PERGAMON

Nonlinear Analysis 47 (2001) 3449–3454

Nonlinear Analysis

www.elsevier.nl/locate/na

# Computing families of periodic orbits through optimization methods

M.N. Vrahatis [a,b], O. Ragos [a,b], G.S. Androulakis [a,b]

[a] *Department of Mathematics, University of Patras, GR-26110 Patras, Greece*

[b] *University of Patras Artificial Intelligence Research Center (UPAIRC)*

**Abstract**

The techniques used for the numerical computation of families of periodic orbits of dynamical systems are based on predictor-corrector schemes. These schemes usually depend on solving systems of approximate equations involving the solutions of the equations of motion and variation. In this contribution we apply some well-known unconstrained optimization methods in obtaining the solutions of these approximate equations and we compare their efficiency on a specific problem of Celestial Mechanics.

## 1 Introduction

Let us consider a dynamical system expressed by the equations $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{f} = (f_1, f_2, \ldots, f_n) : \mathbb{R}^{n+1} \to \mathbb{R}^n$ and $t$ is the independent variable. Any solution $\mathbf{x}$ of this system is periodic of period $T$ if it satisfies the condition $\mathbf{x}(\mathbf{x}_0, t = 0) = \mathbf{x}(\mathbf{x}_0, t = T)$, where $\mathbf{x}_0$ is the initial point of the orbit at $t = 0$.

In many dynamical systems periodic solutions form families, i.e. groups of such solutions whose coordinates in phase space vary continuously while their properties change smoothly. If a periodic solution which belongs to a specified family is known, then we are able to compute the whole family by calculating successive orbits of it. This can be done as follows: we predict an approximation of a nearby periodic orbit, which is also member of this family, and then, by correcting this prediction we are able to calculate the successive orbit.

0362-546X/01/$ - see front matter 
PII: S0362-546X(01)00462-X

The most classical predictor-corrector schemes used for this purpose are the so-called linear predictor-corrector algorithms. These algorithms are based on the solution of linear systems of equations whose variables represent the proper modifications that must be applied to the parameters of a given orbit in order to obtain a new member of the family. Recently, it has been shown that unconstrained optimization techniques can be used in obtaining these modifications [6,10]. In this contribution we apply and test the efficiency of several such techniques on a specific problem of Celestial Mechanics.

## 2 Description of the problem

Let us consider the Radzievskii's model [5] for the description of the motion of a small particle under the influence of the gravitation and the radiation pressure due to the light emission from the members of a binary star. The system of equations that expresses this motion is the following:.

$$\ddot{x}_1 - 2\dot{x}_2 = \frac{\partial U}{\partial x_1}, \qquad \ddot{x}_2 + 2\dot{x}_1 = \frac{\partial U}{\partial x_2}, \qquad \ddot{x}_3 = \frac{\partial U}{\partial x_3}, \tag{1}$$

where

$$U = \frac{1}{2}(x_1^2 + x_2^2) + \frac{q_1(1-\mu)}{\sqrt{(x_1+\mu)^2 + x_2^2 + x_3^2}} + \frac{q_2\mu}{\sqrt{(x_1+\mu-1)^2 + x_2^2 + x_3^2}}, \tag{2}$$

while $1-\mu$, $\mu$ represent the masses of the two main bodies and $q_1$, $q_2$ are parameters expressing the relations between the gravitation attraction and the radiation pressure of each one of them ($0 < \mu \leqslant 0.5$, $q_i \leqslant 1, i = 1, 2$). This system, under suitable conditions, has, among others, two equilibrium points, named $L_6$ and $L_7$, which lie on the $Ox_1x_3-$plane and are symmetrical with respect to the $Ox_1x_2-$plane. For certain combinations of the parameters of the problem these points are stable and then two families of periodic orbits are emanating from each one of them, named $L_6^1$, $L_6^2$ and $L_7^1$, $L_7^2$, respectively [7]. The orbits of these families are symmetrical with respect to the $Ox_1x_3-$plane. In the sequel we deal with the family $L_6^1$.

A first approximation of initial conditions of a small periodic orbit belonging to $L_6^1$ in the vicinity of $L_6$ can be obtained by the use of first or second order expansions of Eqs (1) around this point. Then this approximation must be corrected to a desired accuracy to give the exact periodic solution. After this, other members of the family have to be estimated and corrected successively. The classical linear predictor-corrector that can be used for this purpose is as follows [3,8].

