

Adaptive Saddle Crossings in Diverse Landscapes

Artur Chorążyczewski¹, Roman Galar²

¹Wrocław University of Technology, Institute of Computer Engineering, Control and Robotics,
Wrocław, Poland, e-mail: artur.chorazyczewski@pwr.wroc.pl

²Wrocław University of Technology, Institute of Computer Engineering, Control and Robotics,
Wrocław, Poland, e-mail: roman.galar@pwr.wroc.pl

Abstract: There is always a question as to how sound are assessments of evolutionary algorithms capabilities based on the tests made in the specific adaptive landscapes. We present some results of numerical experiments concerning the ability of phenotypic evolution to cross saddles between adaptive peaks of different shapes. Asexual reproduction, proportional selection and normally distributed mutations are used. Some related comments on the benchmarking of evolutionary algorithms are also given.

1 Introduction

The essence of evolutionary optimization lies in overcoming the effect of search processes being trapped in the local optima. The precise identification of these optima is of a lesser importance, as there are often better tools to achieve it. The most ambitious agenda is to identify, i.e. to find the value and position of the global optimum. The more practical program, which a lot of evolutionary algorithms strives to realize, is to find such optimum in a constrained space of search without too many local optima. The general idea is to localize and to explore all the attraction basins of all prospective optima, starting from covering the search area with the large flock of randomly generated trials. The main obstacle is the tendency of evolutionary processes to converge into relatively compact clusters faster than the parallel exploration is done. There are a number of ideas how to counter such “premature” convergence.

It might be stressed that neither constrained space of search nor the widely dispersed initial populations should be treated as consistent with the Darwinian idea of evolution of species. The usual counter convergence measures distort the similarities between simulated evolution and Darwinian evolution still more. Darwinian evolution is a path dependent process in which clusters of pretty similar individuals move through the space of traits. The possibilities of exploration are always confined to the immediate vicinity of the presently existing phenotypes. This quashes the claims that evolutionary algorithms of the type described inherit adaptive virtues of real evolutions. These algorithms should be assessed on their own rights only and they shed little light on the dynamic of evolution proper.

The less ambitious approach to evolutionary optimization is to concentrate on the incidents of saddle crossing, which allows for passages between the neighboring optima (quantitative evolution). Such incidents are necessary to progress when (1) no information on the possible constraints exist so the multistart tactics cannot be used, or (2) the counter convergence measures

failed, and the search get trapped at some local optimum. The saddle crossing approach does not ever attempt to localize the best solution, the “Holly Grail” of global optimization. It is a “locally global” optimization, concentrated on getting something different and better than the presently dominant solution. In this respect it corresponds to the real business goals [1], paradigm changes [2] and punctuated equilibria [3] of real evolution.

For some time our general concern has been with the dynamics of saddle crossing in real spaces of few or several dimensions (phenotypic evolution). We were interested in wide saddles (in terms of the standard deviation of mutation σ) and investigated the influence of the softness of selection, reproduction mode, population size, environmental variance, etc [e.g. 4,5,6]. The one factor untested so far was the sensitivity of search to the actual shape of the saddle area. Our experiments were with the saddles between two Gaussian bell-shaped adaptive hills only. This time we present experiments involving hills in the form of elliptic paraboloids, cones, pyramids, cubes and roofed cubes, and Gaussian hills covered with small pyramids. We are aware of similar experiments with one-dimensional cones [7].

