ENGINEERING OPTIMIZATION

A projekt keretében elkészült tananyagok:

Anyagtechnológiák

Materials technology

Anyagtudomány

Áramlástechnikai gépek

CAD tankönyv

CAD book

CAD/CAM/CAE elektronikus példatár

CAM tankönyv

Méréstechnika

Mérnöki optimalizáció

Engineering optimization

Végeselem-analízis

Finite Element Method

Budapest University of Technology and Economics
Faculty of Mechanical Engineering

Óbuda University
Donát Bánki Faculty of Mechanical and Safety Engineering

Szent István University
Faculty of Mechanical Engineering

ENGINEERING OPTIMIZATION

Course bulletin

Editor:
GÁBOR KÖRTÉLYESI

Authors:
CSILLA ERDŐSNÉ SÉLLEY
GYÖRGY GYURECZ
JÓZSEF JANIK
GÁBOR KÖRTÉLYESI

TYPOTEX

2012

READERS: Prof. Dr. Károly Jármai

ISBN 978-963-279-686-4

PREPARED UNDER THE EDITORSHIP OF Typotex Publishing House

RESPONSIBLE MANAGER: Zsuzsa Votisky

GRANT:

Made within the framework of the project Nr. TÁMOP-4.1.2-08/2/A/KMR-2009-0029, entitled „KMR Gépészmérnöki Karok informatikai hátterű anyagai és tartalmi kidolgozásai" (KMR information science materials and content elaborations of Faculties of Mechanical Engineering).

Nemzeti Fejlesztési Ügynökség
www.ujszechenyiterv.gov.hu
06 40 638 638

MAGYARORSZÁG MEGÚJUL

A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

KEYWORDS:

shape optimization, topology optimization, sensitivity analysis, Design of Experiments, engineering optimization, robust design, optimization of production processes, constrained optimization, gradient methods, multiobjective optimization, evolutionary optimization methods

SUMMARY:

The handout presents the design optimization tasks occurring in the mechanical engineering practice and clarifies the basic concepts needed to set up optimization model.

The optimization algorithms are explained from simple univariate and unconditional techniques up to the multivariate and conditional methods based on these, taking into account widely usable direct methods, gradient methods which have favorable convergence properties and random procedures having chance to find the global optimum. Encountering with complex engineering optimization problems in the sensitivity analysis helps to reduce the size of the task, but in given time to solve them the design of experiments methods are the natural choice. The uncertainties in engineering systems can be treated by the robust design. The available optimization softwares are shown as practical implementation of these methods and techniques to support the mechanical engineering design. After the methods assisting the planning process (optimization machine parts and components), we are looking at the optimization possibilities of production processes. Applying a complex systems-based model on enterprise level, we learn how to be utilized high-quality machines in the production process and objective decision-making, supplying information to complex technical-economic qualification.

CONTENTS

1. INTRODUCTION

This electronic book was supported by TÁMOP-4.1.2/08/2/A/KMR. Basically this curriculum is addressed to the B.Sc. students but the authors hope that the other readers will interested in this book.
To understand the main fields of this ebook, some mathematical, basic informatics, mechanical, CAD and finite element background knowledge is required. The necessary CAD and finite element knowledge can be found in two other electronic books, which are also developed in the frame of the TÁMOP-4.1.2/08/2/A/KMR project so we refer to these.
In the libraries and on the internet there are lots of books, educational themes in this topic, which all show that this field is very popular and important. The industrial specialists show interest in the optimization researches so it is important that the results of research should be applied in wide range. Nowadays the optimization software's are not only a research tool, but they are included into the commercial software packages becoming basic part of the design process. The other trend of the development is that every industrial finite element software packet has optimization module.
In the light of this development, the following question arises: what is the difference between this electronic book and the hundreds of optimization books can be found in the libraries? The authors trying to introduce the solution method of real world CAD based optimization problems, arising in the mechanical engineering practice. Overcome to the simple analytical optimization problems we will focus on the different numerical solution techniques which can be advantageously applied during the optimization of the real machine parts.

1.1. The structure of the book

In the introduction section of this electronic book the formulation of the optimization problem will be discussed, clarifying the basic definitions. Then we present some typical optimization problems, showing a wide range of applicability of this theme.

In the first part of the educational material we introduce the most important basic optimization methods. Of course, there is no opportunity to introduce all type of optimization methods in details, so we are going to focus on the procedures used by later on solving the CAE based optimization problems.

In the next chapter, the problem classes (and the suggested solution techniques) will be discussed in the case of solving mechanical engineering problems. The different methods and usage of the sensitivity analysis is also presented, and the role of the topology-and shape optimization in the whole design process will be shown. In this section we are dealing with the analysis of complex practical problems, and briefly describe the capabilities of industrial systems, which can be illustrated with an example.

The optimization tasks are not only arises in the design phase of the products, but they have a great importance in the production phase too. The best designed machine can be unsalable, without an economic production technology. In this chapter, the various processes presented an analysis of the optimal interaction, there are also highlighted the importance of optimum and suboptimum. Complex system-levels model through the use of company will show that how the high-quality machines can be economically used in the production process. We will present typical manufacturing process planning optimization examples.

1.2. Optimization in the design process

In the design process (from the basic ideas to the detailed digital mockup of the product), many aspects should be taken into account.
These requirements are usually connected to safety or economy of the product, but also the manufacturing, using repairing considerations come to the fore. For example the designer has to develop cheapest product, while a large number of additional conditions should be met.

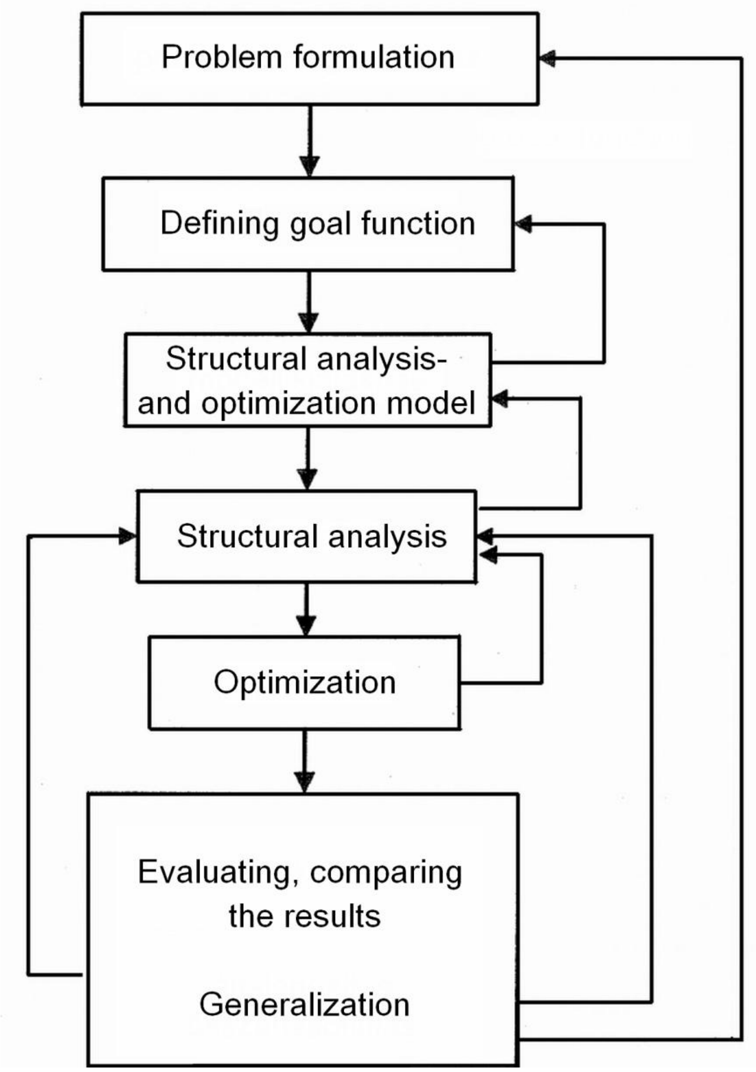


Figure 1.1.: The main steps of solving engineering optimization problem [1.1]

A real design process is usually not a linear sequence, but some steps are repeated, to obtain better solution. The main steps of solving engineering optimization can be seen in Figure 1. It is worth to remark, that the result of the design process, is depending on the accuracy of the introduced steps. For example, if a shape optimization should be taken and the structural responses are calculated numerically, than the error of the structural analysis, can result inaccuracy in the optimized shape too. The error can come from mesh density or form the incorrect modeling of the real boundary conditions as well.

Last time the usage of 3D CAD systems and numerical structural analysis techniques has been increased in the industrial applications. This is because; using these tools wide range of design processes can be covered. On the other hand the “time to market” became first objective so the computer-aided technique arises in the early stages of the design process. The modern numerical simulation tools provide an opportunity to decrease the number of the costly and time consuming physical experiments. Although the advanced numerical simulation tools and optimization procedures has a significant role in the product development in the reduction of the time, but the whole design process cannot be automated.
In case of solving structural optimization task we can choose from analytical and numerical methods. The analytical methods are often not suitable for the analysis of complex engineering problems, so this educational material mainly deals with the numerical techniques of the structural optimization. The numerical procedures based on iterative searching techniques which lead to an approximate solution. The searching process continues until the so-called convergence condition is satisfied, so we are sufficiently near the optimal solution.

1.3. The basic elements of an optimization model

1.3.1. Design variables and design parameters

A structural system can be described by a set of quantities, some of which are viewed as variables during the optimization process. Those quantities defining a structural system that are fixed during the automated design are called preassigned parameters or design parameters and they are not varied by the optimization algorithm. Those quantities that are not preassigned are called design variables. The preassigned parameters, together with the design variables, will completely describe a design.

<u>From the mathematical point of view three types of design variables can be distinguished:</u>

The design variables can be considered as continuous or discrete variables. Generally it is easier to work with continuous design variables, but a part of the real world problems contains discrete type of design variables. An intermediate solution, when we know that a large number of discrete design variable values should be considered, then it will be categorized as pseudo discrete. In this case, we solve the task considering this variable a continuous design variable and after the solution the closest possible discrete values will be checked.

<u>From the physical point of view there are four types of the design variables:</u>

1. Mechanical or physical properties of the material (Material design variable)
Material selection presents a special problem. Conventional materials; have discrete properties, as for example a choice is to be made from a discrete set of materials. If there have a few numbers of materials, the task is easier, if we perform the structural analysis for each material separately and comparing the results to choose the optimum material. A typical application is for reinforced composite materials to determine the angles of the reinforcements. Such type of design variables can be considered to be continuous ones.

2. topology of the structure, connecting members of the scheme, or the number of members of the interface schema; (Topology Design Variables)
The topology of the structure can be optimized automatically in certain cases when members are allowed to reach zero size. This permits elimination of some uneconomical members during the optimization process. An example of a topology design variable is if we looking for the optimal truss structure considering one design variable for each truss element (1 if the member exists or 0 if the member is absent). This type of design variables, according to the mathematical classification is not continuous.

3. The shape of the structure (Configurational or Geometric Design Variables)
This type of design variable leads us to the field of shape optimization. In case of machine design application, the geometry of the part should be modified close to the stress concentration areas, in order to reduce the stresses. On other hands, the material can be removed in the low stress areas, in order to make the structure lighter. So we are looking the best possible shape of the machine part. For example the variable surface of the structure can be described by B-Spline surfaces and the control nodes of such splines can be chosen as a design variable. This is a typical example for shape optimization, and these types of design variables are usually belong to the continuous category.

4. Cross-Sectional Design Variables or the dimensions of the built-in elements
Mainly for historical reasons, size-optimization is a separate category, which has got the simplest design variables. For example, the cross-sectional areas of the truss structure, the moment of inertia of a flexural member, or the thickness of a plate are some examples of this class of design variable. In such cases the design variable is permitted to take only one of a discrete set of available values. However, as discrete variables increase the computational time, the cross- sectional design variables are generally assumed to be continuous.

The design variables and design parameters are together clearly define the structure. If the design variables are known in a given design point, this completely defines the geometry and other properties of the structure. In order to guarantee this, the chosen design variables must be independent to each other.

1.3.2. Optimization constraints

Some designs are useful solutions to the optimization problem, but others might be inadequate in terms of function, behaviour, or other considerations. If a design meets all the requirements placed on it, it will be called a feasible design. In most cases, the starting design is a feasible design. The restrictions that must be satisfied in order to produce a feasible design are called constraints.
From a physical point of view the constraints can be separated into two groups:

Constraints imposed on the design variables and which restrict their range for reasons other than behaviour considerations will be called design constraints or side constraints. The geometrical optimization constraints describes the lower the upper limit of the design variables. These are expressed in an explicit form of the design variables. These could be for example a minimum and maximum thickness of the plate.

Constraints that derive from behaviour requirements will be called behaviour constraints. Limitations on the maximum stresses, displacements, or buckling strength are typical examples of behaviour constraints. This type of constraints based on a result of a structural analysis. Explicit and implicit behaviour constraints are both encountered in practical design. Explicit behaviour constraints are often given by formulas presented in design codes or specifications.

From the mathematical point of view, in most cases, constraints may usually be expressed as a set of inequalities:

$$g_j(x_i) \le 0 \quad (j=1,\ldots,m\ ;\ i=1,\ldots,n), \tag{1.1}$$

where m is the number of inequality constraints and x_i is the vector of design variables. In a structural design problem, one has also to consider equality constraints of the general form:

$$h_j(x_i) = 0 \quad (j=m+1,\ldots,p), \tag{1.2}$$

where p-m is the number of equalities. In many cases equality constraints can be used to eliminate variables from the optimization process, thereby reducing their number.

The equality-type constraints can be used to reduce the number of design variables. This type of constraints may represent also various design considerations such as a desired ratio between the width of a cross section and its depth.

We may view each design variable as one dimension in a design space and any particular set of variables as a point in this space. In case of two design variables the design space reduces to a plan, but in the general case of n variables, we have an n-dimensional hyperspace. A design which satisfies all the constraints is a feasible design. The set of values of the design variables that satisfy the equation $g_j(x_i) = 0$ forms a surface in the design space. It is a surface in the sense that it cuts the space into two regions: where one where $g_j(x_i) > 0$ and the other $g_j(x_i)<0$. The design space and the constraint surfaces for the three-bar truss example are shown in Figure 1.2. The set of all feasible designs form the feasible region. Points on the surface are called constrained designs. The j^{th} constraint is said to be active in a design point for which $g_j(x_i) = 0$ and passive if $g_j(x_i) <0$. If $g_j(x_i) > 0$ the constraint is violated and the corresponding design is infeasible.

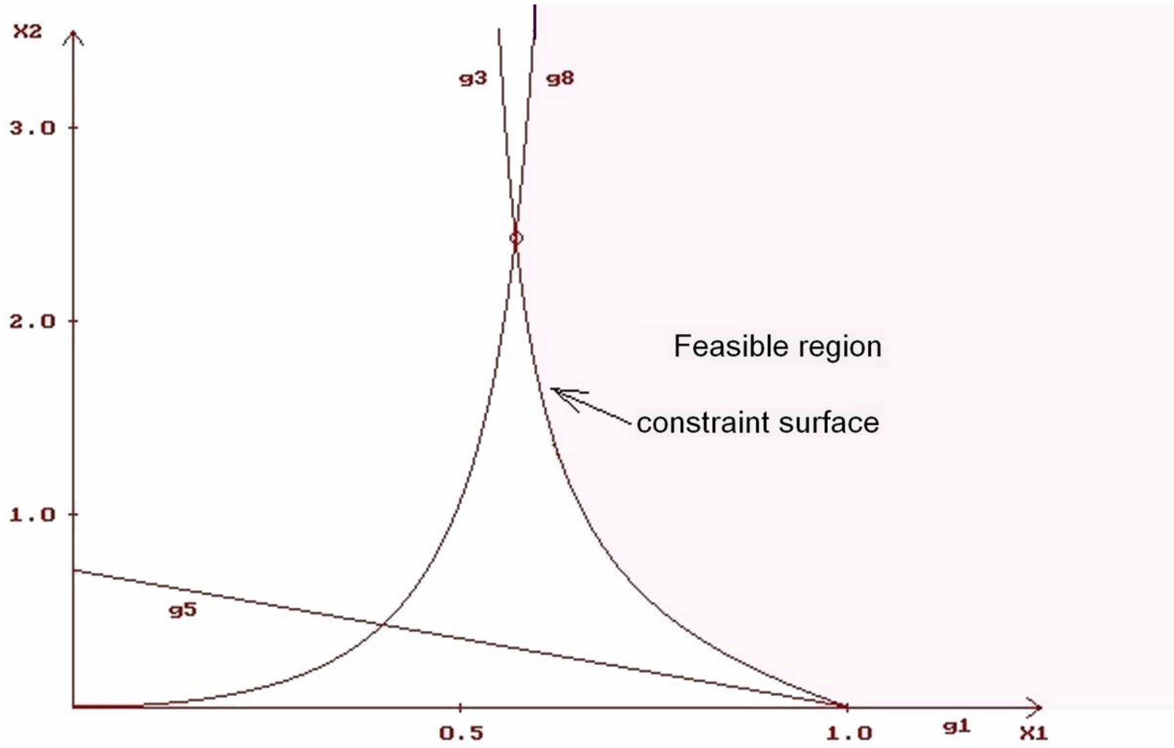


Figure 1.2.: Optimality constraints in the space of the design variables

1.3.3. The objective function

There usually exist an infinite number of feasible designs. In order to find the best one, it is necessary to form a function of the variables to use for comparison of feasible design alternatives. The objective function (or cost function) is the function whose least value is usually-wanted in an optimization procedure. It is a function of the design variables and it may represent the weight, the cost of the structure, or any other criterion by which some possible designs are preferred to others. We always assume that the objective function ($Z = F(x_i)$), is to be minimized, which entails no loss of generality since the minimum of $-F(x_i)$ occurs where the maximum of $F(x_i)$ takes place (see Figure 1.3). The selecting the objective function has got a significant impact on the entire optimization process. For example, if the cost of the structure is assumed to be proportional to its weight, then the objective function will represent the weight. The weight of the structure is often of critical importance, but the minimum weight is not always the cheapest. In general, the objective function represents the most important single property of a design, but it may represent also a weighted sum of a number of properties. A general cost function may include the cost of materials, fabrication, transportation, operation, repair, and many other cost factors. In this case, large numbers of members are considered in the form of the objective function, where it is appropriate to analyze the impact of certain members of the product price. Special attention should be paid to the components, which can result a "nearly constant" objective function. They are not worth to take into account. It is not true, that the most complex objective function gives the best results. In general, the objective function is a nonlinear function of the design variables.

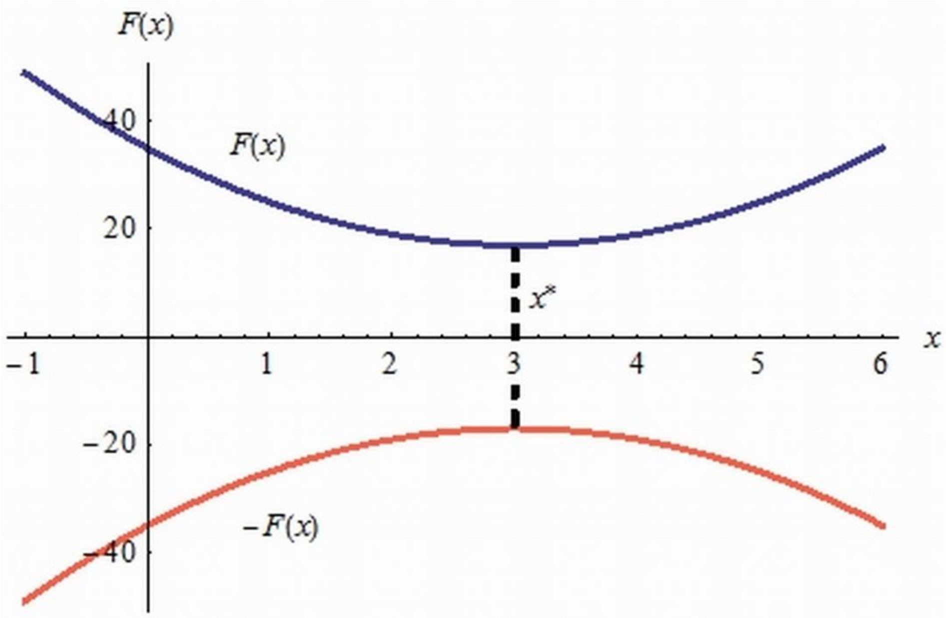


Figure 1.3.: The local extremum of the goal function

It is also possible to optimize simultaneously for multiple objective function, but this is only recommended if dominant objective function could not be selected, or the objective functions are in contradiction. This is the field of multi-objective optimization. The simplest solution technique is if we create a weighted sum of the objective functions and solve the problem as a standard optimization problem with only one objective function.
Pareto has developed the theory of the multi-objective optimization in 1896. It is important to note that the solution of an optimization problem with one goal function is generally a design point in the space of the design variables, while the solution of the Pareto-optimization problem is a set of design points, called Pareto front. A Pareto optimal solution is found if there is no other feasible solution that would reduce some objective function without causing a simultaneous increase in at least one other objective function.

We illustrated this complicated field with an example about „Optimization of Geometry for Car Bag" (this example was solved by company Sigma technology using the IOSO optimization software: www.iosotech.com), which can be found in the attached database for examples.

1.3.4. Formulating the optimization problem

The general formulation of the constrained optimization problem in the n dimensional Euclidian space is the following:

$$\begin{gathered} Z = F(x_i) \rightarrow \min \\ g_j(x_i) \leq 0 \; i = 1,\ldots,n \quad j = 1,\ldots,m \\ h_j(x_i) = 0 \; j = m+1,\ldots,p\,. \end{gathered} \tag{1.3}$$

This is an optimization problem with one goal function ($F(x_i)$), with inequality ($g(x_i)$) and equality ($h(x_i)$) constraints is formulated for n design variable. The numbers of the inequality constraints are m and the numbers of the equality constraints are (*p-m*). In the engineering problem definition, we often highlight the side constraints, defining the searching domain in the *n* dimensional space. This could be formulated as follows:

$$\breve{x}_i \le x_i \le \hat{x}_i \quad i = 1,\dots,n \ , \tag{1.4}$$

where $\breve{x}_i$ and $\hat{x}_i$ are the lower and upper limits of the searching domain.

1.4. Optimization examples

Large number of the optimization problem types made impossible to introduce all relevant fields, as optimization problems can be formulated on every side of the life, they are not only related to engineering problems.

The field of economics also often uses the optimization methods for performing economic analysis and supporting decision. Maximizing the profit or optimizing the elements of a given stock portfolio is an important task.
In the production technology the maximum capacity utilization of the different machines plays a very important role, considering different raw materials, which is not available in unlimited quantity.

The known travelling salesman is also an optimization problem: finding the shortest way between the cities. The solution for 50 cities can be seen on Figure 1.4.

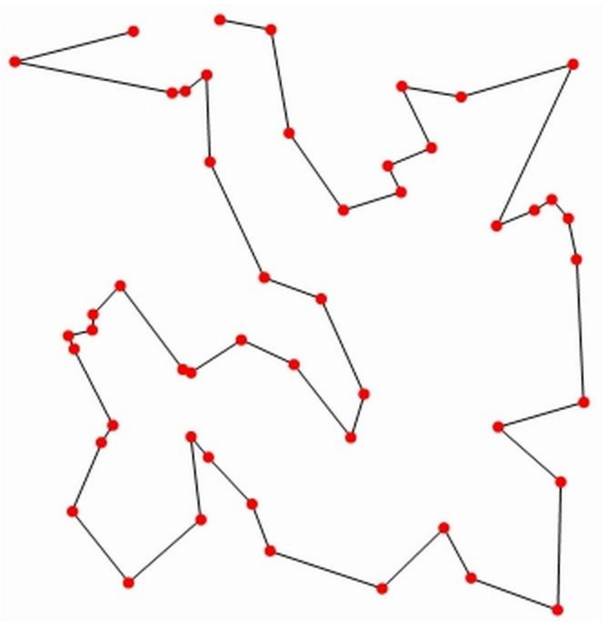


Figure 1.4.: The solution of the travelling salesman problem for 50 cities (Wolfram Mathematica)

In the area of the image and audio processing, for example, an optimization problem is to reduce the noise. A large field of the mathematics is the game theory, uses the optimization techniques very intensively too.

In the mechanical engineering practice the first industrial applications were in the field of military and space industry. In the Figure 1.5 the outline of a modern military airplane (F22) can be seen, by the design, beside the minimum weight, also the minimization of the radar cross section was an important task. In the space industry the weight minimization problem has got a very important role. An example in the field of space industry is a weight minimization of support in a rocket, which includes both topology and shape optimization problem (see Figure 1.6).

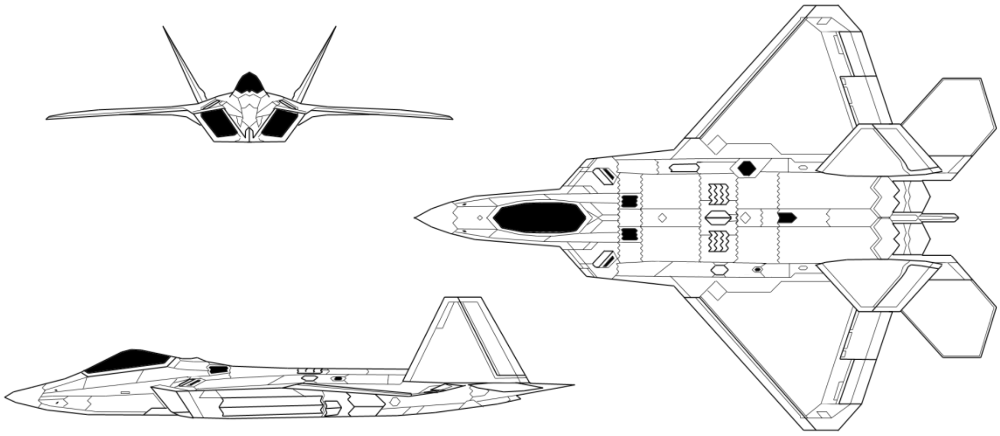


Figure 1.5.: Military airplane

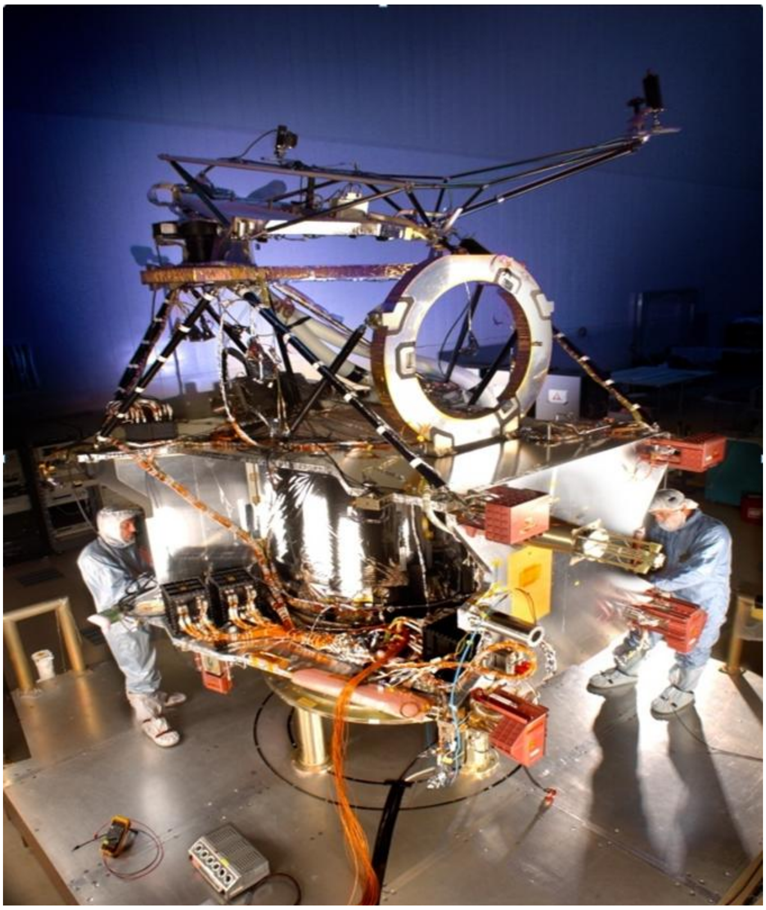


Figure 1.6.: Support in a rocket

Designing trucks, buses and cars different optimization techniques were used also very intensively. Nowadays most parts of these vehicles are undergone some level of optimization process even if the weight of the part is neglectable comparing the weight of the whole car. Because of the heavy price competition by the suppliers in the car industry, the 5-10% weight reduction which can be reached by using shape optimization techniques, gives significant advantages to the company.
Nowdays several optimization processes can be combined with different structural analysis modules. Not only with static load cases can be considered, but for example dynamic, temperature, flow, nonlinear systems can be optimized too. Additionally not only parts, but whole subassemblies using contact conditions between the parts can be solved.
But we have to know, that because the optimizer call lot of times the structural analysis module, the complicated problems often very time consuming if we have only one PC. Generally this type of problem formulation leads to the field of supercomputing; so supercomputers or computer clusters should be used to solve such type of industrial problems.

1.5. References

[1.1] Jármai Károly, Iványi Miklós: Gazdaságos fémszerkezetek analízise és tervezése. Műegyetemi kiadó, 2001

1.6. Questions

1. Define the general form of an optimization problem!
2. Define the design variable, design parameter and goal function!
3. What kind of classes are avaiable as design variables from phisical and mathematical point of view?
4. What is the main difference between the side constraints and behaviour constraints?
5. Explain some relevant optimization examples from the different field of the industry?

2. SINGLE-VARIABLE OPTIMIZATION

Only single-variable functions are considered in this chapter. These methods will be used in later chapters for multi-variable function optimization.

2.1. Optimality Criteria

Sufficient conditions of optimality:
Suppose at point x^*, the first derivative is zero and the first nonzero higher order derivative is denoted by n

(i) If n is odd, x^* is a point of inflection
(ii) If n is even, x^* is a local optimum
(a) If that derivative is positive, x^* is a local minimum
(b) If that derivative is negative, x^* is a local maximum.

2.2. Bracketing Algorithms

Finds a lower and an upper bound of the minimum point

2.2.1. Exhaustive Search

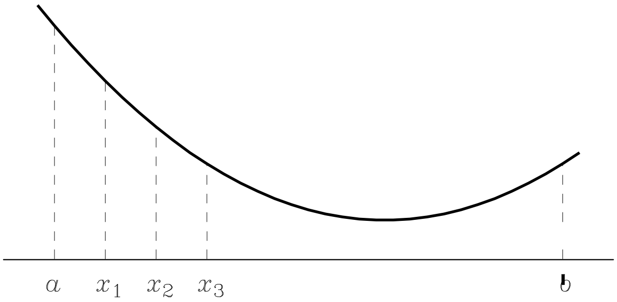


Figure 2.1 The exhaustive search method uses equally spaced points.

Function values are evaluated at n equally spaced points (Figure 2.1).

Algorithm:

Step 1: Set $x^{(0)} = a$, $x^{(n+1)} = b$, $x^{(j)} = a + \frac{b-a}{n+1} j$, $1 \le j \le n$. Set $k = 1$.

Step 2: If $f\left(x^{(k-1)}\right) \ge f\left(x^{(k)}\right) \le f\left(x^{(k+1)}\right)$, the minimum lies in $\left(x^{(k-1)}, x^{(k+1)}\right)$,
Terminate.
Else go to Step 3.

Step 3: Is $k = n$? If no, set $k = k+1$, go to Step 2.

If yes, no minimum exists in (a, b) Accuracy in the result: $\frac{2}{n+1}(b-a)$.

Average number of function evaluations to get to the optimum is $\left(\frac{n}{2} + 2\right)$.

2.2.2. Bounding Phase Method

This method is used to bracket the minimum of a function.
Algorithm:

Step 1: Choose an initial guess $x^{(0)}$ and an increment Δ. Set $k=0$.

Step 2: If $f\left(x^{(0)}-|\Delta|\right)\geq f\left(x^{(0)}\right)\geq f\left(x^{(0)}+|\Delta|\right)$ then Δ is positive.

Else if $f\left(x^{(0)}-|\Delta|\right)\leq f\left(x^{(0)}\right)\leq f\left(x^{(0)}+|\Delta|\right)$ then Δ is negative.

Else go to Step 1.

Step 3: Set $x^{(k+1)}=x^{(k)}+2^k\Delta$.

Step 4: If $f\left(x^{(k+1)}\right)< f\left(x^{(k)}\right)$, set $k=k+1$ and go to Step 3.
Else The minimum lies in the interval $\left(x^{(k-1)},x^{(k+1)}\right)$ and **Terminate.**

If Δ is large, poor bracketing.
If Δ is small, more evaluations.

2.3. Region-Elimination Methods

Fundamental rule for Region-elimination methods:
For $x_1<x_2$ where x_1, x_2 lie in (a, b).

1. If $f(x_1) > f(x_2)$ then minimum does not lie in (a, x_1).
2. If $f(x_1) < f(x_2)$ then minimum does not lie in (x_2, b).

2.3.1. Golden Section Search

Interval is reduced according to *golden* rule.
Properties: For only two trials, spread them equidistant from the center. Subinterval eliminated should be of the same length regardless of the outcome of the trial. Only one new point at each step is evaluated, other point remains from the previous step.

Algorithm:

Step 1: Choose a lower bound *a* and an upper bound *b*. Also choose a small number ε. Normalize the variable *x* by using the equation, $\omega=(x-a)/(b-a)$. Thus, $a_\omega=0$, $b_\omega=1$, and $L_\omega=1$. Set $k=1$.

Step 2: Set $\omega_1=a_\omega+(0.618)L_\omega$ and $w_2=b_\omega-(0.618)L_\omega$. Compute $f(\omega_1)$ or f $f(\omega_2)$ depending on whichever was not evaluated earlier. Use the fundamental region-elimination rule to eliminate a region. Set new a_ω and b_ω.

Step 3: Is $|L_\omega|<\varepsilon$ small? If no, set $k=k+1$, go to Step 2.
If yes, **Terminate.**

Interval reduces to $(0.618)^{n-1}$ after n evaluations. One new function evaluation at each iteration.
Consider the following function:

$$f(x) = x^2 + 54/x$$

Step 1: We choose $a=0$ and $b=5$. The transformation equation becomes $\omega = x/5$. Thus, $a_\omega = 0$, $b_\omega = 1$, and $L_\omega = 1$. Since the golden section method works with a transformed variable ω, it is convenient to work with the transformed function:

$$f(\omega) = 25\omega^2 + 54/(5\omega)$$

In the ω-space, the minimum lies at $\omega^* = 3/5 = 0.6$. We set an iteration counter $k=1$.

Step 2: We set $\omega_1 = 0 + (0.618)1 = 0.618$ and $\omega_2 = 1 - (0.618)1$ or $\omega_2 = 0.382$. The corresponding function values are $f(\omega_1) = 27.02$ and $f(\omega_2) = 31.92$. Since $f(\omega_1) < f(\omega_2)$, the minimum cannot lie in any point smaller than $\omega_2 = 0.382$. Thus, we eliminate the region (a, ω_2) or (0, 0.382). Thus, $a_\omega = 0.382$ and $b_\omega = 1$. At this stage, $L_\omega = 1 - 0.382 = 0.618$. The region being eliminated after this iteration is shown in Figure 2.2. The position of the exact minimum at $\omega = 0.6$ is also shown.

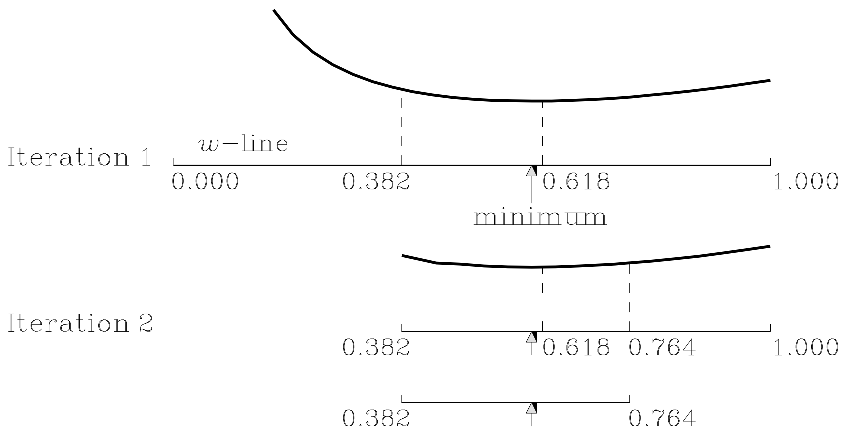


Figure 2.2: Region eliminations in the first two iterations of the golden section search algorithm.

Step 3: Since $|L_\omega|$ is not smaller than ε, we set $k=2$ and move to Step 2. This completes one iteration of the golden section search method.

Step 2: For the second iteration, we set

$$\omega_1 = 0.382 + (0.618)0.618 = 0.764$$
$$\omega_2 = 1 - (0.618)0.618 = 0.618.$$

We observe that the point ω_2 was computed in the previous iteration. Thus, we only need to compute the function value at ω_1: $f(\omega_1) = 28.73$. Using the fundamental region-elimination rule and observing the relation $f(\omega_1) > f(\omega_2)$, we eliminate the interval (0.764, 1). Thus, the new bounds are $a_\omega = 0.382$ and $b_\omega = 0.764$, and the new interval is $L_\omega = 0.764 - 0.382 = 0.382$, which is incidentally equal to $(0.618)^2$! Figure 2.2 shows the final region after two iterations of this algorithm.

Step 3: Since the obtained interval is not smaller than ε, we continue to proceed to Step 2 after incrementing the iteration counter k to 3.

Step 2: Here, we observe that $\omega_1 = 0.618$ and $\omega_2 = 0.528$, of which the point ω_1 was evaluated before. Thus, we compute $f(\omega_2)$ only: $f(\omega_2) = 27.43$. We also observe that $f(\omega_1) < f(\omega_2)$ and we eliminate the interval (0.382, 0.528). The new interval is (0.528, 0.764) and the new range is $L_\omega = 0.764 - 0.528 = 0.236$, which is exactly equal to $(0.618)^3$!

Step 3: Thus, at the end of the third iteration, $L_\omega = 0.236$. This way, Steps 2 and 3 may be continued until the desired accuracy is achieved.

We observe that at each iteration, only one new function evaluation is necessary. After three iterations, we have performed only four function evaluations. Thus, the interval reduces to $(0.618)^3$ or 0.236.

2.4. Methods Requiring Derivatives

- Use gradient information.
- At local minimum, the derivative of the function is equal to zero.

Gradients are computed numerically as follows:

$$f'(x^{(t)}) = \frac{f(x^{(t)}+\Delta x^{(t)}) - f(x^{(t)}-\Delta x^{(t)})}{2\Delta x^{(\neq)}} \tag{2.1}$$

$$f''(x^{(t)}) = \frac{f(x^{(t)}+\Delta x^{(t)}) - 2f(x^{(t)}) + f(x^{(t)}-\Delta x^{(t)})}{(\Delta x^{(t)})^2} \tag{2.2}$$

The parameter $\Delta x^{(t)}$ is usually taken to be a small value. In all our calculations, we assign $\Delta x^{(t)}$ to be about 1 per cent of $x^{(t)}$:

$$\Delta\, x^{(t)} = \begin{cases} 0.01|x^{(t)}| & if\ \ |x^{(t)}| > 0.01 \\ 0.0001 & otherwise. \end{cases} \tag{2.3}$$

According to Equations (2.1) and (2.2), the first derivative requires two function evaluations and the second derivative requires three function evaluations.

2.4.1. Newton-Raphson Method

A linear approximation to the derivative is made to zero to derive the transition rule.

Algorithm:

Step 1: Choose initial guess x_1 and a small number ε. Set $k = 1$. Compute $f'(x_1)$.

Step 2: Compute $f''(x_k)$.

Step 3: Calculate $x_{k+1} = x_k - \dfrac{f'(x_k)}{f''(x_k)}$. Compute $f'(x_{k+1})$.

Step 4: If $|f'(x_{k+1})| < \varepsilon$, **Terminate.**

Else set $k = k+1$ and Go to step 2.

Convergence of the algorithm depends on the initial point and the nature of the objective function.

2.4.2. Bisection Method

Both function value and sign of derivative are used to derive the transition rule.
Algorithm:

Step 1: Choose two points a and b such that $f'(a) < 0$ and $f'(b) > 0$, and a small number ε. Set $L = a$ and $R = b$.

Step 2: Calculate $z = (R+L)/2$ and evaluate $f'(z)$.

Step 3: If $|f'(z)| \le \varepsilon$, **Terminate.**

Else if $f'(z) < 0$ set $L = z$ and Go to step 2; else if $f'(z) > 0$ set $R = z$ and Go to step 2.

Consider again the function:

$$f(x) = x^2 + 54/x.$$

Step 1: We choose two points $a = 2$ and $b = 5$ such that $f'(a) = -9.501$ and $f'(b) = 7.841$ are of opposite sign. The derivatives are computed numerically using Equation (2.1). We also choose a small number $\varepsilon = 10^{-3}$.

Step 2: We calculate a quantity $z = (x_1 + x_2)/2 = 3.5$ and compute $f'(z) = 2.591$.

Step 3: Since $f'(z) >$ 0, the right-half of the search space needs to be eliminated. Thus, we set $x_1 = 2$ and $x_2 = z = 3{,}5$. This completes one iteration of the algorithm. At each iteration, only half of the search region is eliminated, but here the decision about which half to delete depends on the derivatives at the mid-point of the interval.

Step 2: We compute $z = (2+3.5)/2 = 2.750$ and $f'(z) = -1.641$.

Step 3: Since $f'(z) < 0$, we set $x_1 = 2.750$ and $x_2 = 3.500$.

Step 2: The new point z is the average of the two bounds: $z = 3.125$. The function value at this point is $f'(z) = 0.720$.

Step 3: Since $|f'(z)| \nless \varepsilon$, we continue with Step 2.

Thus, at the end of 10 function evaluations, we have obtained an interval (2.750, 3.125), bracketing the minimum point $x^* = 3.0$. The guess of the minimum point is the mid-point of the obtained interval or $x = 2.938$. This process continues until we find a point with a vanishing derivative. Since at each iteration, the gradient is evaluated only at one new point, the bisection method requires two function evaluations per iteration. In this method, exactly half the region is eliminated at every iteration; but using the magnitude of the gradient, a faster algorithm can be designed to adaptively eliminate variable portions of search region – a matter which we discuss in the following subsection.

2.4.3. Secant Method

Both function value and derivative are used to derive the transition rule.
Algorithm:

Same as Bisection method except step 2 is modified as follows:

Step 2: Calculate $z = R - \dfrac{f'(R)}{\left(f'(R) - f'(L)\right)/(R - L)}$ and evaluate $f'(z)$.

2.5. References

[2.1] Powell, M. J.D. (1964): **An efficient method for finding the minimum of a function of several variables without calculating derivatives.** Computer Journal. 7, 155-162.

[2.2] Reklaitis, G. V., Ravindran, A., and Ragsdell, K. M. (1983): **Engineering Optimization-Methods and Applications**. New York: Wiley.

