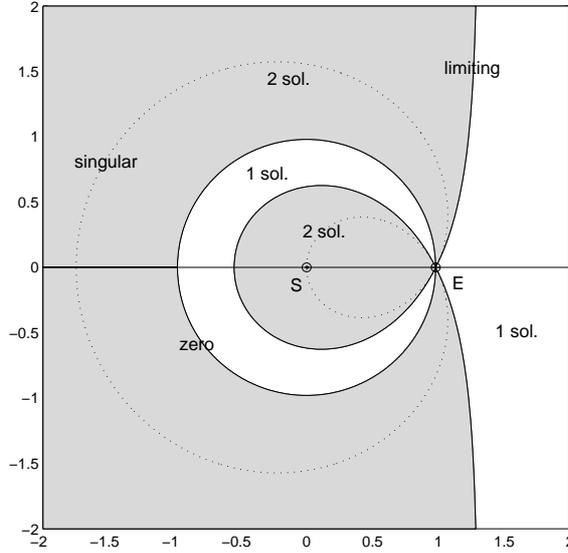


Figure 1: The limiting curve and the zero circle divide the reference plane into four connected regions, two with a unique solution of (3) and two with two solutions (shaded in this figure). The singular curve (dotted) divides the regions with two solutions into two parts, with one solution each. The Sun and the Earth are labeled with S and E respectively. We use heliocentric rectangular coordinates, and Astronomical Units (AU) for both axes.

we conclude that there are always three positive roots of $P(r)$, counted with multiplicity. By (6) at least one of the other two positive roots r_1, r_2 is not spurious: if either r_1 or r_2 is spurious the solution of (3) is unique, otherwise we have two non-spurious solutions.

To detect the cases with two solutions we write $P(r) = (r - q)P_1(r)$, with

$$P_1(r) = \mathcal{C}^2 r^6 (r + q) + (r^2 + qr + q^2) [q^5 - (2\mathcal{C} \cos \epsilon + 1)q^2 r^3] ,$$

so that

$$P_1(q) = 2q^7 \mathcal{C} (\mathcal{C} - 3 \cos \epsilon) .$$

From the relations

$$P_1(0) = q^7 > 0; \quad \lim_{r \rightarrow +\infty} P_1(r) = +\infty$$

it follows that if $P_1(q) < 0$ then $r_1 < q < r_2$, while if $P_1(q) > 0$ then either $r_1, r_2 < q$ or $r_1, r_2 > q$. In the first case the dynamical equation gives us two values ρ_1, ρ_2 with $\rho_1 \rho_2 < 0$, so that one root of $P_1(r)$ is spurious. In the second case both roots give rise to meaningful solutions of (3). If $P_1(q) = 0$ there is only one non-spurious root of $P(r)$.

We introduce two algebraic curves³ in geocentric polar coordinates (ρ, ψ) ,

³in bipolar coordinates (r, ρ) they are defined by polynomials

with $\psi = 0$ towards the opposition direction (corresponding to $\epsilon = 0$), by

$$\begin{aligned} \mathfrak{C}^{(1)}(\rho, \psi) &= 0 && \text{zero circle,} \\ \mathfrak{C}^{(1)}(\rho, \psi) - 3 \cos \psi &= 0 && \text{limiting curve,} \end{aligned}$$

where

$$\mathfrak{C}^{(1)}(\rho, \psi) = \frac{q}{\rho} \left[1 - \frac{q^3}{r^3} \right], \quad r = \sqrt{\rho^2 + q^2 + 2q\rho \cos \psi}.$$

The limiting curve has a loop inside the zero circle and two unlimited branches with $r > q$. By the previous discussion the limiting curve and the zero circle divide the reference plane, containing the center of the Sun, the observer and the observed body at time \bar{t} , into four connected components (see Figure 1), separating regions with a different number of solutions of the orbit determination problem. More precisely, given $(\rho, \psi) \in \mathbb{R}^+ \times [0, \pi] \setminus \{(q, \pi)\}$, describing the position of a celestial body in the reference plane at time \bar{t} , equations (1), (2) with $\epsilon = \psi$, $\gamma = 1$ define a value \mathcal{C} such that the intersection problem defined by $(\gamma, \mathcal{C}, \epsilon)$ has the solution $(r, \rho) = (\sqrt{\rho^2 + q^2 + 2q\rho \cos \epsilon}, \rho)$ and, if the body is situated in a region with two solutions, we can find a second solution in the same region of the first.

Using heliocentric polar coordinates (r, ϕ) , with $\rho^2 = r^2 + q^2 - 2qr \cos \phi$, the limiting curve is given by

$$4 - 3 \frac{r}{q} \cos \phi = \frac{q^3}{r^3} \quad (7)$$

and, in heliocentric rectangular coordinates $(x, y) = (r \cos \phi, r \sin \phi)$, by

$$4 - 3 \frac{x}{q} = \frac{q^3}{(x^2 + y^2)^{3/2}}.$$

Figure 1 shows in particular that, if the celestial body has been observed close to the opposition direction, then the solution of Laplace's method of preliminary orbit determination is unique.

In [4] Charlier introduced the *singular curve*, dividing the regions with two solutions into regions containing only one solution each. The definition of the singular curve will be introduced in Section 4.2 for a generic value of γ .

4 Generalized Charlier's theory

In this section we consider the intersection problem (3) for a generic $\gamma \in \mathbb{R}$.

Given a value of γ and the vector (ρ, ψ) , representing a point of the reference plane in topocentric polar coordinates, equations (1), (2) with $\epsilon = \psi$ define a value of \mathcal{C} such that the intersection problem defined by $(\gamma, \mathcal{C}, \epsilon)$ has the solution $(r, \rho) = (\sqrt{\rho^2 + q^2 + 2q\rho \cos \epsilon}, \rho)$. Therefore in the following we shall speak about the *intersection problem corresponding, or related, to a fixed $\gamma \in \mathbb{R}$ and to a point of the reference plane*.

We introduce the following assumption, that generalizes (6):

$$\text{the parameters } \gamma, \mathcal{C}, \epsilon \text{ are such that the corresponding intersection problem admits at least one solution.} \quad (8)$$

Note that $r = q$ generically is not a root of $P(r)$, in fact

$$P(q) = q^8(1 - \gamma)(2\mathcal{C} \cos \epsilon - (1 - \gamma)),$$

thus we cannot follow the same steps of Section 3 to define the limiting curve. From the dynamical equation we define the function

$$\mathcal{C}^{(\gamma)}(x, y) = \frac{q}{\rho} \left[\gamma - \frac{q^3}{r^3} \right], \quad (9)$$

where

$$\rho = \sqrt{(q-x)^2 + y^2}, \quad r = \sqrt{x^2 + y^2}. \quad (10)$$

If $\gamma > 0$ we can also define the zero circle as $\{(x, y) : r = r_0\}$, with $r_0 = q/\sqrt[3]{\gamma}$. The points of the zero circle fulfill (2) with $\mathcal{C} = 0$.

4.1 THE TOPOLOGY OF THE LEVEL CURVES OF $\mathcal{C}^{(\gamma)}(x, y)$

First we note that for each $\gamma \in \mathbb{R}$

$$\lim_{\|(x,y)\| \rightarrow +\infty} \mathcal{C}^{(\gamma)}(x, y) = 0, \quad \lim_{(x,y) \rightarrow (0,0)} \mathcal{C}^{(\gamma)}(x, y) = -\infty,$$

and

$$\lim_{(x,y) \rightarrow (q,0)} \mathcal{C}^{(\gamma)}(x, y) \begin{cases} = -\infty & \text{for } \gamma < 1 \\ \text{does not exist} & \text{for } \gamma = 1 \\ = +\infty & \text{for } \gamma > 1 \end{cases}.$$

To understand the topological changes in the level curves of $\mathcal{C}^{(\gamma)}(x, y)$ we need the following

Lemma 1. *The stationary points of $\mathcal{C}^{(\gamma)}(x, y)$ lie on the x-axis and depend on γ as follows:*

- (i) for $\gamma \leq 0$ there is only one saddle point, with $x \in (0, \frac{3}{4}q]$;
- (ii) for $0 < \gamma < 1$ there are three points: one saddle point inside the zero circle, one saddle and one maximum point outside.
- (iii) for $\gamma \geq 1$ there is a unique saddle point with $x < -r_0 = -q/\sqrt[3]{\gamma}$.

Proof. We look for the solutions of equations

$$\begin{aligned} \mathcal{C}_x^{(\gamma)}(x, y) &= \frac{qx}{r^5 \rho^3} [3q^3 \rho^2 - r^2(\gamma r^3 - q^3)] + \frac{q^2}{r^3 \rho^3} (\gamma r^3 - q^3) = 0, \\ \mathcal{C}_y^{(\gamma)}(x, y) &= \frac{qy}{r^5 \rho^3} [3q^3 \rho^2 - r^2(\gamma r^3 - q^3)] = 0, \end{aligned}$$

(the subscripts denote the partial derivatives) with ρ, r given by (10); they are the pairs (x, y) with $y = 0$ and x such that

$$(\gamma|x|^3 - q^3)x = 3q^3(x - q) . \quad (11)$$

It is easy to prove that for $\gamma \leq 0$ there is no $x < 0$ solution of (11). For $\gamma > 0$ we define $g_-(x) = \gamma x^4 + q^3(4x - 3q)$. The unique stationary point of $g_-(x)$ is $x = -r_0$, it is a minimum point and $g_-(-r_0) = -3q^4(1 + \sqrt[3]{\gamma})/\sqrt[3]{\gamma} < 0$. From the limits $\lim_{x \rightarrow 0^-} g_-(x) = -3q^4 < 0$ and $\lim_{x \rightarrow -\infty} g_-(x) = +\infty$ we conclude that for $\gamma > 0$ there exists a unique negative solution x_1 of (11), with $x_1 < -r_0$. We can show that $(x_1, 0)$ is a saddle point by examining the behavior of $\mathcal{C}^{(\gamma)}(x, y)$ along the two lines passing through $(x_1, 0)$, parallel to the coordinate axes.

Next we search for the positive solutions of (11). For $\gamma \leq 0$ we can easily prove that there is a unique solution x_1 of (11) with $0 < x_1 \leq \frac{3}{4}q$. For $\gamma > 0$ let us define $g_+(x) = -\gamma x^4 + q^3(4x - 3q)$. The derivative $g'_+(x) = -4(\gamma x^3 - q^3)$ shows that $x = r_0$ is the unique stationary point of $g_+(x)$; it is a maximum point and $g_+(r_0) = 3q^4(1 - \sqrt[3]{\gamma})/\sqrt[3]{\gamma}$. From the limits $\lim_{x \rightarrow 0^+} g_+(x) = -3q^4 < 0$ and $\lim_{x \rightarrow +\infty} g_+(x) = -\infty$ we conclude that if $0 < \gamma < 1$ there are two solutions x_2, x_3 of (11), with $0 < x_2 < r_0 < x_3$; if $\gamma = 1$ we obtain $x = q$, that must be discarded, while for $\gamma > 1$ there is no solution.

An argument similar to the one above shows that the solution of (11) in the interval $(0, r_0)$, present if $\gamma < 1$, gives a saddle point. From $\lim_{\|(x,y)\| \rightarrow +\infty} \mathcal{C}^{(\gamma)}(x, y) = 0$ and from Weierstrass' theorem we have that the stationary point $(x_3, 0)$, present if $0 < \gamma < 1$, must be a maximum point.

□

The qualitative behavior of the level curves of $\mathcal{C}^{(\gamma)}(x, y)$ is sketched in Figure 2 for all the significantly different cases, i.e. $\gamma \leq 0$, $0 < \gamma < 1$, $\gamma = 1$ and $\gamma > 1$. We shall need the following

Proposition 1. (solutions at opposition) *Observing exactly in the opposition direction ($\epsilon = 0$), the number of solutions of (3) is the following:*

- (i) if $\gamma \leq 0$ there is one solution for $\mathcal{C} < 0$, no solution for $\mathcal{C} \geq 0$;
- (ii) if $0 < \gamma < 1$ there is one solution for $\mathcal{C} \leq 0$, two distinct solutions for $0 < \mathcal{C} < \mathcal{C}_\gamma$, one double solution for $\mathcal{C} = \mathcal{C}_\gamma$, no solution for $\mathcal{C} > \mathcal{C}_\gamma$, where \mathcal{C}_γ is the stationary value of $\mathcal{C}^{(\gamma)}(x, y)$ corresponding to the maximum point, i.e. the solution smaller than 3 of $4(\mathcal{C}/3)^{3/4} - \mathcal{C} = \gamma$;
- (iii) if $\gamma = 1$ there is one solution for $0 < \mathcal{C} < 3$, no solution for $\mathcal{C} \leq 0$ or $\mathcal{C} \geq 3$;
- (iv) if $\gamma > 1$ there is no solution for $\mathcal{C} \leq 0$, one solution for $\mathcal{C} > 0$.