2 The Basic Experiment

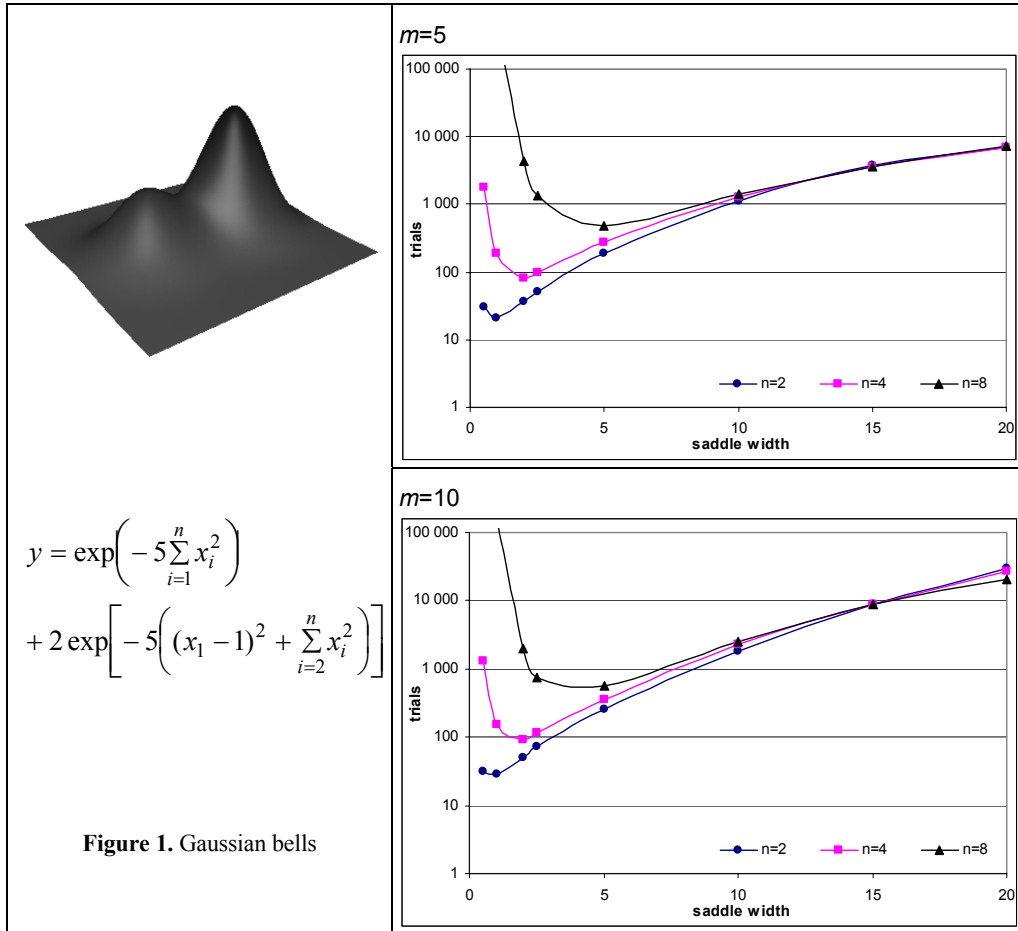
The basic result is presented in Figure 1. Its right panels present simulation based estimates of the expected number of trials necessary to cross n -dimensional saddles. The shape of the adaptive saddle for $n=2$ is shown in the left panel together with the fitness formula. The saddle widths are measured from the top of the lower hill to the closest point elevated above it, and scaled in the units of σ (standard deviation of trait mutation). Instances of evolution of m individuals were simulated using proportional selection and asexual reproduction. Processes started at the local optimum (the initial population of clones put on the top of the lower hill) and ended when a better individual was found on the other side of the saddle. Each marked point gives an average from 400 simulation runs (each evaluation of fitness function counts as a trial).

It might be observed that simulated evolution can cross saddles which are many σ wide. The sensitivity of search to the search space dimension and to the saddle width is small.

This result is striking when compared with chances of getting an improvement by macromutation, i.e. a “direct hit” from the top of the lower adaptive hill [5]. The probabilities of adequate macromutations not only diminish fast with the problem dimensionality but also fall dramatically when the saddle width exceeds 3σ . In the landscape considered, the expected number of trials would be less than 1 000 000 only if the saddle width is less than 4.5σ for $n=1$ and 3.5σ for $n=8$. The best chances exists when the saddle width is close to σ , but even then the minimal values are higher than the ones provided by evolution, and the difference grows strongly with n . This suggests that the often invoked macromutation idea is wrong. In saddle crossings even the finest tuning of mutation rates cannot compensate for the absence of soft selection.

It has to be stressed that — as Wright has already predicted [8] — the ability to cross wide saddles is restricted to very small populations. It might be observed that increasing the population size m from 5 to 10 increases the expected times of saddle crossings about 5 times.

It might be noted that the probabilities of scoring an improvement diminish for very narrow saddles. It is an obvious effect for the search to overshoot the area where the multimodality is expressed, when the saddle is only a fraction of σ wide.