[2.3] Scarborough, J. B. (1966): **Numerical Mathematical Analysis.** New Delhi: Oxford & IBH Publishing Co.

2.6. Questions

1. Explain the sufficient conditions of the optimality!
2. Summarize the main characteristics of the bracketing methods!
3. Explain the fundamental rules of region-elimination methods!
4. What are the advantages of methods using derivatives?

3. MULTI-VARIABLE OPTIMIZATION

Functions of multiple variables (*N* variables) are considered for minimization here. Duality principle can be used to apply these methods to maximization problems.

3.1. Optimality Criteria

A stationary point $\bar{x}$ is minimum, maximum, or saddle-point if $\nabla^2 f(\bar{x})$ is positive definite, negative definite, or $\nabla^2 f(\bar{x}) \lesseqgtr 0$.

Necessary Conditions: For x^* to be a local minimum, $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive semidefinite.

Sufficient Conditions: $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite than, x^* is an isolated local minimum of $f(x)$.

3.2. Direct Search Methods

Use only function values; no gradient information is used.

3.2.1. Simplex Search Method

– A simplex is a set of $(N+1)$ points. At each iteration, a new simplex is created from the current simplex by fixed transition rules.
– The worst point is projected a suitable distance through the centroid of the remaining points.

Algorithm:

Step 1: Choose $\gamma > 1$, $\beta \in (0,\ 1)$, and a termination parameter ε. Create an initial simplex[1].

Step 2: Find x_h (the worst point), x_l (the best point), and x_g (next to the worst point). Calculate $x_c = \frac{1}{N} \sum_{i=1, i \neq h}^{N+1} x_i$.

$$x_j^{(i)} = \begin{cases} x_j^{(0)} + C & if\ j = i; \\ x_j^{(0)} + C\ \Delta, & otherwise, \end{cases} \quad \text{where } \Delta = \begin{cases} 0.25, & if\ N = 3; \\ \left(\frac{\sqrt{N+J}-2}{N-3}\right) & otherwise. \end{cases}$$

Step 3: Calculate the reflected point $x_r = 2x_c - x_h$. Set $x_{new} = x_r$.

If $f(x_r) < f(x_l)$, set $x_{new} = (1+\gamma)x_c - \gamma x_h$ (expansion).

Else if $f(x_r) \geq f(x_h)$, set $x_{new} = (1-\beta)x_c - \beta x_h$ (contraction).

---

[1] One of the ways to create a simplex is to choose a base point x^0 and a scale factor C. Then $(N + 1)$ points are $x^{(0)}$ and for $i, j = 1, 2, \ldots, N$.

Else if $f(x_g) < f(x_r) < f(x_h)$, set $x_{new} = (1+\beta)x_c - \beta x_h$ (contraction).
Calculate $f(x_{new})$ and replace x_h by x_{new}.

Step 4: If $\left\{ \sum_{i=1}^{N+1} \frac{\left(f(x_i) - f(x_c)\right)^2}{N+1} \right\}^{1/2} \leq \varepsilon$, **Terminate**.

Else go to Step 2.

3.2.2. Powell's Conjugate Direction Method

- Most successful direct search method
- Uses history of iterations to create new search directions
- Based on a quadratic model
- Generate N conjugate directions and perform one-dimensional search in each direction one at a time

Parallel Subspace Property: Given a quadratic function $q(x)$, two arbitrary but distinct points $x^{(1)}$ and $x^{(2)}$, and a direction d. If $y^{(1)}$ is the solution to min $q(x^{(1)} + \lambda d)$ and $y^{(2)}$ is the solution to min $q(x^{(2)} + \lambda d)$, then the direction $(y^{(2)} - y^{(1)})$ is C-conjugate to d or $(y^{(2)} - y^{(1)})^T Cd =$ Diagonal matrix.

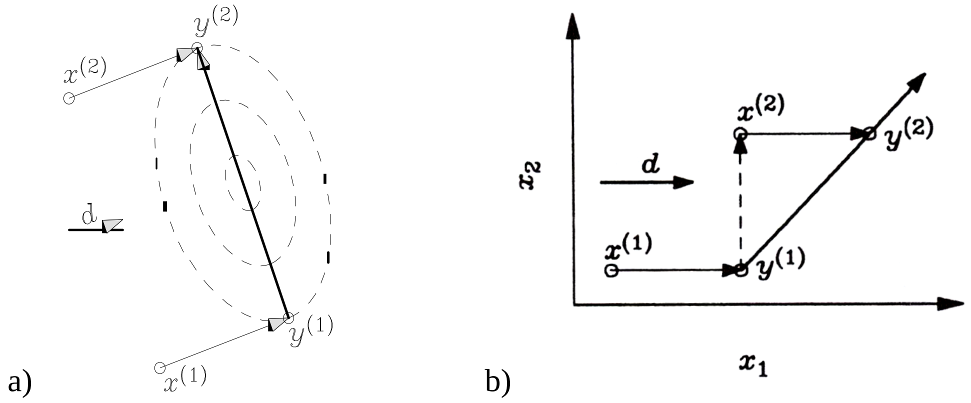


Figure 3.1 Illustration of the paralell subspace property with two arbitrary points and an arbitrary search direction in (a). The same can also be achieved from one point and two coordinate points in (b).

Instead of using two points and a direction vector to create one conjugate direction, one point and coordinate directions can be used to create conjugate directions (Figure 3.1).

Extended Parallel Subspace Property: In higher dimensions, if from $x^{(1)}$ the point $y^{(1)}$ is found after searches along each of $m(< n)$ conjugate directions, and similarly if from $x^{(2)}$ the point $y^{(2)}$ is found after searches along each of m conjugate directions, $s^{(1)}$, $s^{(2)}$,…, $s^{(m)}$, then the vector $(y^{(2)} - y^{(1)})$ will be the conjugate to all of the m previous directions.

Algorithm:

Step 1: Choose a starting point $x^{(0)}$ and a set of N linearly independent directions; possibly $s^{(i)} = e^{(i)}$ for i=1, 2,…, N.

Step 2: Minimize along N unidirectional search directions using the previous minimum point to begin the next search. Begin with the search $s^{(1)}$ direction and end with $s^{(N)}$. Thereafter, perform another unidirectional search along $s^{(1)}$.

Step 3: Form a new conjugate direction d using the extended parallel subspace property.

Step 4: If $\|d\|$ is small or search directions are linearly independent, **Terminate.**
Else replace $s^{(j)} = s^{(j-1)}$ for all $j = N, N-1, ..., 2$. Set $s^{(1)} = d / \|d\|$ and go to Step 2.

A test is required to ensure linear independence of conjugate directions. If the function is quadratic, exactly N loops through steps 2 to 4 are required. If the function is quadratic, exactly $(N-1)$ loops through Steps 2 to 4 is required. Since in every iteration of the above algorithm exactly $(N+1)$ unidirectional searches are necessary, a total of $(N-1)\times(N+1)$ or (N^2-1) unidirectional searches are necessary to find N conjugate directions. Thereafter, one final unidirectional search is necessary to obtain the minimum point. Thus, in order to find the minimum of a quadratic objective function, the conjugate direction method requires a total of N^2 unidirectional searches.
Disadvantages:

- It takes usually more than N cycles for nonquadratic functions
- One-dimensional searches may not be exact, so directions may not be conjugate
- May halt before the optima is reached

Consider the Himmelblau function:

$$\text{Minimize } f(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

in the interval $0 \le x_1, x_2 \le 5$.

Step 1: We begin with a point $x^{(0)} = (0, 4)^T$. We assume initial search directions as $s^{(1)} = (1, 0)^{\mathrm{T}}$ and $s^{(2)} = (0, 1)^{\mathrm{T}}$.

Step 2: We first find the minimum point along the search direction $s^{(1)}$. Any point along that direction can be written as $x^p = x^{(0)} + \alpha s^{(1)}$, where α is a scalar quantity expressing the distance of the point x^p from $x^{(0)}$. Thus, the point x^p can be written as $x^p = (\alpha, 4)^T$. Now the two-variable function $f(x_1, x_2)$ can be expressed in terms of one variable α as

$$F(\alpha) = (\alpha^2 - 7)^2 + (\alpha + 9)^{2}{}_{1} + x_2^2 - 7)^2,$$

which represents the function value of any point along the direction $s^{(1)}$ and passing through $x^{(0)}$. Since we are looking for the point for which the function value is minimum, we may differentiate the above expression with respect to α and equate to zero. But in any arbitrary problem, it may not be possible to write an explicit expression of the single-variable function $F(\alpha)$ and differentiate. In those cases, the function $F(\alpha)$ can be obtained by substituting each variable x_i by x_i^p. Thereafter, any single-variable optimization methods, as described in Chapter 2, can be used to find the minimum point. The first task is to bracket the minimum and then the subsequent task is to find the minimum point. Here, we could have found the exact minimum solution by differentiating the single-variable function $F(\alpha)$ with respect to α and then equating the term to zero, but we follow the more generic procedure of numerical differentiation, a method which will be used in many real-world optimization problems. Using the bounding phase method in the above problem we find that the minimum is bracketed in the interval (1,4) and using the golden section search we obtain the minimum $\alpha^* = 2.083$ with three decimal places of accuracy. Thus, $x^{(1)} = (2.083, 4.000)^T$.

Similarly, we find the minimum point along the second search direction $s^{(2)}$ from the point $x^{(1)}$. A general point on that line is

$$x(\alpha) = x^{(1)} + \alpha s^{(2)} = (2.083.\ (4+\alpha))^{\mathrm{T}}.$$

The optimum point found using a combined application of the bounding phase and the golden section search method is $\alpha^* = -1{,}592$ and the corresponding point is $x^{(2)} = (2.083, 2.408)^T$.

From the point $x^{(2)}$, we perform a final unidirectional search along the first search direction and obtain the minimum point $x^{(3)} = (2.881, 2.408)^T$.

Step 3: According to the parellel subspace property, we find the new conjugate direction.

$$d = x^{(3)} - x^{(1)} = (2.881,\ 2.408\ )^{\mathrm{T}} - (2.083{,}4.000\ )^{\mathrm{T}} = (0.798{-}1.592)^T$$

Step 4: The magnitude of search vector d is not small. Thus, the new conjugate search directions are

$$\begin{aligned} s^{(2)} &= (1,\ 0)^T, \\ s^{(1)} &= (0.798, -1.592)^T / \|(0.798, -1.592)^T\| \\ &= (0.448, -0.894)^T. \end{aligned}$$

This completes one iteration of Powell's conjugate direction method. Figure 3.2 shows the new conjugate direction on a contour plot of the objective function. With these new search directions we now proceed to Step 2.

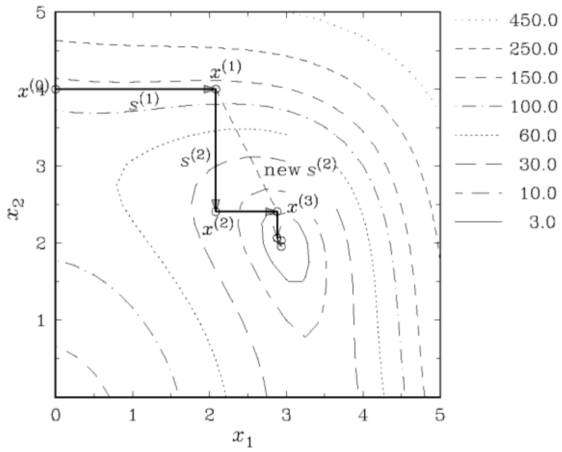


Figure 3.2 Two iterations of Powell's conjugate direction method.

Step 2: A single-variable minimization along the search direction $s^{(1)}$ from the point $x^{(3)} = (2.881, 2.408)^T$ results in the new point $x^{(4)} = (3.063, 2.045)^T$. If the objective function had been a quadratic function, we would have achieved the optimum point at this step. However, it is interesting to note that the solution $x^{(4)}$ is close to the true minimum of the function. One more unidirectional search along $s^{(2)}$ from the point $x^{(4)}$ results in the point $x^{(5)} = (2.988, 2.045)^T$. Another minimization along $s^{(1)}$ results in $x^{(6)} = (3.008, 2.006)^T$ $f(x^{(6)} = 0.004)$.

Step 3: The new conjugate direction is

$$d = (x^{(6)} - x^{(4)}) = (0.055, -0.039)^T$$

The unit vector along this direction is $(0.816, -0.578)^{\mathrm{T}}$.

Step 4: The new pair of conjugate search directions are $s^{(1)} = (0.448, -0.894)^T$ and $s^{(2)} = (0.055, -0.039)^T$ respectively. The search direction d (before normalizing) may be considered to be small and therefore the algorithm may terminate.

We observe that in one iteration of Step 2, $(N+1)$ unidirectional searches are necessary. Thus, computationally this method may be expensive. In terms of the storage requirement, the algorithm has to store $(N+1)$ points and N search directions at any stage of the iteration.

3.3. Gradient-based Methods

- Direct search methods are expensive
- Gradient-based methods assume existence of $f(x)$, $\nabla f(x)$, and $\nabla^2 f(x)$
- All methods employ

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} s(x^{(k)})$$

where $\alpha^{(k)}$ is the step length and $s(x^{(k)})$ is the search direction.

– Usually, $\alpha^{(k)}$ is selected to minimize the function in $s(x^{(k)})$ direction. Update search direction iteratively.

3.3.1. Numerical Gradient Approximation

- In real-world engineering problems, gradients are often difficult to calculate.
- Even if they are available, their computation could be error prone
- Numerically gradients are computed:

$$\frac{\partial f(x)}{\partial x_i}|_{(x)^t} = (f(x_i^{(t)} + \Delta x_i^{(t)}) - f(x_i^{(t)} - \Delta x_i^{(t)}))/2 \Delta x_i^{(t)}$$

$$\frac{\partial^2 f(x)}{\partial^2 x_i^2}|x^{(t)} = (f(x_i^{(t)} + \Delta x_i^{(t)}) - 2f(x^{(t)}) + f(x_i^{(t)} - \Delta x_i^{(t)}))/(\Delta x_i^{(t)})^2.$$

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j}|x^{(t)} = [f(x_i^{(t)} + \Delta x_{i)}^{(t)} x_j^{(t)} + \Delta x_j^{(t)})$$

$$-f(x_i^{(t)} + \Delta x_{i)}^{(t)} x_j^{(t)} - \Delta x_j^{(t)}) - f(x_i^{(t)} - \Delta x_{i)}^{(t)} x_j^{(t)} + \Delta x_j^{(t)})$$

$$+f(x_i^{(t)} - \Delta x_{i)}^{(t)} x_j^{(t)} - \Delta x_j^{(t)})] /(4 \Delta x_i^{(t)} \Delta x_j^{(t)}).$$

The computation of the first derivative with respect to each variable requires two function evaluations, thus totaling $2N$ function evaluations for the complete first derivative vector. The computation of the second derivative $\partial^2 f / \partial x_i^2$ requires three function evaluations, but the second-order partial derivative $\partial^2 f /(\partial x_i, \partial x_j)$ requires four function evaluations. Thus, the computation of Hessian matrix requires $(2N^2 + 1)$ function evaluations (assuming the symmetry of the matrix).

3.3.2. Cauchy's Method (Steepest Descent)

Greatest decrease in $s(x^{(k)}) = -\nabla f(x)$ direction.

Algorithm:

Step 1: Choose a maximum number of iterations M to be performed, an initial point $x^{(0)}$, two termination parameters $\varepsilon_1, \varepsilon_2$, and set $k = 0$.

Step 2: Calculate $\nabla f(x^{(k)})$, the first derivative at the point $x^{(k)}$.

Step 3: If $\left\|\nabla f(x^{(k)})\right\| \le \varepsilon_1$, **Terminate.**

Else if $k \ge M$, **Terminate.**

Else go to Step 4.

Step 4: Perform a unidirectional search to find $\alpha^{(k)}$ using ε_2 such that /($f(x^{(k+1)}) = f(x^{(k)} - \alpha^{(k)}\nabla f(x^{(k)}))$ is minimum. One criterion for termination is when $\left\| \nabla f(x^{(k+1)}) \cdot \nabla f(x^{(k)}) \right\| \le \varepsilon_2$.

Step 5: Is $\left\| \dfrac{x^{(k+1)} - x^{(k)}}{x^{(k)}} \right\| \le \varepsilon_1$? If yes, **Terminate.**

Else set $k = k+1$ and go to Step 2.

Cauchy's method works well when $x^{(0)}$ is far away from x*.

3.3.3. Newton's Method

Use a search direction $s(x^{(k)}) = -\nabla^2 f(x^{(k)})^{-1}\nabla f(x^{(k)})$.

Algorithm: Same as Cauchy's method except step 4 is modified:

Step 4: Perform a line search to find $\alpha^{(k)}$ using ε_2 such that $f(x^{(k+1)}) = f(x^{(k)} - \alpha^{(k)}\nabla^2 f(x^{(k)})^{-1}\nabla f(x^{(k)}))$ is minimum.

Newton's method works well when $x^{(0)}$ is close to x*.

3.3.4. Marquardt's Method

– A compromise between Cauchy's and Newton's method.
– When the point is far away from the optimum, Cauchy's method is used and when the point is close to the optimum Newton's method is used.

Algorithm:

Step 1: Choose a starting point, $x^{(0)}$, the maximum number of iterations, M, and a termination parameter, ε. Set $k = 0$ and $\lambda^{(0)} = 10^4$ (a large number).

Step 2: Calculate $\nabla f(x^{(k)})$.

Step 3: If $\left\| \nabla f(x^{(k)}) \right\| \le \varepsilon$ or $k \ge M$, **Terminate.**

Else go to Step 4

Step 4: Calculate $s(x^{(k)}) = -\left[H^{(k)} + \lambda^{(k)} I\right]^{-1} \nabla f(x^{(k)})$.

Set $x^{(k+1)} = x^{(k)} + s(x^{(k)})$.

Step 5: Is $f(x^{(k+1)}) < f(x^{(k)})$? If yes, go to Step 6.

Else go to Step 7.

Step 6: Set $\lambda^{(k+1)} = \frac{1}{2}\lambda^{(k)}$, $k = k+1$, and go to Step 2.

Step 7: Set $\lambda^{(k)} = 2\lambda^{(k)}$ and go to Step 4.

Need to compute Hessian matrix, which may be cumbersome.

3.3.5. Variable-Metric Method (Davidon-Fletcher-Powell method)

– Uses positive characteristics of Newton's method using only first-order information
– The search direction is

$$S(x^{(k)}) = -A^{(k)}\nabla f(x^{(k)}) \tag{3.1}$$

where $A^{(k)}$ is the Hessian matrix.

– Starting with an identity matrix $(A^{(0)} = I)$, iteratively evaluate the Hessian matrix with $g(x^{(k)}) = \nabla f(x^{(k)})$:

$$A^{(k)} = A^{(k-1)} + \frac{\Delta x^{(k-1)} \Delta x^{(k-1)^T}}{\Delta x^{(k-1)^T} \Delta e(x^{(k-1)})}$$

$$- \frac{A^{(k-1)} \Delta e(x^{(k-1)})(A^{(k-1)} \Delta e(x^{(k-1)}))^T}{\Delta e(x^{(k-1)})^{\tau} A(k-1) \Delta e(x^{(k-1)})} \tag{3.2}$$

– The above relation preserves the positive definiteness of the matrix A.

Algorithm:

Same as Fletcher-Reeves algorithm except the expression for search direction $s(x^k)$ in step 4, which is set according to equation 3.1.

Let us consider the Himmelblau function again:

$$\text{Minimize } f(x_1x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2.$$

The inverse of the Hessian matrix at the minimum point $x^* = (3, 2)^T$ is found to be

$$H^{-1}(x^*) = \begin{pmatrix} 0.016 & -0.009 \\ -0.009 & 0.035 \end{pmatrix}. \tag{3.3}$$

Successive iterations of the DFP method transform an initial identity matrix into the above matrix. This allows the DFP search to become similar to Newton's search after a few iterations. The advantage of this method over Newton's method is that the inverse of the Hessian matrix need not be calculated. Thus, the algorithm has the effect of a second-order search, but the search is achieved only with first-order derivatives.

Step 1: Once again we begin with the initial point $x^{(0)} = (0, 0)^T$. The termination parameters are all set to be 10^{-3}.

Step 2: The derivative vector at the initial point is equal to $\nabla f(x^{(0)}) = (-14, -22)^T$. The search direction is $s^{(0)} = (14, 22)^T$.

Step 3: A unidirectional search from $x^{(0)}$ along $s^{(0)}$ gives the minimum point. We calculate the gradient at this point:

$$\nabla f(x^{(1)}) = (-30.707, 18.803)^T$$

In order to calculate the new search direction, we first compute the parameters required to be used in Equation (3.2):

$$\Delta x^{(0)} = x^{(1)} - x^{(0)} = (1.788, 2.810)^T,$$
$$\Delta e(x^{(0)}) = e(x^{(1)}) - e(x^{(0)}) = (-16.707, 40.803)^T$$
$$A^{(0)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

With these parameters, we compute the next estimate of the matrix A as follows:

$$
\begin{aligned}
A^{(1)} &= A^{(0)} + \frac{\Delta x^{(0)} \Delta x^{(0)T}}{\Delta x^{(0)T} \Delta e(x^{(0)})} - \frac{A^{(0)} \Delta e(x^{(0)}) (A^{(0)} \Delta e(x^{(0)}))^T}{\Delta e(x^{(0)})^T A^{(0)} \Delta e(x^{(0)})}, \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\begin{pmatrix} 1.778 \\ 2.810 \end{pmatrix} (1.788,\ 2.810)}{(1.788, 2.810) \begin{pmatrix} -16.707 \\ 40.803 \end{pmatrix}} \\
&\quad - \frac{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -16.707 \\ 40.803 \end{pmatrix} \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -16.707 \\ 40.803 \end{pmatrix} \right)^T}{(-16.707, 40.803) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -16.707 \\ 40.803 \end{pmatrix}}, \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0.038 & 0.060 \\ 0.060 & 0.093 \end{pmatrix} - \begin{pmatrix} 0.144 & -0.351 \\ -0.351 & 0.856 \end{pmatrix},
\end{aligned}
$$

Knowing $A^{(1)}$, we now obtain the search direction $s^{(1)}$ using Equation (3.1) as follows:

$$
s^{(1)} = -\begin{pmatrix} 0.894 & 0.411 \\ 0.411 & 0.237 \end{pmatrix} \begin{pmatrix} -30.707 \\ 18.803 \end{pmatrix} = \begin{pmatrix} 19.724 \\ 8.164 \end{pmatrix}.
$$

The unit vector along this direction is $(0.924,\ 0.382)^T$. It is interesting to observe that this search direction is similar to the first search direction $s^{(1)}$ obtained by using the Fletcher-Reeves algorithm.

Step 4: Performing a unidirectional search along $s^{(1)}$ from the point $x^{(1)}$, we obtain the minimum point: $x^{(2)} = (2.248, 2.989)^T$ having a function value $f(x^{(2)}) = 26.237$ (Figure 3.2).

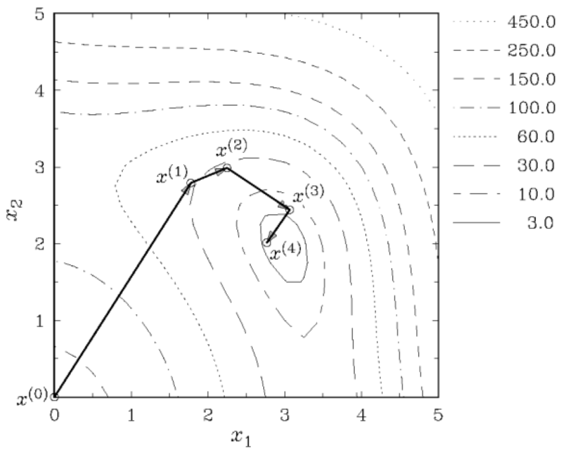


Figure 3.2: Three iterations of the DFP method.

Step 5: We observe that the point $x^{(2)}$ is very different from the point $x^{(1)}$ and calculate the gradient at $x^{(2)}$: $\nabla f(x^{(2)}) = (-18.225, 44.037)^T$. Since $\|\nabla f(x^{(2)})\|$ is not close to zero, we increment k and proceed to Step 4. This completes one iteration of the DFP method. The initial point $x^{(0)}$, the point $x^{(1)}$, and the point $x^{(2)}$ are shown in Figure 3.2.

Step 4: The second iteration begins with computing another search direction $s^{(2)} = -A^{(2)}\nabla f(x^{(2)})$. The matrix $A^{(2)}$ can be computed by calculating $\Delta x^{(1)}$ and $\Delta e(x^{(1)})$ as before and using Equation (3.2). By simplifying the expression, we obtain the new matrix

$$A^{(2)} = \begin{pmatrix} 0.070 & -0.017 \\ -0.017 & 0.015 \end{pmatrix}$$

and the new search direction is found to be

$$s^{(2)} = (0.731, -0.976)^T$$

Note that this direction is more descent than that obtained in the Fletcher-Reeves method after one iteration.

Step 5: The unidirectional search along $s^{(2)}$ finds the best point: $x^{(3)} = (2.995, 1.991)^T$ with a function value equal to $f(x^{(3)}) = 0.003$.
Another iteration updates the A matrix as follows:

$$A^{(3)} = \begin{pmatrix} 0.030 & -0.005 \\ -0.005 & 0.020 \end{pmatrix}.$$

The new search direction is found to be $s^{(3)} = (0.014, 0.095)^T$. The unidirectional search method finds the new point $x^{(4)} = (2.997, 2.000)^T$ with a function value equal to 3×10^4. Another iteration finds the following A matrix:

$$A^{(4)} = \begin{pmatrix} 0.018 & -0.011 \\ -0.011 & 0.036 \end{pmatrix}$$

and the new search direction $s^{(4)} = (0.003, -0.003)^T$. A unidirectional search finds the point $x^{(5)} = (3.000, 2.000)^T$ with a function value equal to zero. One final iteration finds the A matrix:

$$A^{(5)} = \begin{pmatrix} 0.016 & -0.009 \\ -0.009 & 0.035 \end{pmatrix}.$$

which is identical to the inverse of the Hessian matrix of the Himmelblau function at the minimum point (Equation (3.3)). The above calculation shows how the original identity matrix $A^{(0)}$ has transformed into $H^{-1}(x^*)$ matrix in successive iterations. From the fourth iteration onwards, the search direction is very similar to that in the Newton's method, because the matrix $A^{(k)}$ is similar to the inverse of the Hessian matrix. Since, during that iteration the solution is also close to the minimum, the search converges faster to the minimum point. One difficulty with this method is that the matrix A sometimes becomes ill-conditioned due to numerical derivative computations and due to inaccuracy involved in unidirectional searches.

3.4. References

[3.1] Box, G. E. P. and Draper, N. R. (1969): **Evolutionary Operation.** New York: Wiley.

[3.2] Davidon, W. C. (1959): **Variable metric method for minimization** (Technical Report No. ANL-599). AEC Research Development Report.

[3.3] Fletcher, R. and Powell, M. J. D. (1963): **A rapidly convergent descent method for minimization.** Computer Journal. 6. 163-168.

[3.4] Fletcher, R. and Reeves, C. M. (1964): **Function minimization by conjugate gradients.** Computer Journal. 7, 149-154.

[3.5] Kreyszig, E. (1983): **Advanced Engineering Mathematics.** New Delhi: Wiley Eastern.

[3.6] Neider, J. A. and Mead, R. (1965): **A simplex method for function minimization.** Computer Journal. 7, 308-313.

[3.7] Rao, S. S. (1984): **Optimization Theory and Applications.** New Delhi: Wiley Eastern.

[3.8] Reklaitis, G. V., Ravindran, A., and Ragsdell, K. M. (1983): **Engineering Optimization-Methods and Applications.** New York: Wiley.

3.5. Questions

1. Explain the sufficient and the necessary conditions of the optimality!
2. What properties fulfill conjugate directions in Powell's method? Why?
3. What compromise is reached by Marquardt's Method? How?
4. How is the Hessian matrix used in DFP method?

4. CONSTRAINED OPTIMIZATION

Single and multi-variable functions with equality or inequality constraints or both are considered.

4.1. Kuhn-Tucker Conditions

We assume that objective function $f(x)$ (x is an N-dimensional array), inequality constraints $g_j(x),\ j=1,2,...,J$, and equality constraints $h_k(x), k=1,2,...,K$ are all differentiable. The NLP problem is as follows:

$$\begin{array}{lll} \text{Minimize} & f(x) & \\ \text{Subject to} & g_j(x) \geq 0 & j = 1,2 ..., J \\ & h_k(x) = 0 & k = 1,2 ..., K \end{array}$$

The Kuhn-Tucker problem is to find vectors x (size N), u (size J), and v (size K) that satisfy

$$\nabla f(x) - \sum_{j=1}^{J} u_j \nabla g_j(x) - \sum_{k=1}^{K} \nabla h_k(x) = 0 \quad N \text{ equations}$$

$$\begin{array}{rl} g_j(x) \geq 0 & j = 1,2, ..., J \\ h_k(x) = 0 & k = 1,2, ..., K \\ u_j g_j(x) = 0 & j = 1,2, ..., J \\ u_j \geq 0 & j = 1,2, ..., J \end{array}$$

If $g_j(\bar{x}) = 0$, the constraint j is active or binding at $\bar{x}$.
If $g_j(\bar{x}) > 0$, the constraint j is inactive or nonbinding at $\bar{x}$.

Kuhn-Tucker Necessity Theorem:
Consider the NLP problem shown above. Let f, g, and h be differentiable functions and x^* be a feasible solution to NLP. Let $I = \{j \mid g_j(x^*) = 0\}$. Furthermore, $\nabla g_j(x^*)$ for $j \in I$ and $\nabla h_k(x^*)$ for k = 1, 2,…, K are linearly independent (Constraint qualification). If x^* is an optimal solution to NLP, there exists a (u^*, v^*) such that (x^*,u^*,v^*) solves Kuhn-Tucker problem.
If a feasible point does not satisfy constraint qualification, K-T necessity theorem can be used to prove that the point is not optimal but not vice versa.

Kuhn-Tucker Sufficiency Theorem:
Let the objective function be convex, the inequality constraints $g_j(x)$ be all concave functions for j = 1, 2,…, J and equality constraints $h_k(x)$ for k = 1, 2,…, K be linear. If there exists a solution (x^*,u^*,v^*) that satisfies the K-T conditions, then x^* is an optimal solution to the NLP problem.

4.2. Transformation Methods

- Original constrained problem is transformed into a sequence of unconstrained problems via penalty functions
- Methods vary according to the way constraints are handled.

Interior penalty method: Each sequence contains feasible points
Exterior penalty method: Each sequence contains infeasible points
Mixed penalty method: Each sequence contains feasible or infeasible points

4.2.1. Penalty Function Method

- Penalty concept:

$$P(x,R) = f(x) + \Omega(R, g(x), h(x))$$

where *R* is a set of penalty parameters, Ω is the penalty term so selected to favor the selection of feasible points over infeasible points:

1. Parabolic Penalty: $\Omega = R\{h(x)\}^2$. Used for equality constraints and it is an exterior penalty term. A small value of *R* is started with and increased gradually.
2. Infinite Barrier Penalty: $\Omega = 10^{20}\sum_{j\in\bar{j}}|g_j(x)|$. It assigns an infinite penalty to infeasible points.
3. Log Penalty: $\Omega = -R\ln[g(x)]$. A large value of *R* is started with and reduced to zero gradually. It is an interior penalty term.
4. Inverse Penalty: $\Omega = R\left[\frac{1}{g(x)}\right]$. Like Log penalty, the value of *R* starts from large value and reduces to zero. It is also an interior penalty term.
5. Bracket Operator Penalty: $\Omega = R\langle g(x)\rangle^2$, where $\langle\alpha\rangle = \alpha$, when a is negative; zero, otherwise. Here R starts from a small value and increases to a large value. This is an exterior penalty term.

Algorithm:

Step 1: Choose two termination parameters $\varepsilon_1, \varepsilon_2$, an initial solution $x^{(0)}$, a penalty term Ω, and an initial penalty parameter $R^{(0)}$. Choose a parameter c to update *R* such that $0 < c < 1$ is used for interior penalty terms and $c > 1$ is used for exterior penalty terms. Set $t = 0$.

Step 2: Form $P(x^{(t)}, r^{(t)}) = f(x^{(t)}) + \Omega(R^{(t)}, g(x^{(t)}), h(x^{(t)}))$

Step 3: Starting with a solution $x^{(t)}$, find $x^{(i+1)}$ such that $P(x^{(i+1)}, R^{(t)})$ is minimum for a fixed value of $R^{(t)}$. Use ε_1 to terminate the unconstrained search.

Step 4: Is $\left|P(x^{(t)}, R^{(t)}) - P(x^{(t)}, R^{(t-1)})\right| \le \varepsilon_2$?

If yes, set $x^T = x^{(t+1)}$ and **Terminate.**
Else go to Step 5.

Step 5: Choose $R^{(t+1)} = cR^{(t)}$. Set $t = t+1$ and go to Step 2.

Consider the constrained Himmelblau's function:

$$\text{Minimize} \qquad (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

subject to

$$(x_1 - 5)^2 + x_2^2 - 26 \geq 0. \qquad x_1, x_2 \geq 0.$$

The inclusion of the constraint changes the unconstrained optimum point. The feasible region and the optimum point of this NLP is shown in Figure 4.1. The figure shows that the original optimum point $(3,1)^T$ is now an infeasible point. The new optimum is a point on the constraint line that touches a contour line at that point.

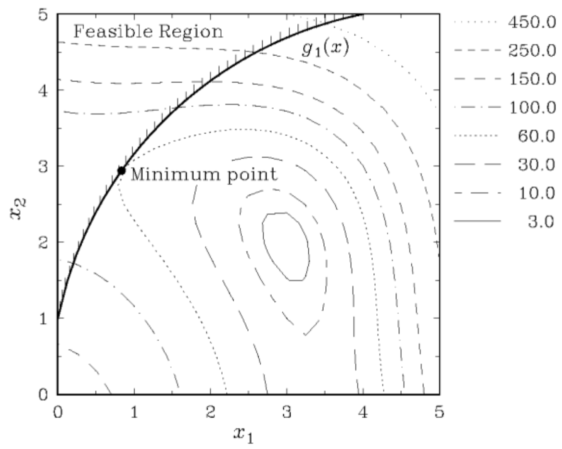


Figure 4.1: The feasible search space and the true minimum of the constrained problem.

Step 1: We use the bracket-operator penalty term to solve this problem. The bracket operator penalty term is an exterior penalty term. We choose an infeasible point $x^{(0)} = (0,0)^T$ as the initial point. We also choose a small value for the penalty parameter: $R^{(0)} = 0{,}1$. We choose two convergence parameters $\varepsilon_1 = \varepsilon_2 = 10^{-5}$.

Step 2: The next task is to form the penalty function:

$$P(x, R^{(0)}) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2 \\ +0.1 \times \langle (x_1 - 5)^2 + x_2^2 - 26 \rangle^2$$

In the above formulation, the variable bounds must also be included as inequality constraints. For clarity and ease of illustration, we have not included the variable bounds in our formulation.

Step 3: The above unconstrained function can now be minimized using one of the methods described in the previous chapter. Here, we use the steepest descent method to solve the above problem. We begin the algorithm with an initial solution $x^{(0)} = (0,0)^T$ having $f(x^{(0)}) = 170{,}0$. At this point, the constraint violation is -1.0 and the penalized function value $P(x^{(0)}, R^{(0)}) = 170{,}100$. Intermediate points obtained by the steepest descent algorithm are tabulated in Table 4.1, and some of these points are shown in Figure 4.2. After

150 function evaluations, the solution $x^* = (2.628, 2.475)^T$ having a function value equal to $f(x^*) = 5.709$ is obtained. At this point, the constraint violation is equal to (-14.248), but has a penalized function value equal to 25.996, which is smaller than that at the initial point. Even though the constraint violation at this point is greater than that at the initial point, the steepest descent method has minimized the penalized function $P(x\ , R^{(0)})$ from 170.100 to 25.996. We set $x^{(1)} = (2.628, 2.475)^T$ and proceed to the next step.

Sequence t	$R^{(t)}$	Solution $x^{(t)}$	$P(x^{(t)}, R^{(t)})$	Constraint violation
1	0.1	$(0, 0)^T$	170.100	−1.000
		$(2.569, 2.294)^T$	27.119	−14.828
		$(2.657, 2.455)^T$	26.019	−14.483
		.	.	.
		$(2.628, 2.475)^T$	25.996	−14.248
2	1.0	$(2.628, 2.475)^T$	208.70	−14.248
		$(1.730, 3.412)^T$	75.140	−3.655
		$(1.166, 2.871)^T$	60.986	−3.058
		.	.	.
		$(1.011, 2.939)^T$	58.757	−1.450
3	10.0	$(1.011, 2.939)^T$	77.591	−1.450
		$(0.906, 3.016)^T$	60.530	−0.143
		.	.	.
		$(0.844, 2.934)^T$	60.233	−0.119

Table 4.1: Tabulation of the Intermediate Points Obtained Using the Steepest Descent Algorithm

Step 4: Since this is the first iteration, we have no previous penalized function value to compare with; thus we move to Step 5.

Step 5: At this step, we update the penalty parameter $R^{(1)} = 10 \times 0{,}1 = 1{,}0$ and move to Step 2. This is the end of the first sequence. It is important here to note that with a different initial point, we could have also converged to the same point. But simulation runs with certain initial points may have taken a longer time to converge than with other points. However, solutions in subsequent sequences will be identical for all simulations.

Step 2: The new penalized function in the second sequence is as follows:

$$P(x, R^{(1)}) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$
$$+1.0 \times \{(x_1 - 5)^2 + x_2^2 - 26\}^2.$$

Step 3: At this step, we once again use the steepest descent method to solve the above problem from the starting point $(2.628, 2.475)^T$. Table 4.1 shows intermediate points of the simulation run. The minimum of the function is found after 340 function evaluations and is $x^{(2)} = (1.011, 2.939)^T$. At this point, the constraint violation is equal to -1.450, which suggests that the point is still an infeasible point. The penalized function and the minimum of the function are both shown in Figure 4.3. The progress of the previous sequence is also shown using dashed lines. Observe that this penalized function is distorted with respect to the original Himmelblau function. This distortion is necessary to shift the minimum point of the current function closer to the true constrained minimum point. Also notice that the penalized function at the feasible region is undistorted.

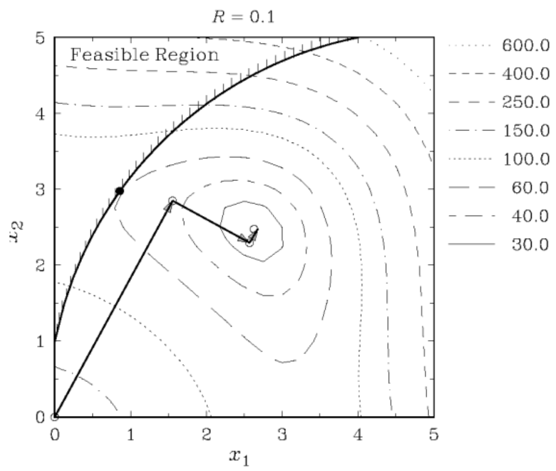


Figure 4.2: A simulation of the steepest descent method on the penalized function with R = 0.1.

Step 4: Comparing the penalized function values, we observe that $P(x^{(2)}, 1{,}0) = 58.664$ and $P(x^{(1)}, 1{,}0) = 25.996$. Since they are very different from each other, we continue with Step 5.

Step 5: The new value of the penalty parameter is $R^{(2)} = 10{,}0$. We increment the iteration counter t = 2 and go to Step 2.

In the next sequence, the penalized function is formed with $R^{(2)} = 10{,}0$. The penalized function and the corresponding solution is shown in Figure 4.4. This time the steepest descent algorithm starts with an initial solution $x^{(2)}$. The minimum point of the sequence is found to be $x^{(3)} = (0.844, 2.934)^T$ with a constraint violation equal to -0.119. Figure 4.4 shows the extent of distortion of the original objective function. Compare the contour levels shown at the top right corner of Figures 4.2 and 4.4. With $R = 10.0$, the effect of the objective function $f(x)$ is almost insignificant compared to that of the constraint violation in the infeasible search region. Thus, the contour lines are almost parallel to the constraint line. Fortunately in this problem, the increase in the penalty parameter R only makes the penalty function steeper in the infeasible search region. In problems with a sufficiently nonlinear objective function and with multiple constraints, a large value of the penalty parameter may create one or more artificial local optima in the search space, thereby making it difficult for the unconstrained search to obtain the correct solution. The advantage of using the unconstrained search method sequentially is that the unconstrained search is always started from the best point found in the

previous sequence. Thus, despite the presence of many local optima in the search space, the search at every sequence is initiated from a point near the correct optimum point. This makes it easier for the unconstrained search to find the correct solution.
After another sequence (iteration) of this algorithm, the obtained solution is

$$x^{(4)} = (0.836, 2.940)^T$$

with a constraint violation of only - 0.012. This point is very close to the true constrained optimum solution. A few more iterations of the penalty function method may be performed to get a solution with the desired accuracy. Although a convergence check with a small difference in the penalized function value at two consecutive sequences is used in this algorithm, any

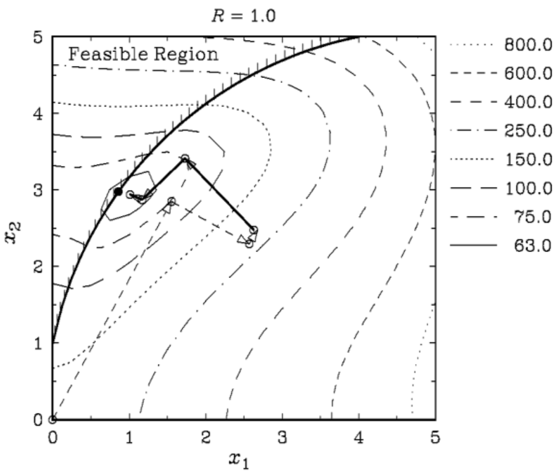


Figure 4.3: Intermediate points using the steepest descent method for the penalized function with R=1.0 (solid lines).

other convergence criteria (for example, a small difference in the *x*-vector of the solutions in two successive sequences) may also be used.
In the presence of multiple constraints, it is observed that the performance of the penalty function method improves considerably if the constraints and the objective functions are first normalized before constructing the penalized function. An inequality constraint $g_j(x) \geq 0$ can be normalized as follows:

$$\frac{g_j(x)}{g_{\max}} \geq 0,$$

where g_{max} is the maximum value of the constraint $g_j(x)$ in the search space. Often, engineering design problems contain constraints restraining resource or capacity of b_j as $g'_j(x) \leq b_j$. The constraint can be normalized as follows:

$$1 - \frac{g'_j(x)}{b_j} \geq 0$$

If an upper bound of the objective function is known, the objective function can also be normalized as shown above, but the normalization of the constraints is more important than that of the objective function.