Proof. Let $f_1(x) = \frac{\mathcal{C}}{q}(x - q)$ and $f_2(x) = \gamma - \frac{q^3}{x^3}$; we search for $x > q$ solutions of equation $f_1(x) = f_2(x)$. For $\mathcal{C} < 0$ the computation of the limits of f_1, f_2 as $x \rightarrow 0^+$ and $x \rightarrow +\infty$ together with the monotonicity of these functions gives

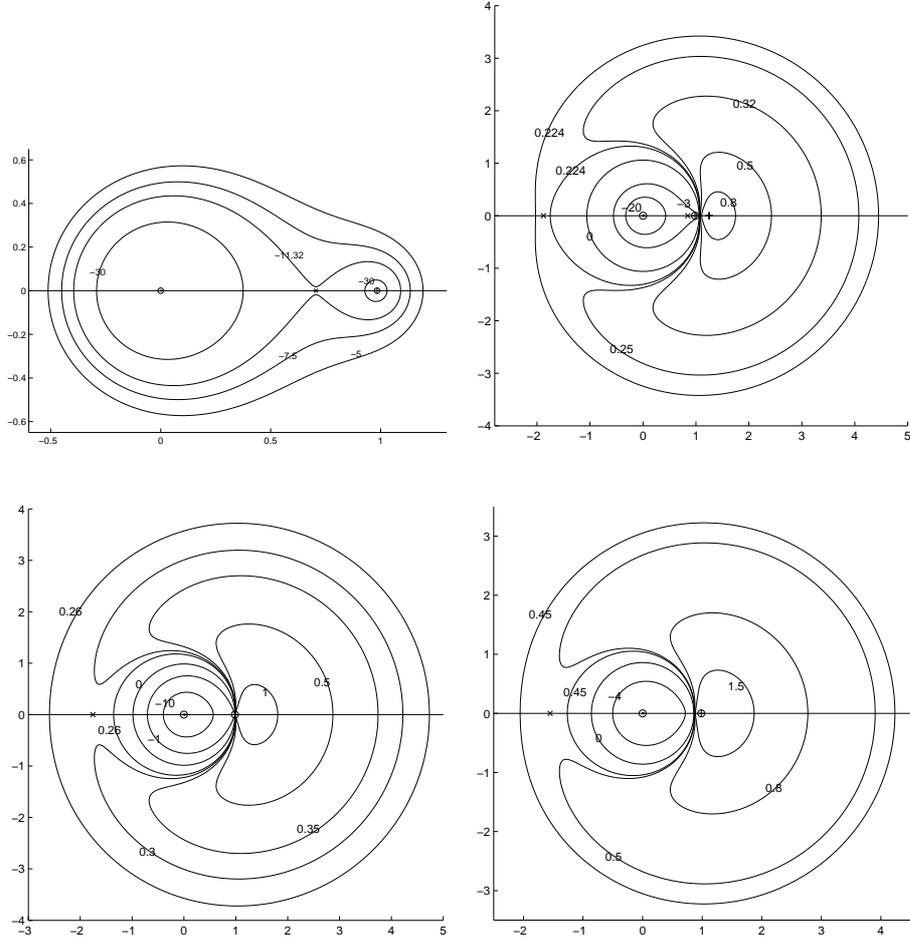


Figure 2: The level curves of $\mathcal{C}^{(\gamma)}(x, y)$. The Sun and the Earth are denoted with \odot and \oplus respectively. The saddle points are marked with x and the maximum point (present only for $0 < \gamma < 1$) with $+$. Top left: $\gamma = -0.5$. Top right: $\gamma = 0.8$. Bottom left: $\gamma = 1$. Bottom right: $\gamma = 1.5$.

immediately a unique positive solution \bar{x} of $f_1(x) = f_2(x)$. From the sign of $f_1(q) - f_2(q)$ we obtain $\bar{x} > q$ for $\gamma < 1$ and $\bar{x} \leq q$ for $\gamma \geq 1$, thus the latter must be discarded.

For $\mathcal{C} = 0$ we find the unique solution $\bar{x} = r_0 = q/\sqrt[3]{\gamma}$, and $r_0 > q$ iff $0 < \gamma < 1$. Let us consider the case $\mathcal{C} > 0$. If $\gamma \leq 0$ we easily verify that there is no solution with $x > q$. If $0 < \gamma < 1$ then $f_1'(\tilde{x}) - f_2'(\tilde{x}) = 0$ iff $\tilde{x} = q\sqrt[4]{3/\mathcal{C}}$; moreover $f_1(\tilde{x}) - f_2(\tilde{x}) = 4(\mathcal{C}/3)^{3/4} - \mathcal{C} - \gamma$. The only stationary point of $g(\mathcal{C}) = 4(\mathcal{C}/3)^{3/4} - \mathcal{C}$ is $\mathcal{C} = 3$, it is a maximum point and $g(3) = 1$; therefore,

using relations $g(0) = 0$ and $\lim_{\mathcal{C} \rightarrow +\infty} g(\mathcal{C}) = -\infty$, we can assert that there are two positive solutions $0 < \mathcal{C}_{\gamma,1} < 3 < \mathcal{C}_{\gamma,2}$ of equation $g(\mathcal{C}) = \gamma$. For $0 < \mathcal{C} < \mathcal{C}_{\gamma,1}$ there are two distinct solutions of $f_1 = f_2$, both greater than q ; for $\mathcal{C} = \mathcal{C}_{\gamma,1}$ we have one double solution; for $\mathcal{C} > \mathcal{C}_{\gamma,2}$ we also have two positive solutions of $f_1 = f_2$, but they are both smaller than q . We set $\mathcal{C}_\gamma = \mathcal{C}_{\gamma,1}$ and note that $\lim_{\gamma \rightarrow 0^+} \mathcal{C}_\gamma = 0$, $\lim_{\gamma \rightarrow 1^-} \mathcal{C}_\gamma = 3$. If $\gamma = 1$ the discussion is similar to the case $0 < \gamma < 1$, but there is always the solution $x = q$ that must be discarded. Finally, if $\gamma > 1$, by using the computation of the limits of f_1, f_2 as $x \rightarrow 0^+$ and $x \rightarrow +\infty$, the monotonicity and convexity properties of f_1, f_2 and computing the sign of $f_1(q) - f_2(q)$, we conclude that there are two solutions x_1, x_2 of $f_1 = f_2$, with $x_1 < q < x_2$, so that x_1 must be discarded. \square

Proposition 2. (solutions at conjunction) *Observing exactly in the conjunction direction ($\epsilon = \pi$), the number of solutions of (3) is the following:*

- (i) *if $\gamma \leq 0$ there are three distinct solutions for $\mathcal{C} < \mathcal{C}_\gamma$, one simple and one double solution for $\mathcal{C} = \mathcal{C}_\gamma$, one solution for $\mathcal{C}_\gamma < \mathcal{C} < 0$, no solution for $\mathcal{C} \geq 0$, where \mathcal{C}_γ corresponds to the unique stationary value of $\mathcal{C}^{(\gamma)}(x, y)$, i.e. the solution of $4(-\mathcal{C}/3)^{3/4} + \mathcal{C} = \gamma$;*
- (ii) *if $0 < \gamma < 1$ there are three distinct solutions for $\mathcal{C} < \mathcal{C}_\gamma^{(1)}$, one simple and one double solution for $\mathcal{C} = \mathcal{C}_\gamma^{(1)}$, one solution for $\mathcal{C}_\gamma^{(1)} < \mathcal{C} \leq 0$, two distinct solutions for $0 < \mathcal{C} < \mathcal{C}_\gamma^{(2)}$, one double solution for $\mathcal{C} = \mathcal{C}_\gamma^{(2)}$, no solution for $\mathcal{C} > \mathcal{C}_\gamma^{(2)}$, where $\mathcal{C}_\gamma^{(1)}$ is the stationary value of $\mathcal{C}^{(\gamma)}(x, y)$ corresponding to the saddle point inside the zero circle, i.e. the solution smaller than -3 of $4(-\mathcal{C}/3)^{3/4} + \mathcal{C} = \gamma$, and $\mathcal{C}_\gamma^{(2)}$ is the stationary value of $\mathcal{C}^{(\gamma)}(x, y)$ corresponding to the saddle point outside the zero circle, i.e. the solution of $4(\mathcal{C}/3)^{3/4} + \mathcal{C} = \gamma$;*
- (iii) *if $\gamma = 1$ there are two distinct solutions for $\mathcal{C} < -3$, one solution for $-3 \leq \mathcal{C} \leq 0$, two distinct solutions for $0 < \mathcal{C} < \mathcal{C}_1$, one double solution for $\mathcal{C} = \mathcal{C}_1$, no solution for $\mathcal{C} > \mathcal{C}_1$, where \mathcal{C}_1 is the unique stationary value of $\mathcal{C}^{(\gamma)}(x, y)$, i.e. the solution of $4(\mathcal{C}/3)^{3/4} + \mathcal{C} = 1$;*
- (iv) *if $\gamma > 1$ there are two distinct solutions for $\mathcal{C} \leq 0$, three distinct solutions for $0 < \mathcal{C} < \mathcal{C}_\gamma$, one simple and one double solution for $\mathcal{C} = \mathcal{C}_\gamma$, one solution for $\mathcal{C} > \mathcal{C}_\gamma$, where \mathcal{C}_γ is the unique stationary value of $\mathcal{C}^{(\gamma)}(x, y)$, i.e. the solution of $4(\mathcal{C}/3)^{3/4} + \mathcal{C} = \gamma$.*

Proof. The proof is similar to that of Proposition 1: we define $f_1(x) = \frac{\mathcal{C}}{q}(q - x)$, $f_2(x) = \gamma - \frac{q^3}{|x|^3}$ and we look for $x < q$ solutions of equation $f_1(x) = f_2(x)$. \square

Remark: The knowledge of the solutions at conjunction is important for a global view of the problem, but of course has no application to Astronomy because we cannot make observations in the direction of the Sun.

4.2 THE SINGULAR CURVE

The function $\mathfrak{C}^{(\gamma)}(x, y)$ in topocentric polar coordinates (ρ, ψ) is given by

$$\mathfrak{C}^{(\gamma)}(\rho, \psi) = \frac{q}{\rho} \left[\gamma - \frac{q^3}{r^3} \right], \quad \text{where } r = \sqrt{\rho^2 + q^2 + 2q\rho \cos \psi} .$$

As the Jacobian of the transformation $(\rho, \psi) \mapsto (x, y) = (q + \rho \cos \psi, \rho \sin \psi)$ has determinant equal to ρ , the stationary points of $\mathfrak{C}^{(\gamma)}(\rho, \psi)$ just correspond to the ones of $\mathfrak{C}^{(\gamma)}(x, y)$, described in Lemma 1.

For a given $\gamma \in \mathbb{R}$ we define

$$F(\mathcal{C}, \rho, \psi) = \mathcal{C} \frac{\rho}{q} - \gamma + \frac{q^3}{r^3}$$

with $r = \sqrt{\rho^2 + q^2 + 2q\rho \cos \psi}$. The derivative of $F(\mathcal{C}, \rho, \psi)$ with respect to ρ is

$$F_\rho(\mathcal{C}, \rho, \psi) = \frac{\mathcal{C}}{q} - 3 \frac{q^3}{r^5} (\rho + q \cos \psi) .$$

Let us consider the equations

$$F(\mathcal{C}, \rho, \psi) = F_\rho(\mathcal{C}, \rho, \psi) = 0 . \quad (12)$$

By the implicit function theorem, for each non-stationary value \mathcal{C} of $\mathfrak{C}^{(\gamma)}(\rho, \psi)$, the tangency points between the level manifolds of $\mathfrak{C}^{(\gamma)}(\rho, \psi)$ and the observation lines fulfill equations (12).

