Using Evolution Strategies to Find a Dynamical Model of the M81 Triplet

Juan Carlos Gomez[1], Olac Fuentes[1] Lia Athanassoula[2], and Albert Bosma[2]

[1] Instituto Nacional de Astrofisica, Optica y Electronica, Luis Enrique Erro 1, Tonantzintla, Puebla, Mexico,
jcgc@inaoep.mx, fuentes@inaoep.mx,

[2] Observatoire de Marseille, 2 Place le Verrier, 13248, Cedex 4, Marseille, France

Abstract. In this work we present Evolution Strategies (ES) as an efficient optimization method for dynamic modelling of the main interacting group of three galaxies in M81. The M81 group is one of the nearest groups of galaxies; its biggest galaxy, M81, sits in the core of the group together with its two companions M82 and NGC3077. The interaction among these three galaxies is very well defined on an image taken in HI. In this first attempt we use non-self-gravitating simulations for modelling dynamically the group; even with this restriction our method reproduces the density distribution of the three galaxies with great precision. Results presented here show that ES is an excellent method to find an accurate model of groups of interacting galaxies, where a global search of a large number of real-valued parameters needs to be performed.

1 Introduction

It is now established that galaxies are not "island universes", but rather interact with each other in pairs or in small or big groups. Interactions can form spirals, bars, warps, rings and bridges. Thus, observing the morphological and kinematic results of an interaction can give us crucial information about the interaction scenario.

In principle, the parameters of a dynamic model that describes the interactions among a group of galaxies can be derived from multiple observations of the system at different times (as was done for planets in the solar system a long time ago). However, given the enormous time scales involved in galactic interactions, this approach is completely unfeasible. Two images of an interacting system taken several centuries apart would be virtually undistinguishable. Thus, the problem of modelling the system consists of finding the set of initial conditions of the system and the interaction time that result in the current observed configuration.

The problem of finding the right parameters for modelling the interaction of a given system of galaxies can be posed as an optimization problem [5]. As input, we have an image of the interacting system and sometimes a velocity field, obtained from spectroscopy. To construct a dynamical model, we need to simulate the system of galaxies, giving certain set of initial conditions (parameters) to a simulator program. These initial conditions are the basis to understand the dynamical nature of the system. Then a simulation program gives a projected surface density map and line-of-sight velocities. These can be compared to the corresponding observed quantities, and then the best model is the one that minimizes the difference between maps. In the end, we have a set of initial conditions that given to a simulator program, can reproduce the interaction among galaxies.

We will here use Evolution Strategies (ES) as the optimization algorithm to find the minimum difference between simulated and observational images [1, 3]; ES is a good method that works efficiently as a global search algorithm and with continuous spaces. Since most of the parameters of interacting systems of galaxies are continuous, this constitutes a clear incentive for trying out ES.

In this work we chose the M81 triplet as the interacting system to be studied [6]. The M81 group is one of the nearest groups of galaxies. Its biggest galaxy, M81, sits in the core of the group together with its two nearby companions M82 (in the upper part of the image in Figure 3) and NGC3077 (in the lower part of the image in Figure 3). This group has a very well defined interaction scenario, the main galaxy has a spiral shape and forms clear tails with its two companions; also, the interaction is only in the outer part of the galaxies, this facilitates the use of non-self-consistent simulation. All of this made this group an ideal candidate to test ES.

The organization of the remainder of this paper is as follows: Section 2 contains a brief description of the method, the implementation is presented in Section 3, the results are given in Section 4 and conclusions and future work are presented in Section 5.

2 The Method

Evolution Strategies (ES) [1, 3] is a technique for finding the global minimum of a function with a large number of variables in a continuous space. We start by choosing K individuals, each characterized by an object parameter vector $\mathbf{O}$ and a corresponding strategy parameter vector $\mathbf{S}$:

$$\mathbf{O}_i = \langle q_{1,i}, q_{2,i}, \cdots, q_{L,i} \rangle \; i = 1, \cdots, K \tag{1}$$

$$\mathbf{S}_i = \langle \sigma_{1,i}, \sigma_{2,i}, \cdots, \sigma_{L,i} \rangle \; i = 1, \cdots, K \tag{2}$$

In the first generation, the elements of the $\mathbf{O}$ and $\mathbf{S}$ vectors, can be chosen either totally at random, or with some help from previous knowledge about the problem to be solved. Each of the K individuals (set of parameters) must be evaluated according to a fitness function. The fitness function is what we need to minimize.