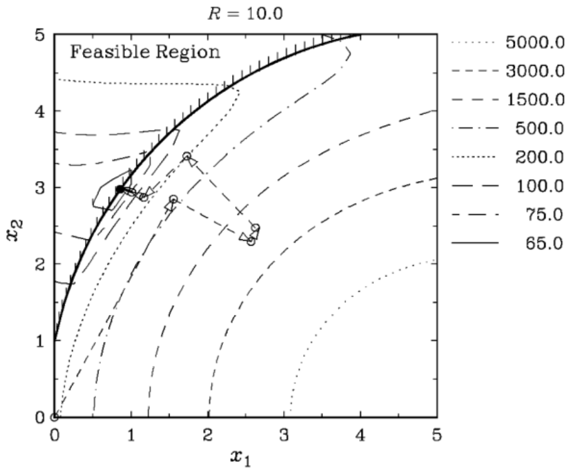


Figure 4.4: Intermediate points obtained using the steepest descent method for the penalized function with R = 10.0 (solid lines near the true optimum). Notice the distortion in the function.

4.2.2. *Method of Multipliers*

- Distortion of functions is avoided by using fixed parameter penalty method
- The following penalty function and update rules are used:

$$P\left(x, \sigma^{(t)}, \tau^{(t)}\right) = f(x) + R \sum_{j1}^{J} \{ (\{g_j(x) + \sigma_j^{(t)}\})^2 - (\sigma_j^{(t)})^2 \}$$

$$+R \sum_{k=1}^{K} \{ (h_k(x) + \tau_k^{(t)})^2 - (\tau_k^{(t)})^2 \}.$$

The final solution x^T is a K-T point. Furthermore, the Lagrange multipliers are computed easily:

$$u_j = -2R\sigma_j^T$$

$$v_k = -2R\tau_k^T$$

Algorithm:

Step 1: Choose a penalty parameter R, termination parameters ε_1 and ε_2. Choose an initial solution $x^{(0)}$. Set multipliers $\sigma_j^{(0)} = \tau_k^{(0)}$ and the iteration counter $t = 0$.

Step 2: Next, form the penalized function:

$$P(x, \sigma^{(t)}, \tau^{(t)}) = f(x) + R \sum_{j1}^{J} \{ (\langle g_j(x) + \sigma_j^{(t)}\})^2 - (\sigma_j^{(t)})^2 \}$$

$$+R \sum_{k1}^{K} \{ (h_k(x) + \tau_k^{(t)})^2 - (\tau_k^{(t)})^2 \}.$$

Step 3: Use an unconstrained optimization technique to solve the penalized function $P(x, \sigma^{(t)}, \tau^{(t)})$ from the starting point $x^{(t)}$ with an convergence factor ε_1. During this optimization, $\sigma^{(t)}$ and $\tau^{(t)}$ are kept fixed; only the vector x is varied. Let us say the solution is $x^{(t+1)}$.

Step 4: Is $\left\| P(x^{(t+1)},\sigma^{(t)},\tau^{(t)}) - P(x^{(t)},\sigma^{(t-1)},\tau^{(t-1)} \right\| \le \varepsilon_2$?

If yes, set $x^T = x^{(t+1)}$ and **Terminate.**
Else go to Step 5.

Step 5: Update $\sigma_j^{(t+1)} = \left\langle g_j(x^{(t)}) + \sigma_j^{(t)} \right\rangle$ for all j = 1,2,…, J and $\tau_k^{(t+1)} = h_k(x^{(t)}) + \tau_k^{(t)}$ for all k = 1,2,…, K. Set $t = t+1$ and go to Step 2.

4.3. Constrained Direct Search

- Detailed structure of constraints are considered,
- Functions may be discontinuous and non-differentiable,
- Algorithms are heuristic in nature,
- Algorithms start from a feasible point. When a new point is created by a fixed rule make sure that the point is feasible. If not, modify by a predefined rule.

4.3.1. Random Search Methods

Points are generated at random.

Luus and Jaakola Algorithm:

Step 1: Given an initial feasible point x^0, an initial range z^0 such that the minimum, x^*, lies in $(x^0 - \frac{1}{2}z^0, x^0 + \frac{1}{2}z^0)$. Choose a parameter $0 < \varepsilon < 1$. For each of Q blocks, initially set $q = 1$ and $p = 1$.

Step 2: For i = 1,2,…, N, create points using a uniform distribution of r in the range (-0.5, 0.5). Set $x_i^{(p)} = x_i^{q-1} + rz_i^{q-1}$.

Step 3: If $x^{(p)}$ is infeasible and $p < P$, repeat Step 2.

If $x^{(p)}$ is feasible, save $x^{(p)}$ and $f(x^{(p)})$, increment p and repeat Step 2.

Else if p = P, set x^q to be the point that has the lowest $f(x^{(p)})$ over all feasible $x^{(p)}$ including x^{q-1} and reset $p = 1$

Step 4: Reduce the range via $z_i^q = (1-\varepsilon)z_i^{q-1}$

Step 5: If $q > Q$, **Terminate.**

Else increment q and continue with Step 2

Suggested values of parameters are $\varepsilon = 0.05$, $p = 100$, and Q is related to the desired reduction in variable uncertainity.

4.4. Method of Feasible Directions

- Rather than relying on solutions to LP problems, use linear approximations to determine a locally good search direction,
- Similar to gradient-based unconstrained search,

– Zoutendijk's method: Have a search direction d that is descent $(\nabla f(x^{(t)}) \cdot d < 0)$ and feasible $(\nabla g(x^{(t)}) \cdot d \geq 0)$.

Algorithm:

Step 1: Set an iteration counter $t = 0$. Choose an initial feasible point $x^{(0)}$ and a parameter for checking the constraint violation, ε.

Step 2: At the current point $x^{(t)}$, let $I^{(t)}$ be the set of indices of active constraints. In other words,

$$I^{(t)} = \{j: 0 \leq g_j(x^{(t)}) \leq \varepsilon, j = 1,2,..)J\}.$$

If $I^{(t)}$ is empty, use $d^{(t)} = \theta^{(t)} = -\nabla f(x^{(t)})$, normalize $d^{(t)}$ and go to Step 4

Step 3: Solve the following LP:

Maximize θ
subject to

$$\begin{aligned}
&\nabla f(x^{(t)})d \leq -\theta, \\
&\nabla g_j(x^{(t)})d \geq \theta, \;\; j \in l^{(t)}, \\
&-1 \leq l \leq 1, \qquad i = 1,2, \ldots, N.
\end{aligned}$$

Label the solution $d^{(t)}$ and $\theta^{(t)}$.

Step 4: If *Q(t) < 0*, **Terminate.**

Else $\bar{\alpha} = \min\left[\alpha : g_j(x^{(t)} + \alpha d^{(t)}) = 0,\; j = 1, 2, \ldots, J, \text{ and } \alpha > 0\right]$.

If no $\bar{\alpha} > 0$ exists, set $\bar{\alpha} = \infty$.

Step 5: Find $\alpha^{(t)}$ such that

$$f(x^{(t)} + \alpha^{(t)} d^{(t)}) = \min\left[f(x^{(t)} + \alpha d^{(t)}) : \; 0 \leq \alpha \leq \overline{\alpha}\right]$$

Set $x^{(t+1)} = x^{(t)} + \alpha^{(t)} d^{(t)}$, $t = t + 1$, and go to Step 2

Comments:

1. Only a subset of constraints are used to define a subproblem, thus LP is smaller,
2. Since only binding constraints are used, zigzag iteration pattern results.

Consider the constrained Himmelblau function again:

$$\text{Minimize } f(x) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

$$\begin{aligned}
&g_1(x) = 26 - (x_1 - 5)^2 - x_2^2 \geq 0, \\
&g_2(x) = 20 - 4x_1 - x_2 \geq 0, \\
&x_1, x_2 \geq 0.
\end{aligned}$$

Let us recall that the minimum point of the above problem lies at $x^* = (3, 2)^T$ with a function value equal to zero.

Step 1: Let us choose an initial feasible point $x^{(0)} = (0,0)^T$ and a tolerance parameter $\varepsilon = 10^{-3}$. We also set the iteration counter $t = 0$.

Step 2: At this step, let us find the active constraints at point $x^{(0)}$. It turns out that only variable bounds are active. Calling these two inequality constraints $g_3(x) = x_1 \geq 0$ and $g_4(x) = x_2 \geq 0$, we update the active constraint set $I^{(0)} = \{3, 4\}$. Since this set is not empty, we continue with Step 3.

Step 3: At this step, we have to find a descent direction which is maximally away from both constraints g_3 and g_4. At the initial point, both active constraints are orthogonal to each other. Thus, the desired search direction may make equal angle with all active constraints. But in any other situation, this may not be true. Thus, we find the optimal search direction by solving an LP problem. Calculating the derivative of the objective function and the active constraints at $x^{(0)}$ numerically and denoting the search direction $d = (d_1, d_2)^T$, we obtain the following LP problem:

Maximize θ
subject to

$$\begin{gathered} -14d_1 - 22d_2 \leq -\theta \\ d_1 \geq \theta, \\ d_2 \geq \theta, \\ -1 \leq d_{1,} d_2 \leq 1. \end{gathered}$$

There exist a number of difficulties with the above formulation to be directly solved using the simplex method of LP technique. First of all, in the above formulation, the variables can take negative values, which are not allowed in a LP technique. Thus, we first substitute $t_i = d_i + 1$ for $i = 1, 2$ such that the variables t_i can take only positive values in the range (0, 2). We rewrite the above LP problem in terms of the new variables:

Maximize θ
subject to

$$\begin{gathered} 14t_1 + 22t_2 - \theta \geq 36 \\ t_1 - \theta \geq 1, \\ t_2 - \theta \geq 1, \\ 0 \leq t_{1)} t_2 \leq 2. \end{gathered}$$

Secondly, the simplex method can handle only equality constraints. Slack variables are usually added or subtracted to convert inequality constraints to equality constraints. Therefore, for each of the above constraints we add a slack variable (y_1 to y_5). Thirdly, we observe that the problem variables and the slack variables do not constitute an initial basic feasible solution. Thus, we introduce three more artificial variables y_6 , y_7, and y_8) to constitute an initial basic feasible solution for the first phase of the dual simplex search method. Thus, the underlying LP problem becomes as follows:

Maximize θ
subject to

$$14t_1 + 22t_2 - \theta - y_1 + y_6 = 36,$$

$$\begin{aligned} t_1 - \theta - y_2 + y_7 &= 1, \\ t_2 - \theta - y_3 + y_8 &= 1, \\ t_1 + y_4 &= 2, \\ t_2 + y_5 &= 2, \\ t_1, t_2, y_1, y_2, t_3, y_4, y_5, y_6, y_7, y_8 &\geq 0. \end{aligned} \tag{4.1}$$

The three-variable problem now becomes an 11-variable problem. At first, we solve the above problem for the objective:

Maximize θ
subject to

$$-(y_6 + y_7 + y_8).$$

Since all artificial variables must also be nonnegative, the solution to the above problem would have $y_6 = y_7 = y_8 = 0$, because the above objective function at this point would be zero. This solution will then be a candidate solution for the initial basic feasible solution of the problem

		0	0	0	0	0	0	0	0	−1	−1	−1			
c_B	Basic	t_1	t_2	θ	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8			
−1	y_6	14	22	−1	−1	0	0	0	0	1	0	0	36	$\frac{36}{22} = 1.64$	
−1	y_7	1	0	−1	0	−1	0	0	0	0	1	0	1	$\frac{1}{0} = \infty$	
−1	y_8	0	$\boxed{1}$	−1	0	0	−1	0	0	0	0	1	1	$\frac{1}{1} = 1$	←
0	y_4	1	0	0	0	0	0	1	0	0	0	0	2	$\frac{2}{0} = \infty$	
0	y_5	0	1	0	0	0	0	0	1	0	0	0	2	$\frac{2}{1} = 2$	
$(\Delta f)_q$		15	23	−3	−1	−1	−1	0	0	0	0	0	$f = -38$		
			↑												

Table 4.2: The First Tableau for the First Phase of the Dual Simplex Search Method

		0	0	0	0	0	0	0	0	−1	−1	−1			
c_B	Basic	t_1	t_2	θ	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8			
−1	y_6	14	0	21	−1	0	$\boxed{22}$	0	0	1	0	−22	14	0.64	←
−1	y_7	1	0	−1	0	−1	0	0	0	0	1	0	1	$\frac{1}{0} = \infty$	
0	t_2	0	1	−1	0	0	−1	0	0	0	0	1	1	−ve	
0	y_4	1	0	0	0	0	0	1	0	0	0	0	2	$\frac{2}{0} = \infty$	
0	y_5	0	0	1	0	0	1	0	1	0	0	−1	1	$\frac{1}{1} = 1$	
$(\Delta f)_q$		15	0	20	−1	−1	22	0	0	0	0	−23	$f = -15$		
							↑								

Table 4.3: The Second Tableau for the First Phase of the Dual Simplex Search Method

presented in Equation (4.1). The successive tables for the first problem of obtaining a feasible starting solution are shown in Tables 4.5 to 4.8.

The objective function value at the Table 4.5 is $f = -38$. (Equation (4.2) is used as the objective.) In the table, it is clear that the nonbasic variable t_2 corresponds to a maximum increase in the function value. (The quantity $(\Delta f)_q$ is larger for t_2.) Thus, we choose t_2 as the new basic variable.

It turns out from the minimum ratio rule that the basic variable y_8 must be replaced by the the variable t_2 in the next iteration. We formulate the next row-echelon matrix. The outcome of the calculation is shown in Table 4.6.

Note that the objective function value has improved considerably from the previous iteration. Here, we also observe that the nonbasic variable y_3 corresponds to the maximum value of the quantity $(\Delta f)_q$. Thus, we choose y_3 as the new basic variable. Using the minimum ratio rule, we also observe that the current basic variable y_6 must be replaced by the variable y_3 (Table 4.7). At the end of the third iteration, we observe that the nonbasic variable t_1 must replace the basic variable y_7. The objective function value at this iteration is $f = -1$.

		0	0	0	0	0	0	0	0	−1	−1	−1		
c_B	Basic	t_1	t_2	θ	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8		
0	y_3	0.64	0	0.95	−0.04	0	1	0	0	0.04	0	−1	0.64	
−1	y_7	$\boxed{1.00}$	0	−1.00	0.00	−1	0	0	0	0.00	1	0	1.00	←
0	t_2	0.64	1	−0.04	−0.04	0	0	0	0	0.04	0	0	1.64	
0	y_4	1.00	0	0.00	0.00	0	0	1	0	0.00	0	0	2.00	
0	y_5	−0.64	0	0.04	0.04	0	0	0	1	−0.04	0	0	0.36	
$(\Delta f)_q$		1.00	0	−1.00	0.00	−1	0	0	0	−1.00	0	−1	$f = -1$	
		↑												

Ratios: 1 (first row), 1 (second row), 2.57 (third row), 2 (fourth row), and −ve (fifth row)

Table 4.4: The Third Tableau for the First Phase of the Dual Simplex Search Method

We form the next row-echelon matrix in Table 4.8. At this stage, we observe that all artificial

		0	0	0	0	0	0	0	0	−1	−1	−1	
c_B	Basic	t_1	t_2	θ	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	
0	y_3	0	0	1.59	−0.04	0.64	1	0	0	0.04	−0.64	−1	0
0	t_1	1	0	−1.00	0.00	−1.00	0	0	0	0.00	1.00	0	1
0	t_2	0	1	0.59	−0.04	0.64	0	0	0	0.04	−0.64	0	1
0	y_4	0	0	1.00	0.00	1.00	0	1	0	0.00	−1.00	0	1
0	y_5	0	0	−0.59	0.04	0.04	0	0	1	−0.04	0.64	0	1
$(\Delta f)_q$		0	0	0.00	0.00	0	0	0	0	−1.00	−1.00	−1	$f = 0$

Table 4.5: The Fourth Tableau for the First Phase of the Dual Simplex Search Method

variables are zero and the objective function is also equal to zero. This is the termination criterion for the first phase of the dual phase method. The solution of the above iteration is $t_1 = 1$, $t_2 = 1$, $y_1 = 0$, $y_2 = 0$, $y_3 = 0$, $y_4 = 1$, and $y_5 = 1$. This solution was not obvious in the formulation of the problem presented in Equation (4.1).
We begin the second phase with the above solution as the initial solution. The objective in the second phase is to maximize the original function: $f(x) = \theta$. Since the artificial variables are no more required, we discontinue with them in subsequent computations.

		0	0	1	0	0	0	0	0		
c_B	Basic	t_1	t_2	θ	y_1	y_2	y_3	y_4	y_5		
0	y_3	0	0	$\boxed{1.59}$	−0.04	0.64	1	0	0	0	$\frac{0}{1.59} = 0$ ←
0	t_1	1	0	−1.00	0.00	−1.00	0	0	0	1	−ve
0	t_2	0	1	0.59	−0.04	0.64	0	0	0	1	$\frac{1}{0.59} = 1.69$
0	y_4	0	0	1.00	0.00	1.00	0	1	0	1	$\frac{1}{1} = 1$
0	y_5	0	0	−0.59	0.04	0.04	0	0	1	1	−ve
$(\Delta f)_q$		0	0	1.00	0.00	0.00	0	0	0		$f(x) = 0$
				↑							

Table 4.6: The First Tableau for the Second Phase of the Dual Simplex Search Method

Now, we get back to Step 3 of the feasible direction search method.

Step 3: **(cont.)** The solution of the LP problem is $t = (2, 2)^T$ or $d^{(0)} = (1, 1)^T$ and $\theta^{(0)} = 1$. This solution implies that the resulting search direction makes equal angles with each of the two active constraints.

Step 4: Since $\theta^{(0)} = 1 > 0$, we do not terminate the algorithm. Instead, we calculate the limits along the direction $d^{(0)}$ before an infeasible point is found. Any generic point along $d^{(0)}$ from $x^{(0)}$ can be written as $x(\alpha) = x^{(0)} + \alpha d^{(0)}$ or $x(\alpha) = (\alpha, \alpha)^T$. The upper limit on a can be calculated by finding points along $d^{(0)}$ that intersect with each constraint. The problem of finding the intersection of a straight line and any generic curve can be posed as a root-finding problem, which can be solved using an optimization algorithm discussed in Chapter 2. We substitute the expression for $x_1 = \alpha$ and $x_2 = \alpha$ in each constraint and then minimize the following problem:

$$\text{Minimize abs } [g_j(x(\alpha))]. \tag{4.3}$$

For example, the upper limit along $d^{(0)}$ can be found for the first constraint by minimizing the unidirectional function:

$$\text{abs } [26 - (\alpha - 5)^2 - \alpha^2]$$

Since the absolute value of the argument is always considered, the above function allows only positive values. Since we are looking for points for which the constraint has a value zero, those points correspond to the minimum value of the above expression. Note that the problem described in Equation (4.3) is a single-variable function. Thus, we first bracket the minimum and then minimize the function. Using the bounding phase method from a starting point $\alpha^{(0)} = 5$ and $\Delta = 1$, we obtain the bracketing interval (4, 6). Next, we use the golden section search in that interval to obtain the minimum point with three decimal places of accuracy: $\alpha = 5.098$. The same solution can also be obtained by solving the quadratic expression $g_1(x(\alpha)) = 0$. Similarly, the limit on the second constraint can also be calculated: $\alpha_2^* = 4{,}0$. Other constraints produce upper limits $\alpha_3^* = \alpha_4^* = 0$, which are not acceptable. Thus, the true upper limit is $\alpha = 4{,}0$.

		0	0	1	0	0	0	0	0		
c_B	Basic	t_1	t_2	θ	y_1	y_2	y_3	y_4	y_5		
1	θ	0	0	1	−0.03	0.40	0.63	0	0	0	−ve
0	t_1	1	0	0	−0.03	−0.60	0.63	0	0	1	$\frac{1}{0} = \infty$
0	t_2	0	1	0	−0.03	0.40	−0.37	0	0	1	−ve
0	y_4	0	0	0	$\boxed{0.03}$	0.60	0.63	1	0	1	$\frac{1}{0.03} = 35$ ←
0	y_5	0	0	0	0.03	−0.40	0.37	0	1	1	$\frac{1}{0.03} = 35$
$(\Delta f)_q$		0	0	0	0.03	−0.4	−0.63	0	0	$f(x) = 0$	
					↑						

Table 4.7: The Second Tableau for the Second Phase of the Dual Simplex Search Method

		0	0	1	0	0	0	0	0	
c_B	Basic	t_1	t_2	θ	y_1	y_2	y_3	y_4	y_5	
1	θ	0	0	1	0	1	0	1	0	1
0	t_1	0	0	0	0	0	0	1	0	2
0	t_2	0	1	0	0	1	−1	1	0	2
0	y_1	0	0	0	1	21	−22	1	0	35
0	y_5	0	0	0	0	−1	1	−1	1	0
$(\Delta f)_q$		0	0	0	0	−1	0	−1	0	$f(x) = 1$

Table 4.8: The Third Tableau for the Second Phase of the Dual Simplex Search Method

Step 5: Once the lower and upper limit on a are found, we perform another one dimensional search with the given objective function to find the minimum point along that direction. Using the golden section search, we obtain the minimum point in the interval (0, 4): $\alpha^* = 2.541$, which corresponds to the new point $x^{(1)} = (2.541, 2.541)^T$. At this point, we increment the iteration counter and go to Step 2. This completes one iteration of the feasible direction method. The progress of this iteration is shown in Figure 4.5.

Step 2: At the new point, we find that no constraints are active. Thus, $I^{(1)} = \emptyset$, which means that the point is not on any constraint boundary and we are free to search in any direction locally. Therefore, we choose the steepest descent direction and the search direction is set according to the negative of the gradient of the objective function at the new point:

$$d^{(1)} = -\nabla f(x^{(1)}) = (16.323, -16.339)^T,$$

which is computed numerically. At this point the function value is $f(x^{(1)}) = 8{,}0$.

Step 4: Once again, we compute the upper limit along the search direction $d^{(1)}$. Posing the root-finding problem as an optimization problem as shown in the previous iteration, we obtain the parameter $\bar{\alpha} = \min[0.162, 0.149] = 0.149$.

Step 5: Performing a unidirectional search along in the domain (0, 0.149), we obtain $\alpha^* = 0.029$. The corresponding point is $x^{(2)} = (3.018, 2.064)^T$ with an objective function value $f(x^{(2)}) = 0.107$.

This process continues until a point with a small derivative of the objective function is found. If the intermediate points fall on the constraint boundary frequently, this method may

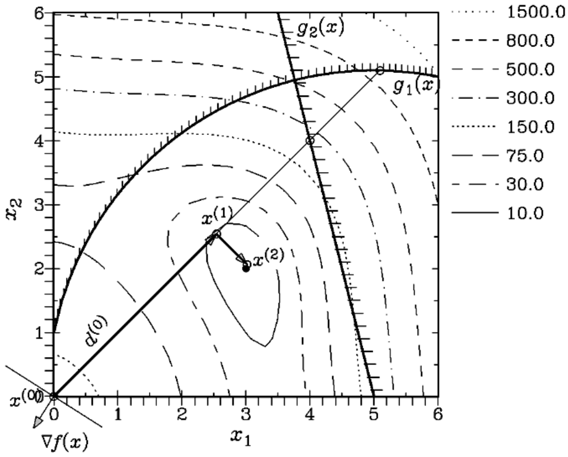


Figure 4.5: A number of iterations of the feasible direction method.

be expensive in terms of the overall computational time required to solve the problem. This situation may happen for problems with a narrow feasible region.

4.5. Quadratic Approximation Method

- Quadratic approximation to objective functions
- Linear approximations to constraints to make the calculations easier

Algorithm:

Step 1: Given $x^{(0)}$ and a suitable method for solving QP problem

Step 2: Formulate QP

Minimize $\nabla f(x^{(t)})^T d + \frac{1}{2} d^T \nabla^2 f(x^{(t)}) d$

Subject to $h_k(x^{(t)}) + \nabla h_k(x^{(t)})^T d = 0$ $\quad k = 1,2 \ldots K$

$g_j(x^{(t)}) + \nabla g_j(x^{(t)})^T d \geq 0$ $\quad j = 1,2 \ldots J$

Step 3: Solve the QP problem and set $x^{(t+1)} = x^{(t)} + d$

Step 4: Check for convergence. If not converged, Go to step 1

– Lead to nonvertex solutions
– The term $\nabla^2 f(x^{(t)})$ in the objective function may be replaced by a Lagrangian term or a variable-metric approximation for Hessian for easier computation

This procedure can be repeatedly used to solve general non-linear programming problems by formulating a QP problem at the current best solution. This methodology is called Sequential Quadratic Programming or SQP. Many commercial optimization software implements this algorithm in which the resulting QP problem is solved using a quasi-Newton multi-variable optimization algorithm.

4.6. References

[4.1] Box, M. J. (1965): A **new method of constrained** optimization **and a comparison with other methods.** Computer Journal. 8, 42-52.

[4.2] Kelly, J. E. (1960): **The cutting plane method for solving** convex **programs.** SIAM Journal. 8, 703-712.

[4.3] Luus, R. and Jaakola, T. H. I. (1973): **Optimization by direct** search **and systematic reduction of the size of search region.** AIChE Journal. 19,760-766.

[4.4] Mangasarian, 0. L. (1969): **Nonlinear Programming.** New York: McGrawHill.

[4.5] Rao, S. S. {1984): **Optimization Theory and** Applications**.** New Delhi:Wiley Eastern.

[4.6] Reklaitis, G. V., Ravindran, A., and Ragsdell, K. M. (1983): Engineering **Optimization-Methods and Applications.** New York: Wiley.

[4.7] Strang, G. (1980): **Linear Algebra and Its Applications.** Orlando: Academic Press.

[4.8] Taha, H. A. (1989): **Operations Research.** New York: MacMillan.

[4.9] Zangwill, W. I. (1969): **Nonlinear Programming.** Englewood Cliffs, New Jersey: Prentice-Hall.

[4.10] Zoutendijk, G. (1960): **Methods of Feasible Directions.** Amsterdam: Elsevier.

4.7. Questions

1. What are the Kuhn-Tucker conditions? How are they used?
2. Explain the concept of penalty function!
3. Explain the method of feasible directions!

5. NONTRADITIONAL OPTIMIZATION TECHNIQUES

This chapter makes a brief description of a nontraditional search and optimization methods. Further details will be supplied by separate sources.

5.1. Genetic Algorithms

Genetic algorithms (abbreviated as GAs) are computerized search and optimization algorithms designed based on the mechanics of natural genetics and natural selection.

5.1.1. Fundamental Differences with Traditional methods

– GAs work on a coding of parameters, instead of parameters. GAs exploit the coding similarities to achieve a parallel search.
– GAs work on a population of points, instead of a single point. That is why GAs are likely to find the global solutions.
– GAs do not require any derivative or auxiliary information. This extends the application of GAs to a wide variety of problem domains. That is why GAs are robust.
– GAs use probabilistic transition rules, instead of deterministic transition rules. This reduces the bias in the search. Initially the search direction is random and as iteration progresses, GAs obtain a directed search adaptively.

Algorithm:

Step 1: Choose a coding to represent problem parameters, a selection operator, a crossover operator, and a mutation operator. Choose population size, N, crossover probability, p_c, mutation probability, p_m. Initialize a random population of strings of size N. Set $t = 0$.

Step 2: Evaluate each string in the population.

Step 3: If $t > t_{\max}$ or other termination criteria is satisfied, **Terminate.**

Else Go to step 4.

Step 4: Reproduction on the population.

Step 5: Crossover on random pairs of strings.

Step 6: Mutation on every string.

Step 7: Set $t = t + 1$ and Go to step 2.

5.1.2. Reproduction Operator

Selects good strings from a population. Some of popular reproduction operators are as follows:

Proportionate Selection Strings are selected according to their fitness. Specifically, a string with fitness f_i is allocated $\dfrac{f_i}{f_{avg}}$ number of copies. Better strings are allocated more copies under this scheme.

Tournament Selection Usually, s strings are selected at random from a population, and the best is chosen. This procedure is continued until the whole population is filled up. This scheme can be performed with and without replacement. When performed without replacement, the best string gets s copies.

Ranking Selection All strings are first ranked from best to worst and ranked in a way so that the best gets s copies and the worst gets zero copies. Each string is then selected with a probability depending on its rank in the population.

5.1.3. Crossover Operator

Exchanges information between two strings selected at random. Some popular operators are as follows:

Single-point Crossover A cross site is chosen at random. The contents on one side of the site are exchanged between parent strings:

$$\begin{array}{cc|ccc} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \Rightarrow \begin{array}{cc|ccc} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array}$$

Two-point Crossover Two cross sites are chosen at random. The contents between two sites are exchanged between parent strings:

$$\begin{array}{c|cc|cc} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \Rightarrow \begin{array}{c|cc|cc} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{array}$$

Uniform Crossover Each bit-position is exchanged between parent strings with a probability 0.5:

$$\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \Rightarrow \begin{array}{ccccc} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{array}$$

The search power of uniform crossover is most among these three operators, but the destruction probability of good partial substrings in parent strings is also more in uniform crossover.

In order to preserve some previously obtained good solutions, all strings in the population is not used in crossover. The proportion of population used in crossover is known as probability of crossover. Usually, a proportion of 0.60-0.95 is used.

5.1.4. Mutation Operator

Alters a bit value to another with a small probability. In a binary string, mutation operation is shown below:

$$11111 \Rightarrow 11101$$

Mutation maintains diversity in the population.

– Other advanced operators exist and tried

The objective is to minimize the function

$$f(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

in the interval $0 \le x_{1,}\ x_2 \le 6$. Recall that the true solution to this problem is $(3,2)^T$ having a function value equal to zero.

Step 1: In order to solve this problem using genetic algorithms, we choose binary coding to represent variables x_1 and x_2. In the calculations here, 10-bits are chosen for each variable, thereby making the total string length equal to 20. With 10 bits, we can get a solution accuracy of $(6-0)/(2^{10}-1)$ or 0.006 in the interval (0,6). We choose roulette-wheel selection, a single-point crossover, and a bit-wise mutation operator. The crossover and mutation probabilities are assigned to be 0.8 and 0.05, respectively. We decide to have 20 points in the population. We set $t_{\max} = 30$ and initialize the generation counter $t = 0$.

	String		x_2	x_1	f(x)	F (x)	Expected	Probability	Cumulative	Random	String	Actual	Mating	
	Substring-2	Substring-1					Count	of copying	probability	number	number	count	pool	
1	1110010000	1100100000	5.349	4.692	959.680	0.001	0.13	0.007	0.007	0.472	10	0	0010100100	1010101010
2	0001001101	0011100111	0.452	1.355	105.520	0.009	1.10	0.055	0.062	0.108	3	1	1010100001	0111001000
3	1010100001	0111001000	3.947	2.674	126.685	0.008	0.98	0.049	0.111	0.045	2	1	0001001101	0011100111
4	1001000110	1000010100	3.413	3.120	65.026	0.015	1.85	0.093	0.204	0.723	14	2	1110011011	0111000010
5	1100011000	1011100011	4.645	4.334	512.197	0.002	0.25	0.013	0.217	0.536	10	0	0010100100	1010101010
6	0011100101	0011111000	1.343	1.455	70.868	0.014	1.71	0.086	0.303	0.931	19	2	0011100010	1011000011
7	0101011011	0000000111	2.035	0.041	88.273	0.011	1.34	0.067	0.370	0.972	19	1	0011100010	1011000011
8	1110101000	1110101011	5.490	5.507	1436.563	0.001	0.12	0.006	0.376	0.817	17	0	0111000010	1011000110
9	1001111101	1011100111	3.736	4.358	265.556	0.004	0.49	0.025	0.401	0.363	7	1	0101011011	0000000111
10	0010100100	1010101010	0.962	4.000	39.849	0.024	2.96	0.148	0.549	0.189	4	3	1001000110	1000010100
11	1111101001	0001110100	5.871	0.680	814.117	0.001	0.14	0.007	0.556	0.220	6	0	0011100101	0011111000
12	0000111101	0110011101	0.358	2.422	42.598	0.023	2.84	0.142	0.698	0.288	6	3	0011100101	0011111000
13	0000111110	1110001101	0.364	5.331	318.746	0.003	0.36	0.018	0.716	0.615	12	1	0000111101	0110011101
14	1110011011	0111000010	5.413	2.639	624.164	0.002	0.24	0.012	0.728	0.712	13	1	0000111110	1110001101
15	1010111010	1010111000	4.094	4.082	286.800	0.003	0.37	0.019	0.747	0.607	12	0	0000111101	0110011101
16	0100011111	1100111000	1.683	4.833	197.556	0.005	0.61	0.030	0.777	0.192	4	0	1001000110	1000010100
17	0111000010	1011000110	2.639	4.164	97.699	0.010	1.22	0.060	0.837	0.386	9	1	1001111101	1011100111
18	1010010100	0100001001	3.871	1.554	113.201	0.009	1.09	0.054	0.891	0.872	18	1	1010010100	0100001001
19	0011100010	1011000011	1.326	4.147	57.753	0.017	2.08	0.103	0.994	0.589	12	2	0000111101	0110011101
20	1011100011	1111010000	4.334	5.724	987.955	0.001	0.13	0.006	1.000	0.413	10	0	0010100100	1010101010

Table 5.1: Evaluation and reproduction of a random population is illustrated.

Step 2: The next step is to evaluate each string in the population. We calculate the fitness of the first string. The first substring (1100100000) decodes to a value equal to $\left(2^9 + 2^8 + 2^5\right)$ or 800. Thus, the corresponding parameter value is equal to $0 + (6-0) \times 800/1023$ or 4.692. The second substring (1110010000) decodes to a value equal to $\left(2^9 + 2^8 + 2^7 + 2^4\right)$ or 912. Thus, the corresponding parameter value is equal to $0 + (6-0) \times 912/1023$ or 5.349. Thus, the first string corresponds to the point $x^{(1)} = (4.692, 5.349)^T$. These values can now be substituted in the objective function expression to obtain the function value. It is found that the function value at this point is equal to $f(x^{(1)}) = 959.680$. We now calculate the fitness function value at this point using the transformation rule: $F(x^{(1)}) = 1.0/(1.0 + 959.680) = 0.001$. This value is used in the reproduction operation. Similarly, other strings in the population are evaluated and fitness values are calculated. Table 5.1 shows the objective function value and the fitness value for all 20 strings in the initial population.

Step 3: Since $t = 0 < t_{\max} = 30$, we proceed to Step 4.

Step 4: At this step, we select good strings in the population to form the mating pool. In order to use the roulette-wheel selection procedure, we first calculate the average fitness of the population. By adding the fitness values of all strings and dividing the sum by the population size, we obtain $\overline{F} = 0.008$. The next step is to compute the expected count of each string as $F(x)/\overline{F}$. The values are calculated and shown in column A of Table 5.1. In other words, we can compute the probability of each string being copied in the mating pool by dividing these numbers with the population size (column B). Once these probabilities are calculated, the cumulative probability can also be computed. These distributions are also shown in column C of Table 5.1. In order to form the mating pool, we create random numbers between zero and one (given in column D) and identify the particular string which is specified by each of these random numbers. For example, if the random number 0.472 is created, the tenth string gets a copy in the mating pool, because that string occupies the interval (0.401, 0.549), as shown in column C. Column E refers to the selected string. Similarly, other strings are selected according to the random numbers shown in column D. After this selection procedure is repeated n times (n is the population size), the number of selected copies for each string is counted. This number is shown in column F. The complete mating pool is also shown in the table. Columns A and F reveal that the theoretical expected count and the true count of each string more or less agree with each other. Figure 5.1 shows the initial random population and the mating pool after reproduction. The points marked with an enclosed box are the points in the mating pool. The action of the reproduction operator is clear from this plot. The inferior points have been probabilistically eliminated from further consideration. Notice that not all selected points are better than all rejected points. For example, the 14th individual (with a fitness value 0.002) is selected but the 16th individual (with a function value 0.005) is not selected.

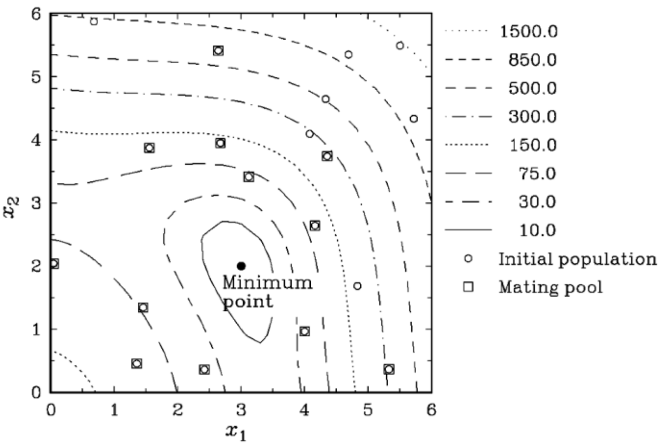


Figure 5.1.: The initial population (marked with empty circles) and the mating pool (marked with boxes) on a contour plot of the objective function. The best point in the population has a function value 39.849 and the average function value of the initial population is 360.540.

Although the above roulette-wheel selection is easier to implement, it is noisy. A more stable version of this selection operator is sometimes used. After the expected count for each individual string is calculated, the strings are first assigned copies exactly equal to the mantissa of the expected count. Thereafter, the regular roulette-wheel selection is implemented using the decimal part of the expected count as the probability of selection.

This selection method is less noisy and is known as the stochastic remainder selection.

Step 5: At this step, the strings in the mating pool are used in the crossover operation. In a single-point crossover, two strings are selected at random and crossed at a random site. Since the mating pool contains strings at random, we pick pairs of strings from the top of the list. Thus, strings 3 and 10 participate in the first crossover operation. When two strings are chosen for crossover, first a coin is flipped with a probability $p_c = 0.8$ to check whether a crossover is desired or not. If the outcome of the coin-flipping is true, the crossing over is performed, otherwise the strings are directly placed in an intermediate population for subsequent genetic operation. It turns out that the outcome of the first coin-flipping is true, meaning that a crossover is required to be performed. The next step is to find a cross-site at random. We choose a site by creating a random number between $(0, \ell - 1)$ or (0, 19). It turns out that the obtained random number is 11. Thus, we cross the strings at the site 11 and create two new strings. After crossover, the children strings are placed in the intermediate population. Then, strings 14 and 2 (selected at random) are used in the crossover operation. This time the coin-flipping comes true again and we perform the crossover at the site 8 found at random. The new children strings are put into the intermediate population. Figure 5.2 shows how points cross over and form new points. The points marked with a small box are the points in the mating pool and the points marked with a small circle are children points created after crossover operation. Notice that not all 10 pairs of points in the mating pool cross with each other. With the flipping of a coin with a probability $p_c = 0.8$, it turns out that fourth, seventh, and tenth crossovers come out to be false. Thus, in these cases, the strings are copied directly into the intermediate population. The complete population at the end of the crossover operation is shown in Table 5.2.

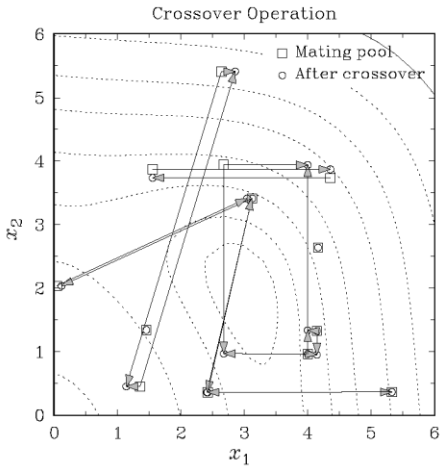


Figure 5.2.: The population after the crossover operation. Two points are crossed over to form two new points. Of ten pairs of strings, seven pairs are crossed.

It is interesting to note that with $p_c = 0.8$, the expected number of crossover in a population of size 20 is $0.8 \times 20/2$ or 8. In this exercise problem, we performed seven crossovers and in three cases we simply copied the strings to the intermediate population. Figure 5.2 shows that some good points and some not-so-good points are created after crossover. In some cases, points far away from the parent points are created and in some cases points close to the parent points are created.

0010100100 1010101010	yes	9	0010100101 0111001000	0010101101 0111001000	1.015	2.674	18.886	0.050
1010100001 0111001000	yes	9	1010100000 1010101010	1010100001 1010101010	3.947	4.000	238.322	0.004
0001001101 0011100111	yes	12	0001001101 0011000010	0001001101 0001000010	0.452	0.387	149.204	0.007
1110011011 0111000010	yes	12	1110011011 0111100111	1110011011 0101100011	5.413	2.082	596.340	0.002
0010100100 1010101010	yes	5	0010100010 1011000011	0010100010 1011000011	0.950	4.147	54.851	0.018
0011100010 1011000011	yes	5	0011100100 1010101010	0011100100 1110101010	1.337	5.501	424.583	0.002
0011100010 1011000011	no		0011100010 1011000011	0011100011 1011100011	1.331	4.334	83.929	0.012
0111000010 1011000110	no		0111000010 1011000110	0101010010 1011000110	1.982	4.164	70.472	0.014
0101011011 0000000111	yes	14	0101011011 0000010100	0101011011 0000010100	2.035	0.117	87.633	0.011
1001000110 1000010100	yes	14	1001000110 1000000111	1001010110 1000000111	3.507	3.044	72.789	0.014
0011100101 0011111000	yes	1	0011100101 0011111000	0011100101 0011111000	1.343	1.455	70.868	0.014
0011100101 0011111000	yes	1	0011100101 0011111000	0011100101 0011111000	1.343	1.455	70.868	0.014
0000111101 0110011101	no		0000111101 0110011101	0000101101 0111011100	0.264	2.792	25.783	0.037
0000111110 1110001101	no		0000111110 1110001101	0000111110 1110001101	0.364	5.331	318.746	0.003
0000111101 0110011101	yes	18	0000111101 0110011100	0000111101 0110011100	0.358	2.416	42.922	0.023
1001000110 1000010100	yes	18	1001000110 1000010101	1001000110 0000010101	3.413	0.123	80.127	0.012
1001111101 1011100111	yes	10	1001111101 0100001001	1001111101 0100001001	3.736	1.554	95.968	0.010
1010010100 0100001001	yes	10	1010010100 1011100111	1010010100 1010100111	3.871	3.982	219.426	0.005
0000111101 0110011101	no		0000111101 0110011101	0000111101 0110011101	0.358	2.422	42.598	0.023
0010100100 1010101010	no		0010100100 1010101010	0010100100 1010101010	0.962	4.000	39.849	0.024

Table 5.2: Crossover and mutation operators are shown.

Step 6: The next step is to perform mutation on strings in the intermediate population.

For bit-wise mutation, we flip a coin with a probability $p_m = 0.05$ for every bit. If the outcome is true, we alter the bit to 1 or 0 depending on the bit value. With a probability of 0.05, a population size 20, and a string length 20, we can expect to alter a total of about $0.05 \times 20 \times 20$ or 20 bits in the population. Table 5.2 shows the mutated bits in bold characters in the table. As counted from the table, we have actually altered 16 bits. Figure 6.3 shows the effect of mutation on the intermediate population. In some cases, the mutation operator changes a point locally and in some other it can bring a large change. The points marked with a small circle are points in the intermediate population. The points marked with a small box constitute the new population (obtained after reproduction, crossover, and mutation). It is interesting to note that if only one bit is mutated in a string, the point is moved along a particular variable only. Like the crossover operator, the mutation operator has created some points better and some points worse than the original points. This flexibility enables GA operators to explore the search space properly before converging to a region prematurely. Although this requires some extra computation, this flexibility is essential to solve global optimization problems.