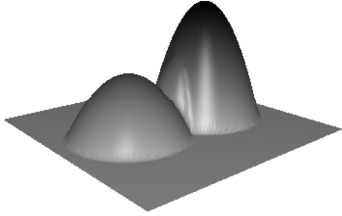


3 Saddle Crossing Between Diversely Shaped Adaptive Hills

To test the sensitivity of evolutionary saddle crossing to the actual shape of the saddle area, a number of experiments were performed involving elliptic paraboloids, cones and pyramids, with the results shown in Figures 2-4.

The results are presented using the same convention as in Figure 1. Simulations, started with the population concentrated on the top of the lower hill, stopped when a better individual was found. Each marked point gives the average from 400 simulation runs. The shapes were set to get the peak elevations and the saddle width the same as in the basic experiment.

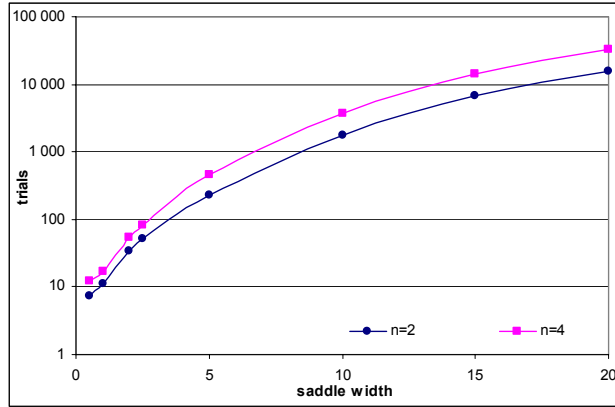
The only difference with the basic experiment is in treating individuals whose traits would render their fitness negative (under the proper functions describing paraboloids, cones and pyramids). Such individuals were interpreted as failed trials and rejected, while reproduction proceeded until m viable individuals were generated. This influenced the results for the very narrow saddles and explains the absence of minima so evident in Figure 1.



$$y = \max\left(1 - 2.7 \sum_{i=1}^n x_i^2; 0\right) + 2 \max\left[1 - 2.7 \left((x_1 - 1)^2 + \sum_{i=2}^n x_i^2\right); 0\right]$$

