Dynamics of Celestial Bodies, 185 - 188 *Contributed paper*

GLOBAL DATA CLUSTERING FOR PERIODIC ORBITS

G. A. TSIROGIANNIS, E. A. PERDIOS and V. V. MARKELLOS
Department of Engineering Sciences, University of Patras, GR–26500 Patras, Greece
E–mail: gtsirogianni@upatras.gr
E–mail: {e.perdios, markellos}@des.upatras.gr

Abstract. We outline a procedure for the classification (clustering) of periodic orbits into families. This complements the traditional grid search method and makes redundant the numerical continuation of the families with predictor–corrector techniques, that would otherwise be required for the determination of their continuously evolving properties (such as period and stability). Thus, substantial economy in computing time and user intervention is gained in achieving the final task of a full description of the families and their properties.

1. INTRODUCTION

The Grid Search Method (GSM) was introduced by Markellos et al. (1974) as a systematic way to compute complete networks of families of periodic orbits of non–integrable systems of two degrees of freedom, contained in a region of the relevant space of initial conditions. Since then it has been used for the numerical treatment of a number of dynamical systems, among which are notably the restricted three body problem (RTBP) and Hill's problem including variants (e.g Markellos et al., 1974; Markellos, 1974, 1975; Hénon, 1997, 2003; Kanavos et al., 2002; Russell, 2005).

We now present a data processing procedure for the clustering–classification of the periodic orbits obtained into families, without further numerical integrations. Addition of this procedure constitutes an improvement which makes the Grid Search Method complete as it makes redundant the numerical continuation of the families with predictor–corrector techniques that would otherwise be required for the determination of their continuously evolving properties (e.g. period and stability). Thus, substantial economy in computing time, and user intervention, is gained in achieving the final task of a complete description of the families and their properties. A full report on the improved Grid Search Method, incorporating the present clustering procedure and other improvements, with applications, will be published elsewhere.

2. GROUPING INTO FAMILIES

Our global classification (clustering) method exploits the elements–properties of symmetric periodic orbits (such as initial conditions and stability parameters) already computed by GSM, in order to group the orbits into one–parameter families. This

classification requires no new numerical integrations. The procedure gets as input the data set of all symmetric periodic orbits computed by GSM and provides as output the distinct families into which they are classified, based on the continuity property of their elements. In combination with GSM it can be seen as a global computation method for the families contained in the region of interest.

We recall that the initial conditions in the form (c, x_0) are enough to describe a symmetric periodic orbit. Using the GSM we compute sixteen values for each orbit: c, x_0, $\dot{y}_0$, $x(T/2)$, $\dot{x}(T/2)$, $\dot{y}(T/2)$, N, $T/2$, a_h, b_h, c_h, d_h, a_v, b_v, c_v, d_v where N is the multiplicity and $a_h, \ldots, d_h$, $a_v, \ldots, d_v$ are the horizontal and vertical stability parameters (see Hénon, 1965, 1973; Markellos, 1976). We suggest a method that produces as output a family (ordered group) of symmetric periodic orbits such that when moving from one end of the final set of orbits to the other, all the above parameters change continuously (except N). If we apply the clustering method several times we are able to compute–identify all the families contained in the region under consideration. The basic idea is the following: starting with one member orbit of a family we include new orbits to the family, one at a time, having the property of the smoothest possible evolution of the parameters among all the other possible (neighboring) orbits. By "orbit" we mean planar symmetric periodic orbit, or the point representing it in the space of initial conditions equivalently the (c, x_0) plane.

3. FIRST PHASE OF DATA CLUSTERING METHOD

Generally, we do not have any information about the beginning or the location of the families in the (c, x_0) plane. We choose one orbit from our data set arbitrarily and seek to find the family this orbit belongs to. It is known that starting from a given orbit of a family, we can move along the family in two and only two directions (see Hénon, 1997, proposition 2.5.1), thus in this first phase we will compute an initial segment of one of the two parts of the family distinguished by the starting orbit p_0.