Step 7: The resulting population becomes the new population. We now evaluate each string as before by first identifying the substrings for each variable and mapping the decoded values of the substrings in the chosen intervals. This completes one iteration of genetic algorithms. We increment the generation counter to $t = 1$ and proceed to Step 3 for the next iteration. The new population after one iteration of GAs is shown in Figure 5.3 (marked with empty boxes). The figure shows that in one iteration, some good points have been found. Table 5.2 also shows the fitness values and objective function values of the new population members.

The average fitness of the new population is calculated to be 0.015, a remarkable improvement from that in the initial population (recall that the average in the initial population was 0.008).

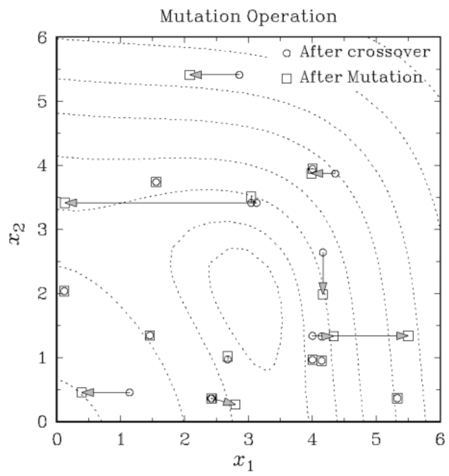


Figure 5.3.: The population after mutation operation. Some points do not get mutated and remain unaltered. The best point in the population has a function value 18.886 and the average function value of the population is 140.210, an improvement of over 60 per cent.

The best point in this population is found to have a fitness equal to 0.050, which is also better than that in the initial population (0.024). This process continues until the maximum allowable generation is reached or some other termination criterion is met. The population after 25 generation is shown in Figure 5.4. At this generation, the best point is found to be $(3.003, 1.994)^T$ with a function value 0.001. The fitness value at this point is equal to 0.999 and the average population fitness of the population is 0.474. The figure shows how points are clustered around the true minimum of the function in this generation. A few inferior points are still found in the plot. They are the result of some unsuccessful crossover events. We also observe that the total number of function evaluations required to obtain this solution is $0.8 \times 20 \times 26$ or 416 (including the evaluations of the initial population).

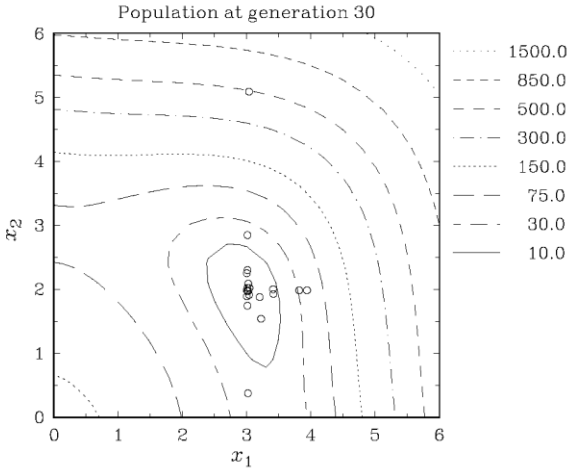


Figure 5.4.: All 20 points in the population at generation 25 shown on the contour plot of the objectivefunction. The figure shows that most points are clustered around the true minimum.

In order to show the efficiency of GAs in arriving at a point close to the true optimum, we perform two more simulations starting with different initial populations. Figure 6.5 shows how the function value of the best point in a population reduces with generation number. Although all three runs have a different initial best point, they quickly converge to a solution close to the true optimum (recall that the optimum point has a function value equal to zero).

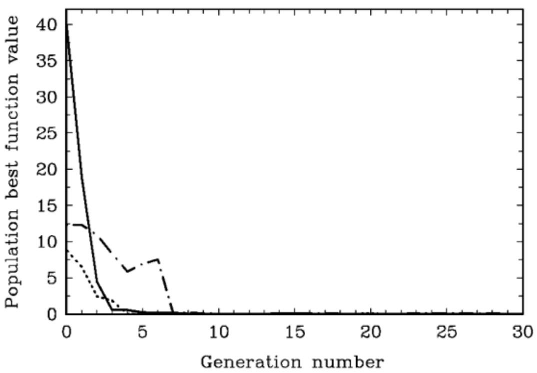


Figure 5.5: The function value of the best point in the population for three independent GA runs. All runs quickly converge to a point near the true optimum.

In order to illustrate the schema processing through genetic operators, we investigate the growth of a particular schema $H = (0 * \ldots * \; * \ldots *)$. This schema represents all points in the range $0 \leq x_2 < 3$. The optimum point lies in this region. With reference to Equation (6.3), we observe that the order, defining length, and the fitness of the schema are such that it is a building block. This schema contains more good points than it's competitor $H^c = (1 * \ldots * \; * \ldots *)$ which represents the range $3 \leq x_2 < 6$. According to Equation (6.3), the schema H must increase exponentially due to the action of genetic operators. We observe that in the random initial population the schema H has nine strings and the schema H^c has 11 strings. At the end of one generation, the population has 14 strings representing the schema H and only six strings representing the schema H^c. We may also investigate other interesting regions in the search space and observe their growth in terms of the number of representative points in the population. Other low-order and above-average schemata are also processed similarly and are combined to form higher-order and good schemata. This processing of several schemata happens in parallel without any extra book-keeping (Goldberg, 1989). Eventually, this processing forms the optimum or a near-optimum point.

5.2. References

[5.1] Goldberg, D. E. (1989): **Genetic Algorithms in Search, Optimization, and Machine Learning.** Reading, Mass.: Addison-Wesley.

5.3. Questions

1. How does a Genetic algorithm work? What are the main operators?
2. What is the aim of the crossover operator?
3. What are the advantages/disadvantages of GAs?

6. MULTI-CRITERION OPTIMIZATION

As the name suggests, a multi-objective optimization problem (MOOP) deals with more than one objective function. In most practical decision-making problems, multiple objectives or multiple criteria are evident. Because of a lack of suitable solution methodologies, an MOOP has been mostly cast and solved as a single objective optimization problem in the past. However, there exist a number of fundamental differences between the working principles of single and multi-objective optimization algorithms. In a single-objective optimization problem, the task is to find one solution (except in some specific multi-modal optimization problems, where multiple optimal solutions are sought) which optimizes the sole objective function.
Extending the idea to multi-objective optimization, it may be wrongly assumed that the task in a multi-objective optimization is to find an optimal solution corresponding to each objective function. In this chapter, we will discuss the principles of multi-objective optimization and present optimality conditions for any solution to be optimal in the presence of multiple objectives.

6.1. Multi-Objective Optimization Problem

A multi-objective optimization problem has a number of objective functions which are to be minimized or maximized. As in the single-objective optimization problem, here too the problem usually has a number of constraints which any feasible solution (including the optimal solution) must satisfy. In the following, we state the multi-objective optimization problem (MOOP) in its general form:

$$\min/\max \quad f_m(x), \qquad m = 1,2 \dots, M;$$

subject to:

$$g_j(x) \geq 0, j = 1,2, \dots, J; \qquad (6.1)$$

$$h_k(x) = 0 \qquad k = 1,2, \dots, K;$$

$$\chi_i^{(L)} \leq \chi_i \leq \chi_i^{(U)} \qquad i = 1,2, \dots, n;$$

A solution **x** is a vector of n decision variables: $\mathbf{x} = (\chi_1, \chi_2, \dots, \chi_n)^T$. The last set of constraints are called variable bounds, restricting each decision variable χ_i to take value within a lower $\chi_i^{(L)}$ and an upper $\chi_i^{(U)}$ bound. These bounds constitute a decision variable space *D*, or simply the decision space. Throughout this chapter, we use the terms point and solution interchangeably to mean a solution vector **x**. Associated with the problem are *J* inequality and *K* equality constraints. The terms $g_j(\boldsymbol{x})$ and $h_k(\boldsymbol{x})$ are called *constraint functions*. The inequality constraints are treated as 'greater-than-equal-to' types, although a 'less-than-equal-to' type inequality constraint is also taken care of in the above formulation. In the latter case, the constraint must be converted into a 'greater-than-equal-to' type constraint by multiplying the constraint function by –l (Deb 1995). A solution ***x*** that does not satisfy all of the $(J + K)$ constraints and all of the *2N* variable bounds stated above is called an infeasible solution. On the other hand, if any solution **x** satisfies all constraints and variable bounds, it is known as a feasible solution. Therefore, we realize that in the presence of constraints, the entire decision variable space *D* need not be feasible. The set of all feasible solutions is called the feasible

region, or *S*. In this script, sometimes we will refer to the feasible region as simply the search space.
There are M objective functions $f(x) = (f_1(x), f_2(x), \dots f_M(x))^T$ considered in the above formulation. Each objective function can be either minimized or maximized. The duality principle (Deb,1995, Rao, Reklaitis et al.) in the context of optimization, suggests that we can convert a maximization problem into a minimization one by multiplying the objective function by -1. The duality principle has made the task of handling mixed type of objectives much easier. Many optimization algorithms are developed to solve only one type of optimization problems, such as e.g. minimization problems. When an objective is required to be maximized by using such an algorithm, the duality principle can be used to transform the original objective for maximization into an objective for minimization.
Although there is a difference in the way that a criterion function and an objective function is defined (Chankong), in a broad sense we treat them here as identical. One of the striking differences between single-objective and multi-objective optimization is that in multi-objective optimization the objective functions constitute a multi-dimensional space, in addition to the usual decision variable space.
This additional space is called the objective space *Z*. For each solution **x** in the decision variable space, there exists a point in the objective space, denoted by $f(x) = z = (z_1, z_2, \dots, z_M)^T$. The mapping takes place between an *n* dimensional solution vector and an M-dimensional objective vector. Figure 6.1 illustrates these two spaces and a mapping between them.

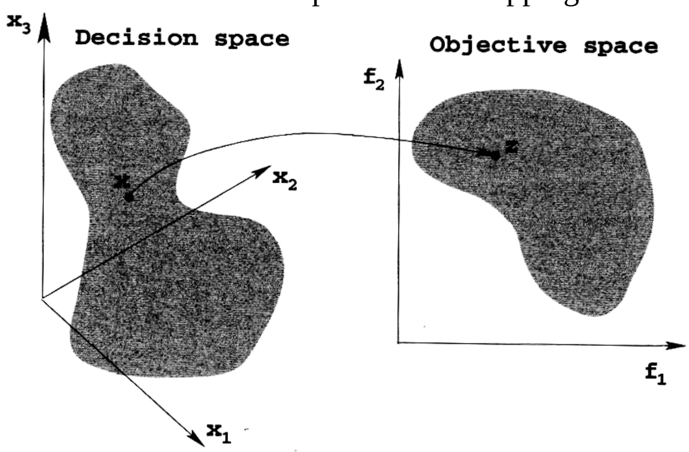


Figure 6.1.: Representation of the decision variable space and the corresponding objective space

Multi-objective optimization is sometimes referred to as vector optimization, because a vector of objectives, instead of a single objective, is optimized.

6.2. Principles of Multi-Objective Optimization

We illustrate the principles of multi-objective optimization through an airline routing problem. We all are familiar with the intermediate stopovers that most airlines force us to take, particularly when flying long distance. Airlines try different strategies to compromise on the number of intermediate stopovers and earn a large business mileage by introducing ‘direct’ flights. Let us take a look at a typical, albeit hypothetical, airline routing for some cities in the United States of America, as shown in Figure 6.2.

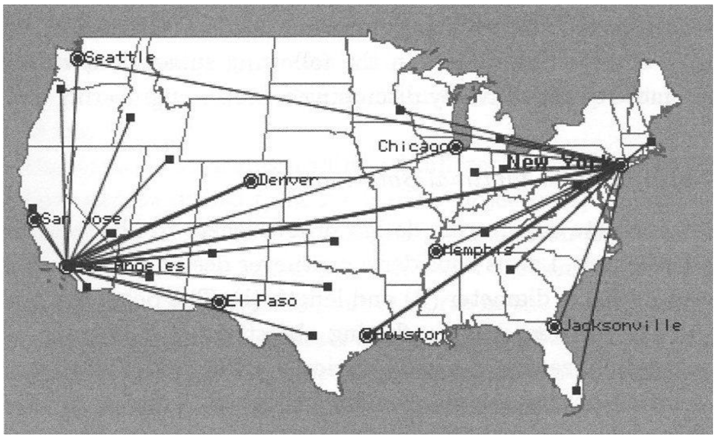


Figure 6.2.: A typical network of airline routes showing hub-like connections

If we look carefully, it is evident that there are two main ‘hubs’ (Los Angeles and New York) for this airline. If these two hubs are one’s cities of origin and destination, the traveler is lucky. This is because there are likely to be densely packed schedules of flights between these two cities. However, if one has to travel between some other cities, let us say between Denver and Houston, there is no direct flight. The passenger has to travel to one of these hubs and then take more flights from there to reach the destination. In the Denver-Houston case, one has to fly to Los Angeles, fly on to New York and then make the final lap to Houston.

To an airline, such modular networks of routes is easiest to maintain and coordinate. Better service facilities and ground staff need only be maintained at the hubs, instead of at all airports. Although one then travels longer distance than the actual geographical distance between the cities of origin and destination, this helps an airline to reduce the cost of its operation. Such a solution is ideal from the airline’s point of view, but not so convenient from the point of view of a passenger’s comfort. However, the situation is not that biased against the passenger’s point of view either. By reducing the cost of operation, the airline is probably providing a cheaper ticket. However, if comfort or convenience is the only consideration to a passenger, the latter would like to have a network of routes which would be entirely different to that shown in Figure 6.1.

A hypothetical routing is shown in Figure 6.2. In such a network, any two airports would be connected by a route, thereby allowing a direct flight between all airports. Since the operation cost for such a scenario will be exorbitantly high, the cost of flying with such a network would also be high.

Thus, we see a trade-off between two objectives in the above problem-cost versus convenience. A less-costly flight is likely to have more intermediate stopovers causing more inconvenience to a passenger, while a high-comfort flight is likely to have direct routes, thus causing an expensive ticket. The important matter is that between any two arbitrary cities in the first map (which resembles the routing of most airlines) there does

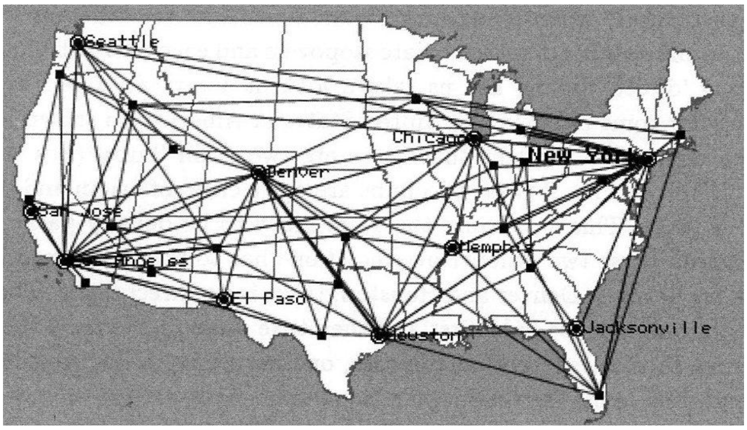


Figure 6.3.: A hypothetical (but convenient) airline routing

not exist a flight which is less costly as well as being largely convenient. If there was, that would have been the only solution to this problem and nobody would have complained about paying more or wasting time by using a 'hopping' flight. The above two solutions of a hub-like network of routes and a totally connected network of routes are two extreme solutions to this two-objective optimization problem. There exist many other compromised solutions which have lesser hub-like network of routes and more expensive flights than the solution shown in Figure 6.1.
Innovative airlines are constantly on the lookout for such compromises and in the process making the network of routes a bit less hub-like, so giving the passengers a bit more convenience. The 'bottom-line' of the above discussion is that when multiple conflicting objectives are important, there cannot be a single optimum solution which simultaneously optimizes all objectives. The resulting outcome is a set of optimal solutions with a varying degree of objective values. In the following subsection, we will make this qualitative idea more quantitative by discussing a simple engineering design problem.

6.3. Illustrating Pareto-Optimal Solutions

We take a more concrete engineering design problem here to illustrate the concept of Pareto-optimal solutions. Let us consider a cantilever design problem (Figure 6.4) with two decision variables, i.e. diameter (d) and l length

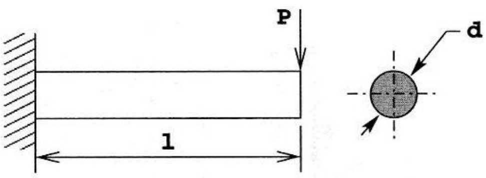


Figure 6.4.: A schematic of a cantilever beam.

The beam has to carry an end load P. Let us also consider two conflicting objectives of design, i.e. minimization of weight f_1 and minimization of end deflection f_2. The first objective will resort to an optimum solution having the smaller dimensions of d and l, so that the overall weight of the beam is minimum. Since the dimensions are small, the beam will not be adequately rigid and the end deflection of the beam will be large. On the other hand, if the beam is minimized for end deflection, the dimensions of the beam are expected to be large, thereby making the weight of the beam large. For our discussion, we consider two constraints: the developed maximum stress σ_{max} is less than the allowable strength σ_y and the end deflection δ is smaller than a specified limit σ_{max}. With all of the above considerations, the following two-objective optimization problem is formulated as follows:

$$\min f_1(d,l) = \rho\frac{\pi d^2}{4}l$$
$$\min f_2(d,l) = \delta = \frac{64Pl^3}{3E\pi d^4}$$

subject to

$$\sigma_{max} \le S_y$$
$$d \le d_{max},$$

where the maximum stress is calculated as follows:

$$\sigma_{max} = \frac{32Pl}{\pi d^3}$$

The following parameter values are used:

ρ = 7800 kg/m^3, P = 1kN, E = 210 GPa,
S_y = 300 MPa, δ_{max} = 5mm.

The left plot in Figure 6.5. marks the feasible decision variable space in the overall search space enclosed by $10 \le d \le 50$mm and $200 \le l \le 1000$mm. It is clear that all solutions in the rectangular decision space are feasible. Every feasible solution in this space can be mapped to a solution in the feasible objective space shown in the right plot. The correspondence of a point in the left figure with that in the right figure is also shown.

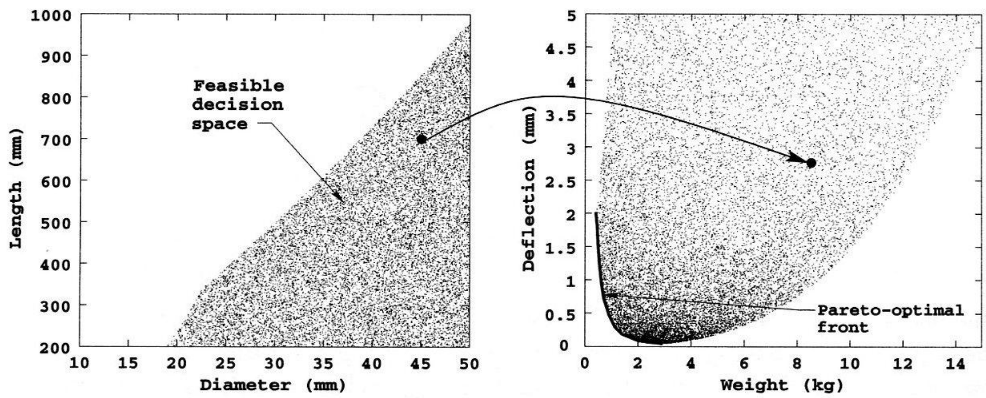


Figure 6.5.: The feasible decision variable space (left) and the feasible objective space (right).

This figure shows many solutions trading-off differently between the two objectives. Any two solutions can be picked from the feasible objective space and compared. For some pairs of solutions, it can be observed that one solution is better than the other objectives. For certain

other pairs, it can be observed that one solution is better than the other in one objective, but is worse in the second objective. In order to establish which solution(s) are optimal with respect to both objectives, let us hand-pick few solutions from the search space. Figure 6.6 is drawn with many such solutions, and five of these solutions (marked A to *E*) are presented in Table 1. Of these solutions, the minimum weight solution (*A*) has a diameter of 18.94 mm, while the minimum deflection solution (*D*) has a diameter of 50 mm. It is clear that solution *A* has a smaller weight, but has a larger end-deflection than solution *D*. Hence, none of these two solutions can be said to be better than the other with respect to both objectives. When this happens between two solutions, they are called *non-dominated solutions*. If both objectives are equally important, one cannot say, for sure, which of these two solutions is better with respect to both objectives. Two other similar solutions (*B* and *C*) are also shown in the figure and in the tables. Of these four solutions (*A* to *D*), any pair of solutions can be compared with respect to both objectives. Superiority of one over the other cannot be established with both objectives in mind. There exist many such solutions in the search space.
For clarity, these solutions are joined with a curve in the figure. All solutions lying on this curve are special in the context of multi-objective optimization and are called *Pareto-optimal solutions*. The curve formed by joining these solutions is known as a *Pareto-optimal front*. The same Pareto-optimal front is also marked on the right plot of Figure 6.5 by a continuous curve. It is interesting to observe that this front lies in the bottom-left corner of the search space for problems where all objectives are to be minimized.

Solution	d (mm)	l (mm)	Weight (kg)	Deflection (mm)
A	18.94	200.00	0.44	2.04
B	21.24	200.00	0.58	1.18
C	34.19	200.00	1.43	0.19
D	50.00	200.00	3.06	0.04
E	33.02	362.49	2.42	1.31

Table 6.1 Five solutions for the cantilever design problem.

It is important to note that the feasible objective space not only contains Pareto-optimal solutions, but also solutions that are not optimal. The entire feasible search space can be divided into two sets of solutions - a Pareto-optimal and a non-Pareto-optimal set. Consider solution *E* in Figure 6.6 and also in Table 1.

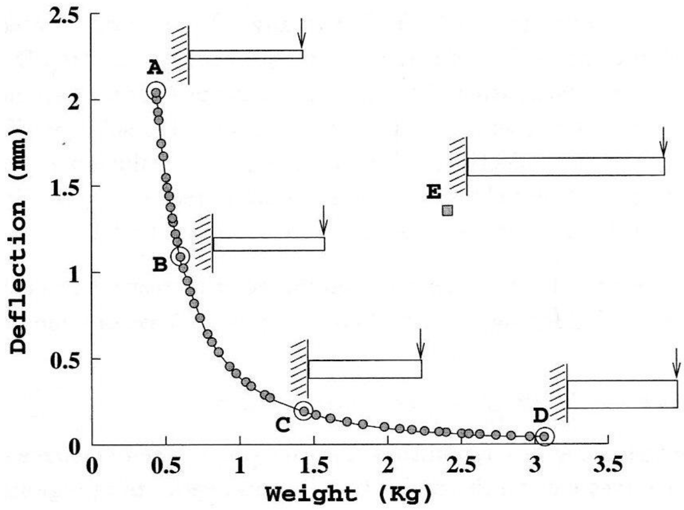


Figure 6.6.: Four Pareto-optimal solutions and one non-optimal solution.

By comparing this with solution *C*, we observe that the latter is better than solution *E* in both objectives. Since solution *E* has a larger weight and a larger end-deflection than solution *C*, the latter solution is clearly the better of the two. Thus, solution E is a sub-optimal solution and is of no interest to the user. When this happens in comparison of two solutions, solution *C* is said to *dominate* solution E or that solution E is *dominated* by solution *C*. There exist many such solutions in the search space which can be dominated by at least one solution from the Pareto-optimal set.

In other words, there exists at least one solution in the Pareto-optimal set, which will be better than any non-Pareto-optimal solution. It is clear from the above discussion that in multi-objective optimization the task is to find the Pareto-optimal solutions.

Instead of considering the entire search space for finding the Pareto- and non-Pareto- optimal sets, such a division based on domination can also be made for a finite set of solutions *P* chosen from the search space. Using a pair-wise comparison as above, one can divide the set *P* into two non-overlapping sets P_1 and P_2, such that P_1 contains all solutions that do not dominate each other and at least one solution in P_1 dominates any solution in P_2. The set P_1 is called the non-dominated set, while the set P_2 is called the *dominated* set. There is an interesting observation about dominated and non-dominated sets, which is worth mentioning here. Let us compare solutions *D* and *E*. Solution *D* is better in the second objective but is worse in the first objective compared to solution *E*. Thus, in the absence of solutions *A*, *B*, *C*, and any other non-dominated solution, we would be tempted to put solution *E* in the same group with solution *D*. However, the presence of solution *C* establishes the fact that solutions *C* and *D* are non-dominated with respect to each other, while solution *E* is a dominated solution. Thus, the non-dominated set must be collectively compared with any solution $\boldsymbol{x}$ for establishing whether the latter solution belongs to the non-dominated set or not. Specifically, the following two conditions must be true for a non-dominated set P_1:

1. Any two solutions of P_1 must be non-dominated with respect to each other.
2. Any solution not belonging to P_1 is dominated by at least one member of P_1.

6.4. Objectives in Multi-Objective Optimization

It is clear from the above discussion that, in principle, the search space in the context of multiple objectives can be divided into two non-overlapping regions, namely one which is optimal and one which is non-optimal. Although a two-objective problem is illustrated above, this is also true in problems with more than two objectives. In the case of conflicting objectives, usually the set of optimal solutions contains more than one solution. Figure 6.6 shows a number of such Pareto-optimal solutions denoted by circles.

In the presence of multiple Pareto-optimal solutions, it is difficult to prefer one solution over the other without any further information about the problem. If higher-level information is satisfactorily available, this can be used to make a biased search. However, in the absence of any such information, all Pareto-optimal solutions are equally important. Hence, in the light of the ideal approach, it is important to find as many Pareto-optimal solutions as possible in a problem. Thus, it can be conjectured that there are two goals in a multi-objective optimization:

1. To find a set of solutions as close as possible to the Pareto-optimal front.
2. To find a set of solutions as diverse as possible.

The first goal is mandatory in any optimization task. Converging to a set of solutions which are not close to the true optimal set of solutions is not desirable. It is only when solutions converge close to the true optimal solutions that one can be assured of their near-optimality properties. This goal of multi-objective optimization is common to the similar optimality goal in a single-objective optimization.
On the other hand, the second goal is entirely specific to multi-objective optimization. In addition to being converged close to the Pareto-optimal front, they must also be sparsely spaced in the Pareto-optimal region. Only with a diverse set of solutions, can we be assured of having a good set of trade-off solutions among objectives. Since MOEAs deal with two spaces - decision variable space and objective space 'diversity' among solutions can be defined in both of these spaces. For example, two solutions can be said to be diverse in the decision variable space if their Euclidean distance in the decision variable space is large. Similarly, two solutions are diverse in the objective space, if their Euclidean distance in the objective space is large. Although in most problems diversity in one space usually means diversity in the other space, this may not be so in all problems. In such complex and nonlinear problems, it is then the task to find a set of solutions having a good diversity in the desired space.

6.5. Non-Conflicting Objectives

It is worth pointing out that there exist multiple Pareto- optimal solutions in a problem only if the objectives are conflicting to each other. If the objectives are not conflicting to each other, the cardinality of the Pareto-optimal set is one. This means that the minimum solution corresponding to any objective function is the same. For example, in the context of the cantilever design problem, if one is interested in minimizing the end-deflection δ and minimizing the maximum developed stress in the beam, σ_{max}, the feasible objective space is different.

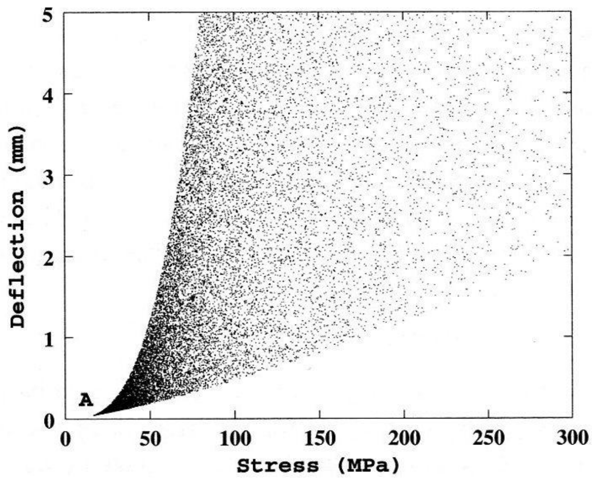


Figure 6.7.: End deflection and developed maximum stress are two non-conflicting objectives leading to one optimal solution (A).

Figure 6.7 shows that the Pareto-optimal front reduces to a single solution (solution *A* marked on the figure). A little thought will reveal that the minimum end-deflection happens for the most rigid beam with the largest possible diameter. Since this beam also corresponds to the smallest developed stress, this solution also corresponds to the minimum-stress solution. In certain problems, it may not be obvious that the objectives are not conflicting to each other. In such combinations of objectives, the resulting Pareto-optimal set will contain only one optimal solution.

6.6. References

[6.1] Deb, K. Optimization for Engineering Design: Algorithms and Examples. New Delhi: Pretience Hall.

[6.2] Reklaitis, G. V., Ravindran, A., and Ragsdell, K. M. (1983): Engineering Optimization-Methods and Applications. New York: Wiley.

[6.3] Rao, S. S. (1984): Optimization Theory and Applications. New Delhi:Wiley Eastern.

[6.4] Chankong, V., Haimes, Y. Y., Thadathil, J. and Zionts, S.: Multiple Criteria optimization: A state-of-the-art review. In Proceedings of the Sixth International Conference on Multiple-Criteria Decision Making, pp. 36-90

6.7. Questions

1. Explain the principles of multi-objective optimization!
2. What are the Pareto-optimal solutions?

7. OPTIMIZATION OF PRODUCTS AND MACHINE COMPONENTS

First we have introduced the basic element and formulations of an optimization problem and then discussed the solution techniques. In this chapter we are coming to the definition of the main problem classes and introducing the role and place of the different optimization techniques in the design process. Through industrial examples illustrates the power of the used methods solving mainly mechanical engineering problems.

The complex industrial optimization problems generally include behavior constraints, which can be evaluated only with numerical structural analysis techniques (Finite Element Method or Boundary Element Method). The studied problems in this chapter can be calculated using numerical techniques both for structural analysis and for optimization.

7.1. The place and role of the optimization in the design process

The design process may be divided into four stages:

1. **Formulation of functional requirements**, which is the first step in any design procedure. In some cases the functional requirements are not explicitly stated beforehand, and the designer has to investigate and take part in formulating these requirements. However, functional requirements are often established already before the engineer enters the design process.

2. The **conceptual design stage**, characterized by creativity, and engineering judgment of the designer, is a critical part of the design process. It deals with the overall planning of a system to serve its functional purposes. At this stage, the designer experiences the greatest challenges as well as chances of success or failure. Selection of the overall topology, type of structure, and materials are some of the decisions made by the designer at the conceptual design stage. In general, this part of the design process cannot be performed by a computer.

3. **Optimization**. Within a selected concept there may be many possible designs that satisfy the functional requirements, and a "trial-and-error" procedure may be employed to choose the optimal design. The computer is most suitable to carry out this part of using optimization methods to search for the optimal solutions. Thus, optimization in the present context is an automated design procedure giving the optimal values of certain design quantities, considering desired criteria and constraints.

4. **Detailing**. After completing the optimization stage, the results obtained must be checked and modified if necessary. In the final detailing stage, engineering judgment and experience are required, and it is again usually necessary for the designer to take part in the decision-making process.

Iterative procedures for the four stages are often required before the final solution is achieved, because the planning process in not a linear sequence. The portion of the structural design process that can be optimized automatically has been considerably increased in recent years. Optimization procedures are usually used to solve specific subproblems and the field of automated design is strongly connected with computer-aided design.

The optimization stage can also be divided into steps:

1. Formulating the optimization problem, choosing the design variables and design parameters.
2. Taking the assumptions.
3. Defining the goal function and the optimization constraints
4. Choosing the suitable optimization algorithm, and defining the convergence conditions.
5. Performing the calculation.
6. Comparing the results with the analytical ones (if exist), taking into the consideration the effect of the assumption.

The available methods of optimization may be subdivided into two categories:

Analytical methods are most suited for such fundamental studies of single structural components, but they are not able to handle larger structural systems. In analytic optimization problems the structural design is represented by a number of unknown functions and the goal is to find the form of these functions.

Numerical methods, which are usually employing a branch in the field of numerical mathematics called programming methods. The recent developments in this branch are closely related to the rapid growth in computing capacities affected by the development of computers. In the numerical methods, a near optimal design is automatically generated in an iterative manner. An initial guess is used as starting point for a systematic search for better designs. The search is terminated when certain criteria are satisfied; indicating that the current design is sufficiently close to the true optimum. Problems solved by numerical methods are called finite optimization problems. This is due to the fact that they can be formulated by a finite number of variables. The modern CAE systems inherit not only the 3D geometry modeler but numerical structural analysis and numerical optimization module too, they are suitable for supporting the numerical optimization techniques.

7.2. The main types of optimization tasks, incorporating them into the design process

The simplest task of the engineering optimization is sizing optimization, such as optimization of truss structures, where the topology and the material of the bars are considered as design parameters (Figure 7.1). The continuous design variables are the cross-sectional dimensions of the bars; they can change between his lower and upper limits because of the manufacturability.

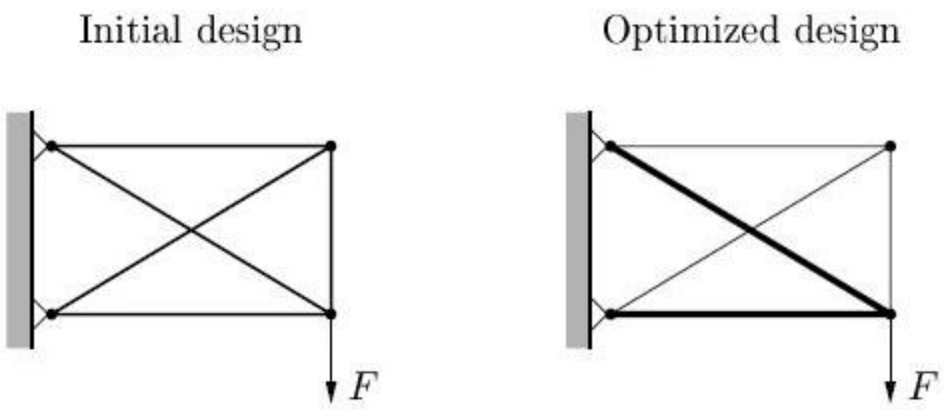


Figure 7.1. ábra: Sizing optimization of the 5 truss bar problem

.

In some cases, partly the surface of the machine part can be varied freely considering the production conditions. This is often happens by casted parts, when the changeable surfaces are not connected to other machine parts. In such cases shape optimization can be used advantageously. Considering a simple supported beam (Figure 7.2), so that the allowable stresses do not exceed the limit. Design variables can only change the bottom contour of the structure and all other properties are set to design parameter. The optimization result (the optimized contour) depends on the number of the design variables. In general, the more design variable means that we can get more information about the optimal shape, but it can lead to numerical instability, which eventually results wavy shape, which is completely useless.

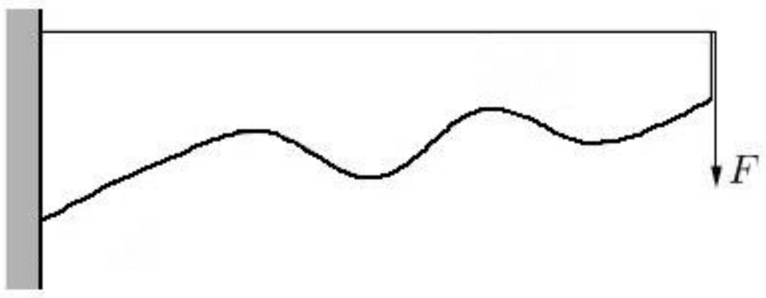


Figure 7.2.: Simply supported beam (starting geometry)

Finally, the most common formulation is the topology optimization. We can apply this technique for optimizing beams and for 2D and 3D solid geometries too. If we allow that the cross sections of the bars become zero, even more bars will disappear in the structure. An example is the design of pillars for high-voltage transmission lines (Fig. 7.3). The figure shows a ruined structure.

Figure 7.3.: Ruined pillar for high-voltage transmission lines

Solving 2D and 3D topology optimization problem first the discretized geometry (Finite element mesh) should be created and the elements separated into two sets: the design elements and the non-design elements. In the non-design region the densities are design parameters and their values either 0 (there is a hole in the structure) or 1 (there should be material). In the design region every element density is coupled with one design variable. After the topology optimization we have a material density for the design domain, which (generally) represents geometry with maximum stiffness (Figure 7.4.). More detailed information about the topology optimization will be presented later.

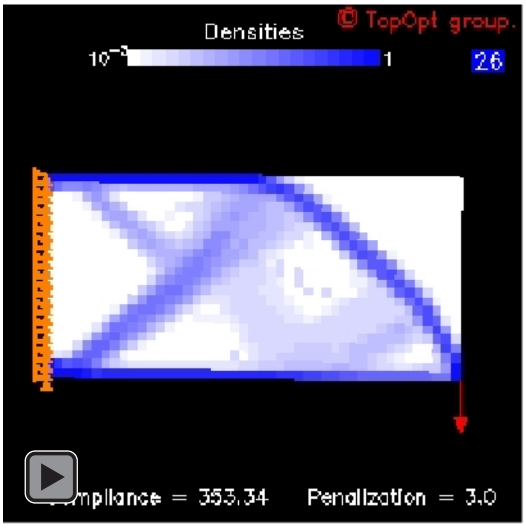


Figure 7.4.: Solving two dimensional topology optimization problems [www.topopt.dtu.dk]

7.3. The optimization examples in the fields of mechanical engineering

Nowdays in the product development process numerical simulation and optimization techniques are shifted into the earlier design stages, in order to coming the products to market earlier, reducing the development costs.
It is also worth to notice, that the number of the different modeles (for example for cars) is much higher than before. It is due to satisfy the changing consumer demands (Figure 7.5.). So the mass production was invented by Henry Ford is now obsolete.

Figure 7.5.: Number of the different models versus time [www.audi.de]

Applying the new CAE systems, with the integrated simulation tools, the designer have the possibility to check the functionality of large number of design ideas parallel. No complicated interfaces and no special simulation tools are needed. Also the number of very costly and time-consuming physical tests can be reduced. (Figure 7.6.).

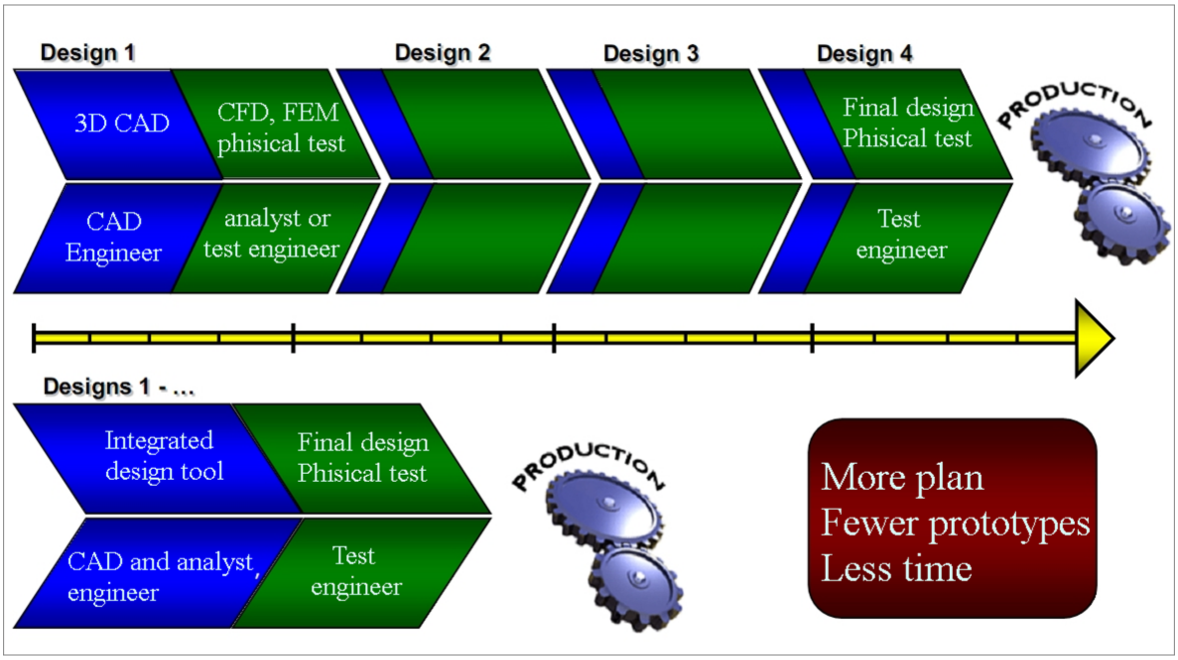


Figure 7.6.: The advantages of using the integrated CAE tools

This technique also allows the designer to use the built in automatic or semi-automatic optimization tools, which increases the probability of finding a better design In this way the structural optimization is linked deeply into the design process.

In the industrial praxis, first the topology optimization techniques are used in the early design stages (Figure 7.7. upper left: defining topology optimization problem showing the design and non-design regions. Figure 7.7. upper right: the result of the topology optimization step). Based on the results of the topology optimization a detailed geometry (CAD model) will be produced by the design engineer which undergoes a shape optimization procedure (Figure 7.7. lower left: the starting geometry for shape optimization, Figure 7.7. lower right: result after shape optimization).

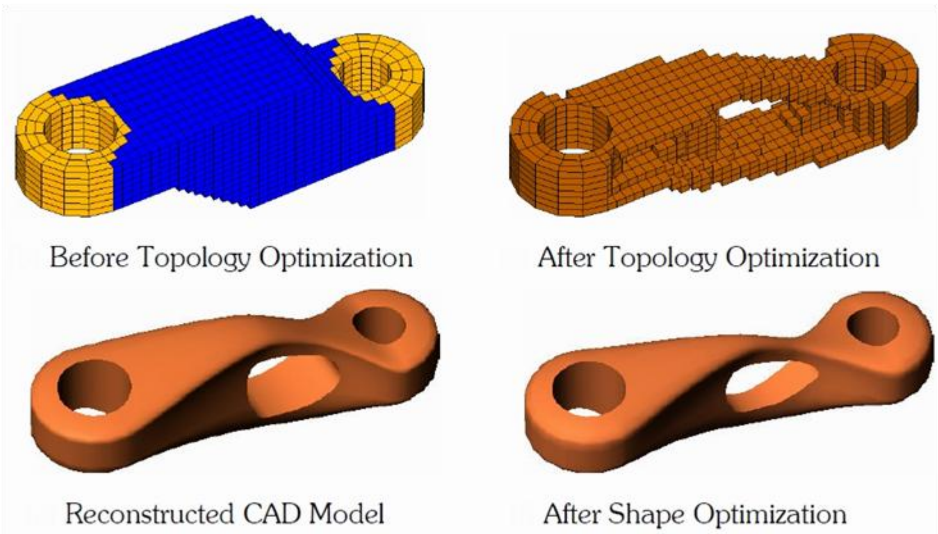


Figure 7.7.: Applying topology and shape optimization of a machine part

Similar optimization application was solved by FE-DESIGN Company using the Tosca system (Figure 7.8.). Using Tosca optimization system large numbers of problems have been solved in the field of automotive industry and also lot of other problems has been solved efficiently.

