Generating optimal 2D structural designs using simulated annealing

N. Gunduz, N. Akbulut and F. O. Sonmez

Department of Mechanical Engineering,
Bogazici University, Turkiye

Abstract

The aim of this study is to generate global optimal shapes for two-dimensional structures subject to given in-plane loads and restraints. The design constraints are the maximum allowable stress and model connectivity, while the objective function to be minimized is the volume (or the weight) of the structure. Optimisation is achieved by a stochastic search algorithm called simulated annealing, which seeks the global minimum through randomly generated configurations. In order to obtain random configurations, a boundary variation technique is used. In this technique, initially, a set of design points is chosen and connected by cubic splines to describe the boundary of the structure. Whenever the position of a randomly chosen keypoint is changed in a random direction, a new configuration is obtained. A FE structural analysis of these randomly generated configurations is carried out in order to determine the value of the cost function to be minimized. If an improvement is made in the cost, the algorithm accepts the new configuration. But, it also allows occasional increases in the cost function to avoid getting stuck into a local minimum point, and thus to converge into the globally minimum point. In this study, a general computer code was developed using ANSYS parametric design language to solve the shape optimisation problems involving two-dimensional structures. A number of cases were examined to check the effectiveness of the optimisation procedure. The results show that this technique can be applied with high accuracy to the two dimensional shape optimisation problems.
1 Introduction

The general purpose of shape optimisation is to find the best shape (and thus the most efficient and effective use of material) for a structure under various constraints imposed by the requirements of the design.

Typical structural optimisation problems may contain many locally minimum configurations. As a result, a deterministic approach may get stuck into any local minimum point rather than globally minimum solution depending on the starting point. In order to find the absolute optimum of an objective function, a global optimisation method has to be employed. Stochastic optimisation techniques are quite suitable in this respect. Among their advantages, they are not sensitive to starting point, they can search a large solution space, and they can escape local optimum points because they allow occasional uphill moves. The simulated annealing (SA) algorithm is one of the most popular stochastic optimisation techniques. SA is relatively easier to implement and can be configured for any application. It has less control parameters than those of genetic algorithms.

Kirkpatrick et al. [1] first proposed SA as a powerful stochastic search technique. The method gets its name from the physical process whereby the temperature of a solid is raised to a melting point and then slowly cooled. The method attempts to model the behaviour of atoms in forming arrangements in liquid or solid material during annealing. When a material is crystallized from the liquid phase, it must be cooled slowly if it is to assume its highly ordered, lowest energy state. However, if the cooling is too rapid, the material does not have time to reach equilibrium, and assumes a less ordered higher energy state.

There is an analogy between this phenomenon and an optimisation process. Different configurations of the problem correspond to different arrangements of the atoms. The cost of a configuration corresponds to the energy of the system. Optimal solution corresponds to the lowest energy state. Just in the same way the atoms find their way to build a perfect crystal structure through random movements, global optimum is reached through a search within randomly generated configurations. In the SA algorithm, a random initial point is selected and systematically updated until a stopping criterion is satisfied. Updating is based on an iterative procedure. In each iteration, a random point is generated in the neighbourhood of the current configuration. If the new point has a smaller value for the cost function compared to that of the current record, the point is accepted. This point replaces the old one. If the new cost function has a bigger value than that of the current record, the acceptability of the point is decided according to the probability of Boltzman distribution. The calculation of this probability depends on a control parameter, c, which plays a similar role as temperature the physical annealing process. The control parameter is reduced slowly so as not to get trapped at a local minimum point. At initial stages of optimisation (at high temperatures), the probability of accepting worse designs is higher but later on at low temperatures, it becomes smaller and smaller so that in the end the designs having higher cost are almost never accepted.

Application of SA to the shape optimisation of structures is quite rare. Anagnostou and his coworkers [2] used a SA method to design a thermal fin
with minimum material. Shim and Manoochehri [3] used the SA algorithm to obtain optimal designs for two-dimensional structures with a minimum use of material. The constraints of the problem were model connectivity and allowable stress. They used an element removal technique to generate random configurations through which the SA algorithm searched the optimum shape. In their approach, a structure was first divided into a number of small finite element blocks and then these blocks were randomly removed or restored to obtain a new shape. One difficulty with this method is that removing and restoring an element may violate model connectivity. Whenever a new configuration is generated its connectivity should be checked, by a complicated algorithm with questionable reliability. Another difficulty is the roughness of the resulting boundaries. If smooth boundaries are desired, small elements, much smaller than the accuracy of the finite element analysis requires, should be used.

