

Simultaneous least-squares adjustment of visual and spectroscopic observations of binary stars

D. Pourbaix

Université de Liège, boulevard du Rectorat, 7 B31 bte 53, B-4000 Liège, Belgium

Received January 15; accepted March 18, 1998

Abstract. Double-lined spectroscopic visual binaries allow the hypothesis-free estimation of their individual masses and of their distances. Instead of two independent determinations of some orbital parameters, we propose a simultaneous adjustment of visual and spectroscopic observations. This paper presents the different steps of such a simultaneous adjustment: (a) to transform the problem into a quest for the global minimum of a nonlinear function; (b) to obtain that minimum using simulated annealing and the Broyden-Fletcher-Goldstrab-Shanno local search method.

Key words: binaries: spectroscopic; visual — methods: data analysis; numerical

1. Introduction

More than 20 years ago, Morbey (1975) wrote a paper about the simultaneous adjustment of visual and spectroscopic observations. Obviously, the last two decades have not brought as many such systems as Morbey expected in 1975. Favorable conditions to apply a simultaneous solution are still rare even if, nowadays, it is becoming important due to new instruments and observation techniques. Unfortunately, adjustments are still frequently realized separately Docobo et al. (1992); Hummel et al. (1994).

More recently, Barlow et al. (1993) got the orbital parameters of Capella using a simultaneous solution. Unfortunately, the paper does not clearly mention how the authors build the initial solution. The final solution is effectively simultaneous but that is not necessary the case in the first part of the process. In other words, if the function that is minimized in the final step take all aspects of the orbit into account, nothing indicates that the components of the initial solution (used as the minimization starting point) do not come from disjoint solutions.

The aim of this paper is to present a different computational way to combine visual and spectroscopic data

to get only one estimate of the parameters of the orbit: the orbit which simultaneously minimizes a certain function of the residuals of the different data sets. We present the sub-parts of such a procedure: the set of parameters, the function to minimize, the procedure to get the global minimum and the local improvement.

To reach this goal, different steps are required. The first one is to select the most suitable set of orbital parameters among the different equivalent ones (e.g., Thiele-Innes constants or semi-major axis and Eulerian angles). The set fixes the working space. The definition of what we call the objective function (that is the function to be minimized by an appropriate choice of the parameters) is also very important. To find the global minimum of that function, the procedure is divided into two parts: the identification of a neighborhood of the global minimum and then the improvement of this solution with a local search.

We illustrate our analysis with HR 466 (WDS 01376–0924) using speckle observations Hartkopf et al. (1996) and radial velocities Tokovinin (1993). This example is intended to show that, starting from scratch, the proposed method provides results consistent with the three sets of measurements.

2. Parameter set and objective function

2.1. Parameter set

Suppose we have three sets of measurements, namely

1. $\{(t_j, \overset{\circ}{x}_j, \overset{\circ}{y}_j) : j = 1, \dots, N_v\}$: a set of measured relative rectangular coordinates of the fainter (B) with respect to the brighter (A) component
2. $\{(t'_j, \overset{\circ}{V}_{A,j}) : j = 1, \dots, N_{sA}\}$: a set of radial velocity measurements for the spectroscopic component A
3. $\{(t''_j, \overset{\circ}{V}_{B,j}) : j = 1, \dots, N_{sB}\}$: a set of radial velocity measurements for the spectroscopic component B

for a double-lined spectroscopic visual binary. We reckon with radial velocities expressed in km s^{-1} and positions in arcsec ($''$).

For the visual orbit, we have

$$x = AX + FY \quad (1)$$

$$y = BX + GY \quad (2)$$

$$X = \cos E - e \quad (3)$$

$$Y = \sqrt{1 - e^2} \sin E, \quad (4)$$

where X and Y (x and y) are the angular rectangular coordinates, in the orbital (tangential) plane, of the fainter component with respect to the brighter one; A , B , F and G are the Thiele-Innes constants, expressed in terms of a'' , i , ω and Ω as

$$A = a''(\cos \omega \cos \Omega - \sin \omega \sin \Omega \cos i) \quad (5)$$

$$B = a''(\cos \omega \sin \Omega + \sin \omega \cos \Omega \cos i) \quad (6)$$

$$F = a''(-\sin \omega \cos \Omega - \cos \omega \sin \Omega \cos i) \quad (7)$$

$$G = a''(-\sin \omega \sin \Omega + \cos \omega \cos \Omega \cos i). \quad (8)$$

e is the eccentricity and E is eccentric anomaly, determined unambiguously by Kepler's equation

$$E - e \sin E = \frac{2\pi}{P}(t - T). \quad (9)$$

P is the period and T is the epoch of periastron passage. Seven parameters (a'' , i , ω , Ω , e , P and T) are necessary.

