TECHNICAL UNIVERSITY OF CLUJ-NAPOCA

The 2nd International Conference
ADEMS'09
Advanced Engineering in Mechanical Systems 2009

MAKING THINGS WORK
Machine Elements and Tribology

24 - 25 September 2009

NOTES ON THE CONNECTIONS BETWEEN SHAPE DEFINITION AND THE OBJECTIVE FUNCTION LANDSCAPE

András SÓBESTER

Abstract: *The key to effective shape optimization is the selection of the appropriate mathematical formulation for the parametric description of the geometry of the artifact being optimized. It is widely understood that a good parameterization scheme is concise, mathematically well-posed, robust and flexible. What is less clear, however, is the way in which the choice of parameterization approach influences the features of the resulting objective function landscape. In this article we examine the issue through a simple, four-variable design problem.* ***Key words:*** *optimal design, shape optimization, geometry design, brachistochrone, modality*.

1. PARAMETRIC GEOMETRIES IN OPTIMAL DESIGN

The cornerstone of effective shape optimization is the mathematical formulation that describes the geometry of the object whose shape we seek to optimize. There are a number of criteria such parameterization schemes are usually expected to satisfy.

The first, and arguably the most important, requirement is *conciseness*. In other words, the number of parameters needs to be kept to a minimum in order to reduce the dimensionality of the resulting design space (it is impossible to over-emphasize the importance of this requirement: the cost of exploring a design space increases exponentially with the number of its dimensions).

Fulfilling the second requirement is generally made difficult by the fact that it often conflicts with the first one: the geometry model has to be *flexible*. It has to be able to cover a broad range of possible shapes, especially if it is likely to see action early in the design process (this conceptual phase usually requires the greatest amount of flexibility).

Further, one usually expects these formulations to be robust, mathematically well-posed, CAD-compatible and, if possible, they should have parameters that have some measure of intuitive significance (this enables the manual 'tweaking' of designs and generally permits all operations based on engineering knowledge).

In this paper we suggest an additional consideration: the parameterization should lead, once an objective functional (or function) is assigned to the geometry, to a landscape with benign characteristics from an optimization standpoint. We have already seen that dimensionality is one of the main drivers here, but let us look one step further. Given a certain dimensionality, can our choice of parameterization scheme have an effect on other features, such as modality or ease of approximation?

The easiest route towards gaining an insight into this question is via a low dimensionality toy problem, which we describe next.

2. AN ILLUSTRATIVE PROBLEM – THE BRACHISTOCHRONE

In order to gain an insight into the impact of the choice of parameterization method on the shape of the resulting objective function landscape, let us consider a simple problem, that of the *brachistochrone*. This is formulated as follows. A ball rolls down a track, starting with zero velocity. What is the shape of the track that will minimize the time the ball takes to roll down it? (we ignore the effects of friction and drag on the ball).

We know the answer since 1696, courtesy of Newton [1]: it is a cycloid segment. We shall not delve into the details of his solution (it is one of the standard results of the calculus of variations); instead, we will consider numerical approximations of this optimum shape using a series of parameterizations of the shape of the track.

Let us define the track T as a function of four design variables x_1, x_2, x_3 and x_4. To each track shape $T(x_1, x_2, x_3, x_4)$ we can now attach an objective functional value $\mathcal{T}(T)$, which is the time it takes a ball to roll down the track T. Computationally, this is easy to approximate by discretizing the track into a number n of straight segments over which the equations of motion are solved (a convergence study of experiments with increasing resolution will reveal the appropriate choice of n) and it therefore gives us an inexpensive way of charting the objective functional across the entire design space.

Perhaps the simplest possible parameterization of T is shown in Figure 1. The space between the abscissas of the starting point and the endpoint is filled with four interpolation points controlling a standard b-spline – these points are equally spaced along the horizontal axis. The vertical coordinates of the points determine the shape of the track.