a) Given the initial conditions $\mathbf{x}_0 = (x_{10}, x_{20} = 0, x_{30}, \dot{x}_{10} = 0, \dot{x}_{20}, \dot{x}_{30} = 0)$ and the period T of an already known orbit of the family, we can predict suitable modifications $\delta\mathbf{x}_0 = (\delta x_{10}, 0, \delta x_{30}, 0, \delta\dot{x}_{20}, 0)$ and $\delta T$ of these parameters in order to predict another orbit of the family by considering a constant deviation $d$ to one of the $\delta$'s and, then, by calculating the others from the solution of the

system:

$$g_1 = \frac{\partial x_2}{\partial x_{10}}\delta x_{10} + \frac{\partial x_2}{\partial x_{30}}\delta x_{30} + \frac{\partial x_2}{\partial \dot{x}_{20}}\delta \dot{x}_{20} + \frac{\partial x_2}{\partial t}\delta T = 0,$$

$$g_2 = \frac{\partial \dot{x}_1}{\partial x_{10}}\delta x_{10} + \frac{\partial \dot{x}_1}{\partial x_{30}}\delta x_{30} + \frac{\partial \dot{x}_1}{\partial \dot{x}_{20}}\delta \dot{x}_{20} + \frac{\partial \dot{x}_1}{\partial t}\delta T = 0, \tag{3}$$

$$g_3 = \frac{\partial \dot{x}_3}{\partial x_{10}}\delta x_{10} + \frac{\partial \dot{x}_3}{\partial x_{30}}\delta x_{30} + \frac{\partial \dot{x}_3}{\partial \dot{x}_{20}}\delta \dot{x}_{20} + \frac{\partial \dot{x}_3}{\partial t}\delta T = 0.$$

b) If the prediction is not satisfactory then we have to alter the so obtained initial conditions and period to get a better approximation. Supposing that these corrections are $\delta \mathbf{x}_0 = (\delta x_{10}, 0, \delta x_{30}, 0, \delta \dot{x}_{20}, 0)$ and $\delta T$, we may consider one of the $\delta$'s equal to zero and, then, find the rest of them by solving the equations:

$$h_1 = x_2 + \frac{\partial x_2}{\partial x_{10}}\delta x_{10} + \frac{\partial x_2}{\partial x_{30}}\delta x_{30} + \frac{\partial x_2}{\partial \dot{x}_{20}}\delta \dot{x}_{20} + \frac{\partial x_2}{\partial t}\delta T = 0,$$

$$h_2 = \dot{x}_1 + \frac{\partial \dot{x}_1}{\partial x_{10}}\delta x_{10} + \frac{\partial \dot{x}_1}{\partial x_{30}}\delta x_{30} + \frac{\partial \dot{x}_1}{\partial \dot{x}_{20}}\delta \dot{x}_{20} + \frac{\partial \dot{x}_1}{\partial t}\delta T = 0, \tag{4}$$

$$h_3 = \dot{x}_3 + \frac{\partial \dot{x}_3}{\partial x_{10}}\delta x_{10} + \frac{\partial \dot{x}_3}{\partial x_{30}}\delta x_{30} + \frac{\partial \dot{x}_3}{\partial \dot{x}_{20}}\delta \dot{x}_{20} + \frac{\partial \dot{x}_3}{\partial t}\delta T = 0.$$

The above mentioned coordinates of the orbits and the partial derivatives correspond to the known orbit and are evaluated at $t = T$.

To use optimization techniques in order to calculate these *delta*'s, we assume that the family is described by a function of the initial conditions and the period of its orbits, $f(\mathbf{x}_0^p, T) = 0$. This means that all these orbits are zeros of $f$. Then, the previously mentioned predictor-corrector steps can be interpreted as follows. Suppose that one of these zeros, $\mathbf{x}_0^p$, is known. Then a new orbit can be predicted by minimizing the function:

$$\phi = g_1^2 + g_2^2 + g_3^2 + g_4^2 + \left((\delta x_{10})^2 + (\delta x_{30})^2 + (\delta \dot{x}_{20})^2 - \theta\right)^2, \tag{5}$$

where $g_1$, $g_2$, $g_3$ are the functions appearing in Eqs (3), $\theta$ denotes a proper small deviation along the family and

$$g_4 = \frac{\partial x_3}{\partial x_{10}}\delta x_{10} + \left(\frac{\partial x_3}{\partial x_{30}} - 1\right)\delta x_{30} + \frac{\partial x_3}{\partial \dot{x}_{20}}\delta \dot{x}_{20} + \frac{\partial x_3}{\partial t}\delta T. \tag{6}$$

The purpose of this construction of $\phi$ is to ensure that $g_1 = g_2 = g_3 = g_4 = 0$ are fulfilled and to force the estimation of the new orbit to be at a distance $\theta$ from the known one. If the approximation is not "good enough" it can be corrected by minimizing the following function:

$$\psi = h_1^2 + h_2^2 + h_3^2, \tag{7}$$

where the functions $h_1$, $h_2$, $h_3$ are those appearing in System (4).

