Approximation of Artificial Satellites Preliminary Orbits: the efficiency challenge ⋆

Víctor Arroyo, Alicia Cordero, Juan R. Torregrosa *

Instituto de Matemática Multidisciplinar,
Universidad Politécnica de Valencia,
Camino de Vera, s/n, 46022
Valencia, Spain

Abstract

In this paper the problem of the determination of the preliminary orbit of a celestial body is studied. We compare the results obtained by the classical Gauss's method with those obtained by some higher-order iterative methods for solving nonlinear equations. The original problem of the determination of the preliminary orbits was posed by means of a nonlinear equation. We modify this equation in order to obtain a nonlinear system which describes the mentioned problem and we derive a new efficient iterative method for solving it. We also propose a new definition of optimal order of convergence for iterative methods for solving nonlinear systems.

Key words: orbit determination, Gauss's method, nonlinear systems, order of convergence, efficiency index

1 Introduction

Finding the simple roots of a nonlinear equation $f(x) = 0$ or a nonlinear system $F(x) = 0$ are common and important problems in science and engineering. In recent years, many modified iterative methods have been developed to improve the local order of convergence of some classical methods such as Newton, Potra-Pták, Chebyshev, Halley and Ostrowski's methods.

⋆ Supported by Ministerio de Ciencia y Tecnología MTM2010-18539
* Corresponding author
Email addresses: `vicarmar@teleco.upv.es` (Víctor Arroyo), `acordero@mat.upv.es` (Alicia Cordero), `jrtorre@mat.upv.es` (Juan R. Torregrosa).

As the order of an iterative method increases, so does the number of functional evaluations per step. The efficiency index (see [1]), gives a measure of the balance between those quantities, according to the formula $p^{1/d}$, where p is the order of the method and d the number of functional evaluations per step. Kung and Traub [2] conjecture that the order of convergence of any multipoint method without memory cannot exceed the bound 2^{d-1}, (called the *optimal order*). Ostrowski's method [1], Jarrat's method [3] and King's method [4] are some of optimal fourth order methods.

More recently, some optimal eight order methods have been proposed, see for example [5] and [6]. In [7] the authors derive an optimal eighth order method, denoted MOP8, starting from the well known third order Potra-Pták's method, by composing it with modified Newton's iterations and approximating several function evaluations in order to improve the efficiency.

In the multidimensional case, it is also important to take into account the number of operations performed, since for each iteration a number of linear systems must be solved. We recall that the number of products/quotiens that we need for solving m linear systems with the same matrix of size $n \times n$, by using LU factorization, is $\frac{1}{3}n^3 + mn^2 - \frac{1}{3}n$, $n = 2, 3, \ldots$. For this reason, in [8] the authors defined the *Computational Efficiency Index* as $CI = p^{1/(d+op)}$, where op is the number of products/quotients per iteration. For example, the computational efficiency index of Newton's method is $CI_N = 2^{\frac{1}{(1/3)n^3+2n^2+(2/3)n}}$.

One of the most used methods to solve nonlinear systems is Jarratt's method [3], of fourth order of convergence. It uses two functional evaluations of the jacobian matrix and one of the nonlinear function, per step. So, its efficiency index is $I_J = 4^{\frac{1}{2n^2+n}}$ and its computational efficiency index is $CI_J = 4^{\frac{1}{(2/3)n^3+5n^2+(1/3)n}}$. The HMT method described in [9] uses four functional evaluations per step, but only one of them involves the jacobian matrix, so the computational effort made is lower than in Jarratt's method. So, its efficiency index is $I_{HMT} = 4^{\frac{1}{n^2+3n}}$ and its computational efficiency index is $CI_{HMT} = 4^{\frac{1}{(1/3)n^3+4n^2+(8/3)n}}$.

