Networks of local minima in optical system optimization

Florian Bociort, Eco van Driel, and Alexander Serebriakov

Optics Research Group, Delft University of Technology, Lorentzweg 1, NL-2628 CJ Delft, The Netherlands

Received July 14, 2003

We discuss a surprising new feature of the merit function landscape in optical system design. When certain conditions are satisfied, the local minima form a network in which all nodes are connected. Each link between two neighboring minima contains a saddle point with a Morse index of 1. For a simple global optimization search (the symmetric Cooke triplet), the network of the corresponding set of local minima is presented.

OCIS codes: 080.2720, 220.2740.

Global optimization methods are currently of major importance in optical system design and in many other fields in which multiple local minima can be found during optimization.[1–5] A limitation of these methods, however, is that they do not provide information about the merit function topography near the individual local minima or in the space between them.

In this Letter we show that, when certain general conditions are satisfied, the merit function landscape has a remarkable property: The local minima then form a network in which they are all connected by links that contain a special type of saddle point. Although our primary interest is optical system design, we believe that the present results are of much broader interest.

We consider an optimization problem with an arbitrary number N of continuous variables and a merit function f that is defined as the rms of image defects, computed with ray tracing. Each set of optimization variables defines a point in an N-dimensional solution space. A point in this space for which the gradient of f vanishes is called a critical point.

An important characteristic of critical points for which the Hessian of f (i.e., the matrix of the second-order derivatives of f with respect to the optimization variables) has a nonzero determinant is the number of negative eigenvalues of the Hessian (the so-called Morse, or Hessian, index).[6] A negative eigenvalue means that along the direction defined by the corresponding eigenvector of the Hessian the critical point is a maximum. Thus minima have Morse index 0, maxima have Morse index N, and for saddle points the Morse index has values between 1 and $N - 1$.

As will be shown below, if we are interested in the detection of local minima, the saddle points with a Morse index of one (SPMI1) play a special role. For the present purpose it is sufficient to keep in mind that a SPMI1 is a maximum in one direction (the downward direction) and a minimum in an $N - 1$-dimensional hyperplane orthogonal to that direction. If, for instance, $N = 2$, then every saddle point is a SPMI1. Intuitively, the downward direction of a SPMI1 is similar to the downward direction of a two-dimensional saddle point, and each of the $N - 1$ upward directions is similar to the upward direction of a two-dimensional saddle point.

We consider the N-dimensional solution space to be divided into infinitely thin sheets in which f is constant (the equimagnitude surfaces). Note first that an equimagnitude surface (EMS) with some value of $f = f_0$ of the merit function encircles a region in the solution space for which $f < f_0$. If, for instance, f_0 is the merit function of a local minimum, then the EMS (or the part of it situated near the minimum) reduces to one point, the minimum itself. For slightly larger values of f_0 (a part of) the EMS encircles a small ellipsoidal region around the local minimum. If the value of f_0 increases, then the encircled volume also increases.

We will show now in an intuitive way that the local minima within an arbitrary EMS form a connected network. Figure 1 shows two local minima, M_1 and M_2, in an N-dimensional solution space. We assume the existence of an EMS with $f_0 = f_a$ that encircles both minima and has no critical points on it, labeled a in the figure. For a sufficiently small value of $f_0 = f_b < f_a$ (for instance, for f_b slightly larger than the larger of the two merit function values corresponding to the local minima) the EMS consists of two separate parts, labeled b in Fig. 1. In a merit function landscape free of pathologies, for some value of f_S with $f_b < f_S < f_a$, we

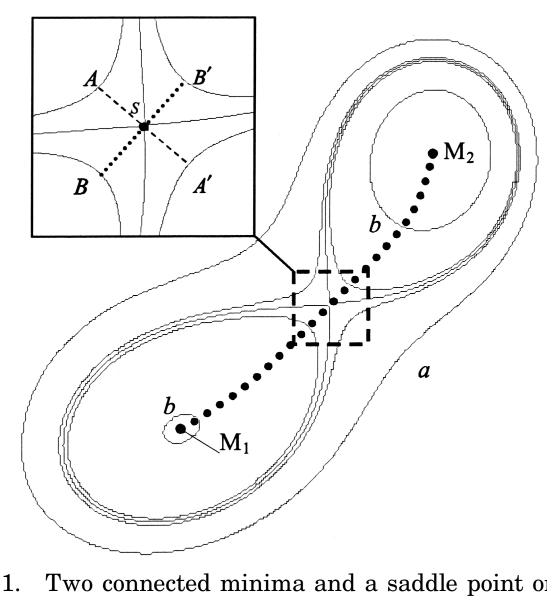

Fig. 1. Two connected minima and a saddle point on the link.

