Determining Orbital Elements of Extrasolar Planets by Evolution Strategies

Andreas M. Chwatal and Günther R. Raidl*

Vienna University of Technology
Vienna, Austria
{chwatal|raidl}@ads.tuwien.ac.at

Abstract. After the detection of the first extrasolar planet (exoplanet) more than one decade ago we currently know about more than 200 planets around other stars and there are about twenty multi-planet systems. Today's most commonly used technique for identifying such exoplanets is based on radial velocity measurements of the star. Due to the increasing time span and accuracy of the respective observations, the measured data samples will more and more contain indications of multiple planetary companions. Unfortunately, especially in multi-planet systems, the determination of these exoplanets' orbital elements is a rather complex and computationally expensive data analysis task. It is usually formulated as an optimization problem in which the orbital parameters yielding minimal residues w.r.t. the measured data are to be found. Very likely, improved algorithms for (approximately) solving this problem will enable the detection of more complex systems. In this paper we describe a specialized evolution strategy for approaching this problem.

1 Introduction

One of the most popular methods for planet discovery is based on radial velocity (RV) measurements of the central star, performed by high resolution spectroscopy. The *stellar wobble* induced by the gravitational influence of planetary companions manifests in Doppler shifts of the spectral absorption lines. A series of RV-measurements in principle makes it possible to determine most of the orbital elements of existing planets: the semi-major axis ("distance") a, the minimal planetary mass $\tilde{m}$ [1], the orbital eccentricity e, the argument of perigee ω, and the periastron time t_0. Only the right ascension of the ascending node Ω and the inclination i are impossible to derive by this method.

To compute the RV of a star resulting from a specific set of orbital elements, the Kepler equation $E - e\sin E = \frac{2\pi(t+t_0)}{T}$ needs to be solved numerically in advance. Using the equation $\tan\frac{\upsilon}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{E}{2}$, the radial velocity of the planet is given by $\frac{dz}{dt} = K\big[\cos(\upsilon+\omega) + e\cos\omega\big]$, where $K = \frac{2\pi}{T}\frac{a\sin i}{\sqrt{1-e^2}}$.

* This work is supported by the European RTN ADONET under grant 504438.

[1] minimal due to the unknown inclination i. Therefore $\tilde{m}$ is the actual parameter to be determined, but we denote this term by $\tilde{m}$ for simplicity.

T denotes the revolution period of the planet and is according to Kepler's third law given by $T^2 = \frac{4\pi^2 a^3}{G\cdot(M+m)}$ (where G denotes the gravitational constant). The radial velocity of a star (with stellar mass M) originating from the considered planet is then given by $v_z^* = -\frac{1}{M} \cdot \frac{dz}{dt}$.

Now, a simple RV-model can be formed by summing up the terms resulting from the respective planets, but the discrepancies to a two-body system can be better approximated by the use of Jacobi-coordinates (as for instance explained in [1]). For outer planets this approach also takes into account the mass of the more interior planets. Thus the total RV of the star is given by $v_z^* = \sum_i \xi_i v_i$ where $\xi_i = \frac{m_i}{M+\sum_{j=0}^{i} m_j}$. By N denoting the total number of planets in a system, the determination of a fit to the observed radial velocities is basically a minimization problem in roughly $P = N \cdot 5$ parameters. Additional parameters arise from offsets in the data-samples (see sec. 3.1). Having RV-Measurements at K distinct points, the objective function is given by

$$\chi^2 := \sum_{k=0}^{K} \left(\frac{\triangle_k}{\sigma_k}\right)^2. \tag{1}$$

where $\triangle_k$ denotes the difference between the model and the observed radial velocity at time i. Sometimes the relative value $\chi^2_{\text{red.}} = \chi^2/\nu$, where $\nu = K - P - 1$ (P denoting the total number of fitting-parameters) is also used.

1.1 Traditional Data Analysis

The traditional data analysis is largely based on the Fourier transform. From this one gets the frequency spectrum of the RV-signal, which is a good starting point for sinusoidal data fits which can later on be improved to Keplerians. In the case of multiple systems this method is typically applied iteratively. After determining the characteristics of one planet, its contribution to the RV data is subtracted. With this residual data, the next planet is considered. This approach is often referred to as *cleaning algorithm*, see [2]. Obviously, such an iterative strategy can be highly misleading. Furthermore, the Fourier-based analysis has disadvantages like the initial assumption of circular orbits. Consequently a simultaneous fitting of Keplerians is a desired and more promising alternative.