The goal of this phase is to find the k first additional member orbits contained in the family whose characteristic family curve passes through the initial point p_0, where k is a small integer (e.g. $k = 5$). Since we have no information about the family except the initial orbit p_0, we use a weak criterion for the inclusion of the first k orbits: the next included orbit will be the orbit that has the minimum distance from the current orbit in $(c, x_0, \dot{y}_0)$. The choice of subspace $(c, x_0, \dot{y}_0)$ was found to give excellent results with very few floating point operations for the computation of distances.

Starting from point p_0 we compute all the distances in $(c, x_0, \dot{y}_0)$ for those orbits that lie inside a sphere with center p_0 and radius a small constant MD, where MD is taken to be sufficiently larger than the step of the mesh of GSM to insure that some member orbits will lie inside the sphere. We choose as next orbit the orbit with the minimum distance from p_0. We set this orbit as the second member orbit of the family, named p_1. We repeat with new center the point p_1 to find p_2. Working in this way we compute a segment of the family consisting of orbits $\{p_0, p_1, \ldots p_k\}$.

4. SECOND PHASE OF DATA CLUSTERING METHOD

In this phase we compute the remaining part of the family up to its terminations (in both directions). We assume that a family terminates according to Strömgren's

termination principle, that is by closing upon itself or when: $\lim(M + T + 1/d) = \infty$, where M is the maximum dimension of an orbit, T is the period and d the minimum distance from the major bodies and equilibrium points. The main idea is to add new member orbits to the already computed k members of the first phase, having as criterion the smoothest evolution of the properties–elements of the member orbits. The partial knowledge of the family from the first phase allows us to include orbits (points) that have minimum distance from the expected characteristic curve based on extrapolation of its already established part comprising the included orbits.

If we plot one of the parameters characterizing the orbit properties versus one other of these parameters we get (discrete point plots of) diagrams as shown in Fig. 1. In order to make the first step to apply the above criterion, we need to make an estimation of the expected characteristic curve of the family based on the already computed part. We need a way to transform all the characteristic curves into diagrams of (uni–valued) functions. To this end we employ an approximation of the arc length:

$$F(p_i) = \sum_{j=1}^{i} D_j^{(c,x)}, \tag{1}$$

where $D_j^{(c,x)}$ is the Euclidean distance in the (c, x_0) plane of two neighboring member orbits p_{i-1}, p_i. The geometrical meaning of this function $F(p_i)$ is that it computes the total arc length from member orbit p_0 up to orbit p_i as a sum of all the internal sub–distances $\{p_0, p_1\}, \{p_1, p_2\}, \ldots, \{p_{i-2}, p_{i-1}\}, \{p_{i-1}, p_i\}$. This function is strictly increasing, thus if we plot any of the parameters that describe an orbit (except N) as a function of $F(p_i)$ we get a diagram of a single–valued function (Fig. 3).

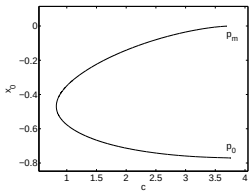

Figure 1: Characteristic curve of a family.

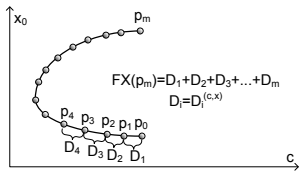

Figure 2: Arc–length function.

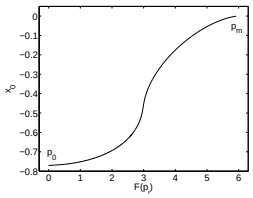

Figure 3: Characteristic curve after transformation.