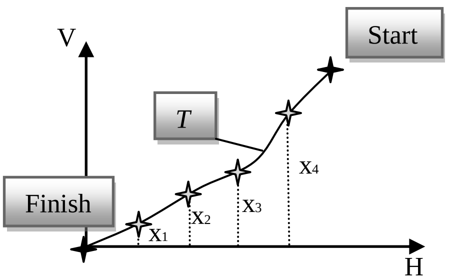

Fig. 1 Formulation ONE. Track shape defined by four interpolation points with equally spaced abscissas.

We assume that the ordinate of the starting point is one, so here (as in the case of all the other formulations that follow) $x_1, x_2, x_3, x_4 \in [0,1]$.

Let us now look at the resulting objective function landscape, as shown in Figure 2. This is a nested contour plot of the time it takes the ball to roll down the track $T(x_1, x_2, x_3, x_4)$, starting from the point (1.57,1) to the origin. Each tile of the plot represents the objective value versus x_3 and x_4, while the values of x_1 and x_2 can be read off the main axes.

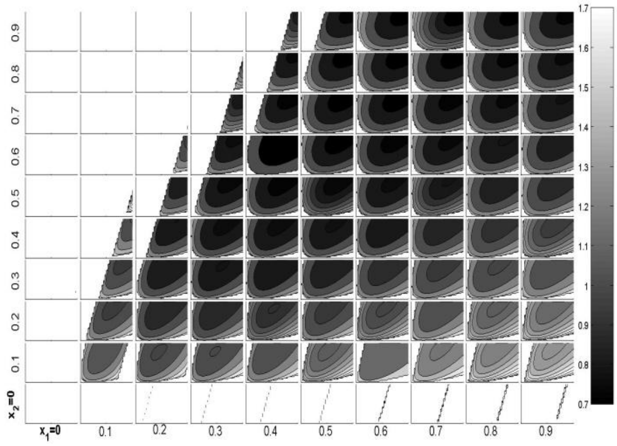

Fig. 2 Objective function contour plot for Formulation ONE (four interpolation points with equally spaced abscissas).

This is not a bad first effort. The surface is unimodal (has a single, global minimum) and the shape of its basin of attraction is close to spherical (a feature that makes it amenable to a quasi-Newton-type local search). Note the blank regions in the plot – these correspond to nonsensical tracks (for example tracks where the ball gets stuck partway down).

Let us now consider an alternative parameterization, where we allow the abscissas of the interpolation points to vary too, but, to keep the dimensionality the same, we need to sacrifice two of them. A naïve implementation of this idea is shown in Figure 2.

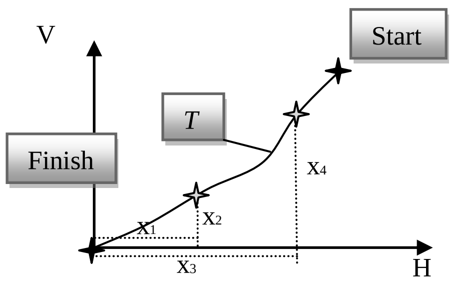

Fig. 3 Formulation TWO. Track shape defined by a b-spline with two interpolation points with variable abscissas and ordinates.

Both abscissa variables (x_1 and x_3) sweep the entire range from the projections of the starting and finish points, so we define the geometry in such a way that the spline goes through them in the order of their values (that is, it does not loop around if the points happen to overtake each other).

Figure 4 shows the corresponding objective function (rolling time) landscape.