For a spectroscopic orbit ($j = A$ or $j = B$), we have

$$V_A = V_0 - K_A(\cos(\omega + v) + e \cos \omega) \quad (10)$$

$$V_B = V_0 + K_B(\cos(\omega + v) + e \cos \omega) \quad (11)$$

$$K_j = \frac{2\pi a_j^{(\text{km})} \sin i}{86400 \cdot 365.242198781 P \sqrt{1 - e^2}} \quad (12)$$

$$\tan \frac{v}{2} = \sqrt{\frac{1+e}{1-e}} \tan \frac{E}{2}, \quad (13)$$

where again the symbols have their usual meaning.

We switch from the angular separation to the linear one using

$$a^{(\text{km})} = \frac{a''}{\varpi} \cdot 1.49598 \cdot 10^8 \quad (14)$$

where ϖ is the parallax of the system. By introducing κ , the ratio of a_A to $a_A + a_B$, we can now write

$$a_A^{(\text{km})} = \kappa a^{(\text{km})} \quad (15)$$

$$a_B^{(\text{km})} = (1 - \kappa) a^{(\text{km})}. \quad (16)$$

The system as a whole thus needs the following ten parameters for its complete specification:

- a'' : the semi-major axis of the relative orbit of the faintest component around the brightest star;
- i : the inclination of the orbital plane with respect to the plane orthogonal to the sight direction;
- ω : the argument of the periastron;
- Ω : the longitude of the ascending node;
- e : the eccentricity;
- P : the period;
- T : the periastron epoch (one of them);
- V_0 : the radial velocity of the system's center of mass;
- ϖ : the parallax of the system;

- κ : the ratio of the semi-major axis (relative to the brighter component) to the sum of the two semi-major axes.

With these parameters, we face a problem of constrained minimization because $e \in [0, 1[$ while $\varpi \in]0, 1[$ and $\kappa \in]0, 1[$.

Other parameter sets are actually possible (e.g., by introducing K_A and K_B , the amplitude of the radial velocity curves, the mass ratio, ...). However, we do not believe one is systematically better than another. Starting from the orbital parameters of visual binaries, the extension with V_0 , ϖ and κ seems natural to us.

One could be tempted to use the variable transformation proposed by Soulié (1986) to convert a constrained minimization problem into an unconstrained one (or, at least, a less constrained one). Soulié's variable transformation is:

$$x = \frac{x'^2}{1 + x'^2},$$

where x is a variable in $[0, 1[$ and $x' \in R$. This could be applied to e , ϖ and κ . Unfortunately, this makes the system even more nonlinear than it already is (the condition number of the objective function increases). We noticed that starting from a uniform distribution of x' , we do not get a uniform distribution of x . We therefore decided to keep the problem in its constrained version. We managed the constraints by returning a huge value for the objective function when one of the constraints is not satisfied.

2.2. The objective function

We are looking for a function that attains its minimum at the best estimate for the orbital parameters. This function has to allow one to combine visual measurements (in the current state of our work, only Cartesian coordinates (in the tangential plane)) and radial velocities. We assume that x and y are observed at the same time, but we assume nothing about the observation time of the three data sets. In fact, we treat these three sets independently.

Assume that none of the observations are correlated and choose the objective function as

$$D(a, i, \dots, \varpi, \kappa) = \sum_{j=1}^{N_v} \left[\left(\frac{\hat{x}_j - \hat{x}_j}{\sigma_{x,j}} \right)^2 + \left(\frac{\hat{y}_j - \hat{y}_j}{\sigma_{y,j}} \right)^2 \right] + \sum_{k=1}^{N_{sA}} \left(\frac{\hat{V}_{A_k} - \hat{V}_{A_k}}{\sigma_{V_{A_k}}} \right)^2 + \sum_{l=1}^{N_{sB}} \left(\frac{\hat{V}_{B_l} - \hat{V}_{B_l}}{\sigma_{V_{B_l}}} \right)^2 \quad (17)$$

where the hat (super) stands for the adjusted (observed) quantity.