We have adapted the definition of optimal order of convergence to the case of iterative methods to solve nonlinear systems. The extension to several variables of the conjecture of Kung and Traub could be done in the following way:

Conjecture 1 *Given a multipoint iterative method to solve nonlinear systems of equations which requires $d = k_1 + k_2$ functional evaluations per step such that k_1 of them correspond to the number of evaluations of the jacobian matrix and k_2 to evaluations of the nonlinear function. We conjecture that the optimal order for this method is $2^{k_1+k_2-1}$ if $k_1 \leq k_2$.*

This concept of optimal order is an important tool to establish a classification between the iterative methods for solving nonlinear systems. In this classification of methods, only Newton's method can be considered as an optimal method of order two. When we look at fourth-order methods, we find that $k_2 > k_1$, as in Jarratt's, or $k_1 + k_2 > 3$, as in HMT. It should be necessary to design fourth-order methods with one functional evaluation of jacobian matrix and only two evaluations of the nonlinear function. As far as we know, this methods does not exist yet. So, further effort must be made in the future to get optimal methods to solve nonlinear systems. In this paper we propose a new iterative method of order five that, not being optimal, has the best efficiency index of the methods we know till now. It is very competitive since it only needs one evaluation of the jacobian matrix per iteration.

The rest of this paper is organized as follows: in Section 2 we present the new method, analyze its convergence order and establish the comparison between known optimal (or not) iterative methods in terms of efficiency indices, in Section 3 we present an application of this analysis with the preliminary orbit determination of a satellite. Finally, in Section 4, different numerical tests confirm the theoretical results.

2 Description and convergence analysis

Following the ideas described in [9], we propose a new iterative method, called M5, that uses the same number of functional evaluations and operations as in the method shown in the mentioned paper, but the order of convergence is higher:

$$\begin{aligned} z^{(k)} &= y^{(k)} - 5[F'(x^{(k)})]^{-1}F(y^{(k)}), \\ x^{(k+1)} &= z^{(k)} - \tfrac{1}{5}[F'(x^{(k)})]^{-1}\left[-16F(y^{(k)}) + F(z^{(k)})\right], \end{aligned} \tag{1}$$

where $y^{(k)}$ is the kth iteration of Newton's method. We show in the next result that the order of convergence of this method is 5. The proof is based on Taylor expansions around the solution, whose notation was introduced in [8].

Theorem 1 *Let $F : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be sufficiently differentiable at each point of an open neighborhood D of $\bar{x} \in \mathbb{R}^n$, that is a solution of the system $F(x) = 0$. Let us suppose that $F'(x)$ is continuous and nonsingular in $\bar{x}$. Then, the sequence $\{x^{(k)}\}_{k\geq 0}$ obtained using the iterative expression (1) converges to $\bar{x}$ with order* 5.

Proof: Taylor's expansion of F and F' around $\bar{x}$ gives

$$F(x^{(k)}) = F'(\bar{x})\left[e^{(k)} + C_2 e^{(k)^2} + C_3 e^{(k)^3} + C_4 e^{(k)^4}\right] + O(e^{(k)^5})$$

$$F'(x^{(k)}) = F'(\bar{x})\left[I + 2C_2 e^{(k)} + 3C_3 e^{(k)^2} + 4C_4 e^{(k)^3} + 5C_5 e^{(k)^4}\right] + O(e^{(k)^5}),$$

where $C_k = (1/k!)[F'(\bar{x})]^{-1}F^{(k)}(\bar{x})$, $k = 2, 3, \ldots$, and $e^{(k)} = x^{(k)} - \bar{x}$. From this expression, we have

$$[F'(x^{(k)})]^{-1} = \left[I + X_2 e^{(k)} + X_3 e^{(k)^2} + X_4 e^{(k)^3}\right][F'(\bar{x})]^{-1} + O(e^{(k)^4}), \tag{2}$$

where $X_2 = -2C_2$, $X_3 = 4C_2^2 - 3C_3$ and $X_4 = -8C_2^3 + 6C_2C_3 + 6C_3C_2 - 4C_4$.