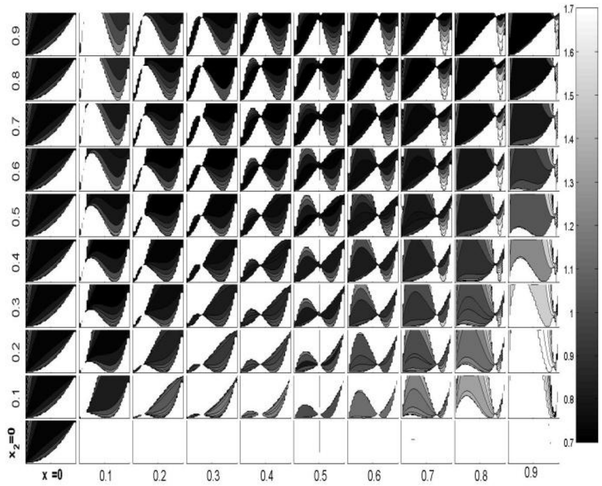

Fig. 4 Objective function contour plot for Formulation TWO (b-spline with two interpolation points sweeping the whole of the abscissa).

Clearly, this is bad news. We can now see a well-known ogre of local optimization: the landscape has two optima of comparable depths. Worse still, their basins of attraction are surrounded by infeasible regions of the design space.

It is worth emphasizing here that multimodality at this scale is not an issue for any but the most basic optimizers. This is merely a toy example meant to illustrate what can happen – the real point of this exercise is to show how relatively easily one can fall into such traps, which can make optimization intractable when there are, say, 40 variables instead of the four shown here and the objective function takes hours to compute, not fractions of a second, as in this case.

Let us now look at a possible way of avoiding the trap of multimodality here. What happens if we repeat the process, but this time we only allow x_1 to sweep half of the distance and x_3 the other half (let us call this Formulation THREE). With the two interpolation points not allowed to overtake each other anymore, the resulting objective function surface can be seen in Figure 5.

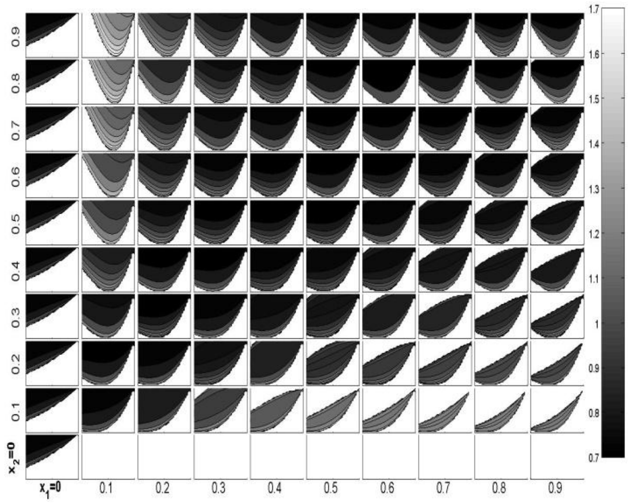

Fig. 5 Objective function contour plot for Formulation THREE (similar to two, but each interpolation point allocated its own half of the full distance).

Unsurprisingly, we are only left with a single optimum now, though its basin of attraction is slightly elongated, a feature that can make optimization slightly more difficult (once again, this is a trivial example, but larger scale problems have the same traps and their cost can be substantial).

What else could we do with two interpolation points? Might it be worth using the (variable) abscissa of the first point as the datum for the abscissa of the second point – as shown in Figure 6?

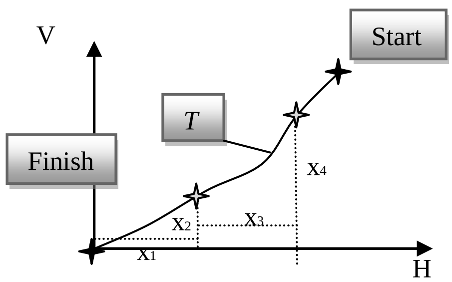

Fig. 6 Formulation FOUR. Two ordinate variables as before, but this time the second abscissa variable has a moving datum.

The resulting (rather similar) objective function landscape is shown in Figure 7.