encounter the limiting case when the two separate parts of the EMS touch each other at one point S. We now show that S is in fact a SPMI1. For $f_B = f_{B'}$ slightly lower than f_S, the corresponding EMS will be split. Then let, B and B′ be the two points on the separate parts of the EMS for which the length of segment BB′ is minimal (see the enlarged detail in Fig. 1). Obviously, along line BB′ the point S is a maximum. We now consider an EMS with a merit function of $f_A = f_{A'}$ slightly higher than f_S. Since this EMS encircles the one with $f_0 = f_S$, any line perpendicular to BB′ and passing through S will intersect the EMS at two points, A and A′. Along AA′ point S is then a minimum. Since this is valid for any choice of line AA′ in an $N - 1$ dimensional hyperplane orthogonal to downward direction BB′, S is a minimum in $N - 1$ directions and is thus a SPMI1. If we now choose points B and B′ as starting points, local optimization will generate two paths in the solution space that will lead to the two minima. Together with the saddle point, these two paths form the link between M_1 and M_2 (the dotted line in the main part of Fig. 1).

Assume now that we have an EMS with some (large) value of $f_0 = f_a$ that encircles an arbitrary number p of local minima and has no critical points on it (in Fig. 2, where $p = 3$, this is the outermost contour.) If we decrease f_0, at some value of $f_{S1} < f_a$ the EMS will split into two surfaces that will now encircle $p_1$and $p - p_1$ local minima. For the same reason as above, point S_1 in Fig. 2 is also a SPMI1. By choosing two points close to the SPMI1, one on each side along the downward direction, and following the paths of local optimizations started at these points, we obtain a link between one local minimum in the group of p_1 encircled local minima and one in the group of $p - p_1$ local minima. By further decreasing f_0, we obtain successive splits of the encircling surfaces. Each such split generates an additional link between two local minima situated in the two different groups resulting from the split. When f_0 has reached a value that is lower than the merit function of the lowest SPMI1 (S_2 in Fig. 2) all the local minima encircled by the EMS with $f_0 = f_a$ are linked in a network by links that each contain a SPMI1.

We have thus shown that the local minima encircled by an arbitrary EMS (without critical points on it) form a connected network. For our purposes it is important to know whether, in typical situations occurring during optical system optimization, we can always find such EMSs that encircle all local minima. At the time of this writing, we have examined only a limited number of situations, and further research is certainly necessary. Our present results make us believe, however, either that this desirable property of the landscape of f is satisfied automatically or that it can be achieved by modifying the optimization problem adequately (e.g., by using inequality constraints). It is well known that outside some useful regions in the solution space the optical system configurations tend to suffer from ray failure because some rays either miss surfaces or suffer from total internal reflection. Close to ray-failure situations, the incidence angles of those rays at the critical surfaces are large; the aberrations and the merit function of the given optical system configuration then tend to be large. Therefore close to the ray-failure borders we can expect in the solution space an EMS with a large f_0. The local minima encircled by these EMSs then form a network.

An algorithm that uses constrained local optimization to detect the SPMI1 has been implemented in the macro language of the commercial optical design program CODE V.[7] Saddle-point detection algorithms that have certain similarities with the one developed independently by us are used to study the energy landscape of systems of many atoms.[8,9] In the case of a triplet in which the optimization variables were the six curvatures of the surfaces, we found 23 SPMI1 (shown by thin-line boxes in Fig. 3). By following the downward paths of local optimization started at these points, we obtained 19 local minima (thick-line boxes). The best two local minima, m_1 and m_2, have the well-known shape of a Cooke triplet. The first 17 of our local minima are identical with the 17 local minima listed in the output of Global Synthesis, the global optimization algorithm of CODE V. Interestingly, the saddle-point configurations s_{i-j} can be viewed as intermediate stages in a continuous transformation of local minimum m_i into minimum m_j.

We chose the specifications (distances between surface and glass types) to be rigorously symmetrical with respect to the aperture stop. For this purpose, the central lens was split by a fictitious stop surface (not shown in Fig. 3). The image plane was placed at its paraxial position and the position of the object plane was controlled such that the transverse magnification was kept equal to -1. Because of an additional equality constraint (the distance between object and image was also kept constant), the search space was effectively five dimensional. The merit function used was the default merit function of CODE V, for which the

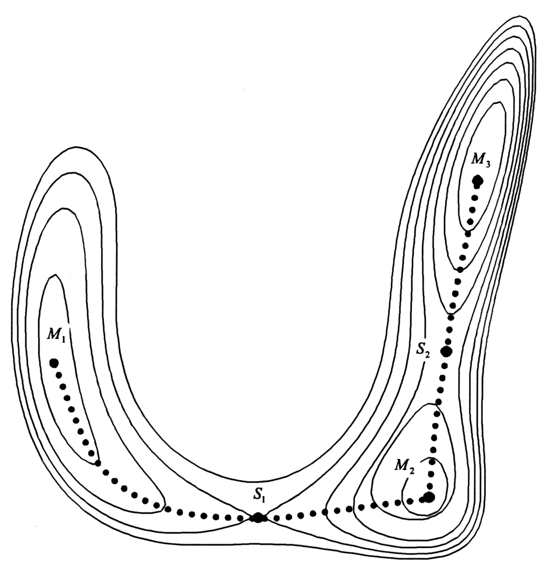

Fig. 2. Three connected minima. The links (dotted curves) are the paths of local optimization started close to the saddle point on both sides along the downward direction. The EMS (continuous curves) have been obtained from a two-dimensional cut through the five-dimensional merit function landscape of a Cooke triplet global search.