Therefore, the expression for $y^{(k)} - \bar{x}$ is

$$y^{(k)} - \bar{x} = C_2 e^{(k)^2} + (2C_3 - 2C_2^2)e^{(k)^3} + (4C_2^3 - 4C_2C_3 - 3C_3C_2 + 3C_4)e^{(k)^4} + O(e^{(k)^5})$$

and

$$F(y^{(k)}) = F'(\bar{x}) \left[C_2 e^{(k)^2} + (2C_3 - 2C_2^2) e^{(k)^3} + (5C_2^3 - 4C_2C_3 - 3C_3C_2 + 3C_4) e^{(k)^4}\right] + O(e^{(k)^5}) \quad (3)$$

In same way, we obtain the expression of $z^{(k)} - \bar{x}$ and

$$F(z^{(k)}) = F'(\bar{x}) \left[-4C_2 e^{(k)^2} + (18C_2^2 - 8C_3) e^{(k)^3} + \right.$$
$$\left. +(-45C_2^3 + 36C_2C_3 + 27C_3C_2 - 12C_4) e^{(k)^4}\right] + O(e^{(k)^5}). \quad (4)$$

Finally, by replacing (2), (3) and (4) in (1), we obtain

$$x^{(k+1)} - \bar{x} = x^{(k)} - \bar{x} - \frac{1}{5}[F'(x^{(k)})]^{-1} \left[-16F(y^{(k)}) + F(z^{(k)})\right] = O(e^{(k)^5}).\square$$

Therefore, the efficiency index of method M5 is $I_{M5} = 5^{\frac{1}{n^2+3n}}$ and its computational efficiency index is $CI_{M5} = 5^{\frac{1}{(1/3)n^3+4n^2+(8/3)n}}$. In Figure 1 and Figure 2 we can see the respective indices of Newton, Jarratt, HMT and M5 methods, for different sizes of the system.

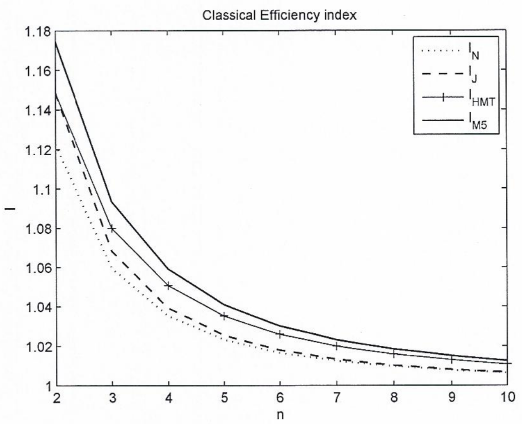

Fig. 1. Classical efficiency indices.

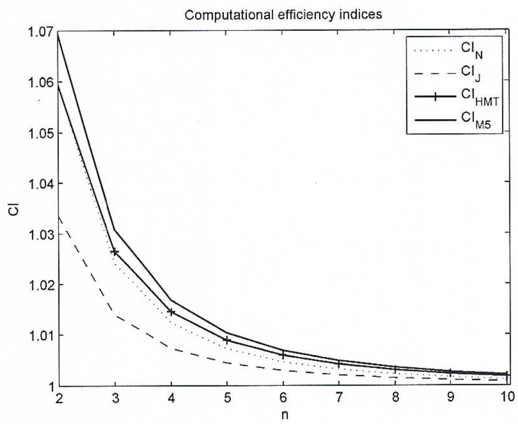

Fig. 2. Computational efficiency indices.

3 An application: the preliminary orbit determination

The first step in orbit determination methods is to obtain preliminary orbits, as the motion analyzed is under the premises of the two bodies problem. It is possible to set a two-dimensional coordinate system (see Figure 3), where the X axis points to the perigee of the orbit, the closest point of the elliptical orbit to the focus and center of the system, the Earth. In this picture the true anomaly ν and the eccentric anomaly E can be observed. In order to place this orbit in the celestial sphere and determine completely the position of a body in the orbit, some elements (called orbital or keplerian elements) must be determined.