The next sept is to produce a new population applying the genetic operators cross-over and mutation. For cross-over, two individuals (parents) are chosen at random, and then we create two new individuals (children) by combining the parameters of the two parents. Mutation is applied to the individuals resulting from the cross-over operation; each element of the new individual is calculated from the old individual using the simple equation:

$$q_{\mathrm{mut}} = q_j + N(0, \sigma_j) \tag{3}$$

where $N(0, \sigma_j)$ is a random number obtained from a normal distribution with a zero mean an a standard deviation σ_j, which is given from the strategy parameter vector. The process of cross-over and mutation is repeated until the population converges to a suitable solution.

3 Modelling the Interacting Group M81

In the beginning we had one image of the group of galaxies taken in HI (neutral hydrogen) by Yun [6], in addition, we had some physical information about the group, also from Yun. We used the HI image to approximate the model because in that image the interaction among galaxies is clearly defined. The HI image was traslated and resized in such a way that M81, the main galaxy, was centered in the image. Then, we calculated the distances (in pixels) between the central points of each pair of galaxies.

The next step was to define the density map of the image. It was done simply by putting a grid over the image and counting the pixels that have a value greater than 0 in each cell of the grid. Then we established a relation between the total amount of pixels in the image and the number of particles we are going to use in the simulations, in such a way that we have a portion of the total mass in each cell. With this we have a density matrix that we can use to measure the fitness of the simulations, as was established by [5]. In this case we have used a 48x48 grid. As an example, Figure 1 shows a 6x6 mass density matrix of an artificial image that represent two interacting galaxies. The values in each cell represent the number of particles.

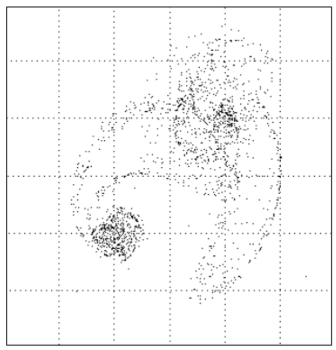
Figure 2 shows a diagram of the solution process. First we create the individuals with ES, then each individual is used as an input for the simulator program, the program returns a mass distribution that is used to be compared with the observational data; if the fitness is good enough or the maximum number of iterations has been reached, the process stops, otherwise we create a new population using genetic operators and return to step 1.

Applying ES to approximate a model for the group of galaxies M81 was done as follows: we use a population with 30 individuals per iteration ($K = 30$), where each individual has the following structure:

$$\mathbf{O}_j = \langle z_2, z_3, V_{x1}, V_{y1}, V_{x2}, V_{y2}, V_{x3}, V_{y3}, i_1, PA_1, i_2, PA_2, i_3, PA_3, \ldots \\ m_1, m_2, m_3, t\rangle, j = 1, 2, \ldots, K$$

$$\mathbf{S}_j = \langle \sigma_{i,1}, \sigma_{i,2}, \sigma_{i,3}, \sigma_{i,4}, \sigma_{i,5}, \sigma_{i,6}, \sigma_{i,7}, \sigma_{i,8}, \sigma_{i,9}, \sigma_{i,10}, \sigma_{i,11}, \sigma_{i,12}, \sigma_{i,13}, \ldots \\ \sigma_{i,14}, \sigma_{i,15}, \sigma_{i,16}, \sigma_{i,17}, \sigma_{i,18}\rangle, j = 1, 2, \ldots, K$$

where z_2 and z_3 represent the distances in the line of sight between the main galaxy in the center of the image and its companions; V_x and V_y the velocities



0	0	2	40	31	0
0	2	59	279	152	0
0	39	53	262	184	4
0	157	167	49	88	1
0	148	159	51	44	1
0	1	0	11	0	0

Fig. 1. Artificial image and its density matrix.