D satisfies the mathematical and statistical requirements for a least-squares adjustment Eichhorn (1993). Equation (17) can be rewritten as

$$D(a, i, \dots, \varpi, \kappa) = \boldsymbol{\xi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\xi}$$

where ξ is a $2N_v + N_{s_A} + N_{s_B}$ -vector composed of $\overset{\circ}{x} - \hat{x}$, $\overset{\circ}{y} - \hat{y}$, $\overset{\circ}{V}_A - \hat{V}_A$ and $\overset{\circ}{V}_B - \hat{V}_B$ and Σ is a diagonal matrix composed of σ_x , σ_y , σ_{V_A} , σ_{V_B} , the variances of the observed quantities in the appropriate order.

3. Minimizations

Up to this point, our approach is not fundamentally different from Morbey's (1975). The difference lies in the way in which we find the minimum of D (Eq. 17). A common procedure is to assume that sufficiently accurate approximations to the orbital parameters were available, as Morbey did. He used these rough approximations to solve the normal equations by linearization and iterations. The possibility that this initial guess might not be in a neighborhood of the global minimum is not considered.

Even with highly precise measurements, the number of local minima of F may be significant. One generally accepts Hoffmann and Salamon (1990) that the number of local minima can be approximated by $O(\exp(N))$ where N is the number of parameters. One should not suppose any initial guess (even when it is based on a previous orbit determination) will lead to the global minimum by using a local search only. An efficient search for (a neighborhood of) the global minimum is mandatory. A survey of methods for global optimization has been edited recently by Horst & Pardalos (1995).

One could think that the more precise the measurements, the lower the number of local minima. If the observations are error free, that only means the objective function reaches 0 at the global minimum. Such a situation does not prevent one from having a huge set of local minima, mainly due to the highly nonlinear nature of D with respect to its components.

The chance to miss the global minimum even when starting with a quite good first guess is not all that small. During the preparation of this paper, we accidentally hit upon such a case. To test our approach, we generated a set of observations with normal errors and we tried to recover the original orbit (or at least the one that minimizes D). Let D_i denote the value of D computed with the orbital parameters used to generate the observations, D_1 the value of D reached with only the local minimization procedure starting from that orbit and, finally, D_g the value of D at the global minimum. We met a case in which

$$D_g < D_1 < D_i$$

where the two inequalities are strict. We hope this example convinces the reader that the probability to miss the global minimum is not as low as current belief would put it.

3.1. Global search

Simulated Annealing Metropolis et al. (1953); Kirkpatrick et al. (1983) has already been successfully applied to the determination of the orbital parameters of visual binaries

Pourbaix (1994); Pourbaix & Lampens (1997). The basically same approach is also applied here. While the basic concepts of simulated annealing have been conserved, the efficiency of the algorithm has been highly improved.

The requirements of simulated annealing are:

1. a working space: each point of this space is potentially a solution;
2. a scalar objective function: one seeks the global minimum of this function;
3. a "temperature" to regulate the "annealing" process and an algorithm to decrease this temperature;
4. a point generator to describe how to go from the current trial to the next one.

The working space and the objective function were defined in the previous sections. For the temperature, two features are required: the initial value and a way to set the value of the temperature after k reductions. After many experiments, we decided to fix the initial temperature to the lowest value of D reached in a sample of points randomly chosen in the working space.

There is a lot of theoretical and experimental work which provides guidance for the reduction of the temperature. We chose the approach suggested by Ingber (1993a,b) and we implemented this algorithm to allow 1000 reductions of the temperature in such a way that the final temperature is one ten-thousandth of the original temperature

$$t_k = t_0 \exp(-c \sqrt[N]{k}),$$

$$c = 4 \exp\left(\frac{-3}{N}\right),$$

where N denotes the dimension of the problem (here, $N = 10$).

The most significant improvement concerns the point generator. Instead of a basic random point generator, we are using a more sophisticated procedure based on the Modified Simplex Method Nelder and Mead (1965) and described by Press et al. (1992).

This algorithm has at least two weaknesses: (a) the algorithm can stop as soon as a local minimum is reached (this is probably implementation dependent); (b) the simplex can degenerate when the dimension of the problem becomes "important" (10 is already important). If (a) happens, we simply have to restart using a new simplex. A way to avoid (b) would be to replace the algorithm of Nelder and Mead by another kind of multidirectional search method such as the one lately proposed by Torczon (1991). Torczon proved there is no chance for her simplexes to degenerate; the cost of this proof is the prohibitive computational effort required for this method. We prefer, as in (a), to start again with a new simplex in cases where the current simplex would degenerate.