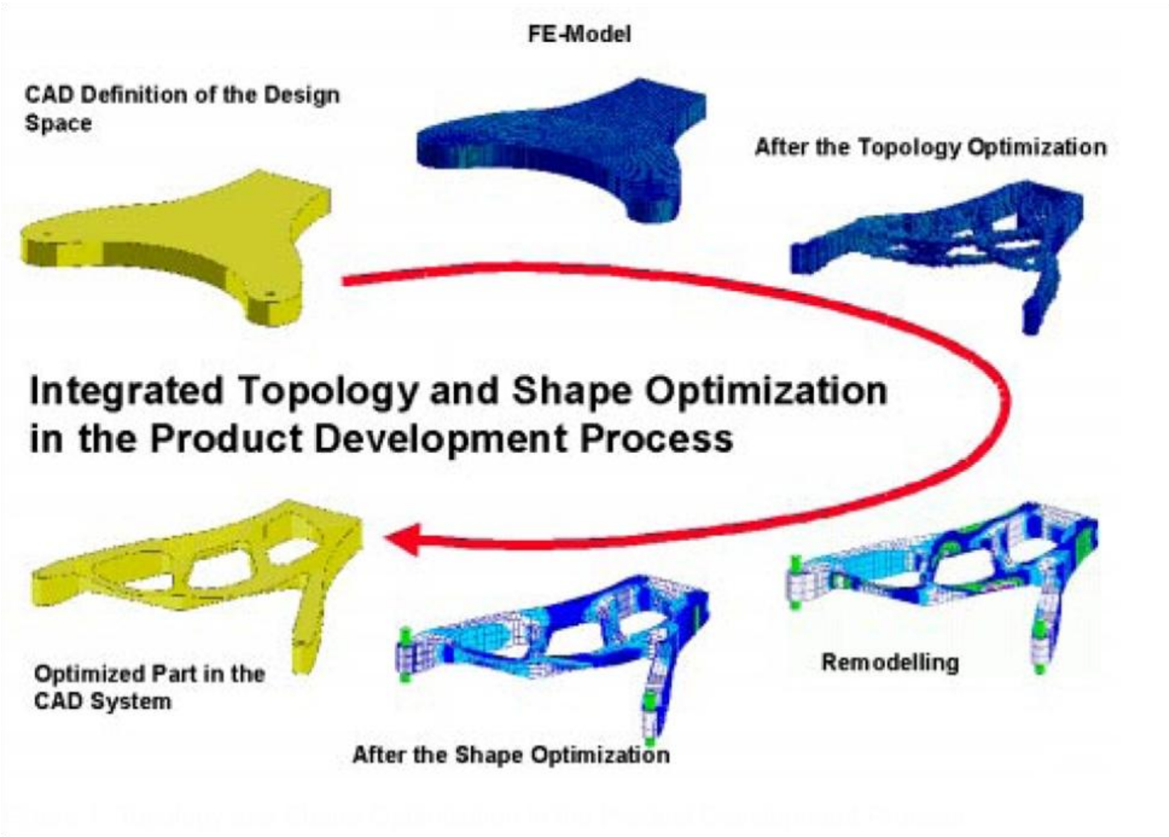


Figure 7.8.: Integration of topology and shape optimization in the product development process with Tosca [FE-DESIGN]

7.4. Questions

1. How can we define a sizing optimization problem?
2 What are the characteristic properties of a sizing optimization problem?
3. What are the characteristic properties of a shape optimization problem?
4. What are the characteristic properties of a topology optimization problem?
5. What is the most relevan design variable in case of solving 2D and 3D topology optimization problem?
6. Why it is important to minimize the “time to market” in the design process?
7. How can we integrate the topology and shape optimization in an industrial design process? What do you think, which approach should be applied in the earlier design stage?

8. GROUPING AND EVALUATION OF THE METHODS FOR SOLVING ENGINEERING OPTIMIZATION TASKS

Because of the diversity of the challenges in engineering, a universal solutions technique cannot be applied. It may be difficult for the reader to find the right optimization method for the problem to be solved. Therefore, in this chapter we will try to review and categorize the presented methods based on the classification presented in the previous chapters, but from the perspective of the tasks to be solved. In some cases, additional methods or examples will be presented too.

8.1. Optimization tasks containing only geometric conditions

We can find such optimization tasks during activities of engineering or product design professionals, where the optimization criteria can also be linked to purely geometric quantities. In this case it's not necessary to solve these computationally intensive real physical processes (for example the numerical solution of partial differential equations).
An optimization criteria is for example, optimizing the external shape of a bottle (or comparing various shape variants), such as the volume of the bottle has to be a certain given value, or between a given lower and upper limit. An equally important geometrical problem is the design of a packaging with minimum material need in case of a complex-shaped part, but with given three-dimensional geometry.
Otherwise, the approximate solution of a problem can be obtained in a way that the simulation is left out. The uniform placement of a car seat's heating element can be approximately examined without the solution of the transient thermal problem, and the solution of certain optical functions (optimization of a simple beam lamp, examination of an object's shadow) can also be approximated on a geometric basis.
The integrated CAE systems are suitable for solving these kind of tasks, where the optimization criteria can be defined by equations, such as the Pro/Engineer, Catia, or SolidWorks systems.
To illustrate the foregoing, we have created an interactive animation, which demonstrates the geometric optimization task through a rotating shaft weight balancing.

8.2. Solving optimization tasks using heuristic optimization procedure

In a design practice sometimes it is necessary to achieve satisfactory results, even in a relatively short period of time, and in the case of a complicated optimization task. The right result means a more favorable result than the initial construction, even if we cannot prove, that the resulting solution is global or local optimization. Such solving methods are called "quick and dirty" procedures. These procedures have a relatively small mathematical apparatus and a wide range of applications. Their application can handle the continuous and discrete design variable problems, and other such cases, when the objective function the optimization criteria is not continuous (for example the optimization on a discontinuous cost function.)
This group also includes a traditional method, also used to minimize volume, that selects the active from the optimization conditions (should be equal to the number of the design variables), and it results in an equation to determine the optimum value of the design variables. In case the active constraints can't be selected, it's possible that in the optimum some con-

straints will not be satisfied, so when this method is used, the optimal point always has to be placed back into the not selected optimization criteria. If, however, all conditions are met, you might want to evaluate the objective function as well, and if it's better than the best so far, then it can be considered an "optimal" solution. If we cannot examine the combination of all constraints, then it won't be sure, that our solution will be the optimum. The method also won't give the optimum, if in the optimum there are fewer active constraints than the number of design variables. The method can be used more favorably in cases of linear objective function and linear constraints, or to solve small tasks; in other cases, the mathematical programming method is recommended. This method is beneficial, if there are relatively few constraints, or the active can easily be selected. The geometric interpretation of the procedure in terms of design variables is the following: through the equations to be solved determine the hyper-plane intersections of the optimization criteria, some of these are on the edge of the allowable range (in which case we accept the solution if the objective function value is positive), others are out of range (in case we reject this point). The former statement is also based on this, that this procedure does not find the extreme if for example the optimum is inside the permissible range, or on the edge, but not in the intersection of the criteria. But in many cases even if the procedure does not find the optimum, it's able to provide a significantly better result than the initial case.

With the increase in the number of variables it is increasingly difficult to determine the global optimization. It's often beneficial using an intermediate step instead of a solution, so we can manage to reduce the complexity of the task and use the "**simple low limit** 'technique. Using this application we are able to estimate the size of each member in the objective function (typically this is used in case of a cost function task consisting of many members), and the ones whose role is small can be neglected. In this way one can create a lower (or upper) limit function, which extreme can be more easily determined than the objective function. The resulting optimum value substituted back into the original and alternative objective function, and the used approximation error can also be estimated. Through the analysis of the **monotonicity test** it's examined if the function respect of the variables is ascending or descending. During the monotonicity test of the objective function, the restrictive conditions also taken into account, the monotonicity of certain variables is often clearly identifiable. Thereafter, it's sufficient to take the general procedures for the remaining design variables, i.e. the dimension of the problem may be reduced.

Another simple optimum calculation approximation method is the definition of the **partial optimum.** In doing so, we can optimize according to the design variables, while we don't change the value of the other design variables. This way always a design variable task needs to be solved, which can be solved using any of the previously explained line side search procedures. As long as we've performed this task for all design variables, we'll get an approximation of the optimum, but if the calculation options allow, the procedure can be repeated in order to clarify the result.

Optimization procedures based on the basics of the probability calculations, for example the Monte-Carlo procedure (the simplest), which randomly selects a point in the *n* dimensional rectangular box located in the space of design variables, and the optimization conditions and the objective function must be evaluated here. The method is applicable to treat continuous and discrete design variables, continuous and discontinuous objective function, and optimization conditions. The method fairly easily can be connected to any type of structure analysis program, since the program is known as a „black box". Another advantage is, that the method with a given probability (which is nearing 100% by increasing the selected points) converges to the global optimum. A disadvantage is however, that the objective function and constraints must be evaluated in many points of the design variables space, therefore it has extensive computational requirements, and however it can be perfectly parallelized on multiprocessor or multicore machines. It should be used in tasks, where the numbers of design variables are small, and where the objective function and the constraints are simple analytic functions of the design variables. The method also applies other optimization procedures, for example in the case of the previously described genetic procedures, and the simulated annealing algorithm can also be considered as the enhancement of the Monte Carlo proceedings.

8.3. Optimality Criteria (OC) method for solving stress concentration problems

It's common to establish criteria in this method, which is met at the optimum level, and that's what we try to achieve. The OC methods are only suitable of solving specific tasks, but at the same time are easily programmable and have quick results. Particularly beneficial for these methods, that in most cases the calculation time doesn't increases with the growth of the design variables, and that it usually doesn't include computationally costly gradient calculations. One example of this is the so-called *Fully Stress Design*, which presumes, that you can reach the optimal solution, if in every element the stress reaches the limit stress value for at least one loading case. Other criteria can include displacement, stability, etc. conditions. The strategy for solving these cases consists of the repeated analysis of the structure, so that at the end of each analysis, the construction is changed by simple rules based on special physical properties. From a variety of OC methods we highlight one, which is probably the widest spread in the industry, and that is the Sauter algorithm, which is part of the TOSCA system.
The Sauter algorithm [1] is based on biological analogy: found out observing in a prevailing wind the growth of trees and branches, that the annual rings are not circularly structured, but that they become thicker where there is greater tension in the tree branch. This way the tree strengthens (adds material) the areas with higher stress, in order to prevent the breakage of tree branches.
This observation provides a basis for development of the geometry modification strategy, perfectly suitable for dismantling stress peaks bound to finite-element net, which modifies the geometry of the structure based on the following algorithm:

- The of node location's changing direction is always perpendicular to the surface, or a predefined direction;
- It's extent is determined by the difference of the limit stress and local stress (where the local stress is higher, material is added to the structure, where smaller, it's taken away):

$$\Delta X_n = S\left|\sigma_{red}^{csp} - \sigma_{ref}\right|^{\kappa} sign\left(\sigma_{red}^{csp} - \sigma_{ref}\right), \qquad (8.1)$$

where ΔX_n is the n-th finite element node location change in the required (surface normal or direct set) direction, S, κ constant parameters, σ_{red}^{csp} state of stress generated in a node characterized by reduced stress based on a stress hypothesis, σ_{ref} the required reference stress.

Based on stress generated in a given node, the rate of location change of the node in case of different S, κ parameters is shown by Figure 8.1.

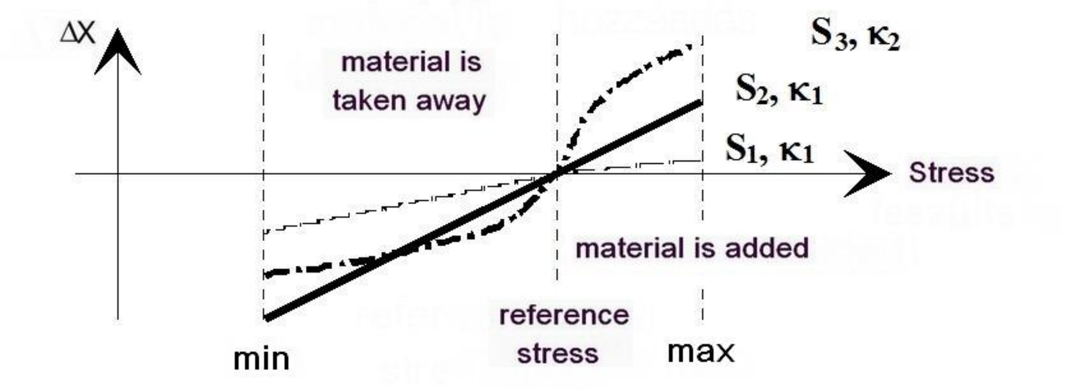


Figure 8.1.: Sauter geometry modification algorithm in case of different constants

When this algorithm is executed in each surface variable node of the examined object, then a new contour is generated. The displacement of the surface nodes leads to deformation of the finite element mesh, that why before any further structural analysis, the mesh needs to be adapted to the new shape, or re-generated.
The efficiency of this method is illustrated on a crank optimization task [8.1], where the test assembly is shown on Figure 8.2. During modeling the bolt pretension, the force from gas pressure, the mass force and the contact relationship between the components should also be considered.

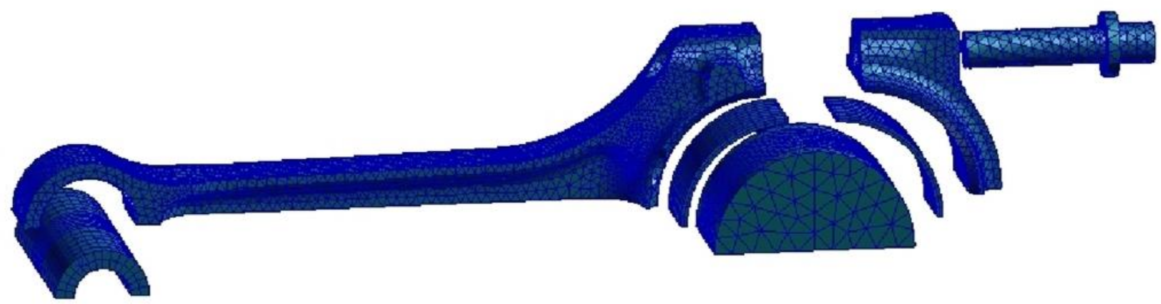


Figure 8.2.: The test rod

The Mises stress distribution on the initial geometry is shown (Figure 8.3.). After five iteration steps, the resulting stress distribution becomes significantly better, and it is maximum value drops to 83% of the original value (Figure 8.4.).

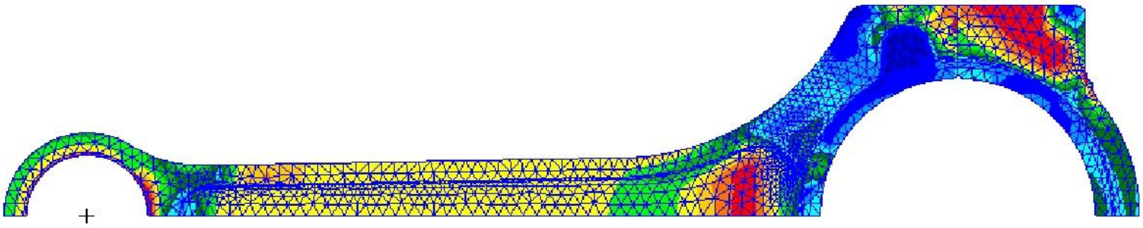


Figure 8.3.: Mises stress distribution in the initial geometry

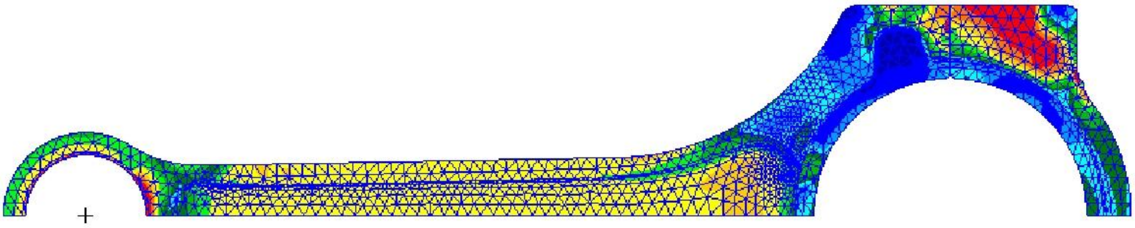


Figure 8.4.: Mises stress distribution on the optimized geometry

The achieved stress decreases during the intermediate steps of the iteration are shown on Figure 8.5.

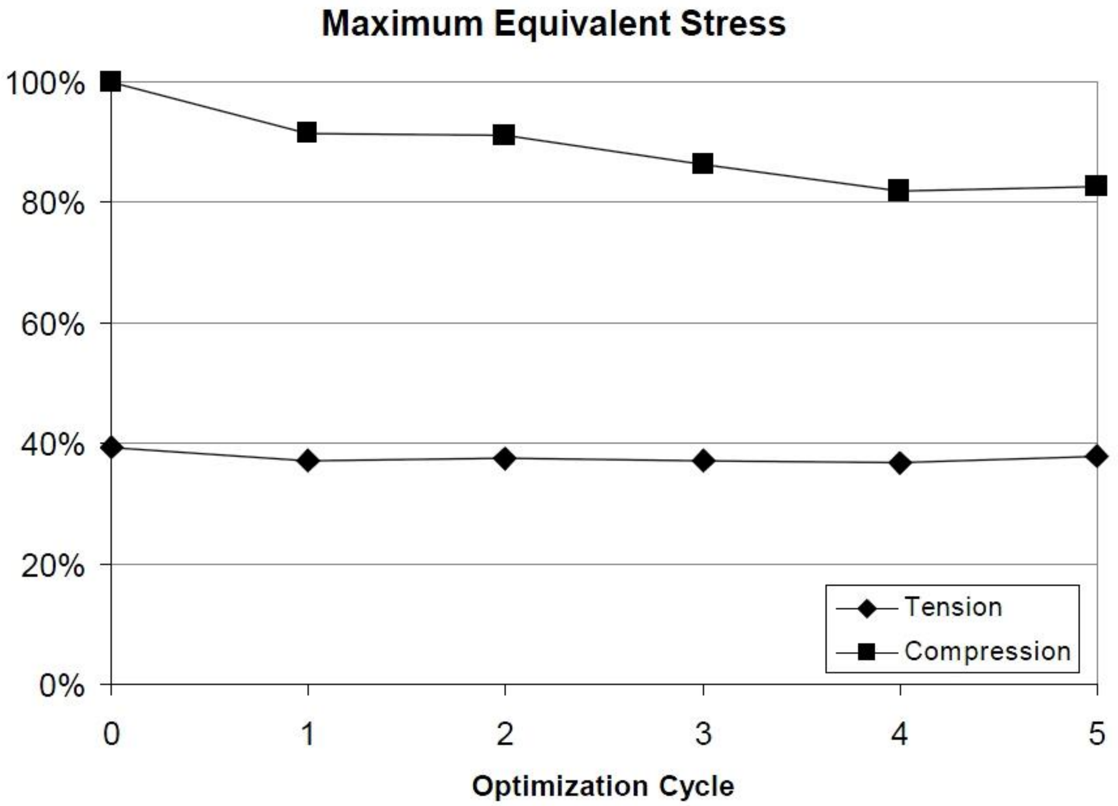


Figure 8.5.: Stress maximum in each optimization step

The advantage of this method is, that approximately during 6 structure analysis CPU time and during a significant peak stress reduction, a new stress distribution was achieved, which is also considerably more favorable in terms of fatigue. However a disadvantage is, that after the finite-element nodes are moved, at the completion of the optimization the changed geometry data is lost and only the surfaces formed by a finite element mesh (for example in STL format) can be exported into a CAD program.

8.4. Mathematical Programming Methods (MP)

The Mathematical Programming Methods (MP) can be widely used for continuous optimization problems. They are commonly used mathematical procedures for the determination of the location of extreme value of multivariable function, also taking into account the constraints (optimization criteria). The simplest such procedure is the Simplex method [2], which can be applied for linear objective function and linear restrictive conditions. Solving non-linear tasks many procedures have been developed, for example the Sequential Linear Programming (SLP), the Sequential Quadratic Programming Procedure (SQP), the Feasible Direction Method, etc. In order to determine the next iteration step, the mathematical programming methods usually require the partial derivatives based on the design variables of the objective function and restrictive conditions. Although capable of solving any problem, at the same time they require quite a large programming effort, and the computational needs are also significant. The MP methods usually guarantee a local extreme value. A great advantage of these procedures is, that they have a mathematically sound shut-down criteria, which is capable of measuring how far away from the optimum is the current iteration step. This convergence criterion is called the Kuhn-Tacker (or Karush-Kuhn-Tacker) criteria.

8.4.1. Lagrange function and the Kuhn-Tucker's criteria

For simplicity consider the optimization task based on the equality of ancillary:

$$Z = F(X) \to \min$$

$$h_j(X) = 0 \quad j = 1, \ldots, k \tag{8.2}$$

Let’s deifne the task’s so called Lagrange function:

$$\Phi(X, \lambda) = F + \sum_{j=1}^{k} \lambda_j h_j \, , \tag{8.3}$$

where λ_j is called the Lagrange's multiplier. Necessary condition of the extreme:

$$\begin{aligned} &\frac{\partial \Phi}{\partial X_i} = 0, \quad i = 1, \ldots, n \ \Rightarrow \ \nabla F = -\sum_{j=1}^{k} \lambda_j \nabla h_j \\ &\frac{\partial \Phi}{\partial \lambda_j} = 0, \quad j = 1, \ldots, k \ \Rightarrow \ h_j(X) = 0, \ j = 1, \ldots, k \end{aligned} \tag{8.4}$$

The meaning of the equations 8.5 can be explained that the gradient of the objective function in the optimum can be written as the linear combination of ancillary gradients.
The majority of optimization tasks can be rather formulated by inequality constraints (for example the displacement or stress to remain below a maximum value). Terms of inequality usually labeled with g_j can be reformulated with the introduction of auxiliary variables into equality terms, and similarly the equations for the optimum conditions can be written.
This corresponds to, that X can be a minimum point, as long as in the test point all $g_j \leq 0$ constraints are met, and there is $\lambda_j \geq 0$, so that

$$\nabla F + \sum_{j=1}^{J} \lambda_j \nabla g_j = 0, \tag{8.5}$$

where $j=1,\ldots,J$ means active constraints.
This is the so called Kuhn-Tucker criteria, which is a necessary condition for local minimum. It's also a sufficient criterion in case of convex programming problems.

8.5. Comparing the different methods

If a complicated task, within a short time has to be solved in a way, that it's enough to propose a better construct than the starting point, a heuristic procedure can be applied. In case of certain tasks, where the derivative is not available, but a gradient-free OC procedure can be well adapted to the task that should be strictly applied, since it's the most effective way. If the derivatives can be calculated, either analytically or semi-analytically or at least can be approximated by a finite differential, then the gradient methods are effective. In many cases because of the large computing time of the gradient methods, a pre-optimization is performed by an OC procedure (or other suitable technique), from its results the mathematical programming method for optimization can be started. By the application of the defined good starting point, the MP algorithm can find the optimum in a few iteration steps. If it's likely, that the found optimum is only local, than the calculation should be repeated starting from more (or randomly selected) starting points. In case of real, large-scale industrial tasks second-order methods using derivatives are rarely used, because although they are rapidly converging, but the calculation of the second-order derivatives is very CPU time intensive.

8.6. References

[8.1] Meske R., Mulfinger F., Warmuth O.: Topology and Shape Optimization of Components and Systems with Contact Boundary Conditions NAFEMS Seminar: Modelling of Assemblies and Joints for FE Analyses April 24 - 25, 2002 Wiesbaden, Germany

[8.2] Haftka, R.T., Kamat, M.P.: *Elements of structural optimization.* Martinus Nijhoff Publishers, 1985

8.7. Questions

1. Give an example for geometric optimization task.
2. What are the characteristics of "quick and dirty" procedures? Give an example of a procedure, that was divided into this group.
3. What are the characteristic of the optimum criterion procedures?

4. Give an example of an optimum criterion procedure!
5. In case of what tasks can the optimum criteria methods be used advantageously?
6. What are the characteristics of the mathematical programming procedures?
7. What is the Kuhn-Tucker criterion used for?

9. THE ROLE AND METHOD OF SENSITIVITY ANALYSIS; CASE STUDY

Sensitivity analysis has three fundamental applications in solving optimization problems in mechanical engineering:

- Certain optimization methods (e.g. gradient methods) mentioned earlier utilize gradient vectors for defining the optimal direction of the iteration
- In contour optimization an import question is what to choose as a design variable. If we choose too many dimensions, then the task will be unmanageable, on the other hand, if too few dimensions are selected then we only stand a chance to find a partial optimum. In the case study presented at the end of the subchapter we demonstrate how local sensitivity analysis can help to decide if the impact of certain design variables rather designates them to be used as design parameters.
- Sensitivity values also give useful guidance if optimization is performed manually, or if a heuristic approach is applied for geometry selection.

Sensitivity calculation is considered as the definition/obtainment of the objective function and its partial derivative functions along its optimality design variables. There is a multitude of methods to obtain these. If the above functions can be given in a closed analytic form, then they can be calculated either manually or using a mathematical programme package (Derive, Wolfram Mathematica, etc.) and evaluated using the actual values of the design variables. This is nevertheless rare in mechanical engineering applications. If these functions are not available in a closed form and if we can modify the source code of the software used for structural analysis then in most of the cases a semi-analytic or numerical calculus is recommended. Basically, as we move from analytic functions to numerical calculus we get less and less precise solutions (resulting in more iteration steps when for example searching for optimal direction), but in these cases we can obtain a general method which can be simply programmed as an external black box to be used for any structural analysis code.
In the case of finite differences based sensitivity analysis first the structural analysis has to be performed for the given $(x_1,x_2,...,x_n)$ design space coordinates. This analysis can be for example a FEM analysis of linear statics, Eigenfrequency analysis, non-linear statics, thermal dissipation or hydrodynamic analysis or the combination of these. In the next step a sufficiently small Δx value has to be chosen with to which we perturbate all of the n coordinates of the design variable in order to get n new design variables to run the structural analysis on. Before using such a technique it is recommended to assure that the design variables are in a similar order of magnitude.
If these results are ready, then the following relationships can define sensitivities of the objective function and the optimality conditions:

$$\begin{aligned}
\frac{\partial F(x_1,x_2,...,x_n)}{\partial x_i} &\approx \frac{\Delta F(x_1,x_2,...,x_n)}{\Delta x_i} = \frac{F(x_1,x_2,...,(x_i+\Delta x),...,x_n) - F(x_1,x_2,...,x_i,x_n)}{\Delta x_i} \\
\frac{\partial g(x_1,x_2,...,x_n)}{\partial x_i} &\approx \frac{\Delta g(x_1,x_2,...,x_n)}{\Delta x_i} = \frac{g(x_1,x_2,...,(x_i+\Delta x),...,x_n) - g(x_1,x_2,...,x_i,x_n)}{\Delta x_i}
\end{aligned} \tag{9.1}$$

We note here that this means that for any analyzed point of location *n+1* structural analyses have to be undertaken, which can be rather time consuming. Nonetheless, we might be aware that industrial optimization tasks are usually solved by super-computing tools. Quite advantageously, developments in microchips resulted in today's state-of-the-art CPU-s being multi-core, multi-thread units, providing a good capability for symmetric, parallel task solutions on an average PC's processor. A further option is to set up a cluster of such PCs which can provide a quite strong means of solving optimization problems. Thus, numerical sensitivity analysis (being the most time consuming task of the optimization process) can be performed efficiently. Then, the *n+1* structural analyses is fully independent and can be solved in the run-time of one structural analysis (assuming the availability of the necessary number of CPU cores).

9.1. The direct and adjoint techniques of sensitivity analysis

Two methods of numerical calculation based sensitivity analysis are common: direct computing or computing with the application of adjoint variables. Solution of the equation is based on the finite element technique (linear flexibility statical calculation) and as the fundamental structural response, the optimality condition for displacement is considered. Nevertheless, the procedure presented here can be generalized for other tasks, which can be traced back/originated in the solution of system of linear equations.

9.1.1. Direct sensitivity calculation

The simulation of a linear static finite element is equivalent to the solution of the following system of linear equations, where **K** is the stiffness matrix of the system (a quadratic symmetrical matrix, its order equaling the degree of freedom of the examined problem – for further details see Finite Element Analysis subject), **u** is the displacement vector of the nodes and **f** is the right-side vector computable from stresses:

$$\mathbf{K}\mathbf{u} = \mathbf{f} \tag{9.2}$$

In direct sensitivity calculation this base equation is derived by the design variables (**X**)

$$\frac{\partial \mathbf{K}}{\partial x}\mathbf{u} + \mathbf{K}\frac{\partial \mathbf{u}}{\partial x} = \frac{\partial \mathbf{f}}{\partial x} \ , \tag{9.3}$$

which after reordering gives the following relationship for displacement sensitivities:

$$\mathbf{K}\frac{\partial \mathbf{u}}{\partial x} = \frac{\partial \mathbf{f}}{\partial x} - \frac{\partial \mathbf{K}}{\partial x}\mathbf{u}\,. \tag{9.4}$$

Note, that this is a similar equation to the one which was solved in the structural analysis problem, only the right side is different. This equation has to be solved for all design variables, which might result in a long computation time in case of multiple design variables.

9.1.2. Adjoint method

The optimality condition can be simply identified from the fundamental result of the structural analysis (in practice **q** is frequently a very simple vector, e.g. almost all but one elements equals zero, non-zero on that node or element where we are interested in the dislocation or stress-sensitivity, and in most cases is independent from the design variables, therefore its partial derivatives with respect to the design variables is zero)

$$g = \mathbf{q}^T \mathbf{u} \tag{9.5}$$

Thus, its sensitivity can be easily concluded from differentiating the above equation:

$$\frac{\partial g}{\partial x} = \frac{\partial \mathbf{q}^T}{\partial x}\mathbf{u} + \mathbf{q}^T \frac{\partial \mathbf{u}}{\partial x} = \frac{\partial \mathbf{q}^T}{\partial x}\mathbf{u} + \mathbf{q}^T \mathbf{K}^{-1}\left[\frac{\partial \mathbf{f}}{\partial x} - \frac{\partial \mathbf{K}}{\partial x}\mathbf{u}\right] \tag{9.6}$$

With the calculated displacement sensitivities we can quickly define sensitivities of even large number of optimality conditions

If we have a large number of design variables it is recommended to introduce adjoint variables (**a**) according to the following equation:

$$\mathbf{Ka} = \mathbf{q}\,, \tag{9.7}$$

then after a short deduction the following relationship is gained for defining the sensitivity of an optimality condition using adjoint variables:

$$\frac{\partial g}{\partial x} = \frac{\partial \mathbf{q}^T}{\partial x}\mathbf{u} + \mathbf{q}^T \frac{\partial \mathbf{u}}{\partial x} = \frac{\partial \mathbf{q}^T}{\partial x}\mathbf{u} + \mathbf{a}^T \mathbf{K}\frac{\partial \mathbf{u}}{\partial x} = \frac{\partial \mathbf{q}^T}{\partial x}\mathbf{u} + \mathbf{a}^T\left[\frac{\partial \mathbf{f}}{\partial x} - \frac{\partial \mathbf{K}}{\partial x}\mathbf{u}\right]. \tag{9.8}$$

Contrary to the direct method, the number of equation systems to be solved equal the number of optimality conditions, and after solving a system of equations the sensitivity of the optimality conditions for all design variables can be obtained using the above equation.

Comparing the two methods, it can be stated that for size and contour optimization problems the direct method is preferred, as the number of design variables is relatively small (e.g. in the range of 5 to 50), but the number of optimality conditions can be large. If we consider a problem of 10 cases of stress with 100,000 elements, where for all elements of the structure a tension criterion is given, then this results in 1,000,000 optimality conditions. If the problem can be formulated with a small number of conditions, then the method of adjoint variables can also be used.

When a topological optimization problem has to be solved with 1-3 design variables per element, thus resulting in 10,000-10,000,000 design variables for the optimization then almost exclusively the adjoint method is used.

9.2. Sensitivity analysis of crank arm– main steps

In this subchapter a component's sensitivity analysis is performed in order to identify which size variable of the potential design variables have the largest impact on the objective function and on the optimality condition. For the analysis any parametric 3D-design framework is capable, which has a built in linear flexibility finite element module (most of the mid-range and high-end CAE systems are such, this being the reason of the selection of this task for demonstration). The following problem is presented by using SolidWorks.

Step 1.: 3D parametric geometry development of the component or composition

Here we don't go through the geometrical design steps as it is relatively simple to develop the examined geometry, and the actual steps depend on the used CAD system (and the background is supported by another educational material). It is important to note nevertheless that

design variables (at least in the first approach) can be selected from the CAD geometry parameters, therefore the development of the geometry (the form feature tree) is very important. The very same physical body can be created with different form feature trees (different order of features), and not all of these are applicable for sensitivity analyses and optimization. There is no generally valid principle for choosing the most proper feature tree, but maintaining a few rules and having adequate experience can aid the choice (unfortunately it is not true that the feature tree used in manufacturing is eligible here, but a tree containing relatively few features can be favorable). It is recommended to consider in advance (at least roughly) which design variables to use and give the initial sizes explicitly. When building the form features the drafts should be fully determined (some frameworks allow for under-determined drafts) since alternations can be better handled this way.
The sensitivity analysis is demonstrated on the example of a crank arm (Figure 9.1 left side), which is a component of sub-assembly (Figure 9.1 right side).

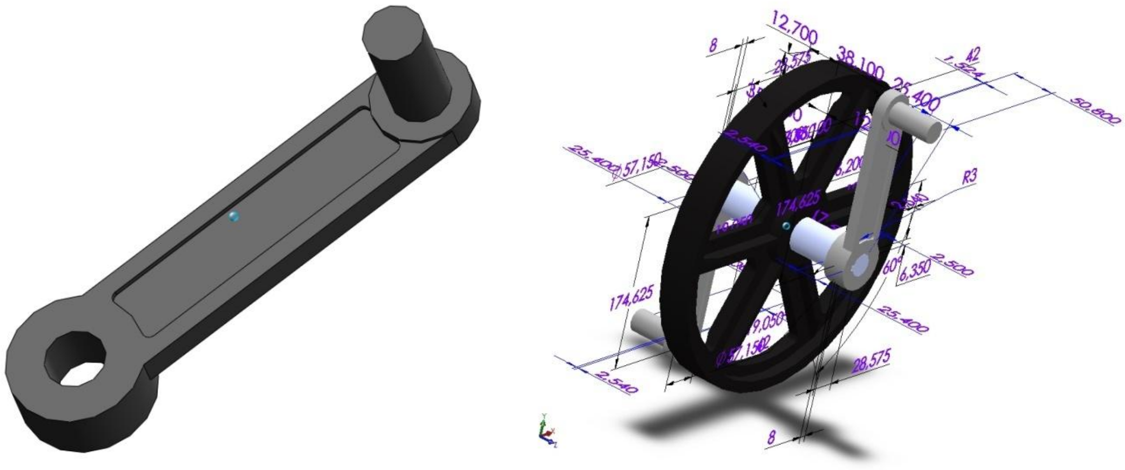


Figure 9.1.: Crank arm and its installation environment

The optimization problem is volume minimization under displacement and Mises-yield criterion optimality conditions. Design parameters are the material characteristics (usual steel characteristics are enumerated), the connector sizes, but beside these still many other sizes can be found on the component (Figure 9.2).

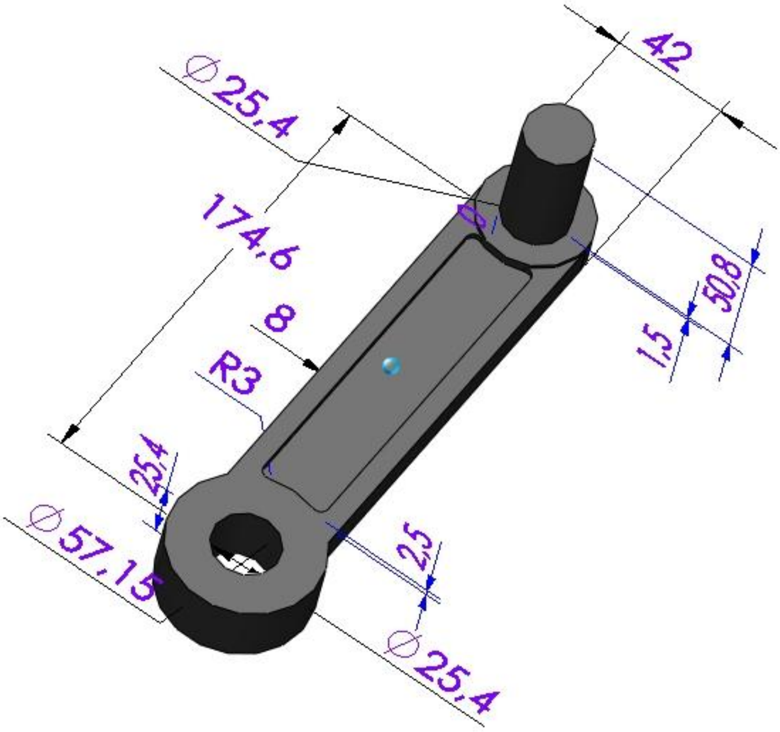


Figure 9.2.: Crank arm and driveshaft size grid

Step 2. Development of a structural analysis model
In the present problem, a linear flexibility statical analysis is performed considering small displacements. The crank arm and the driveshaft are considered as a unit. In our model the arm is fixed to the shaft's axle by rigid coupling, on the driveshaft we assume 1000N of uniformly distributed stress perpendicular to a given plane (Figure 9.3.)

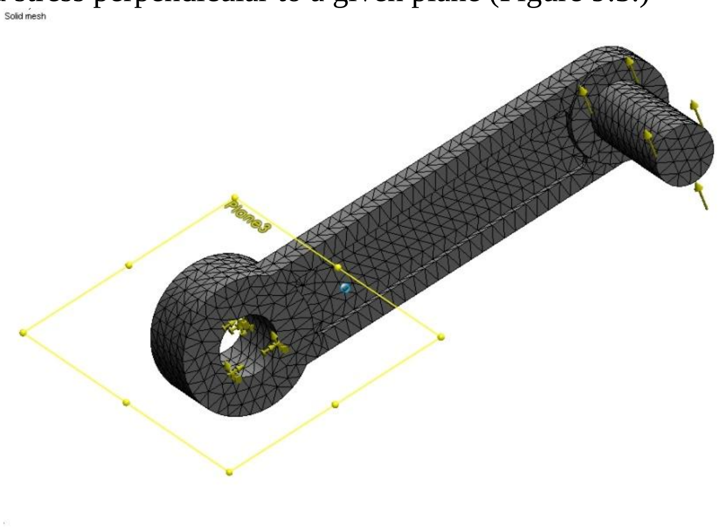


Figure 9.3.: Boundary condition applied at the connecting rod's structural analysis

For the finite element grid an automatic grid generation technique was used utilising second order tetrahedron elements, which is also recommended for evaluation of stress type conditions (if only displacement-type optimality conditions are given, then linear elements can be considered).
The evaluation of the objective function can be realised by a CAD program (in this case 151,88 g), however, for the evaluation of the tension optimality conditions a finite element calculation is necessary, the results of which are shown on Figure 9.4. It is visible that the maximum of tension is at the stem of the hub (110 MPa). In case of a tension optimization condition special attention has to be paid that the necessary level of accuracy in stress calculation is maintained, since the calculation is based on these values. In such cases a convergence analysis should be performed by further refining the finite element grid. In this case it is sufficient to perform a local grid refining in the high tension zone. It can be seen also in this case that the maximal von Mises tensile stress arising in the component grew up to 136 MPa and on the surfaces containing the critical rounding the maximal value grew to 123 MPa. It can be therefore stated that the precise structural analysis model's development is inevitable when preparing for a sensitivity analysis.

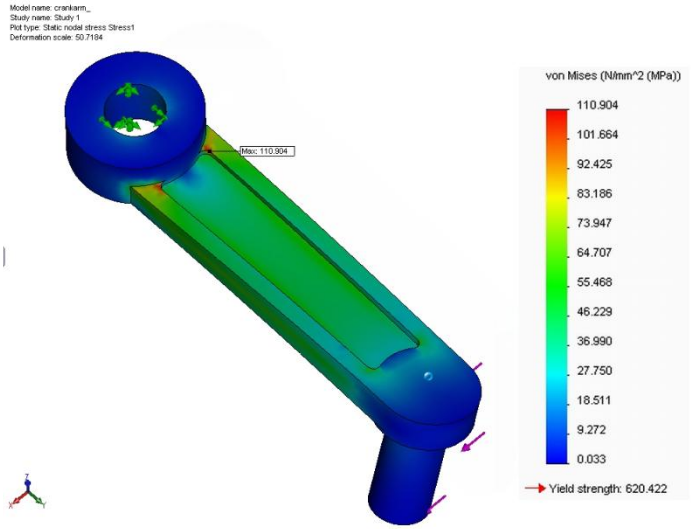


Figure 9.4.: Tensions arising in the crank arm– initial model

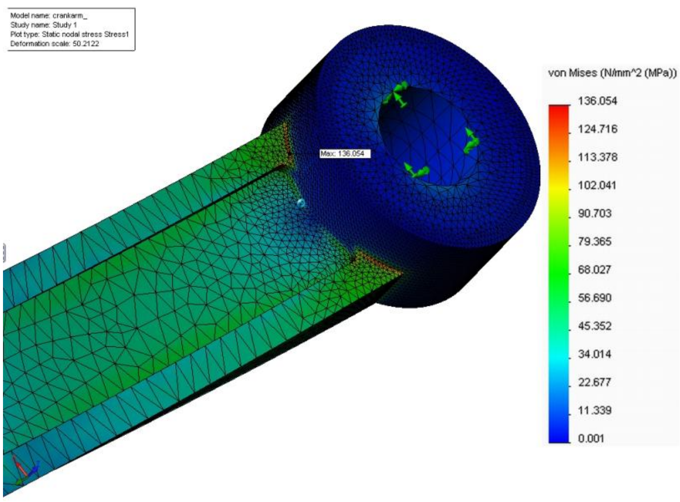


Figure 9.5.: Tensions arising in the crank arm – refined model

Step 3. Selection of design variables
Based on the results the designer formulates an idea which sizes to change on the component in order to save on material requirements while maintaining the required carrying or stress capacity. To avoid too much calculations 4 size variables were defined (Figure 9.6.): width of connecting rod (initial size: 42mm), radius of rounding at alleviation (initial size 2.5mm), considered to be identical on both sides, and the stub size remaining at the edge of alleviation (initial size: 8mm).