We can eliminate the dependence on \mathcal{C} in (12) by considering the difference

$$F(\mathcal{C}, \rho, \psi) - \rho F_\rho(\mathcal{C}, \rho, \psi) = -\gamma + \frac{q^3}{r^3} + 3q^3 \frac{\rho}{r^5} (\rho + q \cos \psi) .$$

The function $r^5(F - \rho F_\rho)$ in heliocentric rectangular coordinates becomes

$$\mathcal{G}(x, y) = -\gamma r^5 + q^3(4r^2 - 3qx), \quad r = \sqrt{x^2 + y^2} . \quad (13)$$

Definition 2. We define the singular curve as the set

$$\mathcal{S} = \{(x, y) : \mathcal{G}(x, y) = 0\} .$$

Note that \mathcal{S} is constituted by all the points whose polar coordinates fulfill (12) plus $(x, y) = (0, 0)$. To describe the shape of the singular curve for different values of γ we need the following results:

Lemma 2. If $\gamma \neq 1$ the singular curve \mathcal{S} meets the x -axis in $(x, y) = (0, 0)$ and in the stationary points of $\mathfrak{C}^{(\gamma)}(x, y)$. If $\gamma = 1$ then \mathcal{S} meets the x -axis also in $(x, y) = (q, 0)$.

In particular these points depend on γ as follows:

- (i) for $\gamma \leq 0$ there are two such points, with $x \in [0, \frac{3}{4}q]$;

- (ii) for $0 < \gamma < 1$ the points are four, two inside the zero circle and two outside;
- (iii) for $\gamma = 1$ the points are three, one outside the zero circle, one inside and one on the circle itself;
- (iv) for $\gamma > 1$ there are two points, one inside the zero circle, the other outside.

Proof. We search for the solutions of

$$\gamma|x|^5 = q^3x(4x - 3q). \quad (14)$$

Note that $x = 0$ is a solution of (14) for each $\gamma \in \mathbb{R}$. Dividing (2) by x we obtain equation (11), whose solutions have been discussed in Lemma 1. \square

Lemma 3. *The points such that*

$$\mathcal{G}(x, y) = \mathcal{G}_y(x, y) = 0 \quad (15)$$

depend on γ as follows:

- (i) for $\gamma \leq 512/625$ they are just the points where $\mathcal{G}(x, 0) = 0$;
- (ii) for $\gamma > 512/625$ besides the points with $\mathcal{G}(x, 0) = 0$ there are two points outside the zero circle fulfilling relations (15), with $y \neq 0$ and positive x coordinate.

Proof. From the expression of the derivative $\mathcal{G}_y(x, y) = y(8q^3 - 5\gamma r^3)$ we find that either $y = 0$ or, if $\gamma > 0$, $r = 2r_0/\sqrt[3]{5} > r_0$, where $r_0 = q/\sqrt[3]{\gamma}$ is the radius of the zero circle.

For $y = 0$ we obtain the points of \mathcal{S} intersecting the x -axis as solutions of (15). If $\gamma > 0$, by inserting $r = 2r_0/\sqrt[3]{5}$ in $\mathcal{G}(x, y) = 0$, we obtain $x = \frac{16q}{5(5\gamma)^{2/3}}$, and $y^2 = r^2 - x^2 = \frac{4q^2}{(5\gamma)^{2/3}} \left[1 - \frac{64}{25(5\gamma)^{2/3}} \right]$ is positive iff $\gamma > 512/625$. \square

We can now prove the following

Proposition 3. (shape of the singular curve) *If $\gamma \neq 1$, the singular curve \mathcal{S} is a compact manifold without boundary:*

- (i) if $\gamma \leq 0$ it has a unique connected convex component;
- (ii) if $0 < \gamma < 1$ it has two connected components, one inside the zero circle, the other outside;
- (iii) if $\gamma > 1$ it has a unique connected component, not convex, intersecting the zero circle only in two points.

If $\gamma = 1$ then \mathcal{S} is not a manifold: it has a self-intersection point at the observer's position $(x, y) = (q, 0)$.

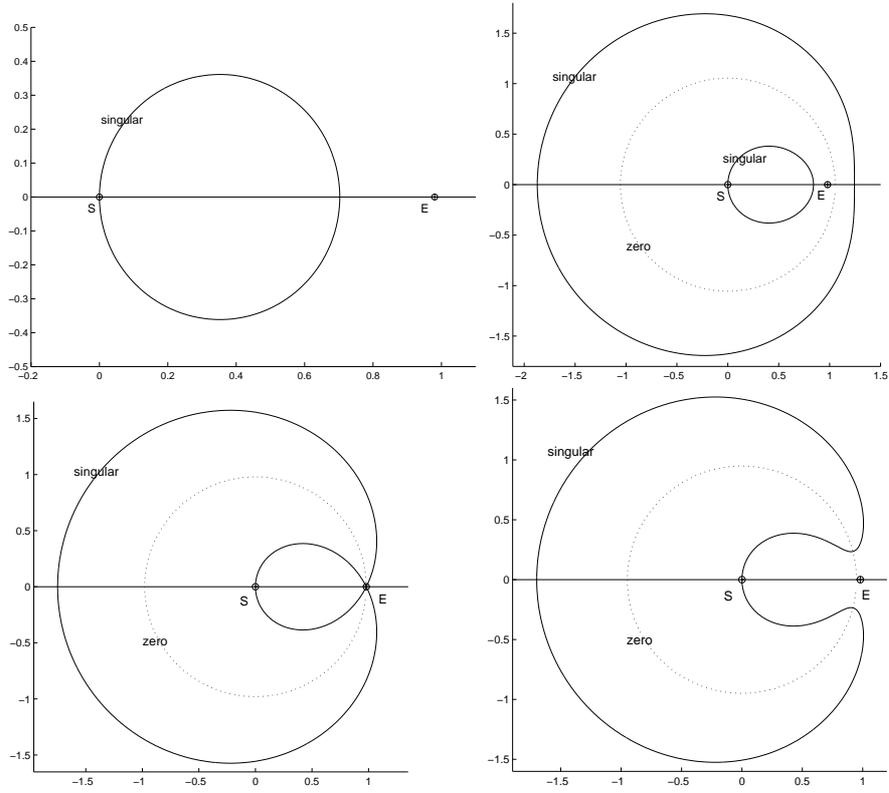


Figure 3: Singular curve (continuous line) and zero circle (dotted). Top left: $\gamma = -0.5$. Top right: $\gamma = 0.8$. Bottom left: $\gamma = 1$. Bottom right: $\gamma = 1.1$. Note that the zero circle does not exist for $\gamma \leq 0$.

Proof. Let us consider the solutions of the system

$$\mathcal{G}(x, y) = \mathcal{G}_x(x, y) = \mathcal{G}_y(x, y) = 0, \quad (16)$$

where

$$\mathcal{G}_x(x, y) = x(8q^3 - 5\gamma r^3) - 3q^4, \quad \mathcal{G}_y(x, y) = y(8q^3 - 5\gamma r^3).$$

From $\mathcal{G}_x(x, y) = \mathcal{G}_y(x, y) = 0$ we find that $y = 0$ and

$$-\gamma|x|^5 + q^3x(4x - 3q) = 0. \quad (17)$$

We observe that $x = 0$ is not a solution of $\mathcal{G}_x(x, 0) = 0$. The equation $5\mathcal{G}(x, 0) - x\mathcal{G}_x(x, 0) = 0$ gives $x = q$, that is not a solution of $\mathcal{G}(x, 0) = 0$ except for $\gamma = 1$. By the previous discussion and by $\lim_{r \rightarrow +\infty} |\mathcal{G}(x, y)| = +\infty$ we can assert that for $\gamma \neq 1$ the singular curve is a compact manifold without boundary.

If $\gamma \leq 0$, from Lemmas 2, 3 and the implicit function theorem it follows that the portion of the singular curve \mathcal{S} with $y > 0$ is the graphic of a function $\sigma(x)$ of the variable x , with

$$\sigma''(x) = - \frac{\mathcal{G}_{xx}\mathcal{G}_y^2 - 2\mathcal{G}_x\mathcal{G}_y\mathcal{G}_{xy} + \mathcal{G}_{yy}\mathcal{G}_x^2}{\mathcal{G}_y^3} \Big|_{(x,y)=(x,\sigma(x))}.$$

From Lemmas 2 and 3 it also follows that for $\gamma \leq 0$ the singular curve has only one connected component, diffeomorphic to a circle. Moreover, for $\gamma \leq 0$,

$$\mathcal{G}_{xx}(x, y) = 8q^3 - 5\gamma r(4x^2 + y^2) > -5\gamma 4|xy|r > -15\gamma r|xy| = |\mathcal{G}_{xy}(x, y)|,$$

and similarly $\mathcal{G}_{yy} > |\mathcal{G}_{xy}|$. It follows that

$$\mathcal{G}_{xx}\mathcal{G}_y^2 - 2\mathcal{G}_x\mathcal{G}_y\mathcal{G}_{xy} + \mathcal{G}_{yy}\mathcal{G}_x^2 > |\mathcal{G}_{xy}|(|\mathcal{G}_x| - |\mathcal{G}_y|)^2 \geq 0.$$

We also have $\text{sign}(\mathcal{G}_y) = \text{sign}(y)$, hence $\frac{d^2\sigma}{dx^2} < 0$ for $y > 0$, that together with the reflection symmetry with respect to the x -axis yields convexity of \mathcal{S} and (i) is proved.

Let us prove (ii). If $0 < \gamma < 1$, using Lemmas 2, 3 and the fact that \mathcal{S} is a compact manifold without boundary, we conclude that it consists of two connected components, each diffeomorphic to a circle. Moreover for $0 < \gamma < 1$ the singular curve does not intersect the zero circle: using heliocentric polar coordinates (r, ϕ) , defined by $x = r \cos \phi, y = r \sin \phi$, and substituting $r = r_0 = q/\sqrt[3]{\gamma}$ in $\mathcal{G}(x, y) = 0$ we obtain $\cos \phi = 1/\sqrt[3]{\gamma} > 1$, that has no real solution. We conclude by Lemma 2 that one of the two components of \mathcal{S} lies inside the zero circle, the other outside.

Now assume $\gamma > 1$. From the results of Lemma 2 there are only two points such that $\mathcal{G}(x, 0) = 0$, with $x \leq 0$. Since in particular $\gamma > 512/625$, by Lemma 3 we obtain two points, solutions of $\mathcal{G}(x, y) = \mathcal{G}_y(x, y) = 0$, with $x > 0$, outside the x -axis and symmetric with respect to it. Thus for $\gamma > 1$ the singular curve \mathcal{S} (also diffeomorphic to a circle) can not be convex. Moreover there are only two points of \mathcal{S} intersecting the zero circle, with coordinates $(x, y) = \frac{q}{\gamma^{2/3}}(1, \pm\sqrt{\gamma^{2/3} - 1})$, and (iii) is proven. \square

In Figure 3 we plot the singular curve in all the significantly different cases.

We prove the following

Lemma 4. (solutions with multiplicity three) *For each $\gamma \neq 1$ there are only two points of the reference plane, outside the x -axis, corresponding to solutions of the related intersection problem with multiplicity three.*

Proof. The solutions with multiplicity three must lie on the singular curve (where solutions with multiplicity ≥ 2 are located). By Lemma 2, for $\gamma \neq 1$, the singular curve meets the x -axis either in $(x, y) = (0, 0)$ (that can not correspond to a solution of (3)) or in the stationary points of $\mathcal{C}^{(\gamma)}(x, y)$. By Propositions 1, 2 the stationary points of $\mathcal{C}^{(\gamma)}(x, y)$ can not correspond to solutions of (3) with

multiplicity three, thus, if any, they must lie outside the x -axis. Due to the symmetry of the level curves of $\mathcal{C}^{(\gamma)}(x, y)$ we can investigate only the region with positive y .