It is possible to build such simplexes to increase the chances to visit the different regions of the working space. The risk to miss (a neighborhood of) the global minimum

Table 1. Orbital parameters and their standard deviations of HR 466 obtained with our approach and those lately published

Element	This work		Tokovinin (1993)		Hartkopf et al. (1996)		SA with older data	
	Value	Std. dev.	Value	Std. dev.	Value	Std. dev.	Value	Standard
a (")	0.324	0.005	0.292	0.009	0.3179	0.0032	0.34	0.01
i (°)	96.0	0.3	97.4	0.2	96.65	0.17	96.6	0.7
ω (°)	251.6	0.7	248.8	0.3	250.71	0.64	253.0	0.7
Ω (°)	159.6	0.7	161.1	1.7	160.35	0.38	165.0	1.8
e	0.798	0.007	0.748	0.005	0.7847	0.0046	0.800	0.004
P (yr)	28.8	0.8	28.83	0.25	29.05	0.60	29.2	0.2
T (Besselian year)	1989.92	0.01	1989.915	0.021	1989.874	0.033	1989.930	0.008
V_0 (km s ⁻¹)	+47.8	0.1	+47.91	0.16			+47.97	0.09
ϖ ("/A.U.)	0.027	0.001	0.0222	0.0008	0.0238		0.027	0.001
κ	0.45	0.013					0.45	0.011
a (A.U.)	12.08	0.071					12.24	0.053
mass A (M_\odot)	1.2	0.2	1.519	0.080			1.2	0.2
mass B (M_\odot)	1.0	0.2	1.241	0.075			1.0	0.1

Table 2. Upper triangle of the correlation matrix of the orbital parameters of HR 466

	a	i	ω	Ω	e	P	T	V_0	ϖ	κ
a	1	-3.44e-1	-1.50e-1	-8.77e-2	+7.30e-1	+6.25e-1	+1.09e-1	-2.77e-1	+3.42e-1	+2.81e-2
i		1	-1.15e-1	+3.36e-1	-2.57e-1	-8.92e-2	-1.34e-1	+5.52e-2	-2.15e-1	-5.62e-3
ω			1	+2.12e-1	+1.03e-1	-7.68e-1	+5.48e-1	+2.28e-1	+4.62e-1	-2.32e-2
Ω				1	+8.86e-2	-5.21e-2	+7.88e-2	+5.76e-2	+1.20e-2	-5.86e-3
e					1	+4.60e-1	+1.88e-1	-2.66e-1	+4.23e-1	+2.71e-2
P						1	-3.21e-1	-3.29e-1	-1.80e-1	+3.34e-2
T							1	-1.89e-1	+4.27e-1	+1.92e-2
V_0								1	-5.94e-2	-1.02e-1
ϖ									1	+6.62e-1
κ										1

is then reduced. At each temperature level, a maximum number (350) of function evaluations is authorized.

An attentive reader could be puzzled by reading that it is possible for SA to fall in a local minimum. It is important to keep in mind that there is, in general, no proof of convergence for that method. The rare existing proofs require an infinite decrease of the “temperature” (i.e., an infinite computation time). Facing such a situation, we have to compromise between the execution time and the confidence we have in the “global” nature of the minimum. That explains the potential relative inefficiency of any implementation of SA.

3.2. Local improvement

Even with the Modified Simplex Method as the point generator, one cannot expect to be at the global minimum at the end of the simulated annealing phase. A local search algorithm has to be used to tune the minimum. Here also, different methods have been proposed: the ones not requiring the gradient (Powell, Modified Simplex Method, etc.), the ones using the gradient (Davidon-Fletcher-Powell, Levenberg-Marquard, Broyden-Fletcher-Goldstrab-Shanno (BFGS), etc.).

We chose the BFGS method for its quite efficient behavior independently of the magnitude of the residuals.

Our implementation is based on the pseudo-code proposed by Dennis & Schnabel (1995), but a complete description of the method can be found in Fletcher (1987).

4. An example

To illustrate the suitability of our approach, we used recently published measurements of HR 466. Our initial plan was to recover an orbit consistent with those already published.

Towards this goal, we use the radial velocities published by Tokovinin (1993) and the interferometric measurements stored in the CHARA database Hartkopf et al. (1997): 28 interferometric observations and 8 radial velocities measurements for each components.

The initial guess is *randomly* chosen in the working space. An initial disjoint determination is required if one wants to check the consistency of different data sets. Such a determination would also help (but is not mandatory at all) to fix a first guess for the standard deviations of the residuals.