In the present study, we adopt a similar approach as Shim and Manoochehri [3]. We also use the SA algorithm to search through randomly generated configurations. However, we utilize another technique called boundary variation to generate random configurations so that the drawbacks mentioned above can be avoided. Although the idea of boundary variation was used before [4, 5], as far as we know, it is the first time to incorporate this idea into the SA algorithm to obtain random shape configurations during the optimisation process.

2 Problem statement

Figure 1 shows a two dimensional structural problem. In-plane loads are acting on a structure with uniform thickness, so that resulting displacements and stresses are in-plane, i.e. their out-of-plane components are zero. The structure should be able to resist the loads without failure. This means stresses should not exceed the yield strength of the material. Also, no part of the structure is to lose its connection to the restraints. The structure should remain in one piece. Therefore, the constraints imposed on the structure are the maximum allowable stress and the model connectivity. Our objective is to minimize the volume (or weight) of the structure, in other words, to obtain a configuration (shape) with the most efficient use of material without violating the constraints.

Figure 1 Representation of a 2D shape optimization problem
3 Problem solution

3.1 The boundary variation technique

SA requires random generation of a new configuration (in our case a new shape) in each iteration. The shape of a 2D-structure having a uniform thickness is defined by its boundary. Accordingly, its shape can easily be described through spline curves passing through keypoints. As illustrated in Figure 2, whenever the position of a randomly chosen keypoint is changed in a random direction, a new boundary, thus a new configuration is obtained.

![Figure 2: Generation of a random configuration by boundary variation.](image)

3.2 Acceptability

Acceptability of the new configuration depends on the value of the cost function as well as constraint violation. In our optimization scheme, when the constraint is violated for a generated configuration, a penalty function is introduced which reduces the cost acceptability of the configuration in the following manner [3]:

\[ P(j) = A_C(j) \cdot A_F(j) \] 

(1)

Here \( P(j) \) is the acceptability of the design for the \( j^{th} \) iteration, \( A_C(j) \) is the cost acceptability, and \( A_F(j) \) is the constraint acceptability. When the cost is lower, \( A_C(j) \) is equal to one. When the constraint is satisfied, acceptability parameter, \( A_F(j) \), takes the value of one. This means every new design having a lower cost is accepted if it is in the feasible region. But, if the cost is higher or the constraint is violated, the new configuration may be accepted depending on the value of \( P(j) \). The values for \( A_C(j) \) and \( A_F(j) \) are calculated as follows [3]:

\[ A_C(j) = \begin{cases} 1 & (\Delta E_j \leq 0) \\ \exp(-\Delta E_j / (c_k \cdot E_0)) & (\Delta E_j > 0) \end{cases} \] 

(2)

\[ A_F(j) = \begin{cases} 1 & (F_j \leq F_{allow}) \\ \exp(-F_j / (c_k \cdot F_{allow})) & (F_j > F_{allow}) \end{cases} \] 

(3)

where \( F_j \) is the value of the constraint (the maximum von Mises stress) in the iteration \( j \), \( F_{allow} \) is the allowable value of the constraint (yield strength of the
material), $E_0$ is the initial value of the cost function (volume of the initial design), and $\Delta E_j$ is the change in the cost function, given by $E_j - E_i$. Here $E_i$ is the cost (volume) of the previously accepted configuration ($i$) and $E_j$ is the cost (volume) of the new configuration ($j$). If model connectivity is violated, $A_M(j)$ is zero.

### 3.3 Markov chain

The SA process consists of first "melting" the system being optimised at a high effective temperature, $c$, then lowering the temperature in slow stages until the system 'freezes' and no further changes occur. At high values of $c$ the probability of acceptance is higher as eqns (2) and (3) imply. At each value of $c$, iterations should be repeated until no further improvement can be achieved. These iterations during which the value of $c$ is kept constant are called Markov chains. The number of iterations is called the length of the Markov chain, $L_k$. As the value of $c$ changes, $L_k$ is changed during the optimisation process. However, a fixed length imposes no major restrictions; it merely simplifies the calculations [6]. In this study, the length of the Markov chains, $L_k$, depends on the equilibrium of the trials in that Markov chain. When the equilibrium is reached, Markov chain is ended. In this study, the equilibrium criterion is defined as [3]

$$\frac{|E_j - E_{j-1}|}{j} \leq \varepsilon_{inner}$$  \hspace{1cm} (4)

where $j$ is the iteration number and $E_j/j$ is the average cost for $j$ trials, and $\varepsilon_{inner}$ is a suitable chosen small number controlling the length of the Markov chain (in our case, its initial value is 0.001). As the value of $c$ decreases, the number of iterations in a Markov chain that should be carried out to reach equilibrium increases, accordingly the value of $\varepsilon_{inner}$ should become smaller. The following equation was suggested [7] to modify $\varepsilon_{inner}$:

$$(\varepsilon_{inner})_k = (\varepsilon_{inner})_{k-1} \left( \frac{c_k}{c_{k-1}} \right)$$  \hspace{1cm} (5)

where $c_k$ and $c_{k-1}$ indicate the present and previous values of the $c$, respectively.