Let us fix $\gamma \neq 1$. For each $\mathcal{C} \in \mathbb{R}$ and for each pair $(\bar{\rho}, \bar{\psi}) \in \mathbb{R}^+ \times (0, \pi)$ such that $F(\mathcal{C}, \bar{\rho}, \bar{\psi}) = 0$, since

$$F_\psi(\mathcal{C}, \bar{\rho}, \bar{\psi}) = 3 \frac{q^4 \bar{\rho}}{\bar{r}^5} \sin \bar{\psi} \neq 0,$$

by the implicit function theorem there exists a function $\rho \mapsto \psi_{\mathcal{C}}(\rho)$ with $\psi_{\mathcal{C}}(\bar{\rho}) = \bar{\psi}$ and $F(\mathcal{C}, \rho, \psi_{\mathcal{C}}(\rho)) = 0$ in a neighborhood of $\rho = \bar{\rho}$. The curve $\rho \mapsto (\rho, \psi_{\mathcal{C}}(\rho))$ is a portion of the level curve $\{(\rho, \psi) : \mathcal{C}^{(\gamma)}(\rho, \psi) = \mathcal{C}\}$.

The solutions with multiplicity three of the intersection problem defined by $(\gamma, \mathcal{C}, \epsilon)$, with $\epsilon \in (0, \pi)$, are the points of the reference plane with topocentric polar coordinates $(\bar{\rho}, \bar{\psi})$ where $\bar{\psi} = \epsilon$ and the map $\rho \mapsto \psi_{\mathcal{C}}(\rho)$ is such that $\psi_{\mathcal{C}}(\bar{\rho}) = \epsilon$ and $\psi'_{\mathcal{C}}(\bar{\rho}) = \psi''_{\mathcal{C}}(\bar{\rho}) = 0$, with

$$\psi'_{\mathcal{C}}(\rho) = -\frac{F_\rho(\mathcal{C}, \rho, \psi_{\mathcal{C}}(\rho))}{F_\psi(\mathcal{C}, \rho, \psi_{\mathcal{C}}(\rho))} \quad \text{and} \quad \psi''_{\mathcal{C}}(\rho) = -\frac{F_{\rho\rho}(\mathcal{C}, \rho, \psi_{\mathcal{C}}(\rho))}{F_\psi(\mathcal{C}, \rho, \psi_{\mathcal{C}}(\rho))}, \quad (18)$$

where

$$F_{\rho\rho}(\mathcal{C}, \rho, \psi) = -3 \frac{q^3}{r^7} [q^2 \sin^2 \psi - 4(\rho + q \cos \psi)^2].$$

Note that in writing the second relation in (18) we have used $F_\rho(\mathcal{C}, \rho, \psi_{\mathcal{C}}(\rho)) = 0$. Thus the solutions with multiplicity three correspond to the solutions of

$$F(\mathcal{C}, \rho, \psi) = F_\rho(\mathcal{C}, \rho, \psi) = F_{\rho\rho}(\mathcal{C}, \rho, \psi) = 0,$$

and it is easy to prove that these are the points where an observation line is tangent to the singular curve \mathcal{S} , in fact we have $F - \rho F_\rho = 0$ and $\frac{\partial}{\partial \rho}(F - \rho F_\rho) = -\rho F_{\rho\rho} = 0$.

The condition of belonging to the singular curve gives

$$F(\mathcal{C}, \rho, \psi) - \rho F_\rho(\mathcal{C}, \rho, \psi) = -\gamma + \frac{q^3}{r^3} + 3 \frac{q^3 \rho}{r^5} (\rho + q \cos \psi) = 0. \quad (19)$$

while from $F_{\rho\rho}(\mathcal{C}, \rho, \psi) = 0$ we obtain

$$\rho + q \cos \psi = \pm \frac{q}{2} \sin \psi. \quad (20)$$

If we choose the positive sign in (20), substituting

$$\rho = q \left(\frac{1}{2} \sin \psi - \cos \psi \right) \quad (21)$$

into (19), since $r^2 = (\rho + q \cos \psi)^2 + q^2 \sin^2 \psi = (5/4)q^2 \sin^2 \psi$, we obtain

$$f_1(\psi) = \frac{25\sqrt{5}}{16} \gamma, \quad \text{where } f_1(\psi) = \frac{4 \sin \psi - 3 \cos \psi}{\sin^4 \psi}, \quad (22)$$

with

$$f_1'(\psi) = \frac{3}{\sin^5 \psi} (\sin \psi - 2 \cos \psi)^2 \geq 0,$$

$\lim_{\psi \rightarrow 0^+} f_1(\psi) = -\infty$ and $\lim_{\psi \rightarrow \pi^-} f_1(\psi) = +\infty$, so that $f_1((0, \pi)) = \mathbb{R}$ and equation (22) has a unique solution $\bar{\psi}_1 \in (0, \pi)$ for each value of γ .

This solution, substituted into (21), must give a positive ρ , so that $\bar{\psi}_1 \in (\arctan 2, \pi)$. For $\psi_{min} = \arctan 2$ we have $\sin \psi_{min} = 2 \cos \psi_{min}$ and, due to the monotonicity of f_1 , we obtain

$$\frac{5}{2 \sin^3 \psi_{min}} = \frac{4 \sin \psi_{min} - 3/2 \sin \psi_{min}}{\sin^4 \psi_{min}} < f_1(\bar{\psi}_1) = \frac{25\sqrt{5}}{16} \gamma,$$

that is true only for $\gamma > \frac{8}{5\sqrt{5} \sin^3(\arctan 2)} = 1$.

If we choose the negative sign in (20), substituting

$$\rho = -q\left(\frac{1}{2} \sin \psi + \cos \psi\right) \quad (23)$$

into (19), we obtain

$$f_2(\psi) = \frac{25\sqrt{5}}{16} \gamma, \quad \text{where } f_2(\psi) = \frac{4 \sin \psi + 3 \cos \psi}{\sin^4 \psi}, \quad (24)$$

with

$$f_2'(\psi) = -\frac{3}{\sin^5 \psi} (\sin \psi + 2 \cos \psi)^2 \leq 0,$$

$\lim_{\psi \rightarrow 0^+} f_2(\psi) = +\infty$ and $\lim_{\psi \rightarrow \pi^-} f_2(\psi) = -\infty$, so that $f_2((0, \pi)) = \mathbb{R}$ and equation (24) has a unique solution $\bar{\psi}_2 \in (0, \pi)$ for each value of γ .

This solution, substituted into (23), must give a positive ρ , so that $\bar{\psi}_2 \in (\pi - \arctan 2, \pi)$. For $\psi_{min} = \pi - \arctan 2$ we have $\sin \psi_{min} = -2 \cos \psi_{min}$ and, due to the monotonicity of f_2 , we obtain

$$\frac{5}{2 \sin^3 \psi_{min}} = \frac{4 \sin \psi_{min} - 3/2 \sin \psi_{min}}{\sin^4 \psi_{min}} > f_2(\bar{\psi}_2) = \frac{25\sqrt{5}}{16} \gamma,$$

that is true only for $\gamma < \frac{8}{5\sqrt{5} \sin^3(\arctan 2)} = 1$. □

4.3 AN EVEN OR AN ODD NUMBER OF SOLUTIONS

We prove the following

Proposition 4. *For $(\gamma, \mathcal{C}, \epsilon) \in \mathbb{R}^2 \times [0, \pi]$, with $\gamma \neq 1$, the following asserts hold:*

- (i) *if $\gamma \leq 0$ the number of solutions of (3) is even (zero solutions) for $\mathcal{C} \geq 0$, odd for $\mathcal{C} < 0$.*

(ii) if $0 < \gamma < 1$ the number of solutions of (3) is even for $\mathcal{C} > 0$, odd for $\mathcal{C} \leq 0$;

(iii) if $\gamma > 1$ the number of solutions of (3) is even for $\mathcal{C} \leq 0$, odd for $\mathcal{C} > 0$;

Proof. For $\gamma \leq 0$ equation (2) has no real solution with $r, \rho > 0$ if $\mathcal{C} \geq 0$. Let us consider the polynomial $P(r)$ defined in (5). If $\mathcal{C} < 0$, from $P(0) < 0$ and $\lim_{r \rightarrow +\infty} P(r) = +\infty$ the number of roots of $P(r)$ in the interval $(0, +\infty)$, counted with their multiplicity, is odd and none of these roots is spurious.

Now assume $\gamma > 0$ and let $r_0 = q/\sqrt[3]{\gamma}$ be the radius of the zero circle: we have

$$P(r_0) = \frac{\mathcal{C}^2 q^8}{\gamma^{8/3}} (1 - \gamma^{2/3}) . \quad (25)$$

If $0 < \gamma < 1$ and $\mathcal{C} \neq 0$, from $P(0) < 0 < P(r_0)$ and $\lim_{r \rightarrow +\infty} P(r) = +\infty$ we obtain that in the interval $(0, r_0)$ the number of roots of $P(r)$ is odd, while in $(r_0, +\infty)$ is even. By relation (2) the roots of $P(r)$ in $(0, r_0)$ are spurious iff $\mathcal{C} > 0$, those in $(r_0, +\infty)$ are spurious iff $\mathcal{C} < 0$. For $\mathcal{C} = 0$ the intersection problem (3) reduces to

$$\begin{cases} D(r, \rho) = q(\gamma r^3 - q^3) = 0 \\ G(r, \rho) = r^2 - q^2 - \rho^2 - 2q\rho \cos \epsilon = 0 \\ r, \rho > 0 \end{cases} , \quad (26)$$

with the only non-spurious solution $(r, \rho) = (r_0, -q \cos \epsilon + \sqrt{q^2 \cos^2 \epsilon + (r_0^2 - q^2)})$.

If $\gamma > 1$ and $\mathcal{C} \neq 0$, from $P(0), P(r_0) < 0$ and $\lim_{r \rightarrow +\infty} P(r) = +\infty$ we obtain that in the interval $(0, r_0)$ the number of roots of $P(r)$ is even, while in $(r_0, +\infty)$ it is odd. Like in case $0 < \gamma < 1$, the roots of $P(r)$ in $(0, r_0)$ are spurious iff $\mathcal{C} > 0$, those in $(r_0, +\infty)$ iff $\mathcal{C} < 0$. Since for $\gamma > 1$ we have $r_0 < q$, there is no solution of (26) if $\cos \epsilon < \sqrt{q^2 - r_0^2}/q$, while if $\cos \epsilon \geq \sqrt{q^2 - r_0^2}/q$ the solutions are two, precisely

$$(r, \rho) = (r_0, -q \cos \epsilon \pm \sqrt{q^2 \cos^2 \epsilon - (q^2 - r_0^2)}) .$$

□

The bound on the solutions of the reduced problem implies that the solutions of the intersection problem (3) can not be more than three. In particular, for $(\gamma, \mathcal{C}, \epsilon)$ fulfilling (8) with $\gamma \neq 1$, when the number of solutions of (3) is even they are two, when it is odd they are either one or three.

4.4 THE LIMITING CURVE

Indeed Charlier's assertion that the occurrence of multiple solutions depends only on the position of the observed body can not be generalized to Gauss' method of preliminary orbit determination or to the modified Laplace's method, taking into account topocentric observations; in fact the position of the body defines a different intersection problem for a different $\gamma \in \mathbb{R}$. However, for

each fixed value of $\gamma \in \mathbb{R}$ we shall divide the reference plane into connected components such that, if a solution of an intersection problem lies in one of these components, then we know how many solutions occur in that problem, and all of them lie in the same component.

Remark: From the point of view of the actual computation of the orbit, the parameters $\gamma, \mathcal{C}, \epsilon$ are computed from the three observations, e.g. from (46), thus there is no guarantee that assumption (8) holds. The failure of this assumption can occur for different reasons: the unavoidable errors in the observations, erroneous consideration of three observations of different objects as belonging to the same, etc.

For $\gamma \neq 1$ we define, with $r = \sqrt{x^2 + y^2}$, the sets

$$\mathcal{D}_2(\gamma) = \begin{cases} \emptyset & \text{if } \gamma \leq 0 \\ \{(x, y) : r > r_0\} & \text{if } 0 < \gamma < 1 \\ \{(x, y) : r \leq r_0\} & \text{if } \gamma > 1 \end{cases}$$

and $\mathcal{D}(\gamma) = \mathbb{R}^2 \setminus (\mathcal{D}_2(\gamma) \cup \{(q, 0)\})$. To use a simpler notation, we shall suppress the dependence on γ in $\mathcal{D}(\gamma), \mathcal{D}_2(\gamma)$. For a fixed $\gamma \neq 1$, if we consider a point in \mathcal{D}_2 and if (8) holds for the parameters $(\gamma, \mathcal{C}, \epsilon)$ of the corresponding intersection problem, then there are two solutions of (3), both contained in \mathcal{D}_2 . We shall also say that \mathcal{D}_2 is a region with two solutions of (3). Our aim is to divide the complementary set \mathcal{D} into two connected regions, each with the same number of solutions of (3).