Figure 2. Elliptic paraboloids

$m=5$



$m=10$

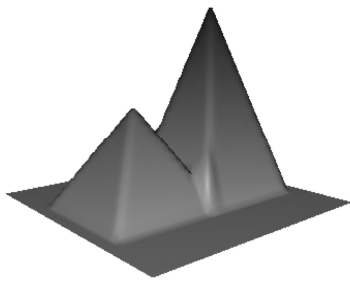
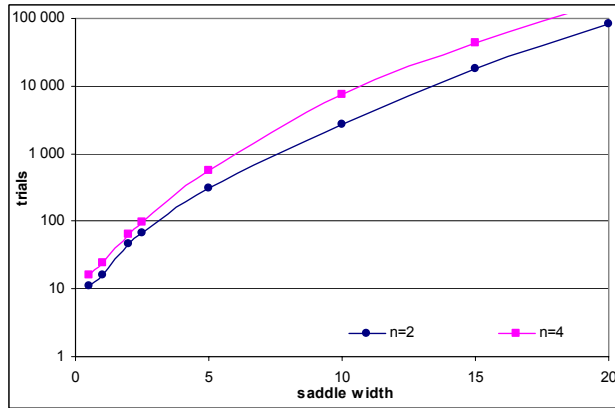
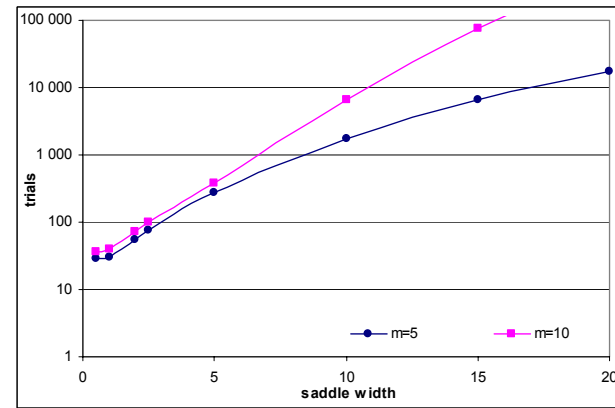
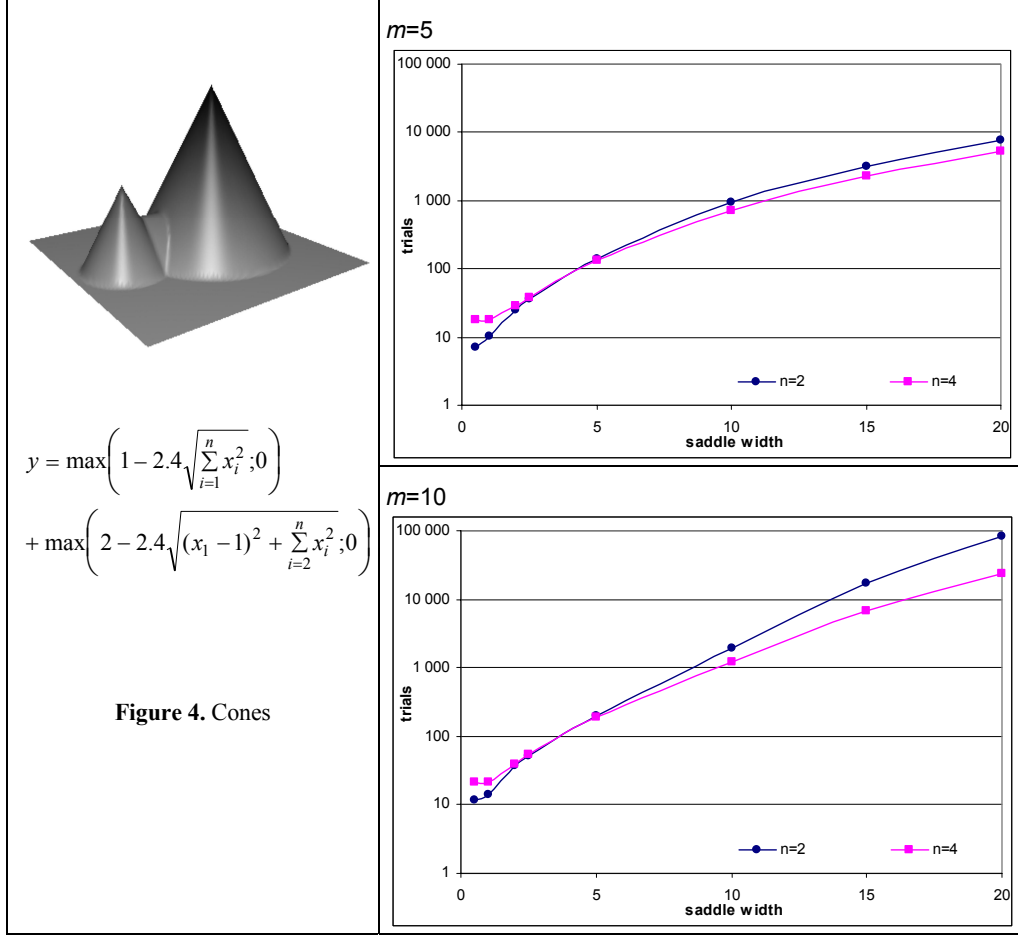


Figure 3. Pyramids

$n=4$



$$y = C([x_1, x_2, \dots, x_n]) + 2C([x_1 - 1, x_2, \dots, x_n]), C([x_1, x_2, \dots, x_n]) = 2 \cdot \max\left(0, \max_{i=1}^n \left(0.5 - \left|\frac{x_i}{d}\right|\right)\right), d = 1.2$$



Comparison of characteristics presented in Figures 2-4 with these in Figure 1 demonstrate the basic similarity in cases involving wider saddles:

1. saddles scores of σ wide might be crossed;
2. the process is faster for smaller population (especially Figure 3.);
3. the impact of dimensionality is moderate, sometimes higher n might be helpful.