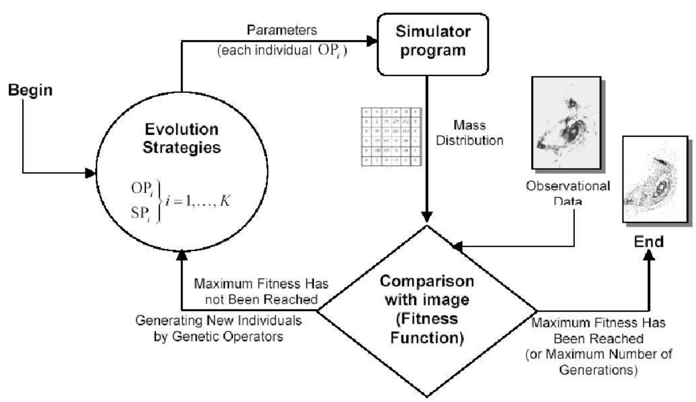


Fig. 2. Block diagram of the solution process.

in the image plane; i's the inclination angles (in the x axis); PA's the position angles (in the z axis); m's are the masses of the galaxies and t represents the total time of interaction. Subindex 1 is for M81 galaxy, 2 for M82 and 3 for NGC3077.

The first generation is created using random values, but with heuristic physical information as reference[6]: first, the masses and the time can not be negative (for physical reasons); since the main galaxy is perturbed only in the outer part, and the two companions are compact, then the main galaxy must have a predominant mass; separations in the line of sight must be small enough to allow perturbations in the galaxies; velocities in the image plane must be also on a range that allows the galaxy discs to be perturbed.

Each individual in the population, in this case each set of initial conditions, is used as input for the simulator program. With the simulation we obtain a surface density map, and multiply it by a mass scale factor. Then the result of each simulation is evaluated with a fitness function. The fitness function compares the density maps of the simulated and original images using the Kullback-Leibler distance [2]:

$$F_d = \left[\sum \left\{ (m_{i,j}^{obs} + m_\epsilon) \left| ln \left(\frac{m_{i,j}^{obs} + m_\epsilon}{m_{i,j}^{sim} + m_\epsilon} \right) \right| \right\} \right]^{-1} \tag{4}$$

where $m_{i,j}^{sim}$ is the total mass in the cell under consideration for the simulation, $m_{i,j}^{obs}$ is the same quantity for the observations, m_ϵ is a very small quantity to avoid problems in regions with zero density, and the sum is carried over all the cells.

Once we have evaluated the first generation, it is necessary to create a new population using the genetic operators as seen in the previous section. Following the ideas described in [1], we use dynamic mutation and first mutate the strategy parameter vector using the equation:

$$\sigma_{mut} = \sigma exp \left[\frac{C_v \delta R_1}{\sqrt{2K}} + \frac{C_v \delta R_2}{\sqrt{2K}} \right] \tag{5}$$

where σ is the value before mutation, R_1 and R_2 are random numbers obtained from a normal distribution with zero mean and a standard deviation equal to the σ before mutation, and C_v and δ are numerical constants. We found the values $C_v = 0.9$ and $\delta = 0.1$ to be the most appropriate for this particular problem.

The cross-over operator is uniform: two individuals are randomly selected from the original population in such a way that each individual is used exactly once, and each parameter in the two individuals has the same probability to be selected to form two new individuals.

The mutation operator is performed by following the classical process: multiplying a random number obtained from a normal distribution with zero mean and standard deviation σ (taken from strategy parameter vector $\mathbf{S}$).

z_2	z_3	V_{x1}	V_{y1}	V_{x2}	V_{y2}	V_{x3}	V_{y3}	i_1	PA_1	i_2	PA_2	i_3	PA_3	m_1	m_2	m_3	t
62.05	11.14	3.17	1.59	-61.91	41.78	-168.75	-0.90	44.90	113.37	38.62	32.58	53.83	232.87	19.47	1.04	1.06	812

Table 1. Parameters to produce the simulation in Figure 1

To form the children population we first apply cross-over to the whole population, and then mutation to the resulting population. Then we merge both populations (parent and children), select the K best individuals from this merged population, and use that set as input for the next iteration.

For the simulation we use the test particle approach [5]. In this approximation, the mass of each galaxy is assumed to be concentrated in a single point in its center, while the disc, which responds most to the interaction, is represented by test particles, initially in co-planar circular orbits around the center of the galaxy. This approach is very fast and thus allows us to run the very large number of simulations necessary for tackling this problem. Furthermore, in our case, the galaxies are not inter-penetrating and thus they are perturbed only in their outer parts, making the test particle approach fairly adequate. We used a total of 4000 particles, 2000 for the main galaxy and 1000 for each companion.

4 Results

After 200 generations of evolution, yielding a total of 6000 simulations, we obtained a good artificial image, that matches very well the original image. The images in Figure 3 show the best simulation reached and the original HI image. Table 1 shows the corresponding parameters for the best simulation. The total time required to reach that model was 12 hours in a Sun Ultra 10 workstation.