Let $\mathfrak{S} = \mathcal{S} \cap \mathcal{D}$ be the portion of the singular curve \mathcal{S} contained in \mathcal{D} . Note that \mathfrak{S} is connected. In \mathcal{D} the solutions of (3) are one or three, and the solutions lying on the singular curve have intersection multiplicity ≥ 2 , therefore for each point $P \in \mathfrak{S}$ the related intersection problem must have three solutions.

There are two cases, sketched in Figure 4 with labels a), b). Case a) is the generic situation: we have $F_{\rho\rho}(\mathcal{C}, \bar{\rho}, \bar{\psi}) \neq 0$ for $(\bar{\rho}, \bar{\psi})$ corresponding to P and \mathcal{C} such that $F(\mathcal{C}, \bar{\rho}, \bar{\psi}) = 0$, thus P corresponds to a solution of (3) with multiplicity two and there is another point $P' \neq P$ corresponding to the third solution of (3). In case b) we have $F_{\rho\rho}(\mathcal{C}, \bar{\rho}, \bar{\psi}) = 0$, so that in P the observation line is tangent to both the singular curve and to the level curve $\mathcal{C}^{(\gamma)}(x, y) = \mathcal{C}$, and it corresponds to a solution with multiplicity three of the related intersection problem.

Definition 3. Let us fix $\gamma \neq 1$ and let $(\bar{\rho}, \bar{\psi})$ correspond to a point $P \in \mathfrak{S}$. If $F_{\rho\rho}(\mathcal{C}, \bar{\rho}, \bar{\psi}) \neq 0$, we call residual point related to P the point $P' \neq P$ lying on the same observation line and the same level curve of $\mathcal{C}^{(\gamma)}(x, y)$ as P (see Figure 4 a)). If $F_{\rho\rho}(\mathcal{C}, \bar{\rho}, \bar{\psi}) = 0$ we call P a self-residual point, i.e. we consider P as a residual point related to itself (see Figure 4 b)).

We agree that the point $(x, y) = (q, 0)$, corresponding to the observer's position, is the residual point related to $(x, y) = (0, 0)$, when the latter belongs to \mathfrak{S} .

We give the following

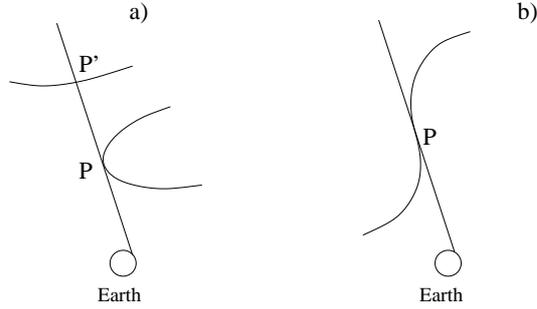


Figure 4: We sketch the tangent intersection between an observation line and a level curve of $\mathcal{C}^{(\gamma)}(x, y)$ giving rise to a residual point in the region \mathcal{D} with an odd number of solutions. Case a) is a generic situation, with P corresponding to a solution with multiplicity two and P' corresponding to the third solution (the residual point). Case b) is non-generic: P is a self-residual point, with intersection multiplicity equal to three.

Definition 4. Let $\gamma \neq 1$. The **limiting curve** is the set composed by all the residual points related to the points in \mathfrak{S} .

Remark: Due to the symmetry of \mathfrak{S} and of the level curves of $\mathcal{C}^{(\gamma)}(x, y)$, the limiting curve is also symmetric with respect to the x -axis. Note that if the point $(q, 0)$ is in \mathcal{L} it is not isolated.

Lemma 5. (separating property) For $\gamma \neq 1$ the limiting curve \mathcal{L} is a connected simple continuous curve, separating \mathcal{D} into two connected regions $\mathcal{D}_1, \mathcal{D}_3$: \mathcal{D}_3 contains the whole portion \mathfrak{S} of the singular curve. If $\gamma < 1$ then \mathcal{L} is a closed curve, if $\gamma > 1$ it is unbounded.

Proof. By Lemma 4 there are only two points A, B in $\mathcal{L} \cap \mathfrak{S}$, outside the x -axis and symmetric with respect to it.

First consider the case $\gamma < 1$. The points C, D $\in \mathfrak{S}$ lying on the x -axis, correspond respectively to the saddle point with $x > 0$ and to the center of the Sun. Take the related residual points $C', D' \in \mathcal{L}$: D' is the observer's position, C' is the solution with $x < 0$ of $\mathcal{C}^{(\gamma)}(x, 0) = \mathcal{C}$, where \mathcal{C} is the value corresponding to the only stationary point of $\mathcal{C}^{(\gamma)}(x, y)$ in \mathcal{D} (see Proposition 2). Then consider the two continuous curves $\widehat{CA}, \widehat{AD}$, that are the portions of \mathfrak{S} connecting C to A and A to D in the reference half-plane with $y \geq 0$. By the continuity of the roots, there exist two continuous curves $\widehat{C'A}, \widehat{AD'}$, connecting C' to A and A to D' composed by the residual points related to the points of \widehat{CA} and \widehat{AD} respectively. Thus $\widehat{C'A}$ and $\widehat{AD'}$ are portions of \mathcal{L} , with $y \geq 0$ (see Figure 5 a)). We can make the same construction in the half-plane with $y \leq 0$ obtaining the two curves $\widehat{C'B}, \widehat{BD'}$, portions of \mathcal{L} symmetric to $\widehat{C'A}, \widehat{AD'}$ with respect to the x -axis. Finally we obtain the whole \mathcal{L} simply by joining these four

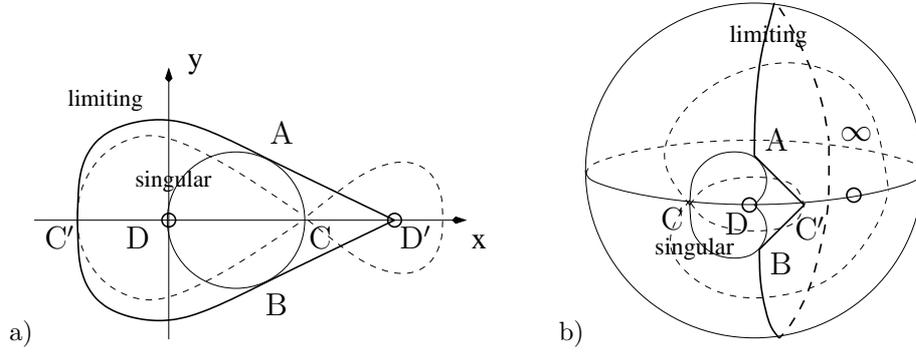


Figure 5: Topological sketch of the limiting curve: a) for $\gamma < 1$, b) for $\gamma > 1$. The dashed curve is the level of $\mathcal{C}^{(\gamma)}(x, y)$ corresponding to the saddle point in \mathcal{D} . The singular curve is drawn with a continuous line, while the limiting curve is enhanced with a thicker line. In case b) we plot the curves in the quotient space $\tilde{\mathcal{A}}(\mathbb{R}^2)$.

curves. Since \mathfrak{S} is a connected closed simple (i.e. without self-intersection points) continuous curve, \mathcal{L} is also a curve with such topological properties, and it divides the reference plane into two connected components \mathcal{D}_1 (unbounded) and \mathcal{D}_3 (bounded). \mathcal{D}_3 is characterized by the fact that it contains the whole \mathfrak{S} inside, in fact there can not be intersections between \mathfrak{S} and \mathcal{L} except for A, B, and the points C' , D' are external to \mathfrak{S} .

Now consider the case $\gamma > 1$. Take Alexandrov's compactification $\mathcal{A}(\mathbb{R}^2)$ of the reference plane by adding a point at infinity (let us call it ∞)⁴, then make the quotient space $\tilde{\mathcal{A}}(\mathbb{R}^2)$ by identifying with a single point D all the points in $\{(x, y) : r \leq r_0\}$. Note that $\tilde{\mathcal{A}}(\mathbb{R}^2)$ is homeomorphic to the 2-sphere S^2 and the region \mathcal{D} with an odd number of solutions corresponds to $S^2 \setminus \{D \cup \infty\}$. We agree that in our representation the x -axis plus ∞ corresponds to the equator of this 2-sphere (see Figure 5 b)).

If we consider the homogenization of the polynomial of degree eight defined in (5), with the homogenizing variable s , we obtain

$$\tilde{P}(r; s) = \mathcal{C}^2 r^8 - q^2 (\mathcal{C}^2 + 2\mathcal{C}\gamma \cos \epsilon + \gamma^2) r^6 s^2 + 2q^5 (\mathcal{C} \cos \epsilon + \gamma) r^3 s^5 - q^8 s^8. \quad (27)$$

The point at infinity ∞ is represented by $(r; s) = (1; 0)$ and $\tilde{P}(1; 0) = \mathcal{C}^2$, thus ∞ is a root of $\tilde{P}(r; s)$ iff $\mathcal{C} = 0$.

We take the compactification $\overline{\mathfrak{S}}$ of \mathfrak{S} in $\tilde{\mathcal{A}}(\mathbb{R}^2)$ by adding D, thus obtaining a closed simple curve. From the discussion above the point at infinity is in the closure of the limiting curve $\overline{\mathcal{L}}$, or better, it corresponds to the boundary $\partial\mathcal{L}$. Now we construct the curve $\overline{\mathcal{L}}$.

Let $C \in \mathfrak{S}$ be the saddle point of $\mathcal{C}^{(\gamma)}(x, y)$ and consider the related residual point C' , i.e. the solution with $0 < x < q$ of $\mathcal{C}^{(\gamma)}(x, 0) = \mathcal{C}$. Consider the

⁴A basis for the topology of $\mathcal{A}(\mathbb{R}^2) = \mathbb{R}^2 \cup \{\infty\}$ is given by the open sets of \mathbb{R}^2 plus the complement in $\mathcal{A}(\mathbb{R}^2)$ of the compact sets in \mathbb{R}^2 .

two continuous curves \widehat{CA} , \widehat{AD} , that are the portions of $\overline{\mathfrak{S}}$ connecting C to A and A to D in the ‘northern’ hemisphere of S^2 . Then take the corresponding curves $\widehat{C'A}$, $\widehat{A\infty}$, portions of $\overline{\mathcal{L}}$, connecting C' to A and A to ∞ , in the same hemisphere (see Figure 5 b)).

We make the same construction in the ‘southern’ hemisphere: we obtain two curves $\widehat{C'B}$, $\widehat{B\infty}$, portions of $\overline{\mathcal{L}}$, symmetric to $\widehat{C'A}$, $\widehat{A\infty}$ with respect to the x -axis. Finally we obtain the whole $\overline{\mathcal{L}}$ simply by joining these four curves and it turns out to be a connected closed simple continuous curve.