Table 9.1: The design variables and their initial values

No. of design variable	Varying size	Initial value[mm]
DV 1.	Width of connecting rod	42
DV 2.	Radius of rounding at alleviation	R3
DV 3.	Depth of alleviations	2,5
DV 4.	Stub size remaining at the edge of alleviation	8

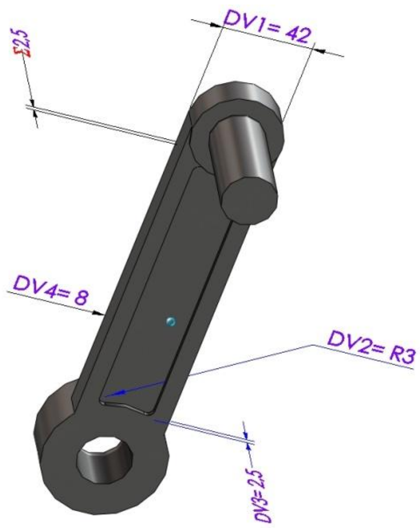


Figure 9.6.: The selected design variables

The sensitivity analysis of the objective function and the optimality criterions (maximal Mises stress on the surface containing the critical rounding) was performed for all four design variables with the presented finite difference technique.
Before turning our attention to the actual problem's results, let's consider some general remarks helping us in the solution.
It is always advisable to assess if the achieved sensitivity results are of usable accuracy, otherwise no decisions can be based on these. Most leading CAE systems, directly or indirectly have such a tool, which help defining sensitivity values. For example, in Pro/Engineer there is a direct option for defining and displaying sensitivity values on a 2D-chart. On the other hand, in SolidWorks a so called „design study" can be defined to calculate sensitivity the easiest way, design variables can be easily given, the objective function and the optimality criterion is done in the geometrical model as sensor.
In solving the task the latter software was used to demonstrate the impact of the respective Δx perturbation values (a separate design study was defined for all perturbation values).

Results can be simply transferred to Excel, where finite differences were evaluated and the sensitivity values were displayed.
It is especially recommended to pay attention in the sensitivity calculation of stress type optimization criterion. This on one hand is the most sensitive to calculation error; on the other hand the maximal tensile stress arising in the structure can sometimes give misleading optimal results. This can be seen in our example (intentionally) where the maximum tension is shown in the environment of an existing, but non-modeled rounding. The value of this „maximum" does not give reference on the structure there it is not recommended to use in the optimization. In these cases rather choose a surface where the maximum tension is a characterizing information (in mainstream CAE systems it is allowed to deviate from only choosing tension arising in the whole structure). It is more favorable if we want to limit the maximum displacement, its accuracy can more easily be guaranteed. An even more favorable situation if we evaluate a condition or an objective function coupled with some geometrical parameter (for example, in the case of volume as objective function no structural analysis is necessary), in this case the sensitivity values can be gained with high precision and independently from the applied discretization.
After these remarks, let's return to the presentation of the problem and the conclusions. For determining the sensitivity values not the direct size values of the geometrical variable were considered as design variables but their values divided with their base value, thus normalising the variables. This technique is especially preferential when there are differences in magnitudes between the variables. In determining sensitivities the value of Δx was chosen to be 10%, 1% and 0.5% respectively.

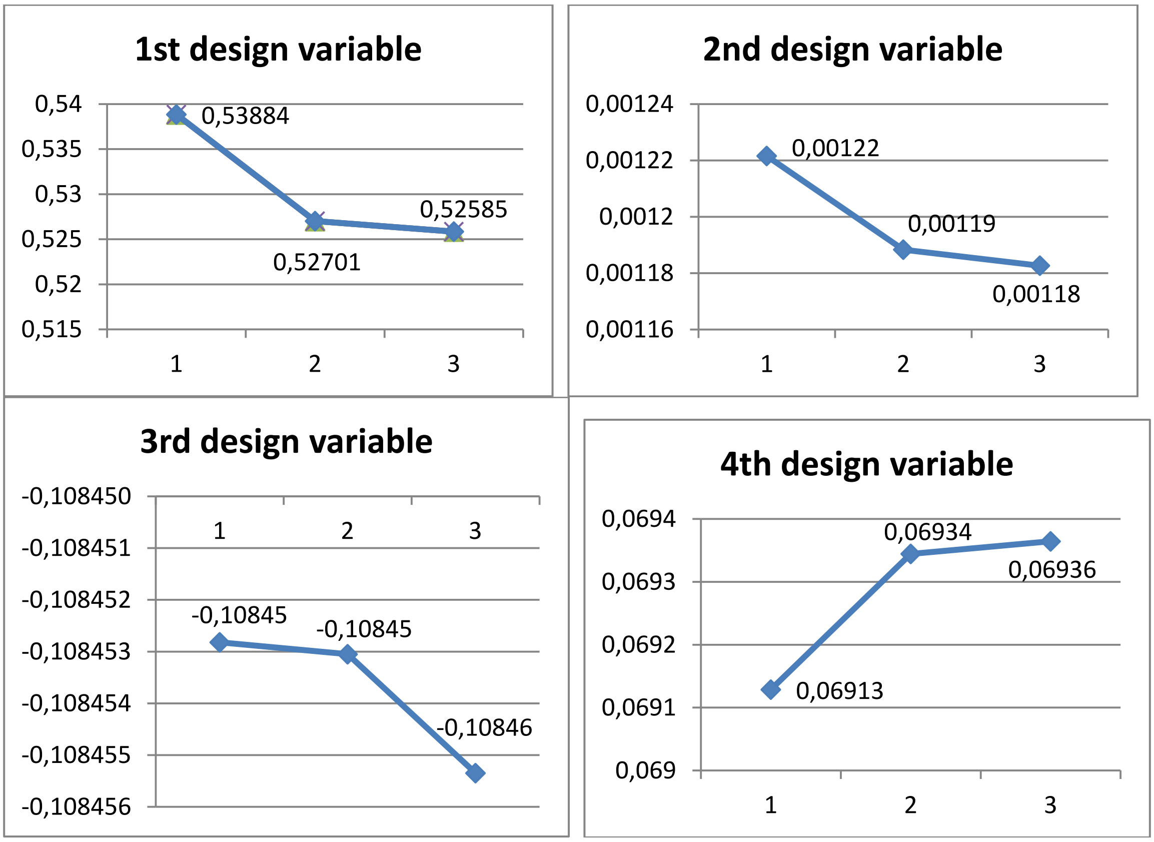


Figure 9.7.: Determination of the sensitivity of the objective function

The determination of the sensitivity of the objective function is relatively easy, and it can be seen (Figure 9.7), that its accuracy hardly depends on the chosen Δx value. The largest impact on volume comes from the first design variable, the width of the crank arm, the smallest impact from the second design variable (radius of rounding). If for example a design variable is chosen which is not connected to the geometry (e.g. the material's modulus of elasticity), then the objective function's sensitivity will be zero.
The reliable determination of sensitivity for stress type optimality conditions is a rather more complex problem. Using an automatic grid generator does not guarantee the identical location of nodes used for the finite element calculation, neither that the structure of the grid in the locality of the analysis remains the same, when we slightly change the geometry and re-grid the whole body. Thus, although the first and third design variable allowed for an easy convergence, but in the second and fourth variable's case an additional smaller (0.1%) point was necessary (Figure 9.8.Figure 9.8.).

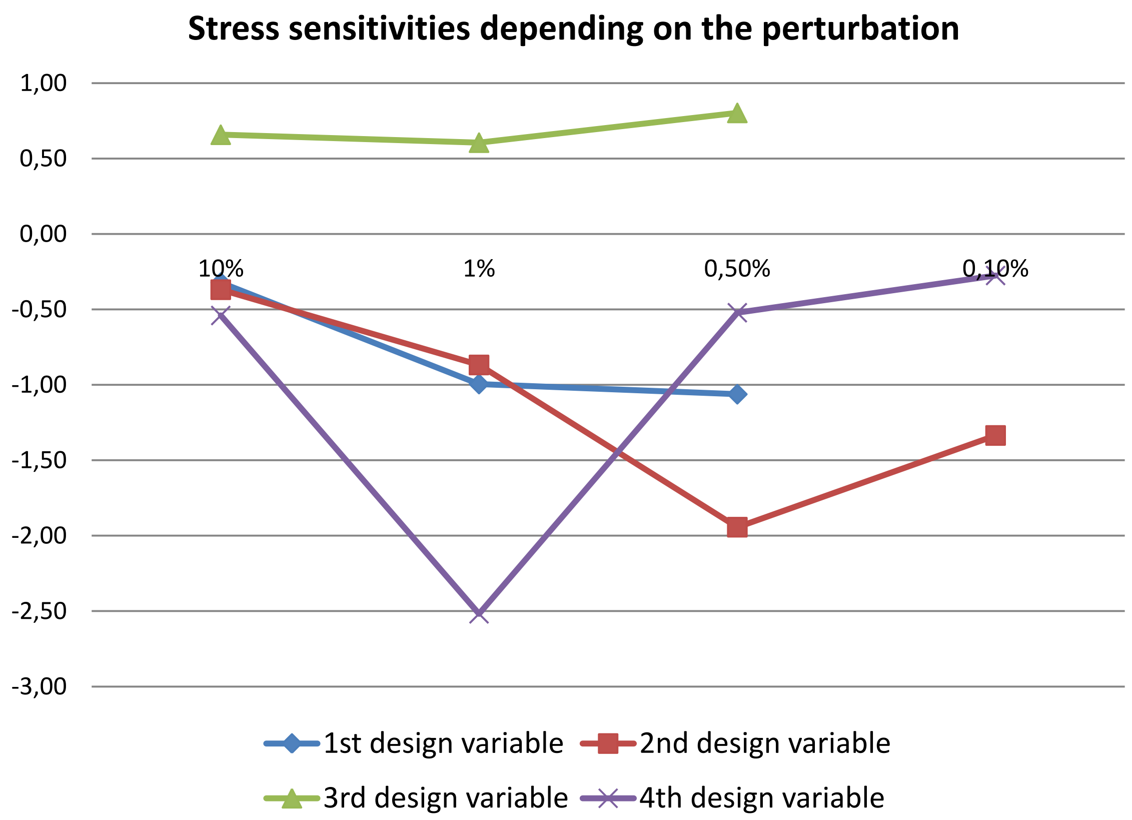


Figure 9.8.: Determination of stress sensitivities

These results, compared with the objective function's sensitivity allow us to draw the conclusion that although the second design variable has the least impact on the objective function but it still has the most significant role in determining maximum tensile stress.
It can be concluded furthermore that changes in all design variables have an impact either on the objective function or on the stress optimality criterion, but if the number of variables would still had to be reduced, then the fourth variable could be omitted, as its sensitivity values are relatively small.

From the engineer's aspect the maximal deformation of a structure can be important, therefore the results for maximal displacement sensitivity calculations are presented (Figure 9.9.) for the sake of completeness. It is well visible from the chart that displacement sensitivity converges more easily than that of tensile stress, thus its determination is easier. According to our expectations, changing the radius of rounding as the second design variable does not have significant impact on maximal displacement.

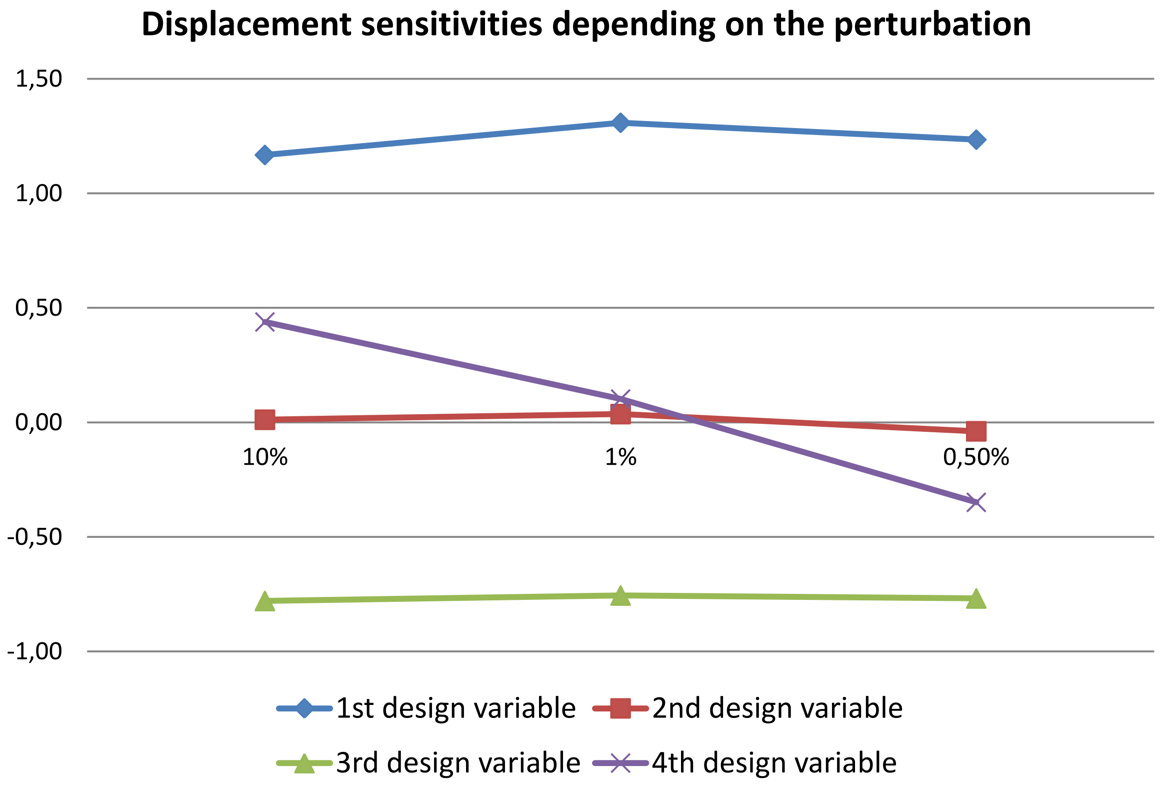


Figure 9.9.: Determination of displacement sensitivity

Summarising the results, in this subchapter we highlighted the importance and applications of sensitivity analysis. We made an overview of the generally applicable finite difference technique, and introduced the fundamental equations and applications of direct and adjoint variables methods. An actual problem was used to demonstrate the method of determining sensitivity values highlighting when and which values can cause numerical stability problems. We showed how to test convergence to ensure the accuracy of sensitivity values and what type of conclusions can be drawn on design variables in such calculations.

9.3. Questions

1. Which areas of applications were mentioned for sensitivity analysis in solving engineering problems?
2. How is finite difference sensitivity calculation technique defined?
3. How does step size influence the precision of finite difference sensitivity calculation?
4. When should direct and when should adjoint sensitivity calculation be rather used?
5. Which are the main steps of sensitivity analysis?

10. SHAPE OPTIMIZATION. GEOMETRIC PARAMETERS AND THEIR IMPACT ON THE OPTIMUM

During shape optimization the data set describing the initial structure is split into two parts: one is a set of design variables, the other is a set of design parameters. During shape optimization the design variables should be selected from the descriptive geometry subset. At the application of parametric 3D CAE systems, the data set describing the geometry consist of the dimensions and order of features, and the definied sketches and specified dimensions and geometric constraints (perpendicular, tangential, pressure, etc...). An examined part can be described by a variety of feature sequence; they will be called geometric representations of the part.

10.1. The effect of the geometry description on the optimization model - a case study

To illustrate the differences based on the variety of describing geometry during optimization, a simple example can be presented: Figure 10.1. shows a rectangular plate with a hole and it's dimenions.

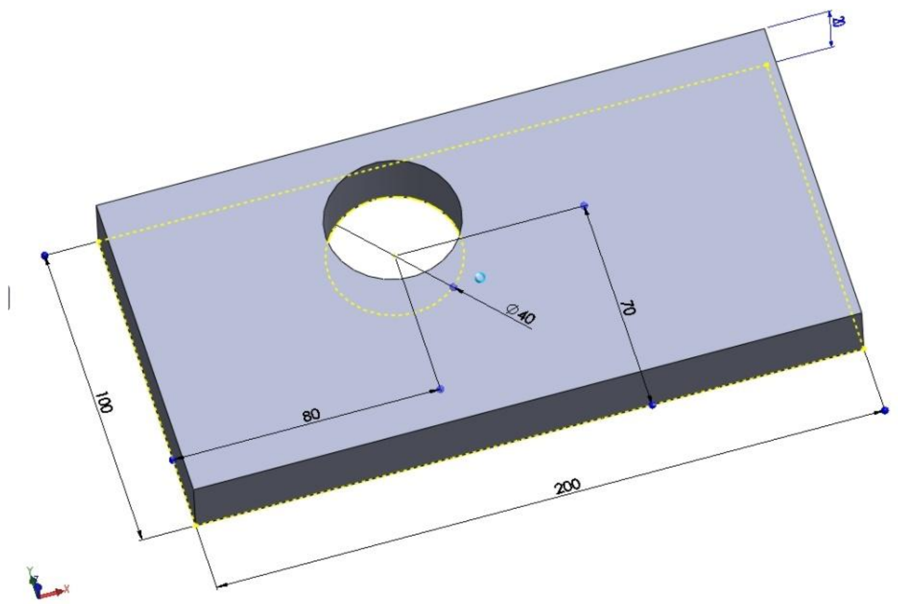


Figure 10.1.: Rectangular plate with a hole - CAD model

Even in the case of this very simple geometry we have multiple opportunities to develop the feature-tree. In the most simple case in which we create the body by one extrude feature from single sketch, containing a rectangle and a circle with given dimensions.
In case the diameter of the hole is increased, the solid geometry disappears when diameter reaching the 60 mms and CAE system is unable to create the body using this simple extrude feature (Figure 10.2. and 1_feature_regen hiba.avi).

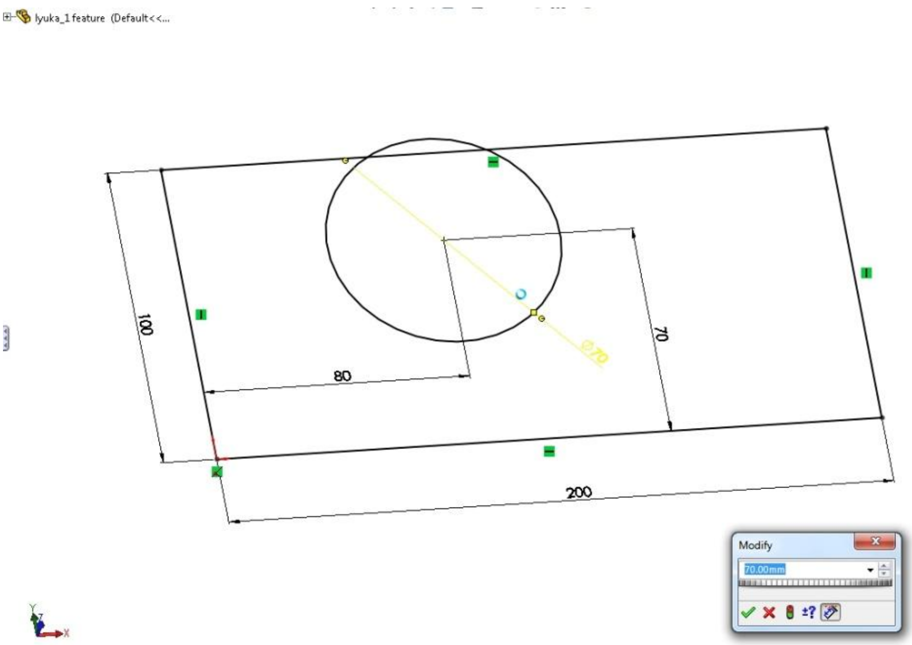


Figure 10.2.: The collapse of the geometry because of the contour intersects

The geometry may also be created by using two features: starting with extraction the rectangular and then drill a hole into (Figure 10.3.). When increasing the size of the drill hole by reaching 60 mm the hole disappears, because the geometric modeler can't construct the cut feature (Figure 10.4. and 2_feature geom .avi).

If we further increase the diameter, the tested CAD system has been able to build up the geometry (Figure 10.5.), but the surface topology of the constructed geometry is different from the initial, which is not recommended in shape optimization. In this case the number of surfaces increased by one. The attached animation (3_FEM changes in boundary conditions.avi) shows, that when on the variable surface finite element boundary condition is given, the FEM modeling in general cannot properly handle the problem.

Therefore, when designing the geometry of the following should be kept in mind:

- the correct structure of the feature tree
- the correct choice of design variables and variation limits
- it should be checked that the structure geometry can be re-build in all of the search range and the surface topology would not change

In spite of the difficulties outlined the advantage of the parametric shape optimization is that parametric surfaces can be easily produced and that the method is easily integrated with CAD systems, as well as the optimized geometry immediately is generated in the usual CAE system.

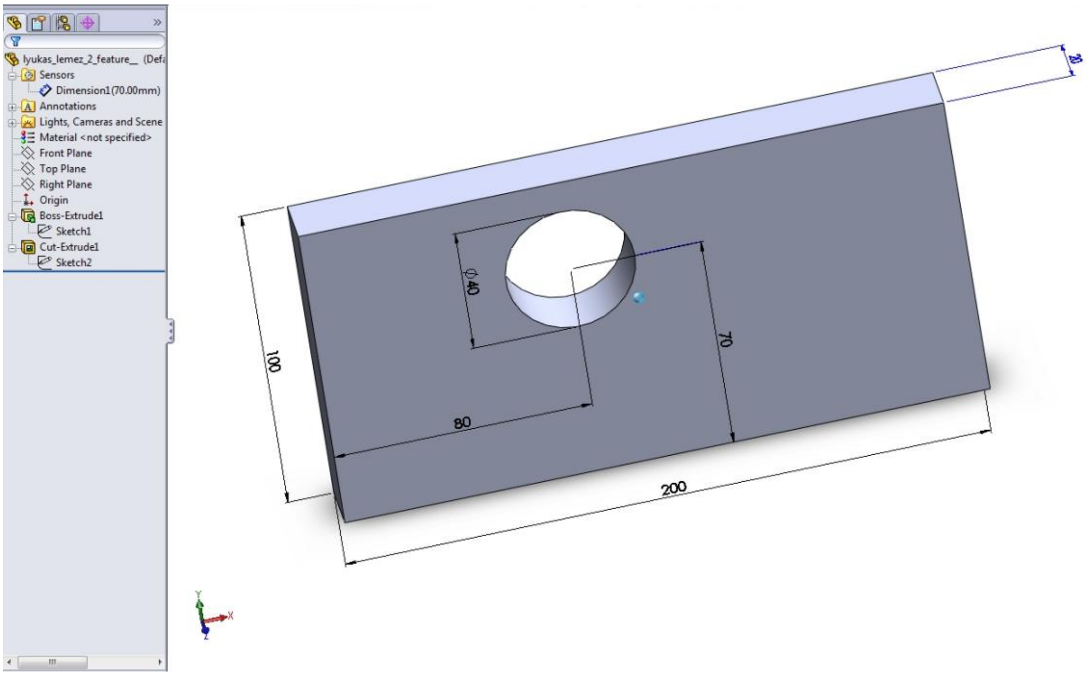


Figure10.3.: The structure of the rectangular with a hole based on two shape features

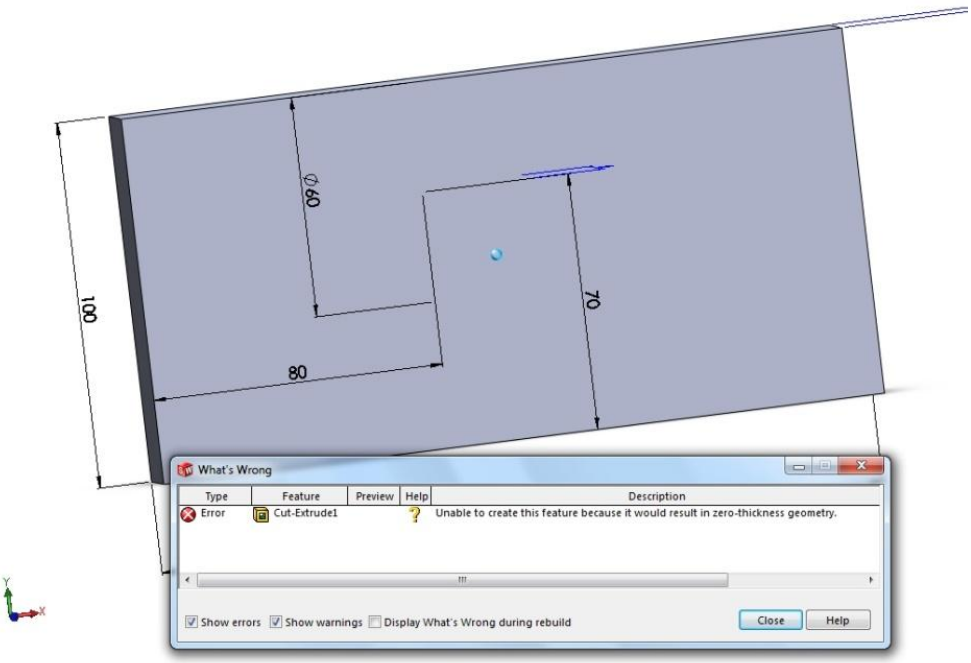


Figure10.4.: The cutout shape feature causes an error, the hole disappears

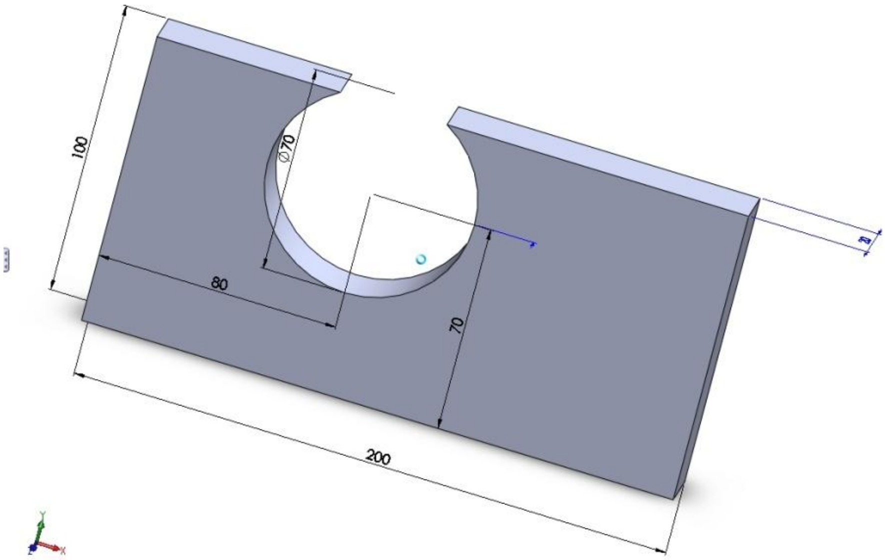


Figure10.5.: Regenerated geometry with the change of the surface number

10.2. Questions:

1. Give examples how the shape feature sequence used in the construction of the body impacts the determination of the optimization range!
2. What kind of problems can cause the changes in topology geometry during reconstruction of the finite element model?
3. Why is it advantages a two-way associative data link between a CAD system and a structural analysis software?
4. What are the most well-known CAE systems, which have both a structural analysis and an optimization module?

11. TOPOLOGY OPTIMIZATION

This chapter shows the place and role of topology optimization methods in the design process, presents the basically different types of methods and demonstrates their benefits and drawbacks. The steps of problem solving will be demonstrated on a simple problem.

11.1. Place of topology optimization in the design process

In a very early stage of the designing we must decide the layout of the designed object for the given operating conditions. In this stage there is possible to achieve great profit on the base of low cost; later on, every small change implies great investment.

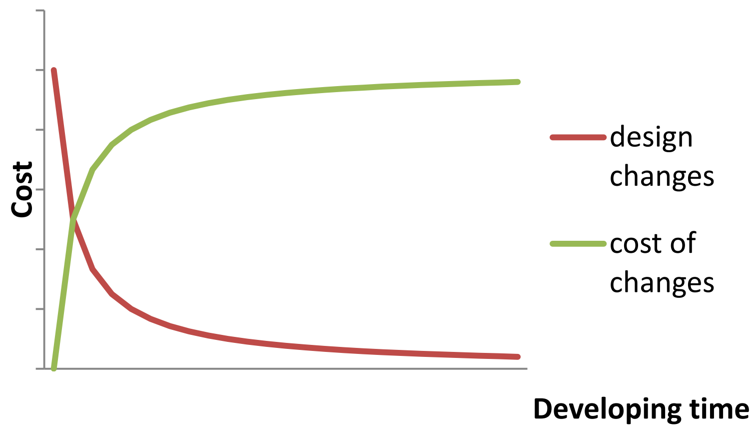


Figure 11.1. Topology optimization can result cost saving in the conceptional phase of the design

On one hand, having a very complicated designing problem, it is worse to recline on the empiric and intuition as in such a case the problem is hard to formalize. On the other hand, the problem can be very new and we are staying there without any specific knowledge about the task, it is worse to search the main effects and we are looking systematically after the best solution. In the case a layout is chosen, we are restricted in the goodness of reachable improvements so we need such a procedure, which supports such a serious decision. The only information about the design is the loading situation, the mathematic model describing the physical behaviour and the connection to the environment. We have no information about the topology of the body at all: how many ribs are needed to enforce the structure o in the contrary how many lightening hole and where we need. The ambition of the topology optimization is to decide which part of the given space should be occupied by our design optimally in a specified view. Nevertheless, the result is rough so it should be further refine having other standpoint, connectivity model and conditions applying in a following size and /or shape optimization loop. Choosing or constructing a topology optimization method the first focus is the characteristics of the problem itself: what the objective(s) is/are; the second one is the way of parameterization to follow the topological changes.
As the evaluation of the solutions follows by solving the physical model through discretization it is important how the parametric representation of design connects to the physical one’s.

Those methods which use a mesh on the changing design, have difficulties as the topology changes. For topology optimization it is more suitable to model the geometry on an implicit way and perform the structural analysis on a fixed, initially given domain. These methods can adapt to the topological changes and there is no need for remeshing the changed body.

11.2. Benchmark problems for testing the algorithms

Mitchell laid down the theoretical base of topology optimization by giving optimality criteria for light structures. Later, in the 70'ths with Rozványi and Prager he extended the theory generally and based the examination of continuous structures. These results serve for comparison to the different topology optimization methods [11.1].
The solution of the here demonstrated examples gilt for one load case, with elastic behaviour. The results for weight minimization are valid for both stress and compliance constrain as a relation exist between the optimal weight W_S, the allowable stress σ_0 and the optimal weight W_C for a given compliance (11.1)

$$W_C = \frac{\sigma_0^2}{\rho E C} W_S^2 \tag{11.1}$$

ρ is the specific weight, E the elastic modulus. The relation between the optimal weight and the length of the structure can be given with dimensionless data $\overline{W}$ and $\overline{L}$ (fig. Figure 11.3.), where h is the height of the beam and P is the load.

$$\overline{W} = \frac{W\sigma_0}{Ph\rho}, \overline{L} = \frac{L}{h} \tag{11.2}$$

Mitchell beam – it is fixed in one end and loaded on the free one. The analytical results can be seen on Figure 11.2.

Figure 11.2. Analitical solution of the Mitchell beam

MBB beam
The problem was originally formulated by the air company of Messerschmidt-Böhlkow-Blohm. In the literature the compliance minimization of the simply supported beam is called MBB problem (Figure 11.3.)

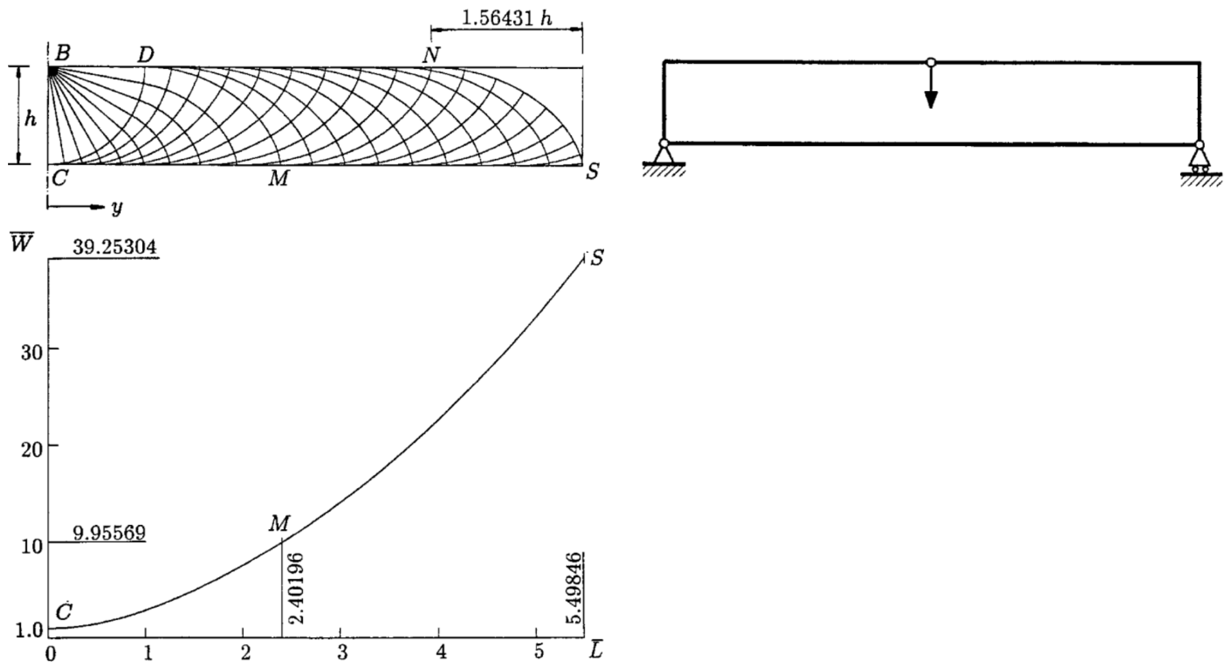


Figure 11.3. Analytical solution of MBB structure

11.3. Methods of topology optimization

In the last decades many topology optimization procedures were developed as the homogenization method, the material distribution method SIMP (Solid Isotropic Microstructure with Penalty), evolutionary techniques (ESO = Evolutionary Structural Optimization), reverse adaptivity, the bubble method, the level set method (LSM), method of topological derivatives and others. The most effective and simple is the SIMP method to solve the basic problem of compliance minimization so it is widely used in the commercial programs and for this reason it will be shown in more detail. Further extension and application can be found in [11.15].

Topology optimization methods can be sorted in three groups:

- methods based on mathematic background which are suitable for global criteria as compliance or volume
  - homogenization method
  - SIMP
  - level set method
  - phase field method
- heuristic methods which stand for homogenization with removing small amount of material from the design space where the governing criteria e.g. stress is low.
  - soft kill (SK) or a hard kill (HK) evolution methods (ESO, AESO, BESO, XESO)
  - reverse adaptivity
  - metamorphic development
- mixed method with combining topology and shape optimization
  - bubble method,
  - isoline method.

In the following sections the basic characteristics of the above methods will be shown.

11.4. Homogenization method [11.2]

By the homogenization method the body is treated to be from a porous material and we are looking for optimal distribution of microscale voids which shape is assumed to be rectangular in a given domain. Applying finite element method, this domain is divided in *N* finite element and design variables are the size and direction of the voids in each of these elements (Figure 11.4).

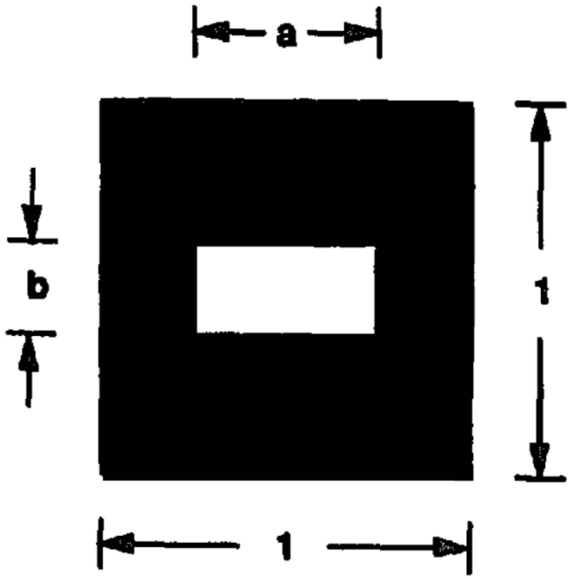


Figure 11.4. Design variables in the case of homogenization method

Rectangular holes are chosen because they can realize the complete void ($a = b = 1$) and solid ($a = b = 0$) as well as generalized porous medium ($0 < a < 1, 0 < b < 1$). Having the variables a and b inside of the intervals, the real structure must be porous or inhomogeneous density which is hardly to carry out. Additionally, in this case we have 3 design variable pro elements resulting a very big optimization problem; it can be slightly reduced for example by applying square voids. However, it was the first method applying fix domain for structural analysis and this main idea was further developed by the SIMP method.

11.5. The SIMP method [11.3]

The SIMP method is the most popular topology optimization method due its simplicity. Solving form finding problems of mechanical engineering, the optimization criteria is mostly the stiffness of a part, for one or more load case constraining the volume. The achieved proposal must be interpreted taking manufacturing viewpoints into account. The reformulated geometry should be further examined refining the loading environment, or other conditions can be considered.

The matter of the method is to search the optimal topology of the structural part as material distribution in a given design domain divided into *N* elements. The design variables are the fictitious ρ_e relative element density, ($e = 1, 2, \cdots, N$). The element e is solid, if $\rho_e = 1$, and void, if $\rho_e = 0$. Avoiding the computational difficulty of the non-continuous description, in place of discrete values $\{0, 1\}$ we use continuous design variable $0 \le \rho_e \le 1$, so

$$E_e = \rho_e^p E_0 \tag{11.3}$$

where p is the penalty power penalizing intermediate densities (usually should be 3 to 5), E_0 is the Young modulus, that means we interpolate with ρ_e^p between 0 and E_0.

Compliance minimization

The topology optimization problem can be given as follows:

we are looking for $\boldsymbol{\rho} = \{\rho_1, \ \rho_2, \ \cdot\cdot\cdot, \ \rho_N\}$

$min : C = \boldsymbol{f}^T{}_k\boldsymbol{u}_k,$

subject to

$\sum_{e=1}^{N} \rho_e v_e \ \leq V^*$,

and $0 < \rho_{\min} \leq \rho_e \leq 1$ (11.4)

where displacement $\boldsymbol{u}$ depends on the vector of design variables $\boldsymbol{\rho}$, $\boldsymbol{f}$ is the vector of outer force, and $\boldsymbol{u}$ is the solution of

$$\mathbf{K}(\boldsymbol{\rho})\,\mathbf{u} = \mathbf{f}, \tag{11.5}$$

where **K** is the global stiffness matrix, composed from the element stiffness matrices which depend on design variable ρ_e threw the above introduced way. V^* is the prescribed volume, v_e is the element volume. The lower limit for density $\rho_{\min}$ was introduced to avoid singularity. Index k denotes the degrees of freedom where the outer load acts.

As the work of the outer load is equal to the energy stored in the deformed body, the objective function can be given as follows:

$$\min\ C = \mathbf{u}^{\mathrm{T}}\mathbf{K}\mathbf{u}, \tag{11.6}$$

that is our problem is to minimize the strain energy.

Introducing the Lagrange function the necessary condition of optimality

$$\mathrm{p}\rho(\mathrm{x})^{\mathrm{p}-1}\mathrm{E}^{0}_{\mathrm{ijkl}}\varepsilon_{\mathrm{ij}}(\mathrm{u})\varepsilon_{\mathrm{kl}}(\mathrm{u}) = \Lambda \tag{11.7}$$

which expresses that the strain energy density-like left-hand side term is constant and equal to Λ for all intermediate densities. This is thus a condition that is similar to the fully stressed design condition in plastic design. As we expect areas with high energy to be too low on stiffness we devise the following fix-point type update scheme for the density [3]:

$$\rho_{K+1} = \begin{cases} max\{(1-\varsigma)\rho_K, \rho_{min}\}, \text{ if } \rho_K B_K^{\eta} \ \leq max\{(1-\varsigma)\rho_K, \rho_{min}\} \\ min\{(1+\varsigma)\rho_K, 1\}, \text{ if } min\{(1+\varsigma)\rho_K, 1\} \leq \rho_K B_K^{\eta} \\ \rho_K B_K^{\eta} \text{ else} \end{cases} \tag{11.8}$$

Here ρ_K denotes the value of the density variable at iteration step K, and B_K is given by the expression

$$B_K = \Lambda_K^{-1} p\rho(x)^{p-1} E^0_{ijkl}\varepsilon_{ij}(u_K)\varepsilon_{kl}(u_K), \tag{11.9}$$

where u_K is the displacement field at at the iteration step K, determined from the equilibrium equation and dependent on ρ_K. Note that a (local) optimum is reached if $B_K = 1$. The update scheme (11.8) adds material to areas with a specific strain energy that is higher than Λ (that is, when $B_K > 1$) and removes it if the energy is below this value; this only takes place if the update does not violate the bounds on p. One can see that Λ is proportional (by a factor p) to the average strain energy density of the part of the structure that is given by intermediate values of the density. The variable η in (11.8) is a tuning parameter and ς a move limit. Both η and ς controls the changes that can happen at each iteration step and they can be made adjustable for efficiency of the method. Note that the update ρ_{K+1} depends on the present value of the Lagrange multiplier Λ, and thus Λ should be adjusted in an inner iteration loop in order to satisfy the active volume constraint. It is readily seen that the volume of the updated values of the densities is a continuous and decreasing function of the multiplier Λ. Moreover, the volume is strictly decreasing in the interesting intervals, where the bounds on the densities are not active in all points (elements of a FEM discretization). This means that we can uniquely determine the value of Λ, using a bisection method or a Newton method. The values of η and ς are chosen by experiment, in order to obtain a suitable rapid and stable convergence of the iteration scheme. A typical useful value of η and ς is 0.5 and 0.2, respectively.

The type of algorithm described above has been used to great effect in a large number of structural topology design studies and is well established as an effective method for solving large scale problems. The effectiveness of the algorithm comes from the fact that each design variable is updated independently of the update of the other design variables, except for the rescaling that has to take place for satisfying the volume constraint. The algorithm can be generalized to quite a number of structural optimization settings, but it is not always straightforward. For cases where for example constraints of a non-structural nature should be considered (e.g., representing geometry considerations), when non-self-adjoint problems are considered or where physical intuition is limited, the use of a mathematical programming method can be a more direct way to obtain results. Typically, this will be computationally more costly, but a careful choice of algorithm can make this approach as efficient as the optimality criteria method.

This optimality criteria method is very effective and can be extended to another problem formulation; it needs few programming effort in case of having a finite element code. In those cases where the application of the optimality criteria is not possible, we can use mathematical programming method.
The prescribed volume V^* can have predominant effect as it can be seen on Figure 11.5 so the designer must be carefully decide by choosing this parameter.