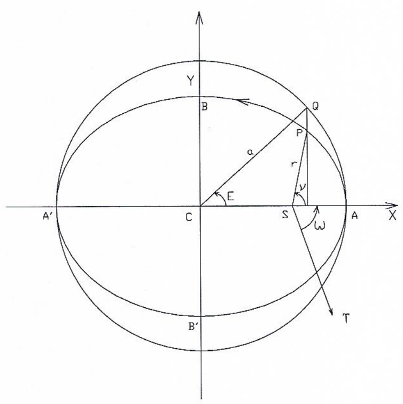

Fig. 3. Size, shape and anomalies in orbital plane 2-dimensional coordinate system.

Some fundamental constants are the Earth gravitational constant, $k = 0.07436574(e.r.)^{\frac{3}{2}}/min$ (see [10]) and the gravitational parameter $\mu = \frac{1}{m_{Earth}}(m_{Earth}+m_{Object}) \approx 1$. Then, modified time variable is introduced as $\tau = k(t_2 - t_1)$, where t_1 is an initial arbitrary time and t_2 is the observation time.

To estimate the velocity we can make use of the closed forms of the f and g series (see [10,11]), $f = 1 - \frac{a}{|\vec{r_1}|}[1 - \cos(E_2 - E_1)]$ and $g = \tau - \frac{\sqrt{a^3}}{\mu}[(E_2 - E_1) - \sin(E_2 - E_1)]$, so we can express the rate respect two positions vectors and time as

$$\dot{\vec{r_1}} = \frac{\vec{r_2} - f \cdot \vec{r_1}}{g}. \qquad (5)$$

So, it is clear that, knowing two position vectors and its corresponding observational instants, the main objective of the different methods that determine preliminary orbits is the calculation of the semi-major axis, a, and the eccentric anomalies difference, $E_2 - E_1$. When they have been calculated, it is possible to obtain by (5) the velocity vector corresponding to one of the known position vectors and, then, to obtain the orbital elements.

From the available input data, two position vectors and times for the observations, τ can be immediately deduced. We can also obtain other intermediate results as the difference of true anomalies, $(\nu_2 - \nu_1)$. This difference is calculated by $\cos(\nu_2 - \nu_1) = \frac{\vec{r_1} \cdot \vec{r_2}}{|\vec{r_1}| \cdot |\vec{r_2}|}$ and $\sin(\nu_2 - \nu_1) = \pm \frac{x_1 y_2 - x_2 y_1}{|x_1 y_2 - x_2 y_1|} \sqrt{1 - \cos^2(\nu_2 - \nu_1}$ with positive sign for direct orbits, and negative for retrograde orbits.

Once the difference of true anomalies is obtained from the position vectors and times, the specific orbit determination method is used. In our particular case, we will introduce in the following section the classical Gauss' method and, thereafter, we will modify it in order to estimate the value of the semi-major axis and eccentric anomalies by means of high-order iterative methods.

4 Modified Gauss' method of orbit determination

Gauss' method calculate a preliminary orbit of a celestial body by means of only two observations (position vectors). It is based on the relation between the areas of the sector of the ellipse and the respective of the triangle delimited by both position vectors, $\vec{r_1}$ y $\vec{r_2}$. The ratio sector-triangle can be expressed as

$$y = \frac{\sqrt{\mu p}\cdot\tau}{r_2 r_1 \sin(\nu_2-\nu_1)} = \frac{\sqrt{\mu}\cdot\tau}{2\sqrt{a}\sqrt{r_2 r_1}\sin\left(\frac{E_2-E_1}{2}\right)\cos\left(\frac{\nu_2-\nu_1}{2}\right)}, \tag{6}$$