The topological space $\tilde{\mathcal{A}}(\mathbb{R}^2)$ is divided into two connected regions by $\overline{\mathcal{L}}$ and one of these regions contains the whole \mathfrak{S} . Going back to the original topological space \mathcal{D} we obtain the result. \square

Proposition 5. (transversality) *The level curves of $\mathcal{C}^{(\gamma)}(x, y)$ cross \mathcal{L} transversely, except for at most the two self-residual points and for the points where \mathcal{L} meets the x -axis.*

Proof. We show this result using topocentric polar coordinates (ρ, ψ) in the reference plane. From the implicit function theorem, if $(\mathcal{C}, \tilde{\rho}, \tilde{\psi}) \in \mathbb{R} \times \mathbb{R}^+ \times [0, \pi]$ is such that $F(\mathcal{C}, \tilde{\rho}, \tilde{\psi}) = F_\rho(\mathcal{C}, \tilde{\rho}, \tilde{\psi}) = 0$, and

$$\det \begin{bmatrix} F_{\mathcal{C}}(\mathcal{C}, \tilde{\rho}, \tilde{\psi}) & 0 \\ F_{\rho\mathcal{C}}(\mathcal{C}, \tilde{\rho}, \tilde{\psi}) & F_{\rho\rho}(\mathcal{C}, \tilde{\rho}, \tilde{\psi}) \end{bmatrix} = \frac{\tilde{\rho}}{q} F_{\rho\rho}(\mathcal{C}, \tilde{\rho}, \tilde{\psi}) \neq 0, \quad (28)$$

then there exists a smooth curve $\psi \mapsto (\mathcal{C}_{\mathcal{S}}(\psi), \rho_{\mathcal{S}}(\psi))$, defined in a neighborhood of $\psi = \tilde{\psi}$, such that $\mathcal{C}_{\mathcal{S}}(\tilde{\psi}) = \mathcal{C}$, $\rho_{\mathcal{S}}(\tilde{\psi}) = \tilde{\rho}$,

$$F(\mathcal{C}_{\mathcal{S}}(\psi), \rho_{\mathcal{S}}(\psi), \psi) = F_\rho(\mathcal{C}_{\mathcal{S}}(\psi), \rho_{\mathcal{S}}(\psi), \psi) = 0,$$

and

$$\begin{pmatrix} \mathcal{C}'_{\mathcal{S}}(\psi) \\ \rho'_{\mathcal{S}}(\psi) \end{pmatrix} = -\frac{1}{F_{\mathcal{C}}F_{\rho\rho}} \begin{bmatrix} F_{\rho\rho} & 0 \\ -F_{\rho\mathcal{C}} & F_{\mathcal{C}} \end{bmatrix} \begin{pmatrix} F_{\psi} \\ F_{\rho\psi} \end{pmatrix} \Big|_{(\mathcal{C}, \rho, \psi) = (\mathcal{C}_{\mathcal{S}}(\psi), \rho_{\mathcal{S}}(\psi), \psi)}$$

where all the derivatives of F are evaluated at $(\mathcal{C}_{\mathcal{S}}(\psi), \rho_{\mathcal{S}}(\psi), \psi)$. In particular the map $\psi \mapsto (\rho_{\mathcal{S}}(\psi), \psi)$ is a local parametrization for the singular curve \mathcal{S} in topocentric polar coordinates and $\mathcal{C}_{\mathcal{S}}(\psi)$ is such that, observing in the direction given by ψ , we meet the level curve $\mathcal{C}^{(\gamma)}(x, y) = \mathcal{C}_{\mathcal{S}}(\psi)$ in a point of \mathcal{S} .

Condition (28) holds for all the points of \mathcal{S} except for the self residual points and, if $\gamma = 1$, the observer position.

If $(\tilde{\rho}, \tilde{\psi})$ corresponds to a point P' of \mathcal{L} that is neither self-residual nor lying on the x -axis, then $F_{\psi}(\mathcal{C}_{\mathcal{S}}(\tilde{\psi}), \tilde{\rho}, \tilde{\psi}) \neq 0$ and relation $F(\mathcal{C}_{\mathcal{S}}(\psi), \rho, \psi) = 0$ implicitly defines a smooth function $\psi \mapsto \rho_{\mathcal{L}}(\psi)$ in a neighborhood of $\psi = \tilde{\psi}$ such that $\rho_{\mathcal{L}}(\tilde{\psi}) = \tilde{\rho}$ and

$$F(\mathcal{C}_{\mathcal{S}}(\psi), \rho_{\mathcal{L}}(\psi), \psi) = 0.$$

The map $\psi \mapsto (\rho_{\mathcal{L}}(\psi), \psi)$ is a local parametrization for the limiting curve \mathcal{L} in topocentric polar coordinates; moreover $\rho_{\mathcal{L}}(\psi) \neq \rho_{\mathcal{S}}(\psi)$ in the range of definition of these functions.

We also have

$$\rho'_{\mathcal{L}}(\psi) = -\frac{F_{\mathcal{C}}(\mathcal{C}_{\mathcal{S}}(\psi), \rho_{\mathcal{L}}(\psi), \psi) \mathcal{C}'_{\mathcal{S}}(\psi) + F_{\psi}(\mathcal{C}_{\mathcal{S}}(\psi), \rho_{\mathcal{L}}(\psi), \psi)}{F_{\rho}(\mathcal{C}_{\mathcal{S}}(\psi), \rho_{\mathcal{L}}(\psi), \psi)}, \quad (29)$$

so that the direction of a vector tangent to the limiting curve in $(\bar{\rho}, \bar{\psi})$ is given by $(\rho'_{\mathcal{L}}(\bar{\psi}), 1)$ in topocentric polar coordinates. Note that the denominator in (29) can vanish only at points of \mathcal{S} .

Given $P' \equiv (\bar{\rho}, \bar{\psi})$ on the limiting curve we can consider the level set $\mathcal{C}^{(\gamma)}(x, y) = \mathcal{C}$, with $\mathcal{C} = \mathcal{C}_{\mathcal{S}}(\bar{\psi})$, passing through P' . By the implicit function theorem the relation $F(\mathcal{C}, \rho, \psi) = 0$ implicitly defines a function $\psi \mapsto \rho_{\mathcal{C}}(\psi)$ in a neighborhood of $\psi = \bar{\psi}$, with $\rho_{\mathcal{C}}(\bar{\psi}) = \bar{\rho}$, such that $F(\mathcal{C}, \rho_{\mathcal{C}}(\psi), \psi) = 0$; hence the tangent vector to the level curve $\mathcal{C}^{(\gamma)}(x, y) = \mathcal{C}$ in P' has the direction of $(\rho'_{\mathcal{C}}(\bar{\psi}), 1)$, with

$$\rho'_{\mathcal{C}}(\psi) = -\frac{F_{\psi}(\mathcal{C}, \rho_{\mathcal{C}}(\psi), \psi)}{F_{\rho}(\mathcal{C}, \rho_{\mathcal{C}}(\psi), \psi)}.$$

Thus the transversality condition in $P' \in \mathcal{L} \setminus (\mathfrak{S} \cup \{(x, y) : y = 0\})$ between the curves \mathcal{L} and $\mathcal{C}^{(\gamma)}(x, y) = \mathcal{C}$, with $\mathcal{C} = \mathcal{C}_{\mathcal{S}}(\bar{\psi})$, is equivalent to

$$\rho'_{\mathcal{L}}(\bar{\psi}) \neq \rho'_{\mathcal{C}}(\bar{\psi}); \quad (30)$$

this condition fails iff $F_{\mathcal{C}}(\mathcal{C}_{\mathcal{S}}(\bar{\psi}), \bar{\rho}, \bar{\psi}) \mathcal{C}'_{\mathcal{S}}(\bar{\psi}) = 0$, that is iff $F_{\psi}(\mathcal{C}_{\mathcal{S}}(\bar{\psi}), \bar{\rho}, \bar{\psi}) = 0$, corresponding to points on the x -axis, excluded *a priori* in the statement of this proposition. \square

The main result of this section is the following:

Theorem 1. (limiting property) *For $\gamma \neq 1$ the limiting curve \mathcal{L} divides the set \mathcal{D} into two connected regions $\mathcal{D}_1, \mathcal{D}_3$: the points of \mathcal{D}_1 are the unique solutions of the corresponding intersection problem; the points of \mathcal{D}_3 are solutions of an intersection problem with three solutions and the additional two solutions also lie in \mathcal{D}_3 .*

Proof. Let us fix a value of $\gamma \neq 1$. Given $\mathcal{C} \in \mathbb{R}$ such that the level curve $\mathcal{M}_{\mathcal{C}} = \{(x, y) : \mathcal{C}^{(\gamma)}(x, y) = \mathcal{C}\}$ is non-empty, either it is all inside \mathcal{D} , or it is all outside, in fact the zero circle (delimiting \mathcal{D} if $\gamma > 0$) corresponds to the particular level curve $\mathcal{C}^{(\gamma)}(x, y) = 0$. Let us consider the case of $\mathcal{M}_{\mathcal{C}}$ lying inside \mathcal{D} .

By Proposition 1 there exists one single solution of (3) in the opposition direction ($\epsilon = 0$) in the following cases: $\gamma \leq 0$ and $\mathcal{C} < 0$, $0 < \gamma < 1$ and $\mathcal{C} \leq 0$, $\gamma > 1$ and $\mathcal{C} > 0$. Thus for each $\gamma \neq 1$ there is exactly one solution at opposition in \mathcal{D} . The limiting curve meets the x -axis only in points with $x \leq q$, hence the solutions at opposition always lie in \mathcal{D}_1 .

If $\mathcal{M}_C \cap \mathfrak{S} = \emptyset$ then $\mathcal{M}_C \subset \mathcal{D}_1$, in fact $\mathfrak{S} \subset \mathcal{D}_3$ and $\mathcal{M}_C \cap \mathcal{L} \neq \emptyset$ iff $\mathcal{M}_C \cap \mathfrak{S} \neq \emptyset$. Thus a necessary condition for \mathcal{M}_C to enter at \mathcal{D}_3 , for increasing values of ψ , is that $\mathcal{M}_C \cap \mathfrak{S} \neq \emptyset$; let us examine this case.

We can exclude that \mathcal{M}_C crosses \mathcal{L} passing from \mathcal{D}_1 to \mathcal{D}_3 at a point $P' \in \mathcal{L} \cap \{y = 0\}$ (that could happen only if \mathcal{C} is a stationary value of $\mathcal{C}^{(\gamma)}(x, y)$) using the symmetry of these curves with respect to the x -axis.

Let \mathcal{C} be a non-stationary value of $\mathcal{C}^{(\gamma)}(x, y)$ and $\psi_P \in (0, \pi)$ be the smallest value of ψ such that the observation direction defined by ψ_P meets $\mathcal{M}_C \cap \mathfrak{S}$; we call P the intersection point. By the symmetry with respect to the x -axis there exists at least another point $Q \in \mathcal{M}_C \cap \mathfrak{S}$. We choose Q so that the angle ψ_Q of the corresponding observation line is such that in the sector of the reference plane defined by $\{(\rho, \psi) : \psi_P < \psi < \psi_Q\}$ there is no point of $\mathcal{M}_C \cap \mathfrak{S}$.

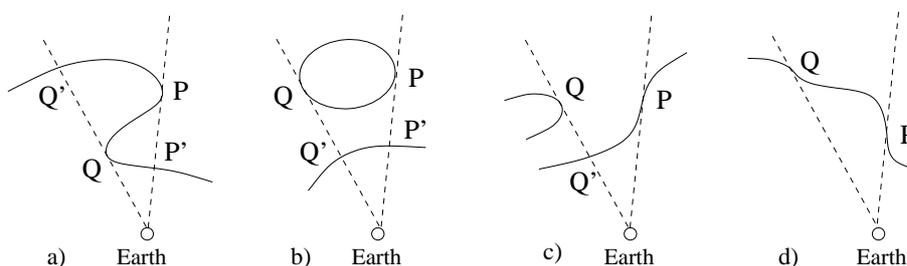


Figure 6: Observation lines crossing a level curve of $\mathcal{C}^{(\gamma)}(x, y)$ in two points of \mathfrak{S} that are consecutive for increasing ψ and lie outside the x -axis. Figures a), b) represent the generic cases, while figures c), d) are the non-generic ones.

The generic cases are sketched in Figure 6 a), b). This because \mathcal{M}_C is a compact manifold without boundary if \mathcal{C} is a non-stationary value of $\mathcal{C}^{(\gamma)}(x, y)$, and the intersections of \mathcal{M}_C with a given direction can not be more than three, counted with multiplicity. By the transversality shown in Proposition 5 \mathcal{M}_C crosses \mathcal{L} entering \mathcal{D}_3 from \mathcal{D}_1 just through the residual point P' , then it returns to \mathcal{D}_1 through the other residual point Q' . In case a) the curve $\widehat{P'Q'}$ is all contained in \mathcal{D}_3 . In case b) the curve $\widehat{P'Q'}$ is in \mathcal{D}_3 , but also Q and P are, and then also the loop containing them, otherwise there should exist a value $\psi^* \in (\psi_P, \psi_Q)$ corresponding to a point in $\mathcal{M}_C \cap \mathcal{L}$, and therefore to a point in $\mathcal{M}_C \cap \mathfrak{S}$. For every $\psi \in [\psi_P, \psi_Q]$ the observation line defined by ψ meets \mathcal{M}_C in three points while for values of ψ slightly less than ψ_P and greater than ψ_Q we have only one intersection. Thus generically \mathcal{M}_C crosses the limiting curve passing from \mathcal{D}_1 to \mathcal{D}_3 and vice versa and, correspondingly, the points of \mathcal{M}_C in \mathcal{D}_1 define intersection problems with only one solution, the points in \mathcal{D}_3 with three solutions.