### 3.4 Cooling schedule

A cooling schedule controls the convergence of the algorithm to the global minimum, and specifies first an initial value of the control parameter $c_0$, secondly decrement function for decreasing the value of $c$, thirdly, the number iterations to be done at each value of the control parameter, and lastly the final value of $c$.

The value of the initial control parameter, $c_0$, should be large enough to allow virtually all trials to be accepted. In the physical analogy mentioned earlier, this corresponds to heating up the solid until all particles are randomly arranged in the liquid phase so that atoms may freely arrange themselves. This is achieved by requiring that the initial acceptance ratio $\chi_0 = \chi(c_0)$ be close to one. In this study a value of 0.9 is chosen for $\chi_0$. By following the procedure described in [6], the value of $c_0$ is determined.
Then, a decrement rule must be established to find the subsequent values of $c$. The probability of reaching the global optimum solution depends on how fast the value of $c$ is lowered. If the cooling rate is fast, the optimisation process will probably end up with one of the high-energy local minima. If the cooling rate is slow, the global optimum solution can be reached. In the polynomial cooling schedule, the decrement rule for $c$ of the $k^{th}$ Markov chain is given by [6]:

$$c_{k+1} = \frac{c_k}{1 + \frac{c_k \ln(1 + \delta)}{3\sigma_c}}$$

$k = 0, 1, 2, 3 \ldots$ (6)

where $\sigma_c$ is the standard deviation of the values of the cost in the $k^{th}$ Markov chain. The amount by which the value of $c$ is decreased by the decrement function of (8) depends on the value of $\delta$, called the distance parameter. Very large values of $\delta$ lead to large decrements in the control parameter, which may cause the optimisation process to be entrapped in a local minimum point. Based on the experience, 0.45 is a suitable value of $\delta$ for the problem at hand.

Termination of the algorithm is based on an extrapolation of the expected cost $(E)_{ck}$ as $c_k$ goes to zero. Accordingly the termination criterion is [6]:

$$\left. \frac{E_{ck} \partial (E)_{ck}}{\partial c} \right|_{c=c_k} < \varepsilon_s$$

(7)

where $\varepsilon_s$ is suitably chosen small positive number (in our case 0.0001), called stopping parameter and $E_0$ is the initial cost value.

Besides, one should determine how wide the neighbourhood, in which the optimum point is searched. Let us call the extent of how far the new configuration is situated from the previously accepted configuration as the step size. Increasing the step size enlarges the neighbourhood to be searched, but decreases the accuracy. If a small step size is chosen, one can get close to the minimum point, but at the cost of large numbers of iterations. Shim and Manoochehri [3] used a constant step size in their study like the other authors Botkin and Yang [8]. In some studies, i.e., Kalivas [9], step size was selected randomly within a given range. Theodoracatos and Grimsley [7] decreased the step size whenever $c$ was decreased. In this study, we adopted a different approach. We decreased the step size, $R$, after the neighbourhood of the present solution was sufficiently searched without finding a better solution.

$$R_n = bR_{n-1} \quad \text{if} \quad r > r_{\text{limit}}$$

(8)

where $n$ is the number of the decreases in step size, $r$ is the number of consecutive rejected moves. The value of its limiting number, $r_{\text{limit}}$, was chosen as 50, while the value of $b$ was determined by trial and error as 0.5. The initial extent of movement of the design points, $R_0$, should be chosen according to the dimensions of the structure. In this study, $R_0$ is taken as

$$R_0 = (W + L)/200$$

(9)

where $W$ and $L$ are roughly the width and length of the initial design.
3.5 Solution procedure

Figure 3 presents the procedure followed in this study to find the optimum shape of 2D structures. Here $k$ and $j$ show the Markow chain number and number of changes in $c$ respectively.

![Solution procedure diagram](image)

Figure 3 The solution procedure.