(with $(\nu_2-\nu_1)\neq\pi$). This method holds also on the first

$$y^2 = \frac{m}{l+x} \tag{7}$$

and second

$$y^2(y-1) = mX, \tag{8}$$

Gauss equations, where the constants of the problem (based on the data and the previously made calculations and the difference of true anomalies), are

$$l = \frac{r_2+r_1}{4\sqrt{r_2 r_1}\cos\left(\frac{\nu_2-\nu_1}{2}\right)} - \frac{1}{2} \quad \text{and} \quad m = \frac{\mu\tau^2}{\left[2\sqrt{r_2 r_1}\cos\left(\frac{\nu_2-\nu_1}{2}\right)\right]^2}. \tag{9}$$

Moreover, also must be determined in the process the value of:

$$x = \sin^2\left(\frac{E_2-E_1}{4}\right) \quad \text{and} \quad X = \frac{E_2-E_1-\sin(E_2-E_1)}{\sin^3\left(\frac{E_2-E_1}{2}\right)}. \tag{10}$$

With these equations we present two different schemes to solve the problem. One of them is the classical method, which reduces first (eq. (7)) and second (eq. (8)) Gauss equations to a unique nonlinear equation, $y = 1 + X(l+x)$, solved by fixed point method. The other one is the proposed modified Gauss scheme, which solve directly the nonlinear system formed by both Gauss equations.

The Gauss method has some limitations as the critical observation angles spread ($\nu_2-\nu_1=\pi$), in which case the denominator of equation (6) vanish. Moreover, it is known that this method is only convergent to a coherent solution if the observation angles spread is less than 70^o. The ratio y grows with the angles spread, leading to an invalid solution, if it converges. So this method is suitable for small spreads in observations, that is, observations which are close to each other.

The first variation proposed is to use high-order schemes in order to solve the unified nonlinear equation in the classical Gauss method. In this case, we will use optimal methods of increasing order: Newton and Ostrowski's methods, and MOP_8 of order 8.

Nevertheless, it is possible to make a different approach to the problem, solving the nonlinear system formed by both Gauss equations, (7) and (8), whose unknowns are the ratio y and the difference of eccentric anomalies, E_2-E_1, with different higher order iterative methods. In particular we will use Newton, Jarratt, HMT and M5 methods.

5 Numerical results

Numerical computations have been carried out using variable precision arithmetic, with 500 digits, in MATLAB 7.1. The stopping criterion used is $\left\|x^{(k+1)}-x^{(k)}\right\|+\left\|F(x^{(k)})\right\|<10^{-250}$, therefore, we check that the iterates succession converge to an approximation to the solution of the nonlinear system. For every method, we count the number of iterations needed to reach the wished tolerance and the elapsed time. The reference or test orbits we use can be found in [10].

1-dimensional			**2-dimensional**		
Scheme	**Iter.**	**e-time**	**Scheme**	**Iter.**	**e-time**
C	133	8.7156			
$N1$	8	3.0753	$N2$	9	3.5372
Os	5	2.7519	J	5	3.4459
			HMT	5	4.8358
$MOP8$	3	3.9366	$M5$	5	4.4452

Table 1
Comparison of different Gauss method schemes for a reference orbit

In Table 1 we show the results obtained by the classical (C), Newton (N), Ostrowsky (Os), MOP8, Jarratt (J), HMT and M5 methods, for one and several variables, in the case of a test orbit with spread of the observations $SP=\nu_2-\nu_1=12.23^o$. Several conclusions can be made:

- In 1-dimensional case, the number of iterations and the elapsed time have been reduced in a great amount. Indeed, the most efficient method is the optimal fourth-order method from Ostrowski.
- In the case of the system of Gauss equations, the number of iterations have also been reduced, but the times of execution are slightly higher, due to the ill-conditioned system. Moreover, as the size of the system is small, the effect of the evaluations and operations made with the jacobian matrix are not very evident. In this case, the new method $M5$ appears to be quite efficient.
- From a global point of view, 1-dimensional Ostrowski's method seems to be the most efficient to solve this particular problem.