2 Previous Work

The first application of Genetic Algorithms (GAs) in the area of exoplanet research has been mentioned in [3], where GAs with classical binary encodings are used to search for the neighborhood of a local minimum. In [4] a GA routine is used to verify the uniqueness of the 3-planet-solution to the υ-And System. In [5] the authors use GAs in combination with the Levenberg-Marquardt algorithm to drive numerical integrations to analyse a strongly interacting system. In [1] and [6] the authors describe the use of a penalty for unstable configurations. They use MEGNO (see [7]) for the fast indication of possible instabilities.

The authors mention, that *"[..]GAs are still not very popular but they have been proven to be very useful for finding good starting points for the precise gradient methods of minimization like e.g. the well known Levenberg-Marquardt scheme"* [1]. To summarize the hitherto efforts, standard GAs with binary encodings have been applied to the problem, and they have been augmented by gradient-based local methods and penalties, but no more problem-adequate representation and variation operators have been studied so far. In fact, a binary representation of continuous parameters is nowadays usually not considered a meaningful encoding for most problems due to its weak locality [8]. In [9], another GA, called *Stakanof* method, is mentioned and applied to the μ Ara system with four planets. To our knowledge, however, no further details are yet published on this GA.

3 Evolution Strategy: ExOD-ES

As, for instance, pointed out in [10] or [8], many researchers consider *evolution strategies* (ES) to be more effective in solving continuous parameter optimization problems than genetic algorithms. This is mainly due to the weak locality of the GAs binary encoding as well as the absence of self-adaptation mechanisms to exploit the topological structure of the fitness-landscapes. In the following we present a problem-specific evolution strategy, essentially following [11].

The approach is based on a (μ, λ)-ES with self-adaptation of strategy parameters [12] where μ denotes the size of the population and λ the number of offsprings created in each generation.

In evolution strategies mutation is considered the primary operator and selection is traditionally done in a deterministic way by choosing the μ best offsprings. Mutation (eq. 2) is performed by adding a Gaussian-distributed random number to each parameter, where the standard deviation is given by a strategy parameter σ_i, associated with each parameter. These strategy parameters are optimized themselves by also undergoing mutation. They are modified by the multiplication with a log-normally distributed random value (eq. 3).

$$x_i' = x_i + N_i(0, \sigma_i') \tag{2}$$

$$\sigma_i' = \sigma_i \cdot e^{N(0,\tau_0)+N_i(0,\tau)} \tag{3}$$

$N(0, \tau_0)$ is a normally-distributed random value that is calculated only once for each candidate solution, while $N_i(0, \tau)$ is sampled individually for each parameter. The corresponding standard deviations are less critical and chosen as usual: $\tau_0 \propto 1/\sqrt{2\sqrt{n}}$. For a more detailed description of these standard elements of evolution strategies, the reader is referred to e.g. [8], or [13]. In the following we describe the special properties of the ES for the determination of the orbital elements of exoplanets, subsequently referred to as ExOD-ES[2]

[2] ExOD-ES: Exoplanet Orbit Determination by Evolution Strategies.

3.1 Encoding

Each candidate solution represents a whole planetary system by orbital element vectors $p_i = (\tilde{m}_i, a_i, \omega_i, e_i, t_{0,i})$, $i = 1, \ldots, N$, where N is the pre-specified number of planets. As several data sets may contain offsets in their radial velocities, which may result from the movement of the star through space or instrument calibration, each sample involved in the fit introduces one additional parameter $rv_{\text{offset},d}$ $d = 1, \ldots, D$, with D denoting the total number of data samples. Thus we altogether have $P = N \cdot 5 + D$ parameters, and for each we additionally maintain a corresponding strategy parameter σ_d.

To avoid undesired symmetries in the representation, which degrade performance of the optimization substantially, we ensure that the orbital element vectors are always sorted according to ascending order of semi-major axes a_i.

3.2 Mutation

The classical (μ, λ)-mutation modifies all parameters of the offspring individuals before selection. In our case it turned out to be better to merely mutate only one parameter-subset corresponding to one planet, which can be seen analogous to the evolution of subsystems in nature. In the following we omit the index i denoting the respective planets for simplicity.