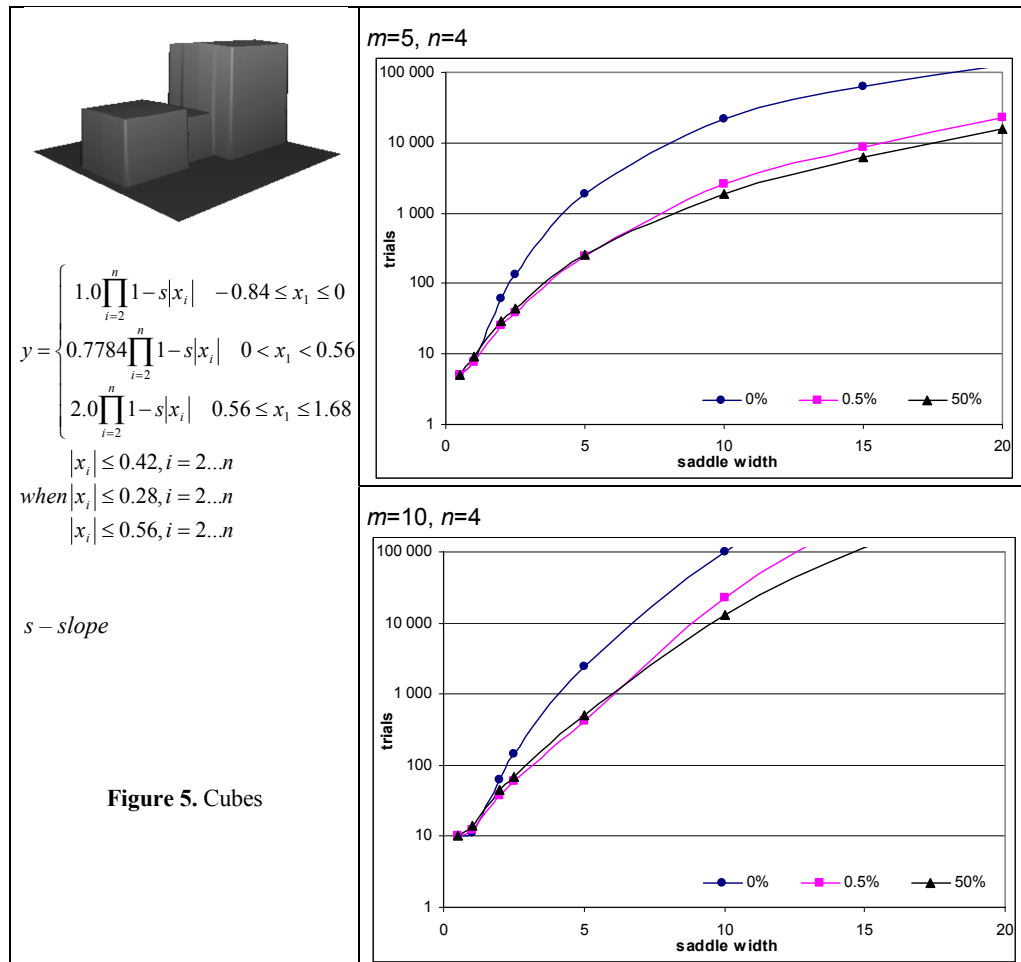
4 Cubes and Leading Edges

Another shape investigated were cubes, with the result presented in Figure 5. Cubes proved to be more difficult. The expected number of trials until improvement in case of saddles of 20σ , $n=4$ and $m=5$ was about 20 times greater than in the basic example; and still worse for $m=10$.

This effect might be easily explained. When the whole population is on the top of a cube, all individuals have identical fitness, so selection provides no information on the proper direction of search. As the clustering tendency still acts [6], such population can only drift on the surface of

the cube, as a more or less compact cluster with the radius of the range of σ . Only when population comes to some edge does some adaptive feedback appear. If a prospective individual falls on the surrounding flats it is rejected, because of the zero fitness. This acts on the process as if the population was repelled from the edges. If a prospective individuals falls on the lowest cube it might reproduce, albeit with odds significantly less than the individuals remaining on the higher level. The chance that the whole population will move to the lowest cube and then drift to its other side (the distance several times it radius), is very slim indeed and quickly diminishes with population size.