Our goal is to obtain an approximating function of each characteristic curve and to use this approximating function for the extrapolation. We are seeking to add new orbits to the already computed part of the family (first phase), which will provide the closest real characteristic curve compared with the estimated. The concept "smooth" evolution is qualitative rather than quantitative. In order to use this "smoothness" criterion on a digital computer we proceed as follows: we make a numerical approximation of the function of the characteristic curve (see Fig. 3) using splines based on each of the parameters (c, x_0, $x_{T/2}$, $\dot{y}_0$, $T/2$) of the already included part of the family. This is a subset of the parameters that was found to give successful results. We call their approximated functions $G^{c}_{approx}(F(p_i))$, $G^{x_0}_{approx}(F(p_i))$, $G^{x_{T/2}}_{approx}(F(p_i))$, $G^{\dot{y}_0}_{approx}(F(p_i))$, $G^{T/2}_{approx}(F(p_i))$, respectively.

The algorithm of the second phase works as follows: Using the k first member orbits of the first phase we compute the above numerical approximations for $i = k$. Having the p_k orbit as center and radius a constant (for example five times the mesh step of GSM) we locate the possible new orbits to be included. For each possible new orbit inside this circle we compute the $G^{c}_{approx}(F(p_i))$, $G^{x_0}_{approx}(F(p_i))$, $G^{x_{T/2}}_{approx}(F(p_i))$, $G^{\dot{y}_0}_{approx}(F(p_i))$, $G^{T/2}_{approx}(F(p_i))$ functional values for $i = k + 1$. Due to the "smoothness" criterion the real member orbit we seek will have the smallest difference between estimations and real values (from GSM), among the other possible orbits.

It is possible that the family has terminated and none of the orbits should be included (we detect this by a special procedure described in the full paper). If we have not reached a termination, in order to locate the next member orbits we use the k last members to construct the approximation functions and a new set of possible orbits based on the last included orbit. To visualize the procedure, in Fig. 4 we show the "moving" group of k orbits for every step. When a segment has terminated we remove the orbits of this segment (included in the second phase) from the input data and, by using the k first members of the first phase, we apply the second phase again, in the opposite direction, to obtain the complete family.

The computational cost of the clustering procedure for a family represented by n orbits in the data set is: $cnm = \mathcal{O}(m)$, $m >> n >> c$, where m is the total number of orbits in the data set and c is a constant. This can be reduced by using spatial data structures such as R trees (see Guttman, 1984).

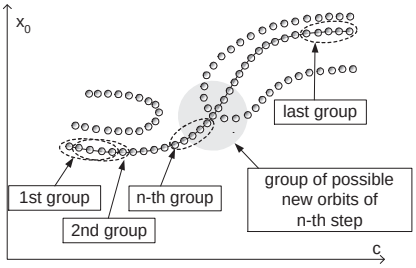

Figure 4: Sub–groups of second phase.

References

Guttman, A.: 1984, R-Trees: A Dynamic Index Structure for Spatial Searching, SIGMOD'84, Proceedings of Annual Meeting, Boston, Massachusetts, June 18-21.

Hénon, M.: 1965, *Annales d' Astrophysique*, **28**, 499.

Hénon, M.: 1973, *Astronomy and Astrophysics*, **28**, 415.

Hénon, M.: 1997, Generating families in the restricted three–body problem, Lecture Notes in Physics, Springer, Berlin.

Hénon, M.: 2003, *Celestial Mechanics and Dynamical Astronomy*, **9**, 223.

Kanavos, S. S., Markellos, V. V., Perdios, E. A., Douskos, C. N.: 2002, *Earth, Moon and Planets*, **91**, 223.

Markellos, V. V., Black, W., Moran, P. E.: 1974, *Celestial Mechanics*, **9**, 507.

Markellos, V. V.: 1974, *Celestial Mechanics*, **10**, 87.

Markellos, V. V.: 1975, *Celestial Mechanics*, **12**, 215.

Markellos, V. V.: 1976, *Astrophysics and Space Science*, **43**, 449.

Russell, R. P: 2005, Global search for planar and three-dimensional periodic orbits near Europa, AAS/AIAA Astrodynamics Specialist Conference, Lake Tahoe, California, August 7-11, 645–672.