Let us consider the non-generic case, sketched in Figure 6 c), d). The curve \mathcal{M}_C can not cross \mathcal{L} passing from \mathcal{D}_1 to \mathcal{D}_3 at a self residual point, otherwise there should be a portion of the curve \mathcal{M}_C inside \mathcal{D}_3 whose points correspond to an intersection problem with one solution only. But then there should exist

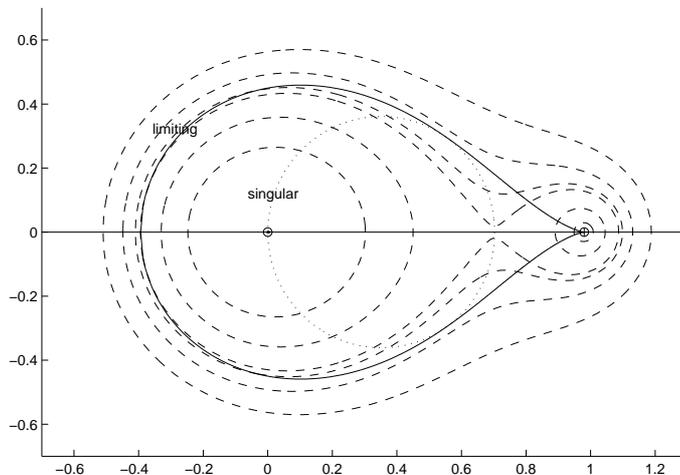


Figure 7: Limiting curve (solid line), singular curve (dotted) and level curves of $\mathcal{C}^{(\gamma)}(x, y)$ (dashed) for $\gamma = -0.5$.

one solution only for the intersection problems corresponding to the points of an entire neighborhood of that curve. □

Remark: in the case $\gamma = 1$ each point of the singular curve has the center of the Earth as residual point, in fact there are always three roots of the polynomial $P(r)$.

In Figures 7, 8, 9 we plot the limiting curve for three significantly different cases.

5 Conclusions

We have introduced a generalization of Charlier's theory of multiple solutions in preliminary orbit determination from three observations: this theory is useful to understand when there are multiple solutions and where they are located.

Figure 10 summarizes the results for all the significantly different cases: we distinguish among regions with a unique solution (white), with two solutions (light grey) and with three solutions (dark grey) of (3). On top-left of Figure 10 we show the results for $\gamma = -0.5$: there are only two regions, with either one or three solutions. On top-right we show the results for $\gamma = 0.8$: in the region outside the zero circle there are two solutions of (3) while the region inside is divided by the limiting curve into two parts, with either one or three solutions. On bottom-left we have Charlier's case ($\gamma = 1$), discussed in Section 3. On bottom-right we show the results for $\gamma = 1.1$: inside the zero circle there are two solutions, while the region outside can contain either one or three solutions.

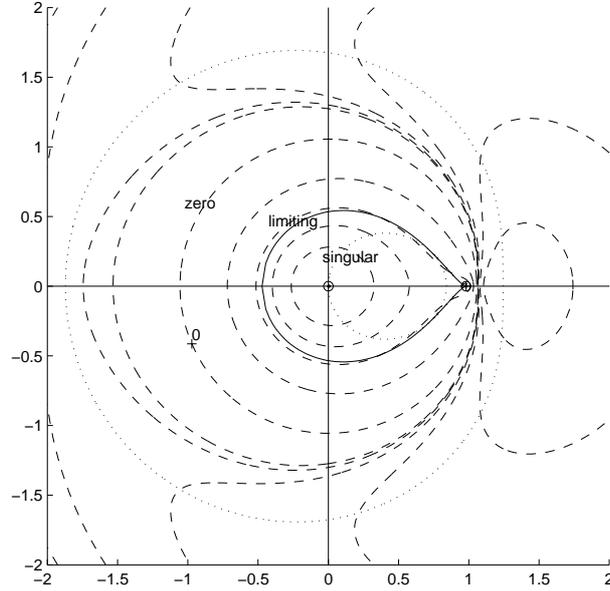


Figure 8: Same as Figure 7 for $\gamma = 0.8$.

Note that in each case the singular curve separates the regions with multiple solutions into parts with only one solution each.

The results on the multiple solutions are generically different from Charlier's: the solutions can be up to three, and up to two close to the opposition direction. Moreover passing to the limit for $\gamma \rightarrow 1^+$ or $\gamma \rightarrow 1^-$ the limiting curve does not tend to Charlier's limiting curve (see Figure 10). Indeed it does neither tend to a portion of Charlier's curve as one could argue at a first glance from Figure 10. Let us first consider the limit $\gamma \rightarrow 1^-$. If $\mathcal{L} = \mathcal{L}(\gamma)$ would tend to the loop of Charlier's limiting curve then there would exist $\bar{\gamma} < 1$ such that $\mathcal{L}(\bar{\gamma})$ has not a cuspidal point with horizontal tangent in $(x, y) = (q, 0)$. But in this case we could not reach the points of \mathfrak{S} close to the center of the Sun along the observation lines passing through points of \mathcal{L} close to the observer. Now consider the limit $\gamma \rightarrow 1^+$. From the computations of Proposition 5 we can check that for each $\gamma > 1$ the tangent vector to the limiting curve, with components $(\rho'_{\mathcal{L}}(\psi) \cos \psi - \rho_{\mathcal{L}}(\psi) \sin(\psi), \rho'_{\mathcal{L}}(\psi) \sin \psi + \rho_{\mathcal{L}}(\psi) \cos(\psi))$, has limit $(0, \lim_{\psi \rightarrow 0} \rho_{\mathcal{L}}(\psi)) \neq (0, 0)$ as $\psi \rightarrow 0$. Hence for each $\gamma > 1$ the limiting curve $\mathcal{L}(\gamma)$ has a vertical tangent when it crosses the x -axis, and passing to the limit for $\gamma \rightarrow 1^+$ it keeps this feature, as opposite to the curve composed by the two unbounded branches of Charlier's limiting curve.

It is important to remark that in case of multiple solutions we should try all of them as first guess for the differential corrections, otherwise we could miss

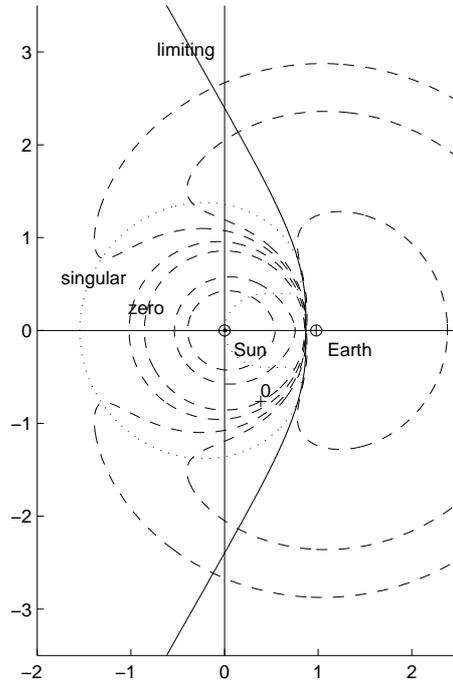


Figure 9: Same as Figure 7 for $\gamma = 1.5$.

the right solution: the convergence of an iterative method is more likely if we start closer to the solution and if we have different preliminary guesses for the solution we must test all of them. Examples of occurrence of multiple solutions with real data of asteroids are discussed in [9]. It is however difficult to find a real case with three solutions such that the one closest to the Earth is the best starting point for differential corrections: in fact the reliability of the heliocentric two-body approximation is questionable for values of the topocentric distance $\rho < 0.01AU$.

6 Acknowledgements

We would like to thank A. Milani for the several interesting discussions on the determination of orbits, that have been among the origins of the present paper. Thanks also to Z. Knežević for a careful reading of the manuscript and for his useful suggestions to improve it.

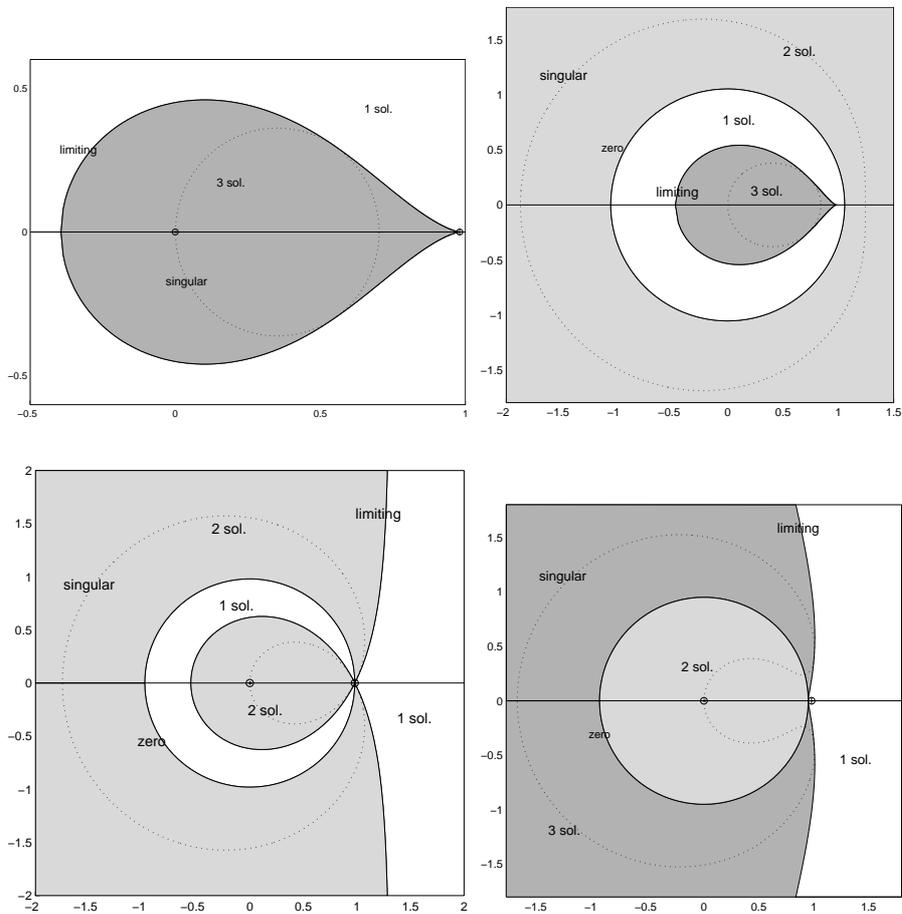


Figure 10: Summary of the results on multiple solutions of (3) for all the qualitatively different cases. The regions with a different number of solutions are enhanced with colours: we use *light grey* for two solutions, *dark grey* for three solutions. Top left: $\gamma = -0.5$. Top right: $\gamma = 0.8$. Bottom left: $\gamma = 1$ (Charlier's case). Bottom right: $\gamma = 1.1$.

7 Appendix

Following [9] we show how to obtain the dynamical equation in Laplace's and Gauss' methods. For a review of these methods see also [2].

7.1 LAPLACE'S METHOD

Let $\boldsymbol{\rho} = \rho \hat{\boldsymbol{\rho}}$ be the *geocentric* position vector of the observed body, with $\rho = \|\boldsymbol{\rho}\|$ and $\hat{\boldsymbol{\rho}} = (\cos \delta \cos \alpha, \cos \delta \sin \alpha, \sin \delta)$, with α, δ the right ascension and

declination. Moreover let $\mathbf{q} = q\hat{\mathbf{q}}$ be the heliocentric position vector of the center of the Earth, with $q = \|\mathbf{q}\|$; then $\mathbf{r} = \mathbf{q} + \boldsymbol{\rho}$ represents the heliocentric position of the body.