These chances of transition might be improved by introducing “roofs” on the cubes; i.e. modifying the fitness function so that it diminishes with traits deflecting from the symmetry line that goes through the bases of the cubes. It was enough to introduce “roofs” declined as little as 0.5% (parameter s) to improve the results 10 times. Making roofs still steeper does not yield much improvement, especially in very small populations.



Explanation of these phenomena is possible by the recourse to the mechanism of ridge detection by evolving populations [4]. Small and compact populations “detect” only the local attraction of the adaptive landscape and selection-mutation mechanisms tend to shift them toward more elevated regions. If there is a ridge, this tendency produces an effect of population being kept close to the ridge but relatively free to drift along it. As a result, an n -dimensional drift is channeled to the one-dimensional drift. This very significantly increases the chances of saddle crossings. It also explains significant invariance of saddle crossing with regard to the landscape dimension.

5 The Perturbed Landscapes

The last case investigated was the basic landscape of Gaussian bells perturbed by multitude of lesser local optima, that produced results presented in Figure 6. Perturbations were made by adding pyramids in each node of the grid superimposed on the space of traits. The base of pyramids and the edge of the grid cells were d and the pyramids height was $2d$. The results for $d=0.1$ and $d=0.2$ are given in the top and bottom panel respectively.

It might be noted that the results for the wider saddles are practically identical as in the case of the basic experiment. The presence of the lesser local optima does not influence the saddle crossings noticeably. This effect might be corroborated by the earlier experiments with evolution in the Gaussian landscapes as in Figure 1., where the values of fitness were recorded with substantial random errors [9]. These errors might be interpreted as accidental fluctuations of the adaptive surfaces. Also in this case there were no discernible differences in the efficiency of saddle crossing.

6 Conclusions

The presented examples show that the saddle crossing effect is largely invariant to the particularities of shape of the adaptive hills investigated. Similarly, quite significant perturbances of adaptive surfaces have little or no effect.

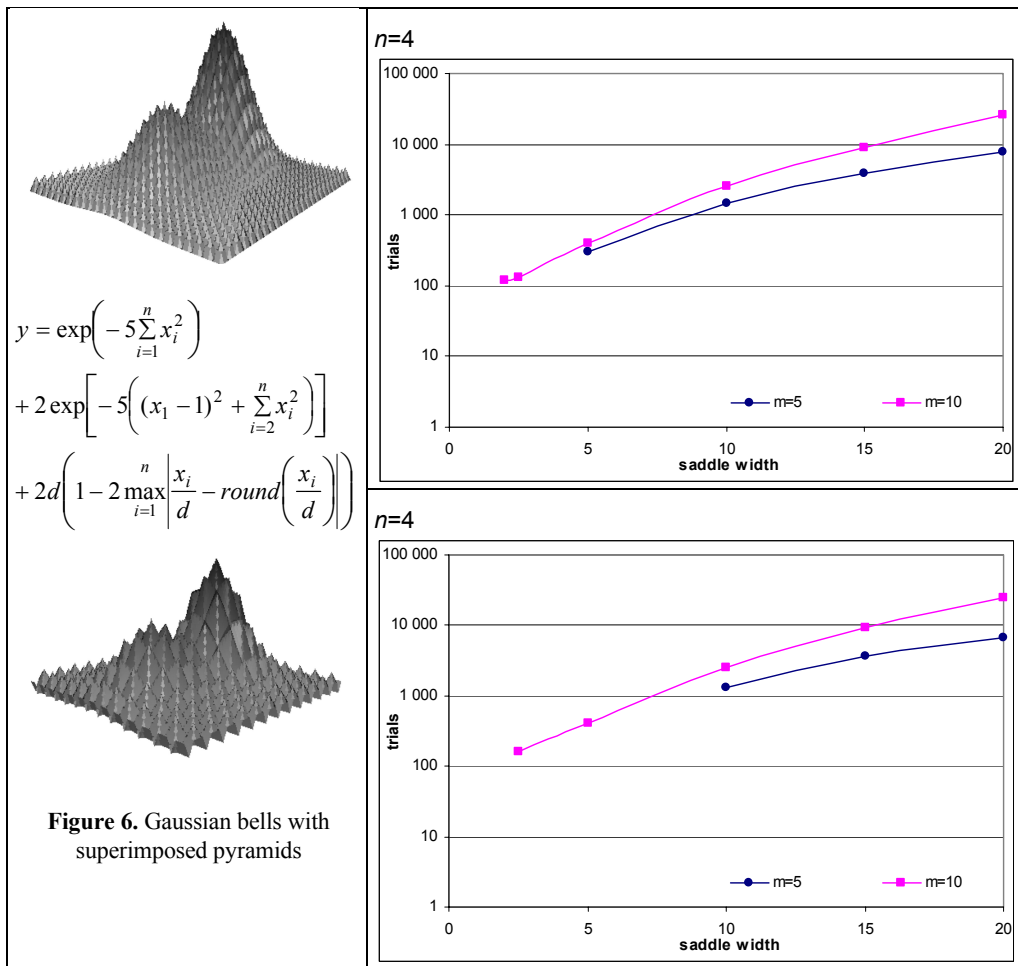
The presented results support the position that the main factor enabling saddle crossing is the attraction of ridges of adaptive landscapes, which channel the drift of small populations. This confirms the Eigen’s view that evolution is a process best suited to adaptive landscapes that are “reminiscent of the ranges of mountains on Earth” [10]. If there are ranges, evolution might follow them, irrespective of irregularities. When formations are more haphazard, e.g. the neighboring hills are elongated in perpendicular directions; the efficiency of evolutionary saddle crossings drops substantially. Unveiling the essence of Eigen’s metaphor seems to be the most challenging purpose of evolutionary algorithms theory. So far little is clear except that the classical optimization approach concerned with continuity and smoothness of response surfaces misses the point.