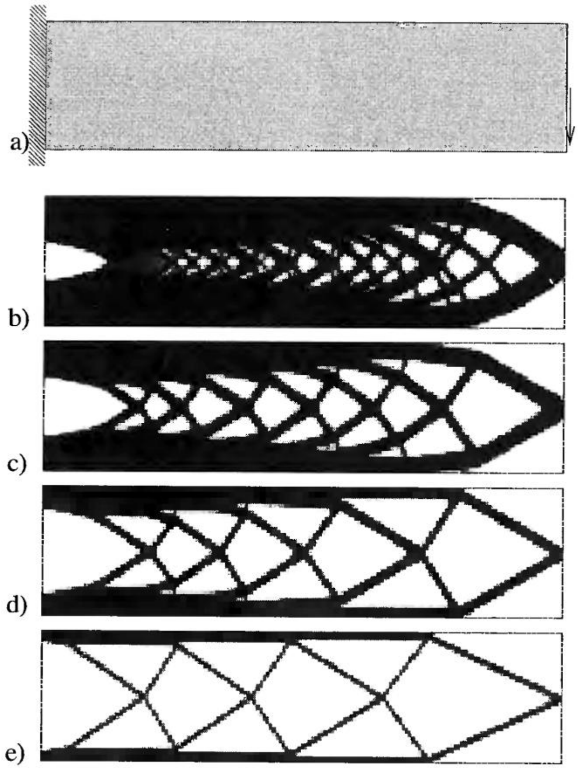


Figure 11.5. Effect of prescribed volume fraction b) 80%; c) 60%; d) 40%; e) 20%

Ensuring convergence with mesh refinement
The basic algorithm does not ensure a unique solution with mesh refinement (Figure 11.6). There is three main, principally different techniques to ensure the convergence in case of mesh refinement:

- control of perimeter or prescribing minimal member size
- reduction of parameter space (coarser mesh on design space)
- filtering methods (sensitivity or density filtering).

Having greater domain with intermediate density can be avoided effectively applying the simple *p* penalty parameter. It is very useful because of interpreting a solution with intermediate density is not easy in every case (in some case we can suppose a porous, or thinner or weaker material).
The method can result a **checker board** like pattern, this has also very few connection to real structure. For solving both problems we can apply projection methods and sensitivity filtering.

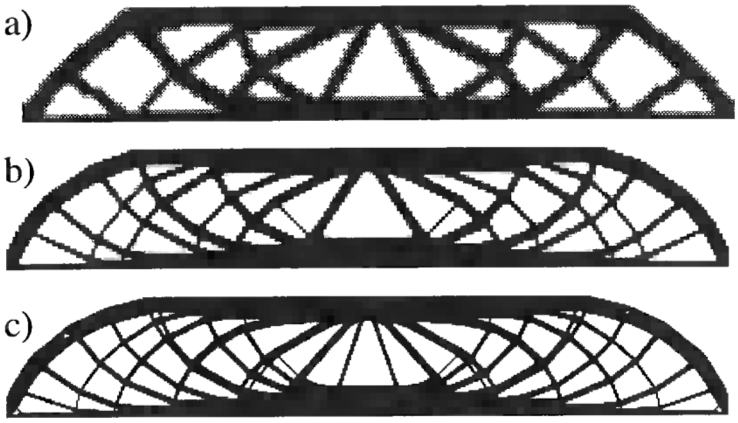


Figure 11.6. Effect of mesh refinement to the optimum a) 2700 b) 4800 c) 17200 elements

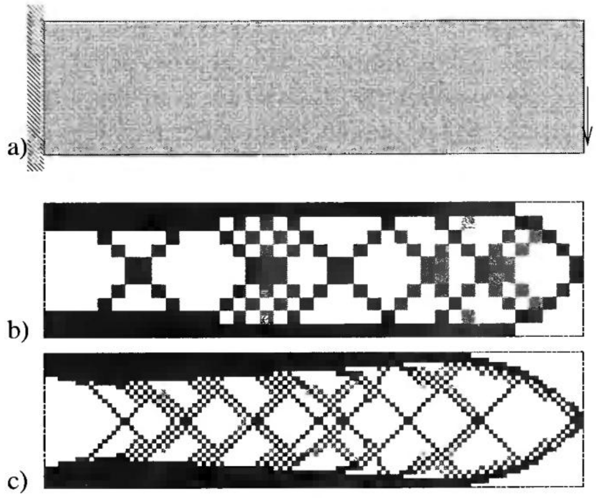


Figure 11.7. Checker board like solution for 400 and 6400 elements

Naturally, applying more than one load case changes the characteristic of the solution (Figure 11.8). Figure c) and d) shows the optimized topologies for all loads in one load case. In Figure e) and f) we can observe the optimized topologies for multiple loading cases. It is seen that single load problems result in instable structures based on square frames whereas multi load case problems results in stable structures based on triangular frames.

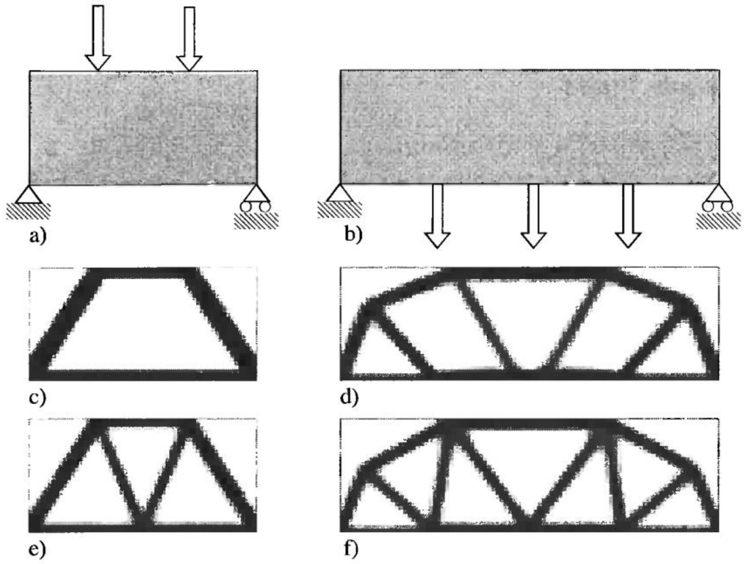


Figure 11.8. Optimum for one and for multiple load cases

The problem solving sequence with SIMP method is as here follows:

- define the initial design domain, the boundary value problem, with homogeneous material distribution,
- compute the displacements and strains for the actual design density with the finite element method,
- compute the compliance for the actual design,
- in case of having no significant changes or achieving the optimum criteria stop else
- compute the new material distribution due to equations (11.7) and (11.8) (and with an inner loop the value of Λ Lagrange multiplicator due to the volume constraint),
- repeat the procedure from the second step.

In the case applying mathematical programming method to compute the optimum, we need also the sensitivities of objective function and optimization constraints. Writing the problem in discrete form after (11.3) and (11.4), the sensitivities can be easily calculated with the adjoint method for the compliance problem as in this case the adjoint variable is identical with the displacement so

$$\frac{\partial c}{\partial \rho_e} = -p\rho_e^{p-1}\, \boldsymbol{u}^{\boldsymbol{T}} \boldsymbol{K_e} \boldsymbol{u} \tag{11.10}$$

The sensitivity does not depend directly on the other elements, only implicit way through the displacements and it can be easily computed. As the sensitivity takes negative values, the intuition as adding material will decrease the compliance and increase the stiffness. Having a fine design, we get high number of design variables (one for each element) which can be handled with Method of Moving Asymptotes (MMA). The algorithm and programming effort is similar to the previously demonstrated optimum criteria method. The utmost time cosuming part of the calculation is the structural analysis.

11.6. Level set methods(LSM) [11.3]

The fast marching method and the level set method are numerical techniques which follows the change of an interface coupling it to a level set function. In level set-based structural optimization methods, complex shape and topological changes can be handled and the obtained optimal structures are free from greyscales, since the structural boundaries are represented as the iso-surface of the level set function. These relatively new structural optimization methods overcome the problems of checkerboard patterns and greyscales. The interface can have unsmooth shape with edges and corners, its topology can be divided or rejoined without having computational problems because the level set function can follow this changes on a smooth way.

Rather than follow the interface itself, the Fast Marching Method makes use of stationary approach to the problem. Let us lay a grid laid down on top of the problem. A time like function *T(x,y)* is ordered to the moving interface: at each grid point *T*, *T(x,y)* gives the time at which the front crosses the point *(x,y).*

As an example by Sethian et al., suppose the initial disturbance is a circle propagating outwards. The original region (the blue one on the left below) propagates outwards, crossing over each of the timing spots. The function *T(x,y)* gives a cone-shaped surface, which is shown on the right. This surface intersects the *xy* plane exactly where the curve is initially. At any height *T* the surface gives the set of points reached at time *T*. The surface on the right below is called the arrival time surface, because it gives the arrival time.

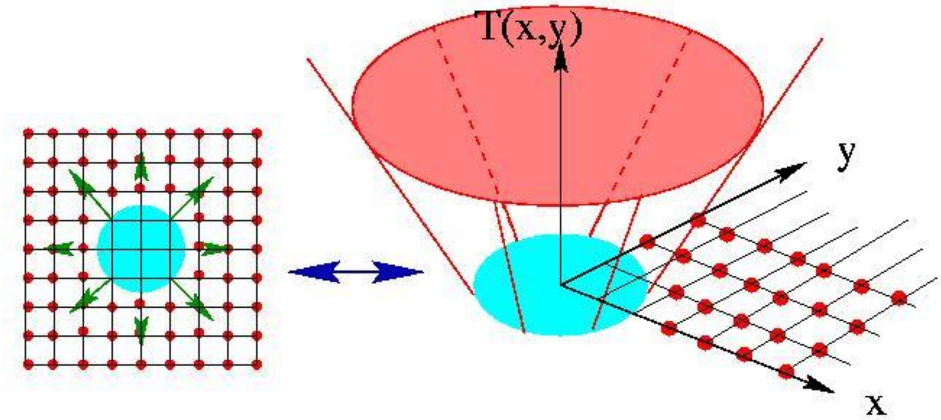


Figure 11.9. Interpretation of the fast marching method

Beneficially, the numerical calculations are made on a fix grid (Eulerian approximation).

The level set function is more general: the interface is ordered to a signed distance function and the initial shape is ordered to the zero level. The initial value problem for the moving of the level set function corresponds to the Hamilton-Jacobi equation.

In the following example a bending plate is given on Figure 11.10. The red arrow denotes the load, green squares the fixation. The design boundary is drawn with broken line. Stress distribution is shown on Figure 11.11.

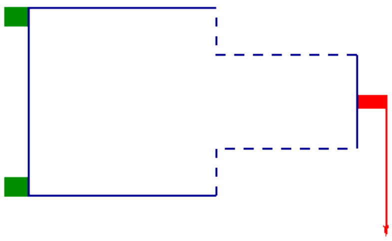


Figure 11.10. Problem definition

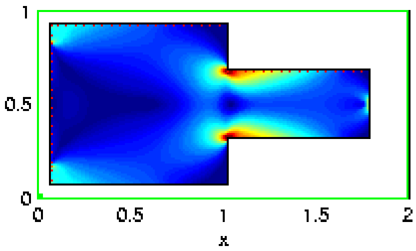


Figure 11.11. Stress distribution of start design

A hole is inserted on the low stressed zone and the stresses are recalculated; the hole is enlarged in the direction to the low stressed zone.

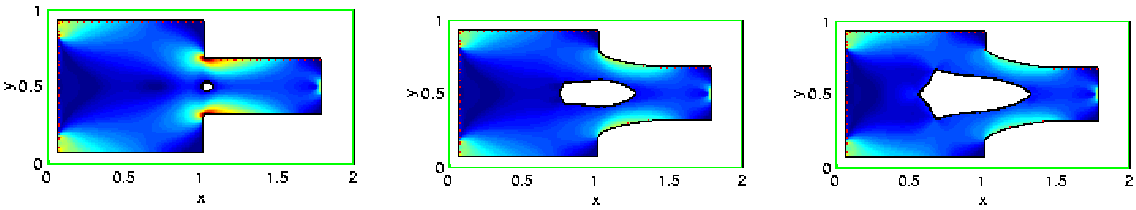


Figure 11.12. Operation of level set method

11.7. Evolutionary Structural Optimization (ESO) [11.7][11.8][11.9]

The evolutionary methods based on engineering intuition: they extract domains not participating in the load transfer and add material to strengthen the weak places (bidirectional methods). Initial finite element model for the **Hard-Kill (HK)** Method can be a structured mesh of a given part; the elimination of elements follows with the reduction to near zero of its material properties. Benefit of the method is there is no need for remeshing and every kind of problem can be solved even for more load cases. Its' drawback is the zigzagging mesh and if the step-width is chosen inappropriately, the supported or the loaded surface can be vanished. Need for mesh refinement is also a problematic task. Metamophic development is a bidirectional version of evolutionary methods, which adapts the amount of added/extracted material dynamically to the structural responses in every iteration steps, accelerating the convergence.

11.8. Nonprofit software tools to obtain the optimal topology

In the internet there are available some free software to demonstrate how to estimate the optimal layout of a design using topology optimization techniques. In this section the capability and problem solving sequence of TOPOPT and TOPOSTRUCT programs will be demonstrated. Both of them search the optimal material distribution (SIMP method) on fixed mesh of an initially given brick shaped domain.

Steps of compliance minimization:

- Choosing between 2D or 3D
- giving the size and resolution of design space
- Prescription of supports and loads
- Volume fraction up 0.1 to 0.2

- Maximal iteration limit and penalty parameter
- solving the optimization
- evaluation the results
- adapt and forward the results.

11.8.1. TOPOPT

TOPOPT is developed on the Technical University of Denmark, in a cooperation of Faculty of Mathematics and Mechanical Engineering for supporting the theoretical developments and solving practical problems. The preprocessing the optimization problem takes place in a Java applet, the calculation runs on a remote computer. Results appear as animated gif on the local machine.

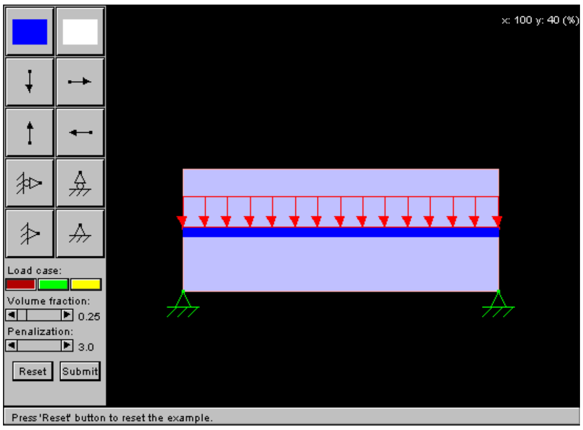


Figure 11.13. Definition of the optimization problem with TOPOPT

In the case program for solving 2 dimensional compliance problem design domain can be given with dragging the size of a light blue rectangle on the black background after its selection between 100x100 units (Figure 11.14.). Prescribing holes for occupancy of connected components or is possible thru selecting, adapting and dragging white rectangles while conservation of supported and/or loaded surfaces can be ensured using darkblue rectangles in arbitrary number.

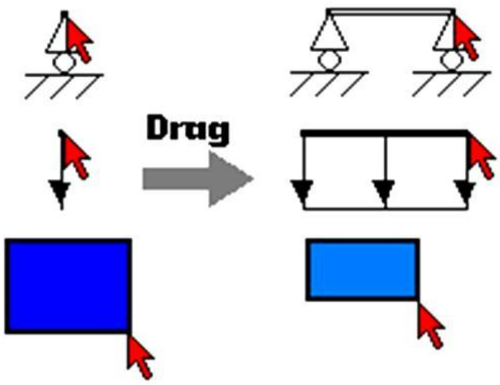


Figure 11.14. Prescribing tools for the boundary value problem

Symmetry should be utilized for achieving finer results. Mechanical problem can be also given by choosing the supports and loads in appropriate direction and applying them on the body thru dragging.

Multiply load case can also be applied up to 3 loadcases.

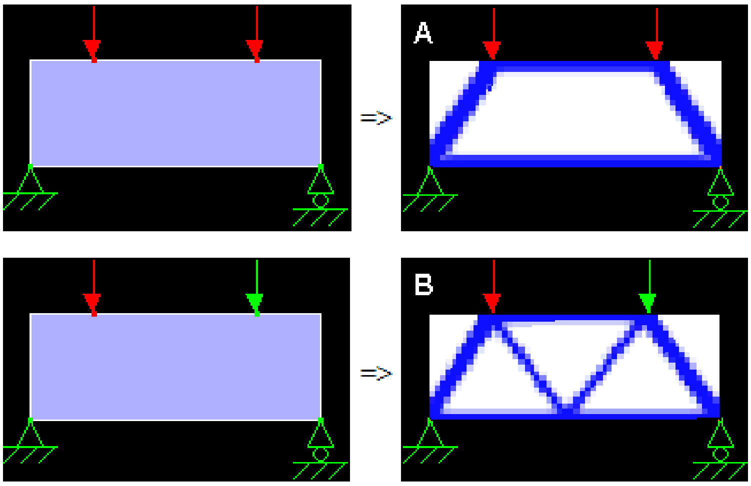


Figure 11.15. Topology optimization for more load cases 1.

On the top of Figure 11.15. the two loads act simultaneously, on the bottom apart from each other. The difference of the solution can be observed on the right hand side.

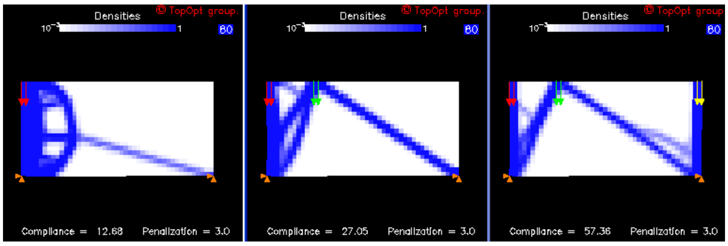


Figure 11.16. Topology optimization for more load cases 2.

Effect of additional load cases can be followed on Figure 11.16.
Volume fraction (up 0 to 1) and penalty parameter (between 1 and 3) can be adjusted with the slides. Limits of the 2 dimensional programs are

- 1000 design variables
- 3 load cases
- 100 iteration

Mechanism design can be made on a similar way with another program. Compliance design can be performed for 3D problem; in that case the work can be saved and results can be transferred to a CAD software in .stl format.

11.8.2. TOPOSTRUCT

Topostruct program was developed by Panagiotis Michalatos and Sawako Kaijima in 2008 for getting acquainted with topology optimization. The model can be given in 2 or 3 dimension, the design space is rectangular or brick.

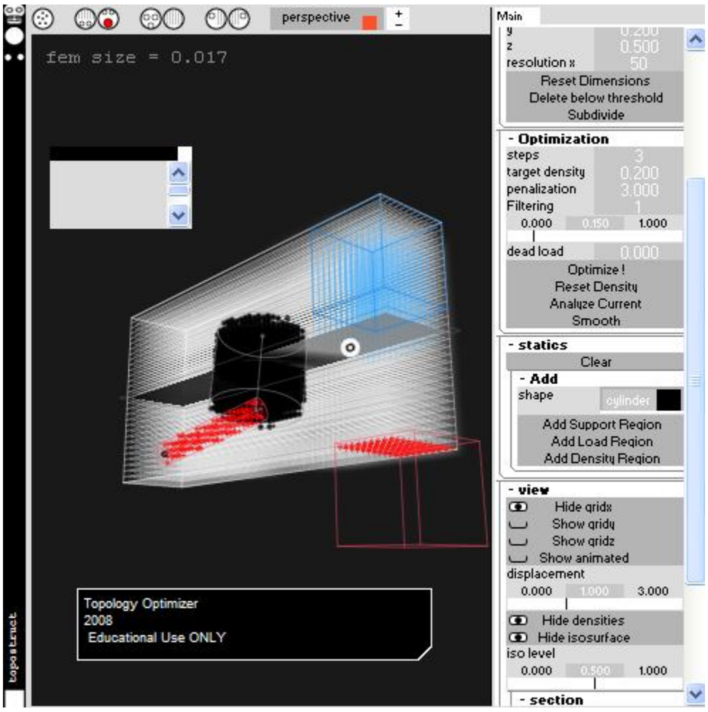


Figure 11.17. Topostruct program

Additional facility to TOPOPT, that the shape of the selected group of nodes for boundary value definition can be sphere or cylinder with axis in arbitrary direction. Model and results of structural analysis can be visually controlled and animated. The optimal result can be manipulated: it can be smoothed and elements can be eliminated under a given density level. Multiply loadcase can not be examined.
More detail can be found in the user manual. Operation can be understood in 2 dimensional és 3 dimensional examples.

11.9. Summary

Topology optimization searches the optimal layout for a global characteristic as compliance or natural frequency of a structural part, for a given material utilization and considering manufacturing constraints as extrusion, symmetry, draw direction by casting. The method should be able to find the optimal solution independent from the initial design guess.
An intuitive way for finding optimal layout is change the design due to engineering decision (cutting out low stressed part and strengthen the high stressed zones). These methods are called evolutionary methods. Such methods highly depend on the choice of parameters, so they are not stable and sometimes delivers unusable solution. The continuous topological changes should rather be followed with methods searching the optimal shape on an extended fix domain. The design variables able to follow these changes continuously could be

- geometrical parameters bonded to microstructure as in the homogenization method
- material like parameter as in the SIMP method
- bonded to a time like parameter after Euler formula, so the moving of the surface will be given in the initial coordinates as in the level set method and in the phase field method.

Numerical problem could arise and should be avoided as

- numerical instability
- dependency from mesh
- dependency from initial design
- checkerboard like solution
- greyscales or porous domains without any practical interpretation
- solution with manufacturing difficulties or with high manufacturing costs.

Solution of the above mentioned problem can be also multitude:

- geometrical limit as perimeter control or minimal member size control
- filtering techniques as sensitivity filtering
- using different mesh for design and analysis
- using manufacturing constraints as draw direction, extrusion direction, circular or axial symmetry
- further stabilizing conditions as giving a fictive limit surface energy or entropy like condition

Specialists are continuously searching the most suitable answers to all of these questions and extending the method for solving further type of problems. Nevertheless, results of a topology optimization should not interpreted directly as a design for manufacturing, interpretation leaves a task of design engineer, especially having a complex design problem. The optimal solution got for global condition should be further refined by an additional shape optimization step for local conditions as stress.

11.10. Literature

[11.1] Rozvany GIN. Exact analytical solutions for some popular benchmark problems in topology optimization. Structural Optimization 1998;15: 42-48

[11.2] Bendsøe MP, Kikuchi, N. Generating optimal topologies in structural design using a homogenization method. Computer Methods in Applied Mechanics and Engineering 1988;71:197-224.

[11.3] Bendsoe MP. Optimal shape design as a material distribution problem. Structural and Multidisciplinary Optimization. 1989;1:193–202.

[11.4] Sethian JA, Wiegmann A. Structural boundary design via level set and immersed interface methods. J Comput Phys 2000;163(2):489–528.

[11.5] Norato J, Bendsøe MP, Haber RB, Tortorelli, DA. Topological derivative method for topology optimization. Structural and Multidisciplinary Optimization 2007;33: 375-386.

[11.6] Takezawa A, Nishiwaki S, Kitamura M: Shape and topology optimization based on the phase field method and sensitivity analysis Journal of Computational Physics 229 (2010) 2697–2718

[11.7] Xie YM, Steven GP. A simple evolutionary procedure for structural optimization, Computers and Structures 1993;49: 885–896.

[11.8] Querin OM, Steven GP, Xie YM. Evolutionary structural optimization (ESO) using a bidirectional algorithm. Eng Comput 1998;15:1031–1048.

[11.9] Sauter J. CAOS oder die Suche nach der optimalen Bauteilform durch eine effiziente Gestaltoptimierungsstrategie. FEM '91, IKOSS CONGRESS, Tagungsband (S. 159-187), Baden-Baden, November 1991

[11.10] Reynolds D, McConnachie J, Bettess P, Christie WC, Bull JW. Reverse adaptivity—a new evolutionary tool for structural optimization. Int. J Numer Methods Engrg 1999;45 :529–552.

[11.11] Liu JS, Parks JT, Clarkson PJ: Optimization of Turbine Disk Profiles by Metamorphic Development. ASME Journal of Mechanical Design 2002;192-200

[11.12] Eschenauer HA, Kobelev HA, Schumacher A. Bubble method for topology and shape optimization of structures. Struct Optim 1994;8:142–151

[11.13] Victoria M, Marti P,Querin OM. Topology design of two-dimensional continuum structures using isolines. Computers & Structures 2009;87(1-2):101-109.

[11.14] Rozvany GIN. A critical review of established methods of structural topology optimization. Struct Multidisc Optim 2009;37:217–237.

[11.15] Bendsøe MP, Sigmund O. Topology Optimization: Theory, Methods, and Applications. Berlin, Heidelberg: Springer 2003

11.11. Questions

1. Which topology optimization methods can be used for more type of problems or for multidisciplinary problems?
2. What are benefits of the mathematical algorithms?
3. What designer decisions should be made before topology optimization?
4. What designer decisions should be made after topology optimization?

12. OPTIMIZATION METHODS OF ENGINEERING PROBLEMS

This chapter demonstrates in what extent and characteristic could a real life engineering optimization problem differ from the simpler model shown until this point. The problems will be highlighted occurring in the course of solving them with nonlinear mathematical programming methods and other types of optimization tools and algorithms will be demonstrated.

12.1. Integration of optimization into the design process

Nowadays the quality, effectiveness, decrease of costs and small time to market are basic demands. Fulfilling them is a challenge task for the designers, mathematicians and software engineers, especially with the will of getting solution of more and more complex, innovative design problem, applying the latest material and manufacturing technology. Design process underlies to the environment defence and sustainability and concentrates no more only to get the strongest design but viewpoints as reassembly, recycling and others occurring in the life-cycle of the product should be taken into account. Design knowledge and information supporting the decision process should be formed clearly, so the producing new design and variances can be supported.
Following design aspects in the course of product lifetime, subsequent places and types of optimization tools can be inserted into the design process:

- completion of project plan (process optimization),
- formulate the aims, required functionalities, restrictions for the probable operating conditions,
- kinematic and geometrical analysis are needed to find the occupancy in the space, connecting/supporting surfaces for load transfer; from that the main geometrical datas conditions and initial geometry can be formulated,
- choosing qualifying parameters as material and safety factor after economical consideration, prescriptions of standards or analysing the concurrent product,
- determination of loads – time dependency, physical characteristic, direction and magnitude with experiments, numerical simulation of multibody systems, prescribed evaluation values),
- **find the optimal design for the given resources and frames - machine, software, knowledge base, experimental equipments, materials, data, costs, deadlines, prescriptions, norms, manufacturability, operating conditions on different models and virtual prototypes,**
- documentation of design process; organizing, filtering and backup of data building and critic of prototype, redesign if needed,
- manufacturing design,
- manufacturing,
- control,
- storage, delivery,
- installation,
- operation,
- maintenance,

- taking out of usage,
- recycling.

In the following we concentrate to the bold faced point, forming the optimal construction. Having a complex design optimization task, the following challenges have to be faced among others (Table 12.1):

Problem	Research area, method, solution proposal
delimiting solution time/cost	• Decreasing the number of variables – on the results of data mining and sensitivity analysis, • meta-models: Design of Experiment and Response Surface Method, kriging meta-model, etc.
optimal construction in variable operating environment	• Robust design, • Reliability Based Design Optimization
finding the global optimum	• Global searching methods, • varying the initial design to avoid local extrema, • algorithms capable to leave local minima as simulated annealing.
multiobjective optimization	• Combining objectives, • hierarchical handling of objectives, • searching Pareto optimal front.
dynamic problems	• Equivalent static load method, • direct methods.
solution of complex problems with great size	• Application of modern distributed computing systems (felhő method, multiprocessor system, supercomputers.), • using parallel algorithms, • dynamic programming.

Table 12.1. Difficulties solving real life problems and proposal to avoid them

12.2. Delimiting the solution time – meta-models, design of experiments

The evaluation of objective functions and optimization conditions in the machine design practice requires complex and time consuming numerical analysis. Application of traditional optimization algorithms is not possible because of the huge numerical effort due to iterative function calls. The computing time can be reduced in one hand with problem adapted modelling, on the other hand with delimiting the numbers of evaluation, optimal choice of evaluation points and building surrogate models (Figure 12.18).

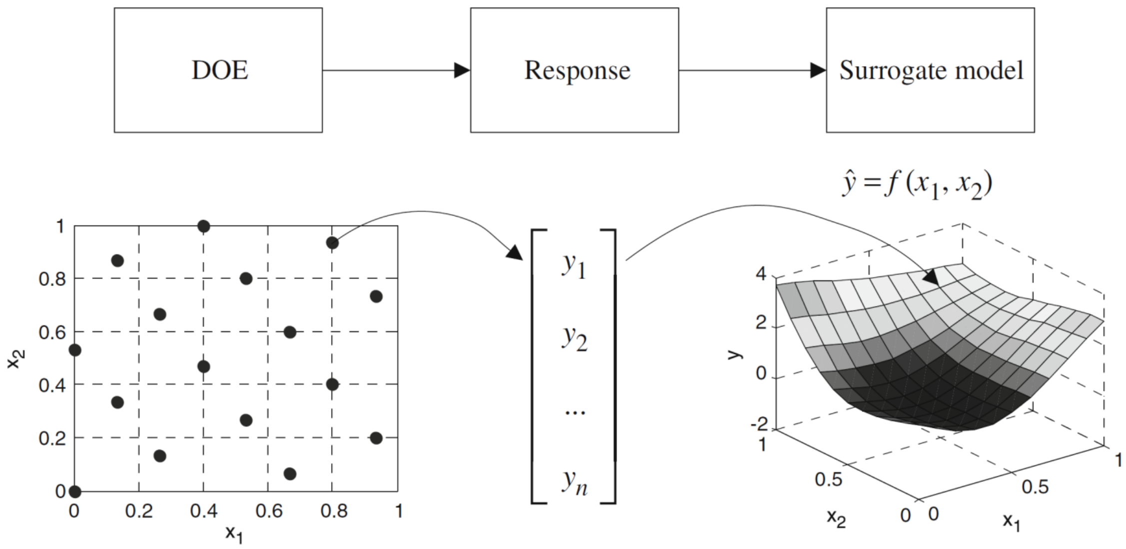


Figure 12.18. Building a surrogate model

Aim of this approach is to give a continuous function which is easy to evaluate and give enough good approximation. It is often called meta model as it is a model of a model itself. Many different meta models exits: simple and adaptive response surface method, kriging method, radial basis function, multivariate adaptive spline regression, neural networks and others. Depending on the way of approximation, the model can be regression or interpolation.
Design of Experiment has the task to establish optimally how many and what type of experiments are needed to gather the **more information, more exactly, with the less cost** about the subject of experiment (in our case the objective function and optimization design constraints) for the given set of variable factors (in our case design variables) In the case of machine design it means the number of structural analysis, and the values of design variables to deliver an enough exact approximation of responses and so the optimization can be established on the surrogate model.

12.2.1. Data mining and sensitivity analysis for reducing the problem size

Data mining helps to extract hidden data connections which cannot be shown with simple data analysis. The elaborated algorithms gives possibility to identify the key characteristics even the in statistics unversed users.
Local sensitivities as correlation factor and partial derivatives can be used only if the connection between input and output variables is linear. In nonlinear cases global sensitivity calculation should be made, for example the variance should be also analyzed. Sensitivity analysis can be helpful by reducing the size of the problem and the reason/effect connection can be explained.
To make a simple estimation for a linear model, we can apply factorial or partial factorial design. It is enough to determine which variable has significant impact on the objective function. The second order approximation can be built in a further step, with a more complex experiment as the central composite design. This will be used then to optimize the objective.

12.2.2. Design of Experiment

Applying Design of Experiment for problem solving, it is worth to execute more experiments: one single experiment with the most base points is not necessarily the best way to become

acquainted to the nature of the problem. It is more effective to make more experiments iteratively, learning from the previous one.
Scientifically, Fischer [12.1] established basis for Design of Experiments. Globally, the steps of the method are the following (with the phrases of structural optimization):

- formulate the problem, deciding its type (it is not the same if we describe a state or a process, aim of the experiment is description, mapping prognosis or testing a hypothesis);
- establish the independent variables or factors (in our case design variables), mapping the connections between them, determine their searching domain;
- assign dependent variables (in our case objective functions);
- set the levels of independent variables a (having more variables, in the praxis it can only be 2 or 3 levels; it means that sample points are set at 2 or 3 different values of each design variables);
- determine the initial (base) value; the chosen point must be near to the expected optimum and should not be at the border of the searching domain;
- set the size of levels; it must be within the searching domain (not too big) and exceed the error limits (not too small);
- eliminate dimensions;
- setting combinations and design matrix (it determines the quality of our model);
- execute the experiments (in the case of real experiments they should be made randomly, in the case of numerical experiments it has no relevance), control them and redesign if required;
- establish mathematical model (response surface);
- control the model (fitting to the effective data);
- verify the significance of coefficients (if every term is needed);
- re-dimensioning;
- interpret the model:
    - e.g. if a design variable rises the value of objective but it acts in the model with negative coefficient, there must be a numerical failure or our assumption were false;
    - if only a part of the coefficients are significant, probably the interval was wrongly selected: in that case we can enlarge the interval, change the initial point, or expand the experiments;
    - if none of the coefficients is good, probably the interval of the experiments is too narrow.

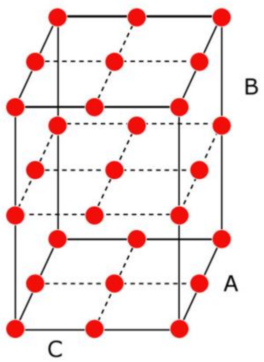


Figure 12.2. Design of full 3 factorial 3 levels

DOE methods can be divided in two main classes: traditional DOE methods based on statistical criteria as full or fractional factorial design, central composite design, random and Latin hypercube design, Placket-Burman, Box-Behnken and Taguchi design. In the other group belongs the optimal DOE methods based on mathematical optimality criteria as I-Optimal, D-Optimal and A-Optimal designs.

12.2.3. Full factorial design

In this experiment all of possible combination will be realized (Figure 12.2.).

Design matrix of experiment
This matrix contains all of the possible combinations of experiments for all levels of factors so all of the possible setting of experiments. The quality of model depends on the characteristics of the design matrix, independent from the number of factors. For example 2^N type full factorial design has the following properties:

- Symmetry: the sum of the elements in each column of factors is null

$$\sum_{j=1}^{N} x_{ij} = 0,\ i = 1 \dots k; \tag{12.1}$$

- Norming condition, namely the square sum of elements in each column equals the number of experimental settings.

$$\sum_{j=1}^{N} x_{ij}^2 = N,\ i = 1 \dots k; \tag{12.2}$$

- The matrix is orthogonal. Scalar product of any two vector column equals zero

$$\sum_{j=1}^{N} x_{ij} x_{lj} = 0, \quad i, l = 1 \dots k \ \ \mathrm{i} \neq \mathrm{l}; \tag{12.3}$$

- Rotatable. The goodness of the prediction is the same in the same distance from the center of the experiment and does not depend on the direction.

The number of possible effects equal to the setting of experiments of a full factorial design namely *2N*. Orthogonality of matrix permits the estimation of coefficients independent of each other in the case of our model has only linear effects and interactions.

12.2.4. Box-Wilson method

Box–Wilson optimum seeking method is an interactive procedure for finding the optimum of a response surface by

- using factorial or fractional factorial experiments to find the best way to change the levels of the factors to search out the region which is close to the optimum,
- using RSM to incorporate curvature into the surface and help you decide whether you have reached the optimum.

Procedure for Box-Wilson method:

- Use a first order model (factorial experiment or fractional factorial) in the neighborhood of the current conditions,

- Test for lack of fit,
- If no significant lack of fit, then locate path of steepest ascent,
- Run a series of experiments along path until no additional increase in response is evident (This a one dimensional search procedure),
- Repeat steps 1 – 4,
- If lack of fit is present, then use response surface design to investigate curvature,
- If curvature is present, use RSM to locate the optimum (either graphically or by setting derivatives = 0, beware of saddle points!),
- Once a maximum has been found, make sure that all excursions from the point result in decreased function values (sensitivity analysis).

The factors have only 2 levels in Box-Wilson method; we are able to compose a linear model on that way. Number of factor is denoted by N. The number of all possible combination of factors is 2N. This model can be appied if

- the response surface is analytical (namely it can be expanded in series around arbitrary point in the feasible domain),
- it has only one local extreme.

This type of experiment has two main version full and fractional factorial design (f the number of combinations in a full factorial design is too high).

12.2.5. Central Composite Design

Central composite design is an experimental design, for building a second order (quadratic) model for the response variable without needing to use a complete three-level factorial experiment.

The design consists of three distinct sets of experimental runs:

- A factorial (perhaps fractional) design in the factors studied, each having two levels;
- A set of center points, experimental runs whose values of each factor are the medians of the values used in the factorial portion. This point is often replicated in order to improve the precision of the experiment;
- A set of axial points, experimental runs identical to the centre points except for one factor, which will take on values both below and above the median of the two factorial levels, and typically both outside their range. All factors are varied in this way.

For example, in a 2 Factor Central Composite Design, each factor has 5 levels (Figure 12.3.):

- extreme high (star point),
- high,
- center,
- low,
- extreme low (star point).

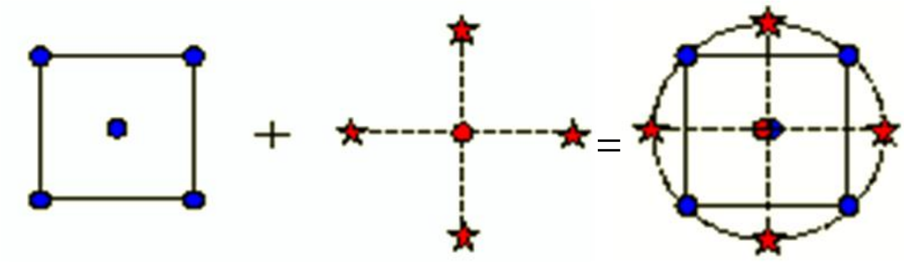


Figure 12.3. Central composite design for 2 factors

Orthogonal CCD's can be constructed by taking $\alpha_1 = \alpha_2 = \ldots = \alpha_n$ and suitably choosing α . Here α is the distance from the star points to the center point. (All star points lie a specific equal distance from the center of the circumscribing sphere.) Orthogonal CCD's assure no correlation among the effects being estimated. The value of α depends on whether or not the design is orthogonally blocked. That is, the question is whether or not the design is divided into blocks such that the block effects do not affect the estimates of the coefficients in the second order model.

12.2.6. Random and Latin hypercube design

Latin hypercube sampling (LHS) is a statistical method for generating a distribution of plausible collections of parameter values from a multidimensional distribution. In two dimension it is the Latin square, namely a $n \times n$ square, filled with different elements and each element occurs in one row and one column exactly only once. The popular Sudoku puzzle is such a sampling. The expression traces back to Leonhard Euler, as he used Latin letters on such a way. A Latin hypercube is the generalisation of this concept to an arbitrary number of dimensions, whereby each sample is the only one in each axis-aligned hyper plane containing it.
When sampling a function of N variables, the range of each variable is divided into M equally probable intervals. M sample points are then placed to satisfy the Latin hypercube requirements; note that this forces the number of divisions, M, to be equal for each variable. Also note that this sampling scheme does not require more samples for more dimensions (variables); this independence is one of the main advantages of this sampling scheme. Another advantage is that random samples can be taken one at a time, remembering which samples were taken so far.
The maximum number of combinations for a Latin hypercube of M divisions and N variables (i.e., dimensions) can be computed with the following formula:

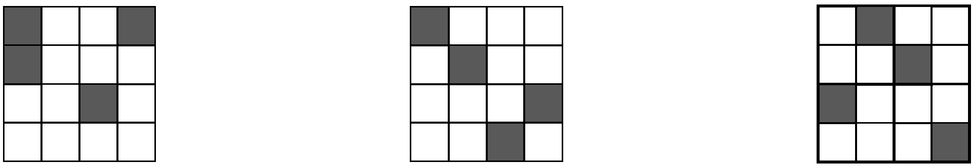
$$\left(\prod_{n=0}^{M-1}(M-n)\right)^{N-1} = (M!)^{N-1} \qquad (12.4)$$

For example, a Latin hypercube of M = 4 divisions with N = 2 variables (i.e., a square) will have 24 possible combinations. A Latin hypercube of M = 4 divisions with N = 3 variables (i.e., a cube) will have 576 possible combinations.
Orthogonal sampling adds the requirement that the entire sample space must be sampled evenly. Although more efficient, orthogonal sampling strategy is more difficult to implement since all random samples must be generated simultaneously.

In two dimensions the difference between random sampling, Latin hypercube sampling and orthogonal sampling can be explained as follows:

- In random sampling new sample points are generated without taking into account the previously generated sample points. One does thus not necessarily need to know beforehand how many sample points are needed.
- In Latin hypercube sampling one must first decide how many sample points to use and for each sample point remember in which row and column the sample point was taken.
- In Orthogonal Sampling, the sample space is divided into equally probable subspaces, Figure 12. showing four subspaces. All sample points are then chosen simultaneously making sure that the total ensemble of sample points is a Latin hypercube sample and that each subspace is sampled with the same density.

Thus, orthogonal sampling ensures that the ensemble of random numbers is a very good representative of the real variability; LHS ensures that the ensemble of random numbers is representative of the real variability whereas traditional random sampling (sometimes called brute force) is just an ensemble of random numbers without any guarantees (Figure 12.4).



random sampling Latin hypercube sampling orthogonal sampling

Figure 12.4. Latin hypercube designs

12.2.7. Box-Behnken design

Box–Behnken designs are experimental designs to achieve the following goals:

- each factor, or independent variable, is placed at one of three equally spaced values (at least three levels are needed for the following goal);
- the design should be sufficient to fit a quadratic model, that is, one containing squared terms and products of two factors;
- the ratio of the number of experimental points to the number of coefficients in the quadratic model should be reasonable (in fact, their designs kept it in the range of 1.5 to 2.6);
- the estimation variance should more or less depend only on the distance from the centre (this is achieved exactly for the designs with 4 and 7 factors), and should not vary too much inside the smallest hypercube containing the experimental points.

The design with 7 factors was found first while looking for a design having the desired property concerning estimation variance, and then similar designs were found for other numbers of factors.
Each design can be thought of as a combination of a two-level (full or fractional) factorial design with an incomplete block design. In each block, a certain number of factors are put through all combinations for the factorial design, while the other factors are kept at the central values. For instance, the Box–Behnken design for 3 factors involves three blocks, in each of which 2 factors are varied through the 4 possible combinations of high and low. It is necessary to include centre points as well (in which all factors are at their central values).

Table 12.2 summarizes the characteristics of the design for different factors;, m represents the number of factors which are varied in each of the blocks. There are further designs, which differs from this table, for example for 16 factors with just 256 sample points, up to 21 factors.

factor	*m*	number of blocks	factorial pts. per block	total with 1 centre point	typical total with extra centre points	no. of coefficients in quadratic model
3	2	3	4	13	15, 17	10
4	2	6	4	25	27, 29	15
5	2	10	4	41	46	21
6	3	6	8	49	54	28
7	3	7	8	57	62	36
8	4	14	8	113	120	45
9	3	15	8	121	130	55
10	4	10	16	161	170	66
11	5	11	16	177	188	78
12	4	12	16	193	204	91
16	4	24	16	385	396	153

Table 12.2: Box-Behnken design

12.3. Metamodels

12.3.1. Response surface method

Essence of response surface method is to optimally fit an approximate response function to a series of sample points. In a smaller part of the design domain a linear

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_N x_N \qquad (12.5)$$

or a quadratical

$$y = \beta_0 + \sum_{j=1}^{N} \beta_j x_j + \sum_{j=1}^{N} \beta_{jj} x_j^2 + \sum_{i=1}^{N} \sum_{j=1}^{i-1} \beta_{ij} x_i x_j \qquad (12.6)$$

polynomial approximation is satisfactory where $x_1, x_2, \ldots, x_n$ are the factors or design variables. Determination of the unknown coefficients made with the function values in the sample points

$$\beta_0 = \frac{\sum_{i=1}^{N} y_i}{N} \qquad (12.7)$$

$$\beta_j = \frac{\sum_{i=1}^{N} x_{ji} y_i}{N} \qquad (12.8)$$

In the case if we have dual interaction term (it is no use to deal with higher ordwer interactions) the coefficients can be computed as follows

$$\beta_{kl} = \frac{\sum_{i=1}^{N}(x_k x_l) y_i}{N} \qquad (12.9)$$

In order to ensure the accuracy of the approximation, we can use adaptive response surface method. For contoll the response surface we compute the function value in additional points until the response surface model become accurate enough or we achieve the limit of allowable number of computations.