For better convergence rates it is profitable to further consider some mutual dependencies. If ω has been changed by an angle ϕ we update t_0 by eq. 4 in order to keep the initial phase information, and thus not bias t_0 by the variation of another parameter:

$$t_0 = t_0 + T(a) \cdot \frac{\phi}{2\pi}. \tag{4}$$

For similar reasons we adjust t_0 after changes of a:

$$t_0 = t_0 + 1/2 \cdot (T(a) - T(a')). \tag{5}$$

This especially assists the development of long period planets as changes of a are distributed more uniformly over the time interval.

Restriction of the solution space: To improve the overall convergence properties of the algorithm we restrict the solution space in such a way that unstable systems, which cannot persist for more than a couple of centuries, will almost never be created by the variation operators.

Of course there remains the low probability of observing such systems that are currently undergoing some change, i.e. there are strong gravitational interactions between the planets and therefore the Keplerian model of the orbits which basically treats the system as a sum of two-body problems is not valid any more. In this case it is necessary to perform numerical integration. To determine meaningful starting values, however, Keplerian models are still useful. Consequently it is important to be able to adjust the impacts of the stability criterion and thus the amount of restriction.

Hill-Stability Criterion: A very simple stability-criterion dates back to G.W. Hill , and states that the toruses defined by the regions around planetary orbits up to distances from the orbital trajectory of Hill-radius

$$r_H = a\left(\frac{m}{3\cdot M}\right)^{1/3} \tag{6}$$

of two planets should not overlap. More formally, for two planets i and $i+1$ this can be written as

$$|a_i - a_{i+1}| > \zeta\big[r_H(a_i, m_i) + r_H(a_{i+1}, m_{i+1})\big], \tag{7}$$

where ζ is a parameter that, for instance, guarantees stability for timescales as the age of the solar system (≈4.5 billion years) when ranging from 11 to 13. The physical meaning of the Hill-radius can be interpreted as the border of the region where the gravitational influence of the planet clearly dominates the effects from the central star. If the toruses of the so called *Hill-spheres* of two planets overlap, it is just a question of time, when a first/next dynamical interaction will occur. ζ from eq. 7 is thus a parameter of ExOD-ES, forcing two distinct orbits to have a specific minimal distance. Moderate values like $\zeta = 3, \ldots, 5$ turned out to be sufficient for the algorithm.

Mutation Operator: The mutation operator is parameterized by the surrounding semi-major axes (in this context denoted by a^- and a^+) and the respective weighted Hill-radii ($\tilde{r}_H^- = \zeta \cdot r_H^-$ and $\tilde{r}_H^+ = \zeta \cdot r_H^+$). If there are no surrounding planets, the parameters $a_{\min}$ and $a_{\max}$, which can be derived from the input data, are used respectively. The mutation of the semi-major axes of one planet is performed as follows:

$$\eta = N(0, \sigma) \tag{8}$$

$$a' = \begin{cases} a + \eta & \text{if} \quad a + \tilde{r}_H + \eta < a^+ - \tilde{r}_H^+ \text{ and } a - \tilde{r}_H + \eta > a^- + \tilde{r}_H^- \\ a & \text{if} \quad a + \eta < a^+ - \tilde{r}_H^+ \text{ and } a + \eta > a^- + \tilde{r}_H^- \\ & \text{but} \quad a + \tilde{r}_H + \eta > a^+ - \tilde{r}_H^+ \text{ and } a - \tilde{r}_H + \eta < a^- + \tilde{r}_H^- \\ a + \eta + 2\cdot\tilde{r}_H^+ & \text{if} \quad a + \eta > a^+ - \tilde{r}_H^+ \\ a + \eta - 2\cdot\tilde{r}_H^- & \text{if} \quad a + \eta < a^- + \tilde{r}_H^- \end{cases} \tag{9}$$

In the first case there is no overlap of the Hill-toruses of the planet under consideration with the ones of the neighboring planets. In this case there is no difference to the conventional mutation. Otherwise there are different grades of violating the Hill-criterion. If the Hill-torus overlaps with the one of a neighbor, then we keep the initial value. The other two cases treat the situations of stronger violation of the criterion, i.e. where the planet itself would intrude one neighboring Hill-torus. Here the mutation operator enables to "tunnel" through the neighboring Hill-torus and therefore takes up another position (according to its distance). This is a very important mechanism to escape local optima. Suppose the situation that the most planets are well determined, but on one position a planet, which is difficult to determine, is still missing in the model. The above mechanism enables a quick adoption of an individual to such a better model.