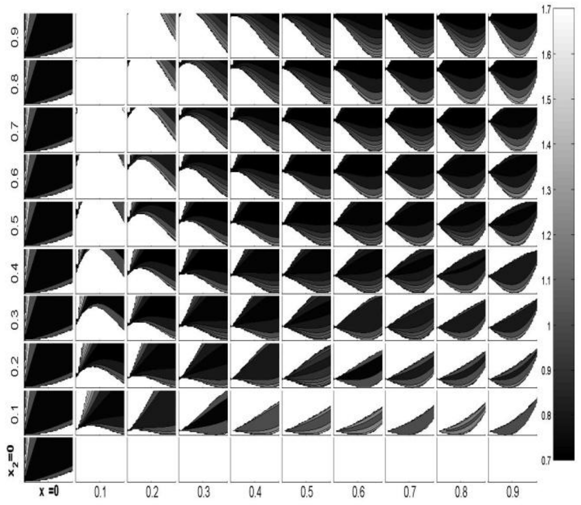

Fig. 7 Objective function contour plot for Formulation FOUR.

Finally, let us consider changing the formulation altogether, by opting for a different type of curve – a Non-Uniform Rational B-Spline (NURBS). The reader interested in the details of shape description via NURBS may wish to consult the excellent text of Piegl and Tiller [2] – here we limit ourselves to stating that these, by comparison to the b-splines discussed earlier, offer an additional means of shape control. The track shape is defined through two so-called control points (instead of the interpolation points used earlier), which also each have a *weight* parameter assigned to them.

Formulation FIVE, then, is a track defined as a NURBS curve clamped at its ends (the starting and finishing points) and controlled by two fixed abscissa control points in-between. The ordinates of the two control points and their corresponding weights make up the four design variables. Figure 8 shows the resulting objective function landscape.

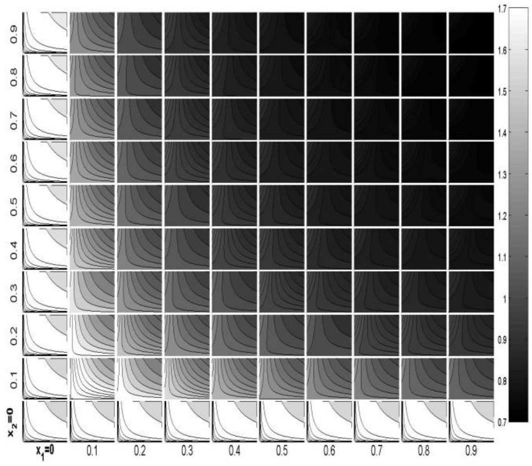

Fig. 8 Objective function contour plot for Formulation FIVE. Note that the entire design space is feasible here, though this is likely to entail a flexibility sacrifice (nevertheless, the optimum – the best approximation of a cycloid – is included in the design spaces of this and all the other cases presented earlier).

3. CONCLUSIONS AND FUTURE DIRECTIONS

The choice of parameterization scheme can have a strong influence on the characteristics of the objective landscape. Even apparently very similar formulations can yield radically different landscapes. Beyond the anecdotal evidence presented here, further research is required to quantify the phenomena discussed above.

4. REFERENCES

[1] Boyer, C. B. and Merzbach, U. C. *A History of Mathematics, 2nd ed.* , Wiley, New York, 1991.

[2] Piegl, L., Tiller, W. *The NURBS Book*, Springer, 1996.

Legaturi intre Descrierea Formei si Forma Functiei Obiectiv

Cheia optimizarii formei unui obiect este alegerea corecta a formei parametrizate a descrierii geometriei sale. Se cunosc multe criterii ce trebuie satisfacute de o asemenea descriere matematica: ea trebuie sa fie concisa, robusta, flexibila, etc. Influenta tehnicii de parametrizare asupra formei functiei obiectiv ce rezulta este, insa, mult mai putin clara. In acest articol vom examina problema prin prisma unei probleme simple, clasice de optimizare a formei.

András Sóbester, PhD, Royal Academy of Engineering Research Fellow, Lecturer, University of Southampton, School of Engineering Sciences, a.sobester@soton.ac.uk, +44 23 8059 2350.