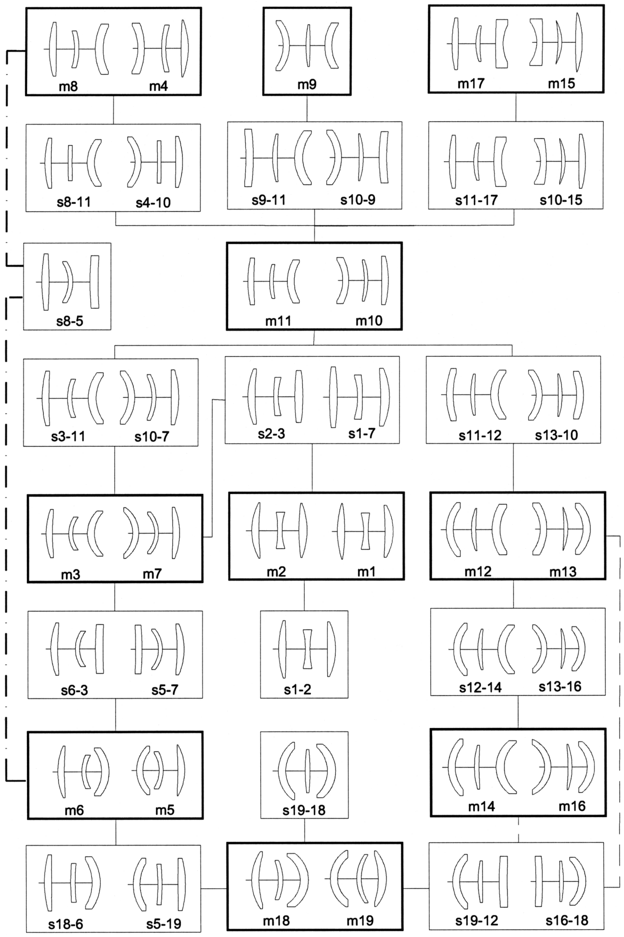

Fig. 3. Network of the global search corresponding to a symmetric Cooke triplet.

image defects are transverse ray aberrations computed with respect to the chief ray.

As expected, the detected network is almost perfectly symmetrical. With one exception, saddle point s_{8-5} (linked to the corresponding local minima through dashed–dotted lines), the configurations in Fig. 3 are either symmetrical with respect to the stop or they have counterparts that look like mirror images. For clarity, the pairs in which one configuration is (almost) the mirror image of the other have been grouped together in the same box. Moreover, with the exception of only two links (dashed line), the detected links display the same symmetry: If a SPMI1 links two minima, then its mirror links the mirrors of the same minima. The minor deviations of the network from perfect symmetry are not surprising since the symmetry between object and image is affected by aberrations. Although we cannot be certain that the present algorithm has detected the entire network, the symmetry detected as expected increases our confidence in the potential of our network idea.

Concluding, this simple but nontrivial example shows that algorithms based on the idea that the local minima form a network could in principle not only reproduce the results of known global optimization algorithms but also provide additional insight into the topography of the merit function landscape. This new insight could be useful in meeting the design challenges encountered in high-quality optics for lithography, microscopy, and space applications.

We thank J. Braat for stimulating discussions and acknowledge the use of an educational license of CODE V. F. Bociort's e-mail address is f.bociort@tnw.tudelft.nl.

References

1. H. Reiner and P. Pardalos, eds., *Handbook of Global Optimization* (Kluwer, Dordrecht, The Netherlands 1995).
2. A. E. W. Jones and G. W. Forbes, J. Global Optim. **6,** 1 (1995).
3. T. G. Kuper and T. I. Harris, Proc. SPIE **1780,** 14 (1992).
4. K. E. Moore, Proc. SPIE **3780,** 40 (1999).
5. M. Isshiki, H. Ono, K. Hiraga, J. Ishikawa, and S. Nakadate, Opt. Rev. **6,** 463 (1995).
6. J. C. Hart, in *Mathematical Visualization*, H.-C. Hege and K. Polthier, eds. (Springer-Verlag, Berlin, 1998), p. 257.
7. E. van Driel, F. Bociort, and A. Serebriakov, "Topography of the merit function landscape in optical system design," Proc. SPIE **5249** (to be published).
8. N. Mousseau and G. T. Barkema, Phys. Rev. E **57,** 2419 (1998).
9. D. J. Wales, J. Chem. Phys. **101,** 3750 (1994).