Using successively global and local search methods, we obtain the elements given in Table 1 (second and third columns) and the corresponding correlation matrix (Table 2). For that matrix, all the components corresponding to i , ω and Ω are computed with these angles expressed

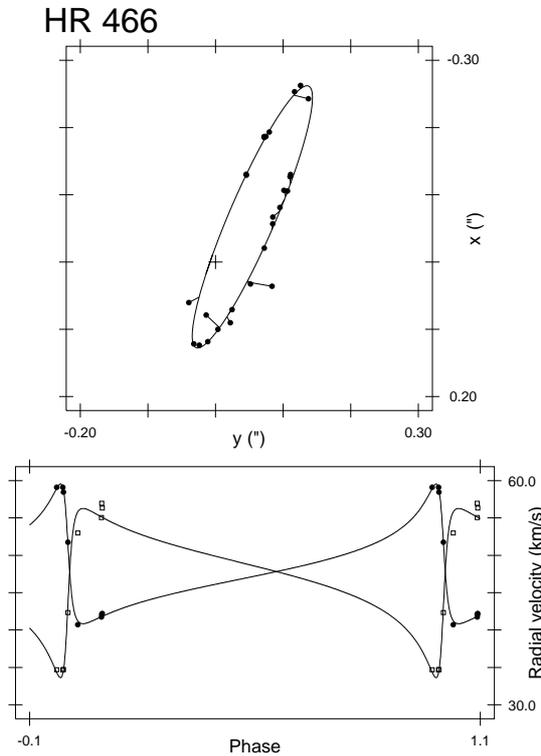


Fig. 1. Visual orbit and radial velocity curves of HR 466. In the spectroscopic plot, the open squares denote the measured radial velocities of B and the filled circles those of A

Table 3. Summary of the residuals using our orbit of HR 466. The hat (super) stands for the adjusted (observed) quantity

$\sigma_{\langle \hat{x} - \hat{x} \rangle}$	$\sigma_{\langle \hat{y} - \hat{y} \rangle}$	$\sigma_{\langle \hat{V}_A - \hat{V}_A \rangle}$	$\sigma_{\langle \hat{V}_B - \hat{V}_B \rangle}$
7.08 mas	9.88 mas	0.28 km s ⁻¹	1.44 km s ⁻¹

in radians. The standard errors of the different residuals are given in Table 3. The difference of the residual variances between the spectroscopic components A and B is confirmed in Fig. 1. The covariance matrix (from which the standard deviations of the parameters and the correlation matrix are deduced) is computed by inverting the Fisher matrix (almost the Hessian matrix) at the minimum Pourbaix (1994); Press et al. (1992).

These residuals are the result of three iterations of the complete minimization scheme. At the first iteration, the values of σ_{x_j} , σ_{y_j} , $\sigma_{V_{A_k}}$ and $\sigma_{V_{B_l}}$ (Eq. 17) had been fixed to a too low value. The residuals at the global minimum were used to estimate new values for these σ , and a new minimization was started. After two iterations of this minimization process, the variances at minimum were close enough to the estimated ones.

The efficiency Eichhorn (1989) is 0.51 which indicates a low correlation between the parameters. Even if we are dealing with few radial velocities, the visual orbit constrains the overall solution well enough.

How are our orbital parameters with respect to those recently published? Table 1 gives the orbital parameters and their standard deviations after Tokovinin (1993) and Hartkopf et al. (1996). With a combined determination of the orbital parameters, Tokovinin gets a parallax of 22.2 mas (equivalent to a distance of 45 pc) and a mass sum of $2.76 M_{\odot}$. In their paper, Hartkopf et al. adopt a distance of 42 pc which yields a mass sum of $2.8 M_{\odot}$.

From the combination of Tokovinin's radial velocities and Hartkopf et al. visual observations, we obtain a parallax of 26.80 mas (corresponding to a distance of 37.31 pc) and a mass sum of only $2.20 M_{\odot}$, which is quite different.

With the same data as used by Tokovinin (1993) for the radial velocities as well as for the visual observations, we obtain the orbital parameters given in the sixth column of Table 1. Here also, we obtain a parallax of 27 mas (which corresponds to a distance of 36.03 pc) and $2.2 M_{\odot}$ for the mass sum. Our two determinations are consistent: in both cases, we get a larger parallax and a lower mass sum than those published recently. A reason for the discrepancy might be that the observations weights are slightly different. We followed Morbey's procedure (1992) to fix the weights. Although it would be possible to associate different weights to different observations of the same type, we fixed a weight for all x , one for all y , one for all V_A and one for all V_B . Discrepant observations do not get lower weights.