Unfortunately it is very likely that the objective value will decrease in such situations, even when the planet moved closer to the correct position. This is because the latter individual can be expected not to be well adapted to the new local minima. To counteract this shortcoming, we modify the conventional selection such that it supports the evolution of such new subsystems. Therefore we introduce a new parameter γ for each individual, which indicates the number of generations this "path of evolution" should survive. During the following iterations only the "new" planet is modified in order to make this individual competitive with the other ones of the population as soon as possible. After such a mutation we initially set $\gamma = \gamma_{\max}$, and then in each generational step the γ-parameter of the respective individuals is decremented by one (if $\gamma > 0$). The related modifications of the selection mechanism are described in sec. 3.4.

There is an additional mechanism that accomplishes the same purpose: the planet that is least important for the model is removed from it, and a new planet is created randomly on another position. Again we set a $\gamma > 0$ to enforce an evolution of this planet.

The mutation of the mass m requires an additional mechanism to guarantee valid solutions (eq. 10). If the mutation of the mass would violate the Hill-criterion we keep the maximal possible value and perform the conventional mutation otherwise.

$$m' = \begin{cases} m + \eta & \text{if} \quad a + \tilde{r}_H(m+\eta) < a^+ - \tilde{r}_H^+ \\ & \text{and } a - \tilde{r}_H(m+\eta) > a^- + \tilde{r}_H^- \\ \text{otherwise:} & \text{argmax}_\iota (m+\iota), \iota \in [0, \mu], \text{ such that} \\ & \text{the above conditions are satisfied.} \end{cases} \tag{10}$$

3.3 Recombination

We apply different variants of recombination, which are performed before the mutation. For each individual we perform an intermediate recombination of all strategy-parameters, which is due to [12] an important prerequisite for a well working self-adaptation:

$$\sigma'_k = u_k \sigma_{a,k} + (1 - u_k)\sigma_{b_i,k} \tag{11}$$

Here u_k is a random variable in the domain $[0, 1]$. For some low percentage ($\approx 10\%$) of the population we additionally perform parameter intermediate recombination.

Furthermore we employ another variant, which is only performed for multiple systems. Hereby for each position one randomly selected planet from randomly selected individuals is inherited. Although this *planetary recombination* is not very effective in the sense that it produces a lot of promising offsprings, it is often the crucial mechanism for finding the global optimum. This works, because the population occasionally consists of various systems where some but not all planets are already identified correctly and in addition maybe reside on a wrong position. By creating new systems by combining planets from existing ones it is very likely to perform one ore more accurate recombinations during the evolution process which finally yield the global optimum. It turned out to be best to perform this operation for about 10% of the individuals of the population and then put them directly into the next generation population (instead of the offspring population).

We found almost no significant improvement of the overall search quality by using the recombination of the strategy parameters and a marginal significance for the other ones. Nevertheless by using the *planetary recombination* the global optimum was often found earlier than without it. Although helpful in some situations, as far as our tests indicate recombination generally did not turn out to be a key factor for a successful search.

3.4 Selection

As already mentioned, the standard deterministic selection mechanism (see e.g. [8]) is modified to support more obtrusively altered individuals to some degree. For this purpose we reserve $\lfloor \mu\pi^\gamma \rfloor$ places of the population for individuals with $\gamma > 0$. Experiments indicated that values $\pi^\gamma \approx 3$ are well suited. Then, in a first step the best $\mu - \lfloor \mu/\pi^\gamma \rfloor$ offsprings are added to the next generation population. The remaining places are successively filled by adding the best $\lfloor \mu/\pi^\gamma \gamma_{\max} \rfloor$ individuals with $\gamma = i$, where $i = 1, \ldots, \gamma_{\max}$. In case of stagnation we re-include the best-so-far solution to the population.