We use the *arc length* s to parametrize the motion: the differential relation between s and the time t is given by

$$\eta = \frac{d}{dt}s = \sqrt{\dot{\alpha}^2 \cos^2 \delta + \dot{\delta}^2} \quad \text{proper motion .}$$

Using the moving orthonormal frame

$$\hat{\boldsymbol{\rho}}, \quad \hat{\mathbf{v}} = \frac{d\hat{\boldsymbol{\rho}}}{ds}, \quad \hat{\mathbf{n}} = \hat{\boldsymbol{\rho}} \times \hat{\mathbf{v}}, \quad (31)$$

we introduce the *geodesic curvature* κ by the relation $\frac{d\hat{\mathbf{v}}}{ds} = -\hat{\boldsymbol{\rho}} + \kappa\hat{\mathbf{n}}$. The acceleration of $\boldsymbol{\rho}$ is given by

$$\frac{d^2}{dt^2}\boldsymbol{\rho} = (\ddot{\rho} - \rho\eta^2)\hat{\boldsymbol{\rho}} + (\rho\dot{\eta} + 2\dot{\rho}\eta)\hat{\mathbf{v}} + (\rho\eta^2\kappa)\hat{\mathbf{n}} .$$

On the other hand we have $\frac{d^2}{dt^2}\boldsymbol{\rho} = \frac{d^2}{dt^2}(\mathbf{r} - \mathbf{q})$ and, from the two-body theory,

$$\frac{d^2}{dt^2}\mathbf{r} = -\frac{\mu}{r^3}\mathbf{r}; \quad \frac{d^2}{dt^2}\mathbf{q} = -\frac{\mu + \mu_{\oplus}}{q^3}\mathbf{q},$$

with $r = \|\mathbf{r}\|$ and μ, μ_{\oplus} the masses of the Sun and of the Earth respectively.

From three observations (α_i, δ_i) of a celestial body at times $t_i, i = 1, 2, 3$ we can interpolate for $\alpha, \delta, \dot{\alpha}, \dot{\delta}$ at a mean time \bar{t} .

Neglecting the mass of the Earth and projecting the equation of motion onto $\hat{\mathbf{n}}$ at time \bar{t} we obtain

$$\mathcal{C}\frac{\rho}{q} = 1 - \frac{q^3}{r^3} \quad \text{where} \quad \mathcal{C} = \frac{\eta^2 \kappa q^3}{\mu(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}})}, \quad (32)$$

where, with a little abuse of notation, we use $\rho, q, r, \eta, \hat{\mathbf{q}}, \hat{\mathbf{n}}, \mathcal{C}$ to denote the values of these quantities at time \bar{t} .

According to [3] equation (32) had already been found by Lagrange in 1778: note that in this equation ρ and r are unknowns, while the other quantities can be obtained at time \bar{t} by interpolation.

Using (32) and the geometric equation

$$r^2 = q^2 + \rho^2 + 2q\rho \cos \epsilon, \quad (33)$$

where $\cos \epsilon = \mathbf{q} \cdot \boldsymbol{\rho}/(q\rho)$ can be interpolated at time \bar{t} from the three co-elongations of the observed body, we can write a polynomial equation of degree eight for r at time \bar{t} by eliminating the geocentric distance:

$$\mathcal{C}^2 r^8 - q^2(\mathcal{C}^2 + 2\mathcal{C} \cos \epsilon + 1)r^6 + 2q^5(\mathcal{C} \cos \epsilon + 1)r^3 - q^8 = 0 . \quad (34)$$

The projection on $\hat{\mathbf{v}}$ gives

$$\rho\dot{\eta} + 2\dot{\rho}\eta = \mu(\mathbf{q} \cdot \hat{\mathbf{v}}) \left(\frac{1}{q^3} - \frac{1}{r^3} \right) . \quad (35)$$

We can use equation (35) to find $\dot{\rho}$ from the values of r, ρ found by (34) and (32).

7.2 GAUSS' METHOD

For the times $t_i, i = 1, 2, 3$, let $\mathbf{r}_i, \boldsymbol{\rho}_i$ denote the heliocentric and topocentric position of the body respectively, and let \mathbf{q}_i be the heliocentric position of the observer. Gauss' method uses 3 observations corresponding to heliocentric positions

$$\mathbf{r}_i = \boldsymbol{\rho}_i + \mathbf{q}_i \quad i = 1, 2, 3 \quad (36)$$

at times $t_1 < t_2 < t_3$. We assume that $t_i - t_j, 1 \leq i, j \leq 3$, is much smaller than the period of the orbit and we write $\mathcal{O}(\Delta t)$ for the order of magnitude of the time differences.

We have the coplanarity condition

$$\lambda_1 \mathbf{r}_1 - \mathbf{r}_2 + \lambda_3 \mathbf{r}_3 = 0 \quad (37)$$

for $\lambda_1, \lambda_3 \in \mathbb{R}$. The vector product of both members of (37) with $\mathbf{r}_i, i = 1, 3$ and the fact that the vectors $\mathbf{r}_i \times \mathbf{r}_j, i < j$ have all the same orientation as $\mathbf{c} = \mathbf{r}_h \times \dot{\mathbf{r}}_h, \forall h = 1, 2, 3$ (that is the angular momentum integral per unit mass at any of the three times) allows us to write

$$\lambda_1 = \frac{\mathbf{r}_2 \times \mathbf{r}_3 \cdot \hat{\mathbf{c}}}{\mathbf{r}_1 \times \mathbf{r}_3 \cdot \hat{\mathbf{c}}}, \quad \lambda_3 = \frac{\mathbf{r}_1 \times \mathbf{r}_2 \cdot \hat{\mathbf{c}}}{\mathbf{r}_1 \times \mathbf{r}_3 \cdot \hat{\mathbf{c}}} \quad \text{triangle area ratios.}$$

From the scalar product of $\hat{\boldsymbol{\rho}}_1 \times \hat{\boldsymbol{\rho}}_3$ with both members of (37), using (36), we obtain

$$\rho_2 [\hat{\boldsymbol{\rho}}_1 \times \hat{\boldsymbol{\rho}}_3 \cdot \hat{\boldsymbol{\rho}}_2] = \hat{\boldsymbol{\rho}}_1 \times \hat{\boldsymbol{\rho}}_3 \cdot [\lambda_1 \mathbf{q}_1 - \mathbf{q}_2 + \lambda_3 \mathbf{q}_3]. \quad (38)$$

The differences $\mathbf{r}_i - \mathbf{r}_2, i = 1, 3$, are expanded in powers of $t_{ij} = t_i - t_j = \mathcal{O}(\Delta t)$ by the f, g series formalism [7]; thus $\mathbf{r}_i = f_i \mathbf{r}_2 + g_i \dot{\mathbf{r}}_2$, with

$$f_i = 1 - \frac{\mu}{2} \frac{t_{i2}^2}{r_2^3} + \mathcal{O}(\Delta t^3), \quad g_i = t_{i2} \left(1 - \frac{\mu}{6} \frac{t_{i2}^2}{r_2^3} \right) + \mathcal{O}(\Delta t^4). \quad (39)$$

Then $\mathbf{r}_i \times \mathbf{r}_2 = -g_i \mathbf{c}, \mathbf{r}_1 \times \mathbf{r}_3 = (f_1 g_3 - f_3 g_1) \mathbf{c}$ and

$$\lambda_1 = \frac{g_3}{f_1 g_3 - f_3 g_1}, \quad \lambda_3 = \frac{-g_1}{f_1 g_3 - f_3 g_1}, \quad (40)$$

$$f_1 g_3 - f_3 g_1 = t_{31} \left(1 - \frac{\mu}{6} \frac{t_{31}^2}{r_2^3} \right) + \mathcal{O}(\Delta t^4). \quad (41)$$

Using (39) and (41) in (40) we obtain

$$\lambda_1 = \frac{t_{32}}{t_{31}} \left[1 + \frac{\mu}{6r_2^3} (t_{31}^2 - t_{32}^2) \right] + \mathcal{O}(\Delta t^3), \quad (42)$$

$$\lambda_3 = \frac{t_{21}}{t_{31}} \left[1 + \frac{\mu}{6r_2^3} (t_{31}^2 - t_{21}^2) \right] + \mathcal{O}(\Delta t^3). \quad (43)$$

Let $V = \hat{\boldsymbol{\rho}}_1 \times \hat{\boldsymbol{\rho}}_2 \cdot \hat{\boldsymbol{\rho}}_3$. By substituting (42), (43) into (38), using relations $t_{31}^2 - t_{32}^2 = t_{21}(t_{31} + t_{32})$ and $t_{31}^2 - t_{21}^2 = t_{32}(t_{31} + t_{21})$, we can write

$$-V \rho_2 t_{31} = \hat{\boldsymbol{\rho}}_1 \times \hat{\boldsymbol{\rho}}_3 \cdot (t_{32} \mathbf{q}_1 - t_{31} \mathbf{q}_2 + t_{21} \mathbf{q}_3) + \quad (44)$$

$$+\hat{\rho}_1 \times \hat{\rho}_3 \cdot \left[\frac{\mu}{6r_2^3} [t_{32}t_{21}(t_{31} + t_{32})\mathbf{q}_1 + t_{32}t_{21}(t_{31} + t_{21})\mathbf{q}_3] \right] + \mathcal{O}(\Delta t^4) .$$

If the $\mathcal{O}(\Delta t^4)$ terms are neglected, the coefficient of $1/r_2^3$ in (44) is

$$B(\mathbf{q}_1, \mathbf{q}_3) = \frac{\mu}{6} t_{32} t_{21} \hat{\rho}_1 \times \hat{\rho}_3 \cdot [(t_{31} + t_{32})\mathbf{q}_1 + (t_{31} + t_{21})\mathbf{q}_3]. \quad (45)$$

Then multiply (44) by $q_2^3/B(\mathbf{q}_1, \mathbf{q}_3)$ to obtain

$$-\frac{V \rho_2 t_{31}}{B(\mathbf{q}_1, \mathbf{q}_3)} q_2^3 = \frac{q_2^3}{r_2^3} + \frac{A(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)}{B(\mathbf{q}_1, \mathbf{q}_3)},$$

where

$$A(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = q_2^3 \hat{\rho}_1 \times \hat{\rho}_3 \cdot [t_{32}\mathbf{q}_1 - t_{31}\mathbf{q}_2 + t_{21}\mathbf{q}_3] .$$

Let

$$\mathcal{C} = \frac{V t_{31} q_2^4}{B(\mathbf{q}_1, \mathbf{q}_3)}, \quad \gamma = -\frac{A(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)}{B(\mathbf{q}_1, \mathbf{q}_3)}; \quad (46)$$

then we obtain the dynamical equation of Gauss' method:

$$\mathcal{C} \frac{\rho_2}{q_2} = \gamma - \frac{q_2^3}{r_2^3} . \quad (47)$$

After the possible values for r_2 have been found by (47) and the geometric equation $r_2^2 = \rho_2^2 + q_2^2 + 2\rho_2 q_2 \cos \epsilon_2$, the velocity vector $\dot{\mathbf{r}}_2$ can be computed by different methods, e.g. from Gibbs' formulas [7, Chap. 8].

References

- [1] R. R. Bate, D. D. Mueller and J. E. White, 'Fundamentals of Astrodynamics', Dover publications (1971)
- [2] A. Celletti, G. Pinzari, 'Four Classical Methods for Determining Planetary Elliptic Elements: A Comparison', *CMDA* **93**, 1–52 (2005)
- [3] C. V. L. Charlier, 'On Multiple Solutions in the Determination of Orbits from three Observation', *MNRAS* **71**, 120–124 (1910)
- [4] C. V. L. Charlier, 'Second Note on Multiple Solutions in the Determination of Orbits from three Observation', *MNRAS* **71**, 454–459 (1911)
- [5] D. A. Cox, J. B. Little, D. O'Shea, 'Ideals, Varieties and Algorithms', Springer (1996)
- [6] C. F. Gauss: 1809, 'Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections', reprinted by Dover publications (1963)
- [7] S. Herrick: 'Astrodynamics', Vol.1, Van Nostrand Reinhold, London (1971)

- [8] P. S. Laplace, Mém. Acad. R. Sci. Paris, in *Laplace's collected works*, **10**, 93–146 (1780)
- [9] A. Milani, G. F. Gronchi, D. Farnocchia, Z. Knežević, R. Jedicke, L. Dennau, F. Pierfederici, 'Topocentric Orbit Determination: Algorithms for the Next Generation Surveys', *Icarus* **195**, 474–492 (2008)
- [10] H. C. Plummer, 'An introductory treatise on Dynamical Astronomy', Cambridge University press, reprinted by Dover publications, New York (1960)