## 3 Applications and concluding remarks

For the optimization of the functions $\phi$ and $\psi$ we can use any optimization method. Of course, there is a large variety of optimization algorithms for computing the minima of an objective function $f: \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}$. These algorithms possess advantages and disadvantages and it is not always evident which one is proper for a given class of applications. Here, we shall apply four well known and widely used algorithms and we shall compare their efficiency on the problem of Section 2.

Let us briefly describe the considered algorithms. There is a class of methods called *nonlinear conjugate gradient methods*, as typified by the *Fletcher–Reeves* (FR) *algorithm* and the closely related *Polak–Ribiere* (PR) *algorithm* [1,2,4]. Conjugate gradient methods require storage of order $n$. On the other hand, they are very sensitive to rounding off errors. The most known methods in this class are expressed by:

$$x^{k+1} = x^k + \lambda^k p^k, \quad k = 0, 1, 2, \ldots, \quad \text{where } \lambda^k \text{ minimizes } f(x^k + \lambda^k p^k),$$

$$p^k = -\nabla f(x^k) + \zeta^k p^{k-1}, \qquad p^0 = -\nabla f(x^0),$$

$$\zeta^k = \frac{\left(\nabla f(x^k) - \gamma \nabla f(x^{k-1})\right)^\top \nabla f(x^k)}{\nabla f(x^{k-1})^\top \nabla f(x^{k-1})},$$

where $\nabla f(x)$ is the gradient of $f$ at $x$. For $\gamma = 0$ we have the FR method while for $\gamma = 1$ we obtain the PR method.

Another efficient class of methods is known under the names *quasi–Newton* and *variable metric methods*, as typified by the *Davidon–Fletcher–Powell* (DFP) *algorithm* or the closely related *Broyden–Fletcher–Goldfarb–Shanno* (BFGS) *algorithm* [1,4].The above mentioned methods are very stable and they converge superlinearly. On the other hand, these methods require storage of order $n^2$. The most known methods in this class are expressed by:

$$x^{k+1} = x^k - \lambda^k B_k \nabla f(x^k), \quad k = 0, 1, 2, \ldots,$$

$$B_{k+1} = B_k + \frac{r^k (r^k)^\top}{(r^k)^\top q^k} - \frac{B_k q^k (q^k)^\top B_k}{(q^k)^\top B_k q^k} + \gamma (q^k)^\top B_k q^k u^k (u^k)^\top,$$

$$u^k = \frac{r^k}{(r^k)^\top q^k} - \frac{B_k q^k}{(q^k)^\top B_k q^k}, \qquad r^k = x^{k+1} - x^k, \qquad q^k = \nabla f(x^{k+1}) - \nabla f(x^k),$$

where $B_0$ is an arbitrary symmetric and positive definite matrix, usually taken to be the identity matrix and $\lambda^k$ is the optimal length in the direction $p^k = -B_k \nabla f(x^k)$. Now, for $\gamma = 0$ we obtain DFP method, while for $\gamma = 1$ we get the BFGS method.

In all cases, we have used an efficient line search procedure described in [4] which is based on Armijo's algorithm and it is applicable to any descent direction $p^k$ (see

also [9]). This procedure uses two parameters $\alpha, \beta \in (0,1)$ and can be implemented in two versions depending on the input value of the parameter $s$.

Next we present a high level description of this stepsize adaptation procedure, in which the corresponding parameters indicate: $x^0$ initial point, MIT the maximum number of iterations required and $\varepsilon$ the predetermined desired accuracy.

**Algorithm 3.1 Algorithm model with Armijo line search**

1. **Input** $\{f; x^0; \alpha, \beta \in (0,1); s \in \{0,1\}; m^* \in \mathbb{Z}; \text{MIT}; \varepsilon\}$.
2. **Set** $k = 0$.
3. **If** $\|\nabla f(x^k)\| \leqslant \varepsilon$ go to Step **6**. **Else**, compute a descent direction $p^k$.
4. **If** $s = 0$, set $M^* = \{m \in \mathbb{Z} \mid m \geqslant m^*\}$, and compute the stepsize
   (a) $\lambda^k = \beta^{m_k} = \arg\max_{m \in M^*} \left\{\beta^m \mid f(x^k + \beta^m p^k) - f(x^k) \leqslant \beta^m \alpha \langle \nabla f(x^k), p^k \rangle\right\}$.
   **Else** $(s = 1)$ compute the stepsize $\lambda^k = \beta^{m_k}$, where $m_k \in \mathbb{Z}$ is any integer such that
   (b) $f(x^k + \beta^{m_k} p^k) - f(x^k) \leqslant \beta^{m_k} \alpha \langle \nabla f(x^k), p^k \rangle$ and
   (c) $f(x^k + \beta^{m_k - 1} p^k) - f(x^k) > \beta^{m_k - 1} \alpha \langle \nabla f(x^k), p^k \rangle$.
5. **Set** $x^{k+1} = x^k + \lambda^k p^k$.
   **If** $k <$ MIT, replace $k$ by $k + 1$, and go to Step **3**. **Else** go to Step **6**.
6. **Output** $\{x^k; f(x^k); \nabla f(x^k)\}$.