This line of thinking might lead to some skepticism with regard concerning advantages coming from the widespread use of typical benchmarking functions. They generate rather formidably looking surfaces assumed to test adaptive abilities of algorithms to the highest. One might raise here a number of doubts:

1. Complication, in terms of multimodality doesn’t necessarily mean that the problem is difficult, as it was with the problem presented in Figure 6.

2. Benchmarks usually prize the best results in terms of the fitness achieved, thus mixing inexorably adaptive exploration (saddle crossing) with local optimization (hill climbing). This often happens to the benefit of the latter and at the expense of the former.
3. In this way benchmarks might patronize over emergence of methods narrowly adapted to do well with the conventional set of benchmarks and not necessarily suited to regularities typical for evolutionary problems.

There is still an overwhelming need to understand, what the real predicaments of the evolutionary search for optima are. Knowing this, it would be possible to make benchmarks suitable to test the specific valours of evolutionary algorithms. This could pave the way to the custom made optimization procedures.



References

- [1] Simon, H. A. Models of Bounded Rationality, MIT Press, 1982.
- [2] Kuhn, T. *The Structure of Scientific Revolutions*. University of Chicago Press, 1962.
- [3] Eldredge, N., Gould, S. J. Punctuated equilibria: an alternative to phyletic gradualism. [In:] *Models in Paleobiology*, 1972.
- [4] Galar, R. Evolutionary search with soft selection. *Biological Cybernetics*, 60: 357–364, 1989.
- [5] Galar, R. Simulation of local evolutionary dynamics of small populations. *Biological Cybernetics* 65:37–45, 1991.
- [6] Chorazyczewski A., Galar R. Evolutionary Dynamics of Population States. [In:] Congress on Evolutionary Computation, Seoul, 2001.
- [7] Arabas, J., Adamski K. Influence of mutation range on dynamic properties of an evolutionary algorithm. In: *Algorytmy Ewolucyjne i Optymalizacja Globalna*. Korbiewów, 2005.
- [8] Wright, S. The roles of mutation, inbreeding, crossbreeding and selection in evolution. [In:] *Proc. 6th Int. Cong. Genet.* 1: 356–366, 1932
- [9] Galar R., The impact of nonheritable variance on simulated evolutionary adaptation. [In:] *Algorytmy Ewolucyjne i Optymalizacja Globalna*. Zielona Góra, 2003.
- [10] Eigen, M. Macromolecular Evolution: Dynamical Ordering in Sequence Space. *Ber Bunsenges Phys Chem* vol. 89, 658-667, 1985.