In the first step of the algorithm, the designer states the initial shape of the structure, the loads applied to the structure, the restraints, the material properties, and the constraints of the design. Initial geometry of the design can also be
generated randomly by the computer. In the initial design, violation of the constraints is permissible. In the next steps, a new random configuration is generated by giving a small change to the position of one of the keypoints in a random direction, and then passing spline curves through them to describe its boundary. Then, the finite element analysis of the design is carried out using automatic mesh generation to calculate its cost function. The corresponding data are recorded for the next step.

The design acceptability of the new configuration is calculated according to eqn (1). If its value is greater than a randomly generated number, the new configuration is accepted, otherwise rejected. If the number of consecutively rejected moves exceeds a limit value, the step size is reduced in accordance to eqn (7) in order to contract the neighbourhood to be searched. If no further improvement can be accomplished for the current value of the control parameter according to eqn (4), the value of the control parameter is reduced as described in eqn (6). Besides, the parameter, $\varepsilon_{\text{inner}}$, controlling the length of the Markov chain is also reduced. When the stopping criterion defined by eqn (7) is satisfied, the optimisation process is ended.

4 Results and discussions

One of the problems that we considered in this study is optimal design of a plate constrained at one end and subjected to forces at the other end (Figure 4). The plate has the dimensions of 100mm×50mm×1mm. It is made of a material having a modulus of 300 N/mm$^2$ and a Poisson’s ratio of 0.33. Allowable stress is 300 MPa. Figure 5 shows the initial configuration with its finite element mesh. Because of the symmetry, only upper half of the plate is analysed. The left hand side is constrained for all degrees of freedom while the bottom is constrained only in the vertical direction. Initial volume is 2323.7 mm$^3$. The number of keypoints used to create the profile of the boundary is seven. The keypoints located on the part of the boundary which is either constrained or subjected to external force are fixed the others are free to move. Figure 6 shows the stress state in the initial model. The maximum stress is 322 MPa. Initial value of the control parameter, $c_0$, is 0.026.

![Figure 4 The design problem](image-url)
Figure 5 The initial configuration with its finite element mesh.

Figure 6 The stress state in the initial configuration.

Figure 7 shows the stress state in the final configuration. The maximum stress is 297 MPa, which is lower than the allowable stress (300 MPa). One may notice that most of the lower stressed regions in the initial configuration were removed and nearly all of the final structure is fully stressed. This is an indication of the effectiveness of the optimisation scheme. The boundary is smooth with no sharp edges leading to stress concentrations. The final volume is 607.7 mm$^3$, which is a considerable improvement in comparison to the initial configuration.

5 Conclusions

A 2D shape optimisation procedure based on the SA method was presented in this study. One shape design problem was chosen. The method successfully found the optimum design.

Despite the advantages of the method, determining suitable parameters for the cooling schedule poses difficulty. Some of these parameters are found by trial and error. If the most appropriate cooling schedule parameters are found for
the particular problem at hand, the resulting configuration will be close to the
global optimum without excessive computational effort. Once suitable
parameters are found, the algorithm can be applied to other 2D shape
optimisation problems by minor modifications of these parameters.

In this study, in order to generate random configurations, we used the
boundary variation technique. The main advantage of this technique is that low
numbers of design variables can be used, so that the time to generate a random
configuration as well as the convergence time to the global optimum solution is
shorter compared to the element removal technique. Moreover, in this technique,
it is easy to handle the model connectivity. In this case it is sufficient to control
whether or not the spline segments of the boundary cross each other. By using
large numbers of keypoints, description of the shapes is more precise, but at the
expense of increased convergence time. Also, another problem of using too
many keypoints is the possibility of severe distortion of the model during the
optimisation process that may lead to unreliable finite element solution. This
difficulty arises when the automatic meshing can not handle very sharp edges.
However, this problem can be avoided by using a finer mesh.

Acknowledgment

This paper is based on the work supported by the Research Fund of Bogazici
University with the code number 99A603.

References

1. Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P. Optimization by simulated
3. Shim, P.Y. & Manoochehri, S. Generating optimal configurations in
procedure applied to the design of gas turbine discs. Computers & Structures,
5. Haftka, R.T. & Grandhi, R.V. Structural shape optimisation. Computer
6. Aarts, E. & Korst, J. Simulated Annealing and Boltzmann Machines, John
7. Theodoracato, V. E. & Grimsley, J. L. The optimal packing of arbitrarily-
shaped polygons using simulated annealing and polynomial-time cooling
8. Yang, R.J. & Botkin, M.E. Three dimensional shape optimization with
9. Kalivas, J. H., Adaptation of Simulated Annealing to Chemical Optimization