The selection $s = 0$ is not very good for the conjugate gradient methods because, on average, it requires considerably more function evaluations than the selection $s = 1$. To this end, we have used $s = 1$ for the FR and PR algorithms. The selection $s = 0$ is normally used with variable metric methods algorithms, with $m^* = 0$ to ensure superlinear convergence. Thus we have used this value for the DFP and BFGS algorithms.

Furthermore, since in our case the objective function $f$ is bounded from below, the following subprocedure is used to find an $m_k$ satisfying Relations (b) and (c) of Step 4 of the Algorithm 3.1. This subprocedure uses the last used step length $\lambda^{k-1} = \beta^{m_{k-1}}$ as the starting point for the computation of the next step [4].

**Stepsize Subprocedure**

1. **If** $k = 0$, set $m' = m^*$. **Else** set $m' = m_{k-1}$
2. **If** $m_k = m'$ satisfies Relations (b) and (c) of Step 4 of Algorithm 3.1, stop.
3. **If** $m_k = m'$ satisfies (b) but not (c), replace $m'$ by $m' - 1$, and go to Step **2**.
   **If** $m_k = m'$ satisfies (c) but not (b), replace $m'$ by $m' + 1$, and go to Step **2**.

In practice, only a very small number of iterations of the above subprocedure are required to compute the Armijo stepsize. When a very small stepsize occurs for several iterations, causing slow convergence, the user can revert to setting $s = 0$ for one or two iterations.

The search strategy of Algorithm 3.1 converges to a minimizer of $f$ for both $s = 0$ and $s = 1$. This can be verified by a convergence theorem due to Polak *et al.* [4]. This theorem requires the search direction $p^k$ to be bounded from above, it imposes

a restriction on the angle between $\nabla f(x^k)$ and $p^k$ and states that Algorithm 3.1 is well defined in the sense that whenever $\nabla f(x^k) \neq 0$, the search for a stepsize $\lambda^k$ is a finite process, whether $s = 0$ or $s = 1$.

The above methods have been applied to the problem of Section 2 for three values of $\theta$: 0.05, 0.1 and 0.15. For these values, all methods have managed to calculate the family $L_6^1$. The behavior of each method is estimated from the total number of integrations of Eqs. (1) and the total number of function evaluations of the method that are necessary to compute the whole family. In Table 1 we present comparative numerical results for all the above mentioned methods.

Table 1

*Numbers of integrations (NI) and numbers of function evaluations (NFE).*

| | | FR | PR | DFP | BFGS |
|---|---|---|---|---|---|
| $\theta = 0.05$ | NI= | 296 | 302 | 278 | 281 |
| | NFE= | 8009379 | 9271024 | 25774 | 26069 |
| $\theta = 0.10$ | NI= | 266 | 276 | 274 | 278 |
| | NFE= | 7271126 | 8516422 | 25684 | 25899 |
| $\theta = 0.15$ | NI= | 288 | 290 | 271 | 284 |
| | NFE= | 7886236 | 9091579 | 31065 | 26744 |

It can be seen that, while the methods are almost equivalent regarding the number of integrations, DFP and BFGS clearly outperform the others in terms of the number of function evaluations. This results to a remarkable superiority for the latter methods in the total computational effort.

## References

[1] G.S. Androulakis and M.N. Vrahatis, J. Comput. Appl. Math. 72 (1996) 41.

[2] J.C. Gilbert and J. Nocedal, SIAM J. Optimization 2 (1992) 21.

[3] C.L. Goudas, Bull. Soc. Math. Gréce 2 (1961) 1.

[4] E. Polak, Optimization: Algorithms and Consistent Approximations, Springer-Verlag, New York, 1997.

[5] V.V. Radzievskii, Astron. Zh. 27 (1950) 250.

[6] O. Ragos, M.N. Vrahatis and G.S. Androulakis, In: Proc. of the Sixth International Colloquium on Differential Equations, D. Bainov, (ed.), 213, VSP International Science Publishers, Zeist, The Netherlands, 1996.

[7] O. Ragos and C.G. Zagouras, Celest. Mech. 44 (1988) 135.

[8] G. Shearing, Computation of Periodic Orbits in the Restricted Problem of Three-Bodies, Ph.D Thesis, University of Manchester, 1960.

[9] M.N. Vrahatis, G.S. Androulakis, J.N. Lambrinos and G.D. Magoulas, J. Comput. Appl. Math. 114 (2000) 367.

[10] M.N. Vrahatis, O. Ragos and G.S. Androulakis, In: Hamiltonian systems with three or more degrees of freedom, Vol. 533, C. Simó (ed.), 642, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.