In Table 1 index 1 corresponds to M81, 2 to M82 and 3 to NGC3077. The quantities are given in program units and we used the following scale to convert to physical units: the unit of length is taken to be 1kpc, the unit of time 1×10^6yr, the unit of mass $1 \times 10^{10} M_\odot$ and the angles are given in degrees.

In the simulated image we can see the spiral shape of the main galaxy M81, the tail that joins M81 with NGC3077 in the lower part of the image, the tail in the upper part of NGC3077, and part of the mass concentration in the upper part of M81.

The density was fairly well reproduced in this way, we reached a maximum of 0.45 with the fitness function (1 corresponds to a perfect match). Obviously, reaching a perfect match between images is not possible because of the limited number of particles used in simulations.

In this first attempt we used only data from the density map, without considering the velocity distribution, so the resulting velocities were inaccurate. To test the velocity match, we took some real velocity data in the line of sight from certain parts of the system, but in a comparison with the simulated velocity field, it does not match very well. So, in future work we are planning to introduce that velocity information to improve the velocity estimation.

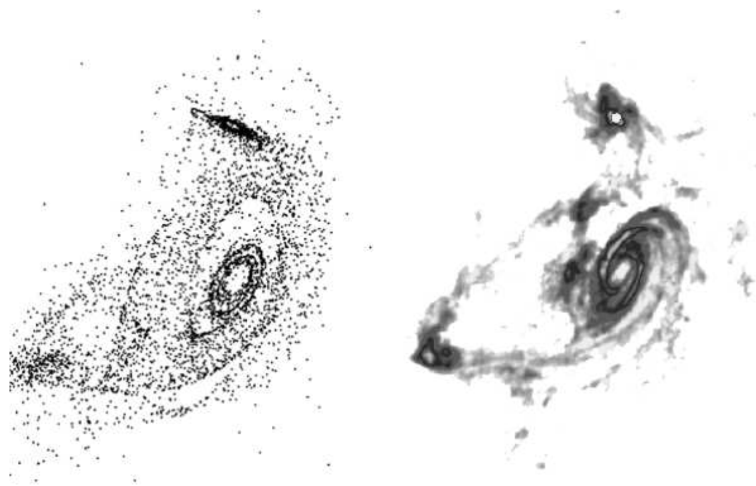


Fig. 3. Simulated and HI images for M81 group.

5 Conclusions and Future Work

In this work we presented an efficient method, based on ES, to approximate a dynamical model for the M81 triplet.

Even with the several simplifying assumptions done in simulations, searching with ES has demonstrated to be an excellent method for optimization problems where a global exploration of continuous parameters spaces is needed; ES could find a good simulation that matches very well the HI density distribution in this problem. We are planning to extend the application of ES to the study of other interacting systems.

On the other hand, and in order to improve the method, the possibility of implementing a parallelization of the ES could be considered with the purpose of reducing the computing time required for simulations. We are planning to include physical information about velocity trying to reproduce the velocity field of the image. Also, methods based on self-gravitating N-body simulations can be used to improve the match between simulations and the HI density distribution.

Also we will implement new hybrid algorithms to accelerate the convergence. Two main algorithms are in consideration: using ES with a traditional optimization algorithm (Quasi-Newton) and combine ES with Locally Weighted Linear Regression. With this we hope the convergence can be reached faster, reducing the total number of simulations. Another possibility is trying to solve this problem with a different alternative algorithm such as Simulated Annealing.

References

1. Back, T., Hoffmeister, F., Schwefel, H. P., A Survey of Evolution Strategies. Proceedings of the Fourth International Conference on Genetic Algorithms. (1991) 2–9
2. Kullback, S., Leibler, On Information and Sufficiency. R. A., Ann. of Math. St. (1951) 79–86
3. Rechenberg, I., Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution. Stuttgart: Fromman-Holzboog. (1973)
4. Wahde, M., Donner, K. J., Determination of the Orbital Parameters of the M51 System Using a Genetic Algorithm. Astronomy and Astrophysics. (2001) 115–124
5. Wahde, M., Determination of Orbital Parameters of Interacting Galaxies Using a Genetic Algorithm. Astronomy & Astrophysics Supplement Series. (1998) 417–429
6. Yun, M., Tidal Interactions in M81 Group. IAU Symp. 186, Galaxy Interactions at Low and High Redshift. (1999) 81