Due to limitations in number of digits and format in observations data, and to the last phase of calculations, some accuracy is lost, but it is hard to determine differences in errors in the presented schemes. In fact, the maximum exact error is round about 10^{-198}.

In Table 2, we can compare the number of iterations needed for different test orbits with different angles spread in observations. Our aim is to realize that the limitation of angles spread is still present, but overall process is made faster, not increasing iterations to find a solution in cases with bigger difference of true anomalies. Nevertheless, in these cases a higher sensitivity is observed in the 1-dimensional case, as methods of order 4 and 8 usually do not converge. In this respect, the Modified Gauss schemes that use the system of Gauss equations (eqs. (7) and (8)) appear to be more stable

Scheme	$SP = 12.23^o$	$SP = 22.06^o$	$SP = 31.46^o$
C	133	188	250
$N1$	8	8	8
Os	5	NC	5
$MOP8$	3	NC	NC
$N2$	9	8	9
J	5	5	5
HMT	5	5	5
$M5$	5	5	5

Table 2
Iterations needed for different spreads

and competitive.

Acknowledgement

The authors would like to thank the referee for the valuable comments and for the suggestions to improve the readability of the paper.

References

[1] A.M. Ostrowski, Solutions of equations and systems of equations, Academic Press, New York-London, 1966.

[2] H.T. Kung, J.F. Traub, Optimal order of one-point and multi-point iteration, Applied Mathematics and Computation, 21 (1974) 643-651.

[3] P. Jarratt, Some fourth order multipoint iterative methods for solving equations, Mathematical Computation, 20 (1966) 434-437.

[4] R. King, A family of fourth order methods for nonlinear equations, SIAM J. Numer. Anal. 10 (1973) 876879.

[5] W. Bi, H. Ren, Q. Wu, Three-step iterative methods with eighth-order convergence for solving nonlinear equations, Journal of Computational and Applied Mathematics, 255 (2009) 105-112.

[6] Sharma, J.R. Sharma, R. A new family of modified Ostrowski's methods with accelerated eighth order convergence, Numerical Algoritms, 54(4) (2010) 445-458.

[7] A. Cordero, J.L. Hueso, E. Martínez, J.R. Torregrosa, New modifications of Potra-Pták's method with optimal fourth and eighth order of convergence, Journal of Computational and Applied Mathematics, 234 (2010) 2969-2976.

[8] A. Cordero, J.L. Hueso, E. Martínez, J.R. Torregrosa, A modified Newton-Jarrattt's composition, Numerical Algorithms, 55 (2010) 87-99.

[9] J.L. Hueso, E. Martínez, J.R. Torregrosa, Third and fourth order iterative methods free from second derivative for nonlinear systems, Applied Mathematics and Computation, 211 (2009) 190-197.

[10] P.R Escobal, Methods of Orbit Determination, Robert E. Krieger Publishing Company, 1975.

[11] M.J. Sevilla, Mecánica Celeste Clásica, Instituto de Astronomía y Geodesia. Facultad de Ciencias Matemáticas. Universidad Complutense de Madrid, 1989.

Table

1-dimensional			2-dimensional		
Scheme	**Iter.**	**e-time**	**Scheme**	**Iter.**	**e-time**
C	133	8.7156			
$N1$	8	3.0753	$N2$	9	3.5372
Os	5	2.7519	J	5	3.4459
			HMT	5	4.8358
$MOP8$	3	3.9366	$M5$	5	4.4452

Table 1
Comparison of different Gauss method schemes for a reference orbit

Scheme	$SP = 12.23^o$	$SP = 22.06^o$	$SP = 31.46^o$
C	133	188	250
$N1$	8	8	8
Os	5	NC	5
$MOP8$	3	NC	NC
$N2$	9	8	9
J	5	5	5
HMT	5	5	5
$M5$	5	5	5

Table 2
Iterations needed for different spreads