When compared with the results obtained by Hipparcos (1997), 26.15 ± 0.81 mas, it seems our result is in very good agreement with the latter value.

In fact, it seems we got more than expected at the beginning. The orbit we propose is probably the best one based on the available observations, even if it does not fully agree with the orbits already published (mainly on the mass sum and the parallax of the system).

5. Conclusions

Due to the progress in the determination of precise radial velocities by CORAVEL type instruments, the lowering of the observable separation limit for visual binaries by interferometric techniques and the concomitant drastic improvement of positional precision, more and more spectroscopic binaries are going to be visually resolved in the coming years. In a close future, new observing techniques (such as speckle and visual interferometry, satellites) are going to add to the set of stars that are resolved both visually and spectroscopically. Efficient reduction techniques will then be even more needed.

To obtain stellar masses and distances is still the main goal of binary studies. To get more powerful instruments and to improve the observation techniques should not be the unique objective of astronomers working in this field. In many instances, powerful reduction methods such as the one we propose can derive significantly improved data from already existing observations.

The approach presented in this paper is not limited to binary stars and other areas of astronomy could take advantage of our method. Any problem that can be considered as a minimization problem may profit from a similar approach. We hope this paper is going to promote SA in the astronomical community.

Acknowledgements. I express my thanks to H. Eichhorn (who also suggested this investigation field) and R.E. Wilson for having critically read the manuscript and for the many fruitful discussions we had about this topic. I also thank P. Lampens and the anonymous referees for their suggestions.

References

- Barlow D., Feckel F., Scarfe C., 1993, PASP 105, 476
 Dennis Jr. J., Schnabel R., 1995, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM, 2 edition
 Docobo J., Ling J., Prieto C., 1992, in: McAlister and Hartkopf (eds.), p. 220
 Eichhorn H., 1989, Bull. Astron. Inst. Czechosl. 40(6), 394
 Eichhorn H., 1993, Celestial Mech. 56, 337
 ESA 1997, The Hipparcos and Tycho Catalogues, ESA SP-1200
 Fletcher R., 1987, Practical Methods of Optimization. John Wiley & Sons, 2 edition
 Hartkopf W., Mason B., McAlister H., 1996, AJ 111, 370
 Hartkopf W., McAlister H., Mason B., 1997, Third Catalog of Interferometric Measurements of Binary Stars, Center for High Angular Resolution Astronomy, Georgia State University, Atlanta, Georgia, 30303-3083
 Hoffmann K., Salamon P., 1990, J. Phys. A: Math. Gen 23, 3511
 Horst R., Pardalos P., 1995, Handbook of Global Optimization. Kluwer Academic Publishers
 Hummel C., Armstrong T., Quirrenbach A., Buscher D., Mozurkevich D., Elias II N., Wilson R., 1994, AJ 107, 1859
 Ingber L., 1993a, Adaptative simulated annealing (asa), 1 Reseach note, Caltech, Lester Ingber Research, P.O. Box 857, Mc Lean, VA 22101
 Ingber L., 1993b, Simulated annealing: Practice versus theory, 1 Reseach note, Caltech, Lester Ingber Research, P.O. Box 857, Mc Lean, VA 22101
 Kirkpatrick S., Gelatt Jr. C., Vecchi M., 1983, Sci 220 (4598), 671
 McAlister H., Hartkopf W., 1992, Complementary Approches to Double and Multiple Star Research, Vol. 32
 Metropolis N., Rosenbluth A., Rosenbluth M., Teller A., Teller E., 1953, J. Chem. Phys. 21(6), 1087
 Morbey C., 1975, PASP 87, 689
 Morbey C., 1992, in: McAlister and Hartkopf (eds.), p. 127
 Nelder J., Mead R., 1965, Computer J. 7, 308
 Pourbaix D., 1994, A&A 290, 682
 Pourbaix D., Lampens P., 1997, in: J. Docobo, A. Elipe, and H. McAlister (eds.), Visual double stars: formation, dynamics and evolutionary tracks. Kluwer Academic Publishers, p. 383
 Press W., Teukolsky S., Vetterling W., Flannery B., 1992, Numerical Recipes in C. Cambridge University Press, 2 edition
 Soulié E., 1986, A&A 164, 408
 Tokovinin A., 1993, SvAL 19, 73
 Torczon V., 1991, SIAM J. Opt. 1(1), 123