4 Results

Parameter values of $\mu = 50$, $\lambda = 5000$ turned out to be adequate for a reliable search and high solution quality. The algorithm converges after some hundreds of iterations most of the time. Assuming about 150 measurements for a 3-planet system it takes about a couple of hours to find high quality solutions.

As currently no benchmark data-sets exist and previous publications are not focused on algorithmic aspects, it is difficult to compare our approach to others quantitatively. Tests with artificially created data-sets clearly demonstrate the ability to solve multiplanet-systems consisting of up to four planets.

As an example for the application to real data-sets we applied our algorithm to the υ-Andromedae and the 55-Cancri systems [4, 14]. In most of our test runs, ExOD-ES was able to obtain the same or very similar configurations as published in these previous works.

5 Conclusions

In this article we described an evolution strategy which has been specifically tailored for effectively fitting Keplerians to RV data. Besides various mechanisms to create promising offsprings the use of the Hill-stability criterion for the mutation operator considerably reduces the size of the configuration space and therefore supports an effective search. We evaluated the proposed algorithm on real and artificially created data. Results indicate that the ExOD-ES appears to be a promising approach to solve complex systems with high accuracy. Our next step will be to apply our algorithm to the μ Ara system [9].

References

1. Goździewski, K., Konacki, M., Maciejewski, A.J.: Where is the Second Planet in the HD 160691 Planetary System? Astrophysical Journal **594** (2003) 1019–1032
2. Charbonneau, P.: Genetic Algorithms in Astronomy and Astrophysics. Astrophysical Journal Supplement **101** (1995) 309 ff.
3. Lazio, T., Cordes, J.: Genetic Algorithms: Searching for Planets Around Pulsars. In: ASP Conf. Ser. 72: Millisecond Pulsars. A Decade of Surprise. (1995) 391 ff.
4. Butler, R.P., Marcy, G.W., Fischer, D.A., Brown, T.M., Contos, A.R., Korzennik, S.G., Nisenson, P., Noyes, R.W.: Evidence for Multiple Companions to υ Andromedae. Astrophysical Journal **526** (1999) 916–927
5. Laughlin, G., Chambers, J.E.: Short-Term Dynamical Interactions among Extrasolar Planets. The Astrophysical Journal **551** (2001) L109–L113
6. Goździewski, K., Konacki, M., Maciejewski, A.J.: Orbital Solutions to the HD 160691 (μ Arae) Doppler Signal. Astrophysical Journal **622** (2005) 1136–1148
7. Cincotta, P.M., Simó, C.: Alternative Tools to Study Global and Local Dynamics-Application to Galactic Dynamics. In Gurzadyan, V.G., Ruffini, R., eds.: The Chaotic Universe, Proceedings of the Second ICRA Network Workshop. Volume 10 of Advanced Series in Astrophysics and Cosmology. (2000) 247 ff.
8. Bäck, T.: Evolutionary Algorithms in Theory and Practice. Oxford University Press, New York (1996)
9. Pepe, F., Correia, A.C.M., Mayor, M., Tamuz, O., Couetdic, J., Benz, W., Bertaux, J.L., Bouchy, F., Laskar, J., Lovis, C., Naef, D., Queloz, D., Santos, N.C., Sivan, J.P., Sosnowska, D., Udry, S.: The HARPS search for southern extra-solar planets. IX. μ Arae, a system with four planets. Astronomy and Astrophysics **462** (2007) 769–776
10. Schöneburg, E., F., H., Feddersen, S.: Genetische Algorithmen und Evolutionsstrategien. Addison-Wesley (1994)
11. Chwatal, A.: Bestimmung der Bahnelemente von extrasolaren Planeten aufgrund von Radialgeschwindigkeitsmessdaten mittels evolutionärer Algorithmen. Master's thesis, Vienna University of Technology, Vienna, Austria (2006)
12. Schwefel, H.P.: Numerical Optimization of Computer Models. Wiley, Chichester (1981)
13. Bäck, T., Fogel, D., Michalewicz, Z., eds.: Handbook of Evolutionary Computation. IOP Publishing Ltd. (1997)
14. Marcy, G.W., Butler, R.P., Fischer, D.A., Laughlin, G., Vogt, S.S., Henry, G.W., Pourbaix, D.: A Planet at 5 AU around 55 Cancri. Astrophysical Journal **581** (2002) 1375–1388