12.3.2. Kriging metamodel

Kriging is a group of statistical techniques to interpolate the value of a random field. Kriging metamodel has greater attention to date for engineering applications because is suitable for approximation of highly nonlinear functions even with uncertainty. The name derives from D. G. Krige who applied it for mining data. The method was applied to engineering design problems in 1989.

The method belongs to the family of linear least squares estimation algorithms. As illustrated in Figure 12.19, the aim of kriging is to estimate the value of an unknown real-valued function *f*, at a point, x * , given the values of the function at some other points, x_1, x_2, ...x_n . A kriging estimator is said to be linear because the predicted value $\hat{f}(x^*)$ is a linear combination that may be written as

$$\hat{f}(x) = \sum_{i=1}^{n} \lambda_i(x) f(x_i). \qquad (12.10)$$

The weights λ_i are solutions of a system of linear equations which is obtained by assuming that *f* is a sample-path of a random process *F(x),* and that the error of prediction

$$\varepsilon(x) = F(x) - \sum_{i=1}^{n} \lambda_i(x) F(x_i) \qquad (12.11)$$

is to be minimized in some sense. For instance, the so-called simple kriging assumption is that the mean and the covariance of *F(x)* is known and then, the kriging predictor is the one that minimizes the variance of the prediction error (Figure 12.19).

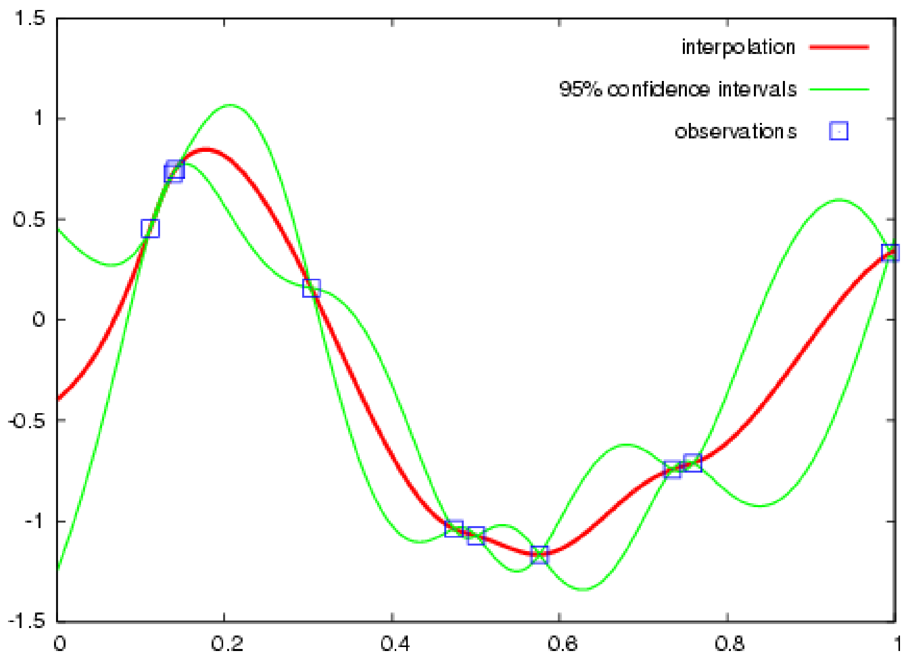


Figure 12.19. One dimensional kriging interpolation with confidence intervals

12.3.3. Metamodel with radial basis function

Radial basis functions are a popular family of methods for multidimensional interpolation problems. The base of the methods is to order a basis function for every sample points and the fitting occurs as a linear combination of them (Figure 12.6.):

$$f(x)=\sum_{j=1}^{N}\lambda_j\Phi_j(x-x_j)\quad, \tag{12.12}$$

where $\Phi_1,\Phi_2,...,\Phi_N:\mathbf{R}^2\to\mathbf{R}$ are given radial or central symmetrical basis functions, that is $\Phi_j(x)$ depends on only from the norm $r:=\|x-x_j\|$. The unknown $\lambda_1,\lambda_2,...,\lambda_N$ coefficients can be determined from the interpolation equations:

$$\sum_{j=1}^{N}\lambda_j\Phi_j(x_k-x_j)=f_k\qquad(k=1,2,...,N)\,, \tag{12.13}$$

where $f_k=f(x_k)$ are the function values in the sample points. Φ_j radial basis function can be chosen on diverse ways as for instance by

multiquadratical method :

$$\Phi_j(r):=\sqrt{r^2+c_j^2}\;, \tag{12.14}$$

where $c_1,c_2,...,c_N>0$ are suitably chosen scaling parameters, as for example:

$$c_k:=\min_{j\neq k}\|x_k-x_j\|\;. \tag{12.15}$$

thin plate method:

$$\Phi_j(r) := r^2 \log r \tag{12.16}$$

Gauss-functions:

$$\Phi_j(r) := e^{-c_j^2 r^2}, \tag{12.17}$$

($c_1, c_2, \ldots, c_N$ scaling factors, too).

Functions used in dual reciprocity:

$$\Phi_j(r) := 1 + r \tag{12.18}$$

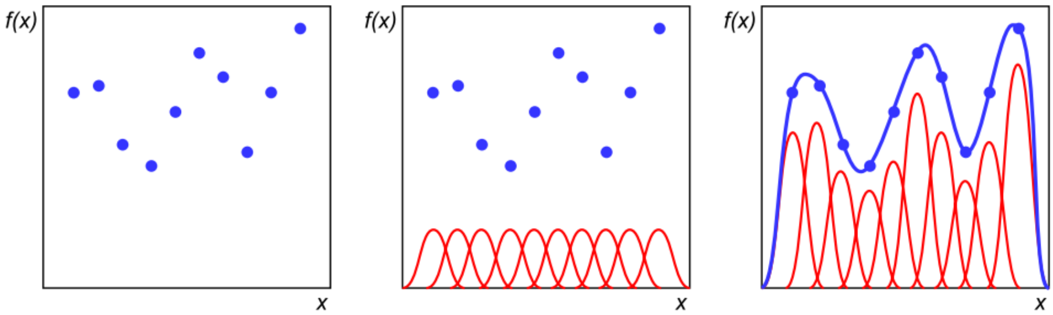


Figure 12.6. Function approximation with radial basis function

12.4. Neural networks [12.4]

Neural networks are also computing methods suitable for approximation of an unkown function or fro optimization. They are built up from parts execute similar operation (called neuron) which allows a parallel distributed computation and they are able to learn and use the acquired knowledge. Output of a neuron can be linear combination of input data and given weights, value of the so called activation function (Figure 12.7.):

$$y = f(x) = \sum_{i=1}^{n} w_i g_i(\mathbf{x}). \tag{12.19}$$

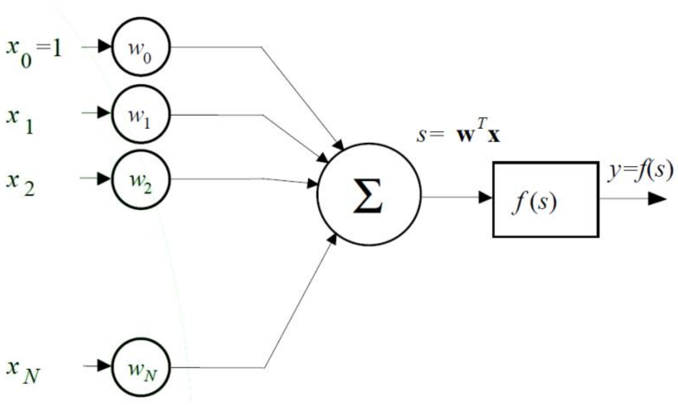


Figure 12.7. Operation of a neuron with activation function

Neuron in a network usually is bounded to neurons of a specific number and these connections are one way, hence networks are divided to different layers accordingly to the bindings.

Each neuron in a layer is bounded to the output of the previous layer, respectively to the input of the next layer. In the most cases there is no feedback between the individual layers: then the network is called feedforward, in opposite cases it is a feedback network. There are more cases for feedback, according as from where the feedback was made:

- **elementary feedback** is when the output of a neuron on a hidden layer can be appended to the own input,
- **feedback between layers** is when it happens to the input of the previous layer,
- **lateral feedback** is when the feedback is made to the input of another neuron within the actual layer.

Consequently, a neural network can be built up in many ways and it can be learnt with diverse methods. These properties guarantee the wide application of them as they can be changed dynamically, depending the structure and the teaching method.

The neural network are designed on the following way:

- collecting representative observations (inputs and expected outputs),
- choosing the adequate neural paradigm specified to the task,
- choosing the system parameters (transfer function, number, teaching method and parameters, initial weightings),
- teaching and testing the system (until we have success or give up).

In the course of teaching process the weightings are varied according to satisfy an error criteria; it should be noticed that it is also an optimization process itself for what traditional optimization methods can be used.

12.5. Global optimum [12.5]

In many cases we are not settled for a good result but we need to know the very best one to the given condition. Searching the global optimum such methods are relevant which can be implemented on computers, find the answer in a given tract of time and are reliable. Regarded to reliability and effectivity optimization methods can be classified on the following way:

- **Incomplete methods** are based on heuristics. Here we don't have any guarantee not to be trapped in a local solution and also do not have information how far we are from the global minimum. For that reason the stopping criteria are also heuristics. As in Table 12.1 is shown, it is an intuitive way to avoid local solution if we start the search from more initial points. Another way is to apply such a method which is able to jump out from a local optimum as for instance simulated annealing.
- **For asymptotically complete methods** it can be proved they are able to find the global optimum with a certain probability, provided unlimited running time and suppsing a prescribed tolerance. Such methods are genetic algorithms and other probabilistic evolutionary algorithm. Stopping criteria is here also a heuristic one as these methods don' have information about whether they have found the global optimum.
- **Complete methods** are able to find the global optimum within the given tolerance in a predictable time limit, supposing exact arithmetic.
- **Rigorous methods** are complete methods able to find the global optimum with a given tolerance even in existence of rounding error.

The full search of the design space can not made in due detail, especial when the responses are calculated with finite element simulation. Applying surrogate models and branch and bound methods (B&B) together can be a suitable methodology to find the global optuimum of engineering systems. Essence of branch and bound is to divide the search space to subproblems recursively (this is branching) and set upper and lower limits on the possible values of the objective function (this is bounding) so the parts not having better solution as the known best one until that time can be eliminated.
More details on these global optimization methods can be found in Global optimization algorithms - Theory and application.

12.6. Multidisciplinary optimization

To date, design engineers have to come to reasonable decision by solving more and more complex problems. Very often different disciplines act in the same time, though treating one after the other manage not necessarily to good results. Without taking interaction into account having a coupled problem, the solution will be poor.
Many methods have been developed to solve multidisciplinary problems. The one level methods as Multidisciplinary Design Feasible (MDF) and Simultaneous Analysis and Design (SAND) cab be simply appied for smaller problem but they are not very suitable to solve big problems or problems with more disciplines. A very often industrial practice is when more group deal with the problem in different point of view. For this type of problem solution methodology the multilevel methods can be adapted beneficially as the Collaborative Optimization (CO), Concurrent Subspace Optimization (CSSO) and the Bilevel Integrated Systems Synthesis (BLISS). The methods can be hard compared in respect of efficiency because it performance can depend on problem itself or implementation even it can be occurred that for a specific problem a method is not able to find any solution.

12.7. Robust design

Variability, insecurity, tolerance and error of engineering systems have great role in the product design process. These can be caused by manufacturing tolerance, process instabilities, environmental effects, wear and human factors and so on. They can be described by stochastic distribution. Deterministic simulation is not able to predict the behaviour of a real system with uncertain inputs because the calculation is accomplished in a single design point. Rather a probabilistic simulation is needed in which the distribution of output data is calculated due to the distribution of inputs, for arbitrary deterministic simulation model. From the output distribution we can deduce for the real behaviour.
To date, in the hard competition, products are set up to the limits its' operating conditions. The variability of the parameters can often lead on failure of the structure. Reliability analysis examines the violence of limit conditions due to variance of input variables. A product is reliable if the deviation is within the limit surfaces. Robust optimization aims at developing a solution that has a good value, is insensitive to variations of the nominal design and is feasible in an uncertainty range around the nominal design. They are applied rather in an early phase of the design process, helping to evaluate the conceptions; in the final design stage, for every parameter it would be too costly. The design is held robust if the sensitivity of the objective function is the smallest to the material, manufacturing and operational changes.
As shown in Figure 12.8. , the x-axis represents the uncertain parameters, including design variables (control factors) and noise factors (uncontrollable factors), while the vertical axis represents the objective function *f(x)* to be minimized. Of these three solutions 1, 2, and 3

pointed, solution 3 is considered robust as a variation of $\pm\Delta x$ in design variables does not alter the objective function too much and maintains the solution within the design constraint when the design variable is perturbed. Although Solution 2 is also within the design space when the design variable varies in $\pm\Delta x$, the perturbation causes a larger change in objective function. Solution 1 is highly sensitive to the parameter perturbation and usually cannot be recommended in practice, though it has the best mean value of all the three solutions.

Real-life engineering problems are typically characterized by a number of quality and/or performance indices, while some of which could be conflicting with each other. To address such a multi-objective robust optimization problem, some attempts have been made.

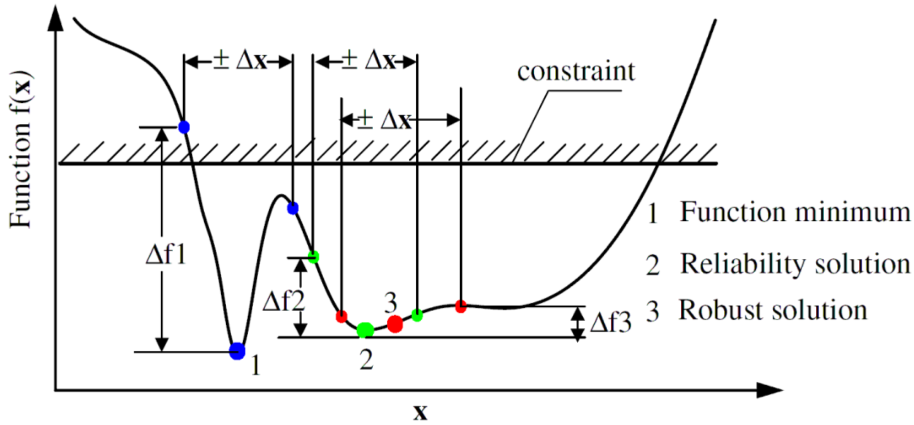


Figure 12.8. Robust optimization [12.6]

In the following, a possible way for robust design will be shown on an example.

Sun et al. [12.6] performed an impact optimization of a vehicle. When car crashing occurs, it is expected that most of impact energy is absorbed by the vehicle structure to reduce risk to occupants. However, increase in energy absorption capacity often leads to unwanted increase in structural weight. Furthermore, the deceleration peak should be restricted to a certain level, for instance, 40 g (g = 9.81 m/s2) for crashworthiness design.

Design variables were thicknesses of the 3 front brackets (Figure 12.20.).

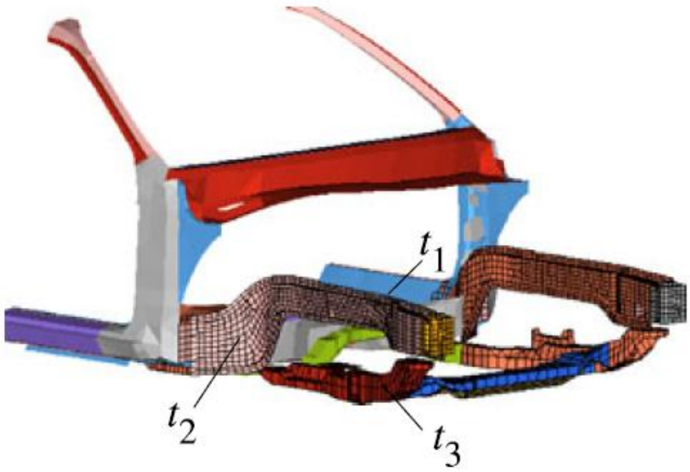


Figure 12.20. Design variables at crashworthiness optimization of a vehicle

The material properties, such as Young's modulus E, density ρ, yielding stress σ_y can be affected by the manufacturing process whose fluctuation are respected (noise parameters). Surrogate models were developed: the mean value and standard deviation of energy absorption U and deceleration peak a were approximated quadratic, surrogate model for mass is enough to be linear as it depends linearly from the thickness parameters. Othogonal Latin hypercube sampling and orthogonal design were integrated to perform DoE analysis. The noise factors sampled with orthogonal design were arranged in an outer array with the sample points of 4, and control factors sampled with OLHS were arranged in an inner array with the sample points of 16, between 1 and 2 mm. Experiments in the inner array are repeated at 4 points corresponding to the outer array to simulate the variability due to the uncertainties of the three noise factors, making the total number of simulations equals to 64. One single crash simulation with LSDYNA for this problem takes 3 hours, so the numbers of simulation should be limited. Next step was to evaluate the accuracies of the surrogate model. The total variance of regression (sum of squares total =SST) can be given with equation (12.20)

$$\mathrm{SST}=\sum_{i=1}^{n}(y_i-\bar{y}_i)^2. \tag{12.20}$$

The deviation of predicted value from its mean value can be divided to deviation from the regression and with regression not explainable parts:

$$(y_i-\bar{y}_i)=(y_i-\tilde{y}_i)+(\tilde{y}_i-\bar{y}_i), \text{így} \qquad (12.21)$$

$$SST=\sum_{i=1}^{n}(y_i-\tilde{y}_i)^2+\sum_{i=1}^{n}(\tilde{y}_i-\bar{y}_i)^2=\mathrm{SSR}+\mathrm{SSE},$$

where SSR a regression variance, SSE is the rest, unexplained variance.

Determination coefficient R^2 gives which fraction of total variance can be explained with regression:

$$R^2=\mathrm{SSR/SST=SST\text{-}SSE/SST=1\text{-}SSE/SST}=1-\frac{\sum_{i=1}^{n}(\tilde{y}_i-\bar{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y}_i)^2} \tag{12.22}$$

From table 12.3, the accuracies of the dual response models are adequate and allow us to carry out the design optimization properly.

	U_μ (%)	$U\sigma$ (%)	M_μ (%)	$M\sigma$ (%)	a_μ (%)	$a\sigma$ (%)
R^2	98.92	99.70	99.99	99.69	98.67	99.95

Table 12.3. Error analysis of response surface models with determination coefficient

After introducing the sigma criterion (η = 3, 6, respectively herein repectively for manufacturing tolerance λ=0.1 *és* λ=0.01), the multiobjective robust optimization for energyconsumption maximization and minimization material usage is thus formulated as (12.23)

$$\begin{gathered}\min(f_1,f_2) \quad f_1=-\lambda U_\mu^2+(1-\lambda)U_\sigma^2,\, f_2=\lambda M_\mu^2+(1-\lambda)M_\sigma^2\\ a_\mu+\eta a_\sigma\le 40\\ t_i^A+\eta t_{\sigma i}\le t_{\mu i}\le t_i^F-\eta t_{\sigma i},\quad i=1,2,3\end{gathered} \tag{12.23}$$

First the deterministic multiobjective optimization was solved without considering the perturbations of design variables and parametric noise for given value of weight, acceleration constants and algorithm parameters with the multiobjective particle swarm optimization (MOPSO). In order to take into account the uncertainties, the design variables are assumed to distribute normally, whose standard deviations are given as [0.01, 0.01, 0.01] from typical manufacturing tolerance. Figure 12.10. gives the Pareto optimal fronts for different sigma levels. A higher sigma level indicates that the perturbations of the design variables and parametric noise have a lower probability to violate constraint; however, the objective functions must sacrifice more.

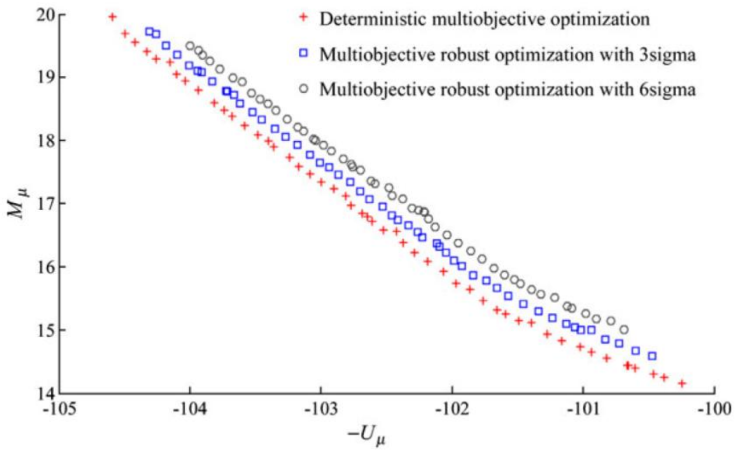


Figure 12.10. Effect of sigma criteria on the Pareto front

Figure 12.11. presents the optimal Pareto fronts for the deterministic multiobjective optimization (λ = 1) and the means of robust multiobjective optimizations with λ = 0.01 and λ = 0.1, respectively. It is interesting to note that the consideration of randomness of the parameters leads to sacrifice of Pareto optimum, i.e. the robust solution is farther to the origin in the Pareto space than the deterministic counterpart. Hence, compromise must be made between the robustness and nominal performance in practice. The optimized Pareto set, which takes into account the perturbations of the design variables and noise parameters, does not violate the constraint, as in Figure 12.11., so the all the solutions are reliable. Finally, decision must be made for the most satisfactory solution (termed as "knee point") from Pareto-set. Conventionally, the most satisfactory solution is often decided by taking the nearest point on the Pareto front to the utopia point as knee point for the conflicting objectives if any further economical aspect emerges to support the decision process(Figure 12.14.). As last step, we should control with the simulation program if our optimum got with the surrogate model is really good enough.

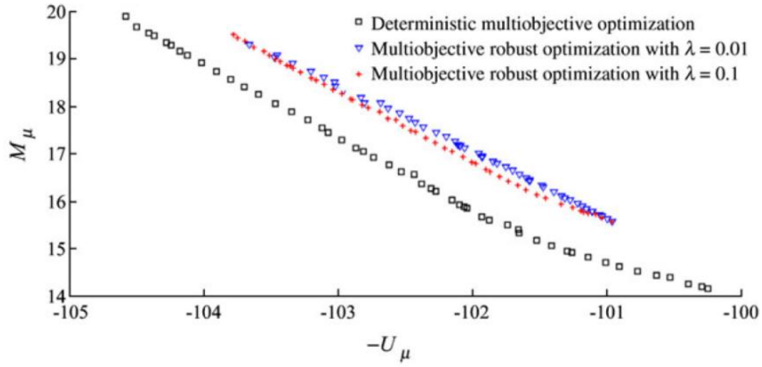


Figure 12.11. A deterministic optimum and Pareto fronts of robust design for the mean values

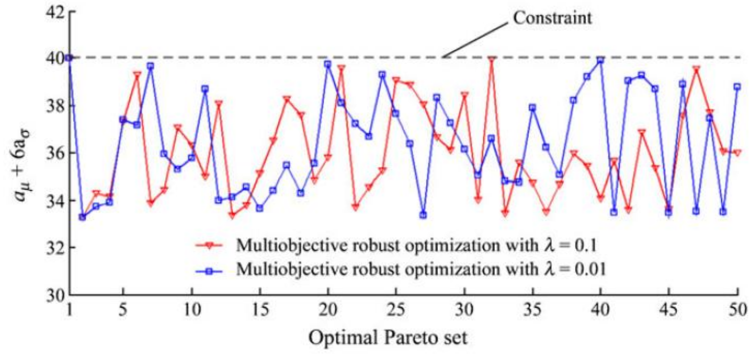


Figure 12.13. Optimization contstraints are fulfilled

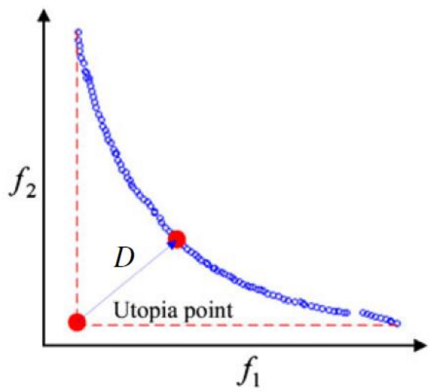


Figure 12.14. Selection of best solution from the Pareto front

To sum up, in this example a multiobjective robust optimization for vehicle design by using dual response surface model and sigma criteria was presented, which allows taking into account the effects of system uncertainties on different objectives. The adoption of a multiobjective particle swarm optimization algorithm no longer requires formulating a single cost function in terms of weight average or other means of combining multiple objectives. The procedure proved fairly effective in a full-scale vehicle crashing model, where energy absorption and weight are taken as the objectives, while the peak deceleration as the constraint. The example demonstrated that the multiobjective particle swarm optimization generates the Pareto points fairly efficient and evenly. The comparison of Pareto optimums between the deterministic design and robust design clearly indicated that improvement in the robustness must sacrifice the Pareto optimum of mean objectives. For this reason, a weight factor can be prescribed to balance the importance between deterministic design and robust design. The more the emphasis on robustness (a higher sigma level), the worse the objective means. The example showed that the energy absorption and weight of the car were improved, at the same time; the robustness of the design is enhanced.

12.8. Dynamic problems

Optimization problems with transient loads can be solved with the method of static equivalent load. The optimal solution is searched through a series of static optimization which means a two level (outer and inner) iteration with the following steps:

- determine the initial design,
- dynamic analysis,
- static optimization with more loadcases of equivalent loads,
- reaching the optimum we stop, else continue from the second step.

Loadcases of static equivalent load correspond to the time steps of dynamic analysis.

12.9. Acceleration of computation

Optimization methods are iterative process from its nature so they have high computation costs. In order to raise the computational effectivity, parallelization is one common approach. Another way is to adapt the computational accuracy to the actual demand.

Basic idea of dynamic programming is to divide a complex problem into more subproblems and from the solutions of the subproblems we can deduce to the original one's. If most of the subproblems are fast the same we can spare many computation time. Procession of solution can be direct or recursive. Modern distributed computational systems (cloud method, multiprocessor systems) assist the solution in a short time of repeated tasks.

12.10. Summary

Engineering optimization problems can be multitude according to what design process is supported by them. Shape design problems are mostly complex, multivariate problems with uncertain variables. Solving then in a real time is possible if the response involved in objective function and optimization constraints (as displacements, temperature distribution, normal modes, stresses and so on) are evaluated with costly numerical simulation only at a limited number and we use surrogate models. This approximation can concern for the whole design domain or divided it into parts we can sweep them, without to be locked to a specific initial point, searching the global optimum. It can be made with interpolation, or we can compute the regression function minimizing he error of the regression function; also radial basis function can be applied.

We can verify its relevance and suit to our expectation thru error function and learning mechanism. Uncertainty of design variable can be followed with surrogate models and robust optimization problem can be treated as a multivariate optimization problem where the aim is to get the solution with the best mean value and with the most insensitive solution to the variability of the not controllable changes within a security domain.

12.11. Literature

[12.1] Ronald A. Fisher: The Design of Experiment. (1935)
[12.2] Jasbir Arora: Optimization of structural and mechanical systems, 2007, Chapter 16: T. H. 446 Lee and J. J. Jung: Kriging Metamodel Based Optimization
[12.3] Santner, T.J., Williams, B.J., Notz, W.I.: The Design and Analysis of Computer Experiment. Springer-Verlag New York 2003
[12.4] Christopher M. Bishop: Neural Networks for Pattern Recognition, Oxford University Press (1995), ISBN 0-19-853864-2
[12.5] R. Horst and P.M. Pardalos (eds.), Handbook of Global Optimization, Kluwer, Dordrecht 1995.
[12.6] Sun G,·Li G, Zhou S,·Li·H, Hou S, Li Q: Crashworthiness design of vehicle by using multiobjective robust optimization. Struct Multidisc Optim DOI 10.1007/s00158-010-0601-z

12.12. Questions

When is it worth to build a surrogate model for searching the optimum?
What effect has the choice of design variable on the optimization results?
How can we choose the proper design of experiment?
Outline the steps of building a surrogate model!
What kind of method is suitable for solving a transient dynamic optimization problem?
What is the purpose of the robust optimization?

13. SOFTWARE TOOLS FOR STRUCTURAL OPTIMIZATION

This chapter demonstrates the need and expectation of a design engineer for optimization software and what tools can be ensured to satisfy these demands. Some practical example will be shown for complex engineering problem. The main softwares, developers and distributors will be presented to the actual stand of market (however these dates are rapidly changing due to economical reasons). Further details about the capability of the programs can be found following the embedded links to them.

Several software tools support design optimization. These are developed by universities and industrial firms, sometimes with financial supports from the governments. The modern design optimization software are connected or embedded in finite element structural analysis softwares. Choosing an appropriate one it is crucial, what kind of expectation the designer have and what type problem should be solved in what circumstances. Naturally these softwares are permanently developing according to the newest stand of computational sciences and state of computer tools. For example, in the beginning of 90'th evolutionary techniques as genetic algorithms were treated suitable only for solving very simple problems as the coupled computational cost are very high, nowadays the rapid structural analysis, the greater computational capacity and the economical demands motivate the breaking through of global optimization procedures.
However, due to the need of specific knowledge of optimization theory the application rate of using optimization tools is low. For that reason developing follows the direction to create easy to survey, simple to use software tools, open for adaption to the special user needs and purposes occurring at the given firm or problem.

13.1. Design cycle, possibility of modelling

Demands of design optimization tools and role of modelling cannot be understood without the knowledge of design cycle (Figure 13.1.). In the conceptual phase different data are available depending on the characteristic of design problem.
Developing a new design there are no initial geometry. In this case the main and auxiliary functionality of the product are identified first and together with the examination of load transfer and analysing the concurrent products designer can identify the main geometrical elements and their role in the construction. So specification of a part construction arises, including the main expectation of load carrying capacity, geometrical sizes and connection to the rest of construction. Skeleton building facilities of CAD systems can be very helpful in this case.
If the designer task is to redesign an existing construction, probably a detailed CAD model is available or a physical model and with reverse engineering tools the detailed geometrical model can be built. Loads can come from measurements of similar product, or from a multibody simulation.
Depending on problem characteristics designer chooses the physical parameters describing the behaviour of the structure.
After model building a topology optimization step can help to identify the main behaviour; in simpler cases results of it can be directly applied as design proposal. In the most case topology optimization results can be multitude interpreted so designer must be included and more detailed conception should be constructed for further analysis about additional standpoints as manufacturability and secondary loading.

This refinement of the design follows in a shape optimization step, so local effects can be taken into account. The proposal for shape optimization should be prepared: some sizes must be enlarged to ensure a wider design domain for the searching algorithms.
Depending on design purpose shape parameterization and optimization can be performed on different ways. In the case of strong dynamical effects, local stress peaks should be decreased so a freeform parameterization is needed and there is not great geometrical changes expected. In other cases, if greater geometrical parts should be reformed, geometrical changes can be guided threw meshless techniques, mesh morphing or CAD spline parameters. Design becomes more detailed with the design progress, luckily with smaller and smaller changes. Getting no acceptable design proposal, initial specification must be modified and design process repeated.

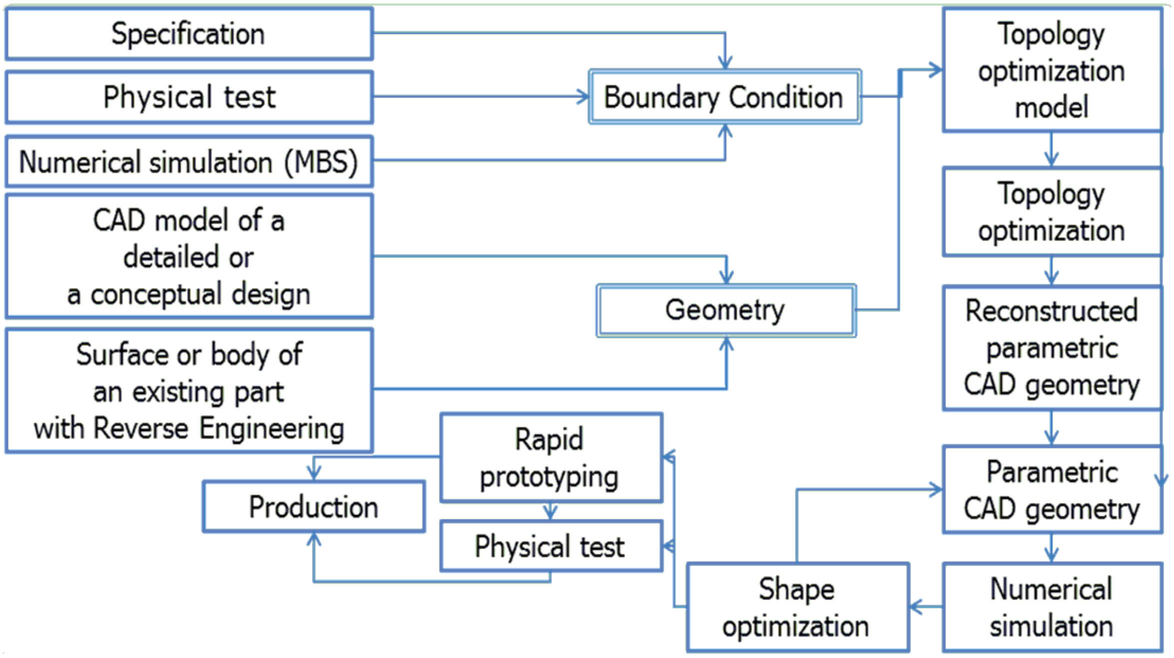


Figure 13.1. Design process

13.2. Structure of software, demands

The modern structural optimization systems utilize the tools of object oriented development and suits to the changing demands as chameleons. Their structure is not fixed as the predecessors' but a graphical interface supplies for the construction of the most suitable optimization process to solve the specific optimization problem (OPTIY). Another basic contradictory demand is the black box characteristic: designer should be able to use the system without special optimization knowledge and much training. The choice is always depend on the application environments, facilities, problem and demand. The basic designer need to an optimization tools can be the following: solution time is limited; DOE

- optimum should be approximated in different degree depending on the problem
- handling of special construction as shells, composites and so on, handling different design variables
- present the steps of the process and the optimum results, give aids to analyse the process

- the whole structure must be optimal, not the parts of it
- some components must stands special loads and circumstances
- it should give help to form a design conception and also the final details
- the optimum must be robust for a given tolerance of parameter changes
- should be suitable for multi-objective and multi constraint optimization
- no need for special skills of software usage (applying standards)
- restart possibility and analysis tools
- visualizing tools for solution
- …

13.3. Commercial software tools in the design optimization

Without being complete, the most spreaded optimization systems are the following (owner or developer can be found in the bracket):

- OptiY (OptiY e.K.)
- GENESIS, VisualDOC, DOT és BIGDOT (Vanderplaats Research and Development)
- MSC/NASTRAN (McNeal and Schwendler Corporation)
- OPTISTRUCT és Hyperstudy (Altair Engineering)
- ALFGAM (ALFGAM Optimering AB)
- TOSCA (FE-DESIGN)
- ANSYS (ANSYS, Inc)
- Pro/Engineer (PTC)
- SOLIDWORKS (Dassault Systèmes SolidWorks Corp.)
- iSIGHT (Dassault Systémes SIMULIA Corp.)
- Hypersizer (Collier Research Corporation)
- BOSS quattro (SAMTECH)
- FEMtools Optimization (Dynamic Design Solutions)
- HEEDS (Red Cedar Technology)
- IOSO (Sigma Technology)
- LS-OPT (Livermore Software Technology Corporation)
- modeFRONTIER modeFRONTIER (Esteco)
- ModelCenter (PhoenixIntegration)
- Optimus(Cybernet Systems ,korábban Noesis Solutions)
- CETOL 6 Sigma (Sigmetrix)

These softwares are partly standalone optimization softwares collecting different algorithms and with several interface to commercial finite element systems. Some of the is embedded in a structural analysis system as GENESIS or OPTISTRUCT, there are in CAD system embedded versions as in the case of SolidWorks and there exist some optimization software directory as , DOT, BIGDOT. To the special need for problem preparation and result procession there exist supplements to the pre/postprocessor of finite element softwares, partly new interface were developed supporting integration of different finite element modules solving multidisciplinary problems and monitoring probability based calculations.

13.4. Summary

Modern structural optimization systems can be sorted into two categories. Engineers with optimization knowledge can advantageously use general purpose systems

- offer different mathematical algorithms and simulation tools of a wide scale. and have graphical interface to adapt optimization process to the actual problem, chose and build the elements of the method, together with data connection;
- have an open systems giving possibility to build special program in script or other languages to solve special problems;
- they have built in methods for
    - handling multi-objective problems,
    - building metamodels of multidisciplinary and other complex problems with the aid of statistics and design of experiment,
    - searching robust and reliable optimum,
    - handling different types of design variables (size, shape topology, topometry, special parameters for composite optimization);
- they have wide range of graphical tools for present and monitor the design space, data connections and optimization results, they can connect arbitrarily to the designer preferred pre and prostprocessor systems for preparation of the structural analysis model;
- they can be installed on any computer platform, with different architectures;
- using the possibility of running environment rising the effectivity (parallel algorithms).

The other group of methods are able to solve a specialized problem with restricted toolkit but very rapidly. They are in usually integrated in CAD/CAE systems, providing effective help for design engineers having no special optimization knowledge.

13.5. Questions

What kind of demand has an engineer for a design optimization software?
What are the characteristics of the integrated CAD/CAE optimization methods?
Which optimization system would you use for solving a multiobjective optimization problem?

14. OPTIMIZING PRODUCTION PROCESSES

„The engineer's most important instrument are the paper the pencil and the eraser"
Quotation from professor Ádám Muttnyánszky:
Its time: 1952. mechanics lecture,
its place Technical University, Budapest Auditórium Maximum.

Preface

The three words quoted remained indelibly in my memory in spite of the fact that the content importance of exceptional sigfinicance of these words became unambiguously only during my later teaching, planning and scientific work. Surely the creative engineer's primary dialogue partner are „the paper, pencil and the eraser" even if I can hear the objection of those who set the nearly immeasurable mine of the possibilities given by computer with the view accepted by me, too. The recognition of the problem originates in our brain and not in the computer. The computer can do that very quickly as against to which it was taught by its constructors. The pencil constitutes the instrument of our ideas in the process of realizing our thoughts in majority cases and only after comes the picture screen. I recommend the above thoughts to all engineering students aspiring to get creative knowledge during their studies for the sake of getting proper knowledge in the fields to recognize and to establish the ability of solving the problem.

14.1. Introduction

According to the special literature [14.1] the scientific foundation of production organization (production management) can be fixed to the end of XIX. century and the first period of XX. century.
Disregarding from detailed survey of the historic events it is important to emphasize those persons – belonging to the period – who so called laid the foundation of the production organization:
- Eli Whitney [14.1] laid the foundation of the idea of the mass production by introducing the interchangeable planning and manufacturing of parts in the arm production.
- Frederick W. Taylor [14.1] carried out sussesful research work concerning the production organization, the operation analyses, the working methods and capacity increase.
- Henry L. Gantt [14.1] plotted such clear diagrams with consecutive organization of the process elements or with their interchangeability among them, on which the place, the condition of preparedness and the time of probable finish of the product can be well followed.
- Henry Ford recognizing the posibility to be in the theoretical basis introduced the straight - line production by which it can be significantly decrease the outlet cycle of time and the market price of the product. It can arise in the reader that the questions listed how connected to optimization. Under optimization in the reality we are looking for not only the maximum or a minimum of a process, but a better, a more favourable solution is also convenient. However in the beginning of the industrial revolution these had got very great importance and nowdays just it is increasing the significance of this attitude to which Ferenc Erdélyi calls attention to

more than very remarkebly in this article “some technological relations of the global crisis.” [14.2].

It has to be mentioned by all means, that Walter Rautenstrauch [14.1] invented the so called cost – covered diagram to qualify the economy of the production process, containing the main componens (fixed costs, variable costs, revenue ..., production volume) of the production process. Based on this the financial situation of the company (process respectively) can be demonstrated in the function of production and what is the most important the amount of profit how is formed (Figure 14.1.)

To study these question very abundant mine of special literature dealing with production is at disposal for the inquirers.

The different production philosophies “satisfactory, optimizing and adaptive” can be studied in very detail in Achoff’s [14.3] book. He points at that to reveal the different questions, problems are unavoidable the immersed knowledge of interdiscipline (logic, mathematics, statistics, physics, mechanics) sciences. The decision made on logical (common sense) basis can be sufficient naturally in simple cases but in case of more complex processes it is needed to do appropriate models. Nowdays the computer aided planning and simulation programs offer very wide – spread varieties shortening the solution of the problem given.

It can be said generally that the optimization, optimum – search is nearly a part of everyday life, only not so form drawn up scientifically but among the manifold executive possibilities of different tasks, naturally on logical basis the most favourable solution is looked for. Accordingly the optimization respectively optimum-search fills a decisive part in organizing our activity process during solving different tasks in majority cases consciously or unconsciously. Naturally this appears not so that we consider thorougly the succession of functional elements needed to solve the task given in every cases and their interchangeability but just we consider on a large scale the executive sequence.

Further I wish to present the process of task organization, optimizing, optimum-search on such examples where it can be followed a favourable or optimal (maximum or minimum) solution of the task by still so called “paper, pencil and eraser” logical steps. The “paper, pencil, eraser” instrument triplicate is the most suitable to represent our thoughts, intuitions then after this can follow the reasonable application of nowdays instrument depositories (computers).

Our aim is basically to make yourselves master of production organization, of the process of thinking philosophy of optimization and to form an opinion of the range of basic instruments needed to determine these and of their practical use.

Let’s take for example a simple everyday case: *x* wants to travel from Budapest to Nyíregyháza. It can be considered several possibilities practically:

a./ by train,
b./ by bus.
c./ by own car,
d./ by taxi,
e./ etc.

The object of decision consideration can be certainly, that:

- at what time it is required to be there,
- at what time it is required to get up,
- by which vehicle of transport is the travel done within the shortest time, on the other hand if there is no own car or the taxi is too expensive then these possibilities don’t constitute the object of consideration,
- which travel is the least tiring,

- is it needed to review the possible material of discussion during travel,
- which travel is the cheapest,
- etc.

The task is to determine beside the most favourable – more exactly according to some kind of consideration the most favourable – solution, that has to be determined that standpoint which seems optimal by means of neglecting the other certain competent solution. Let’s suppose that the shortest time is set as a goal and so it is natural that it has to be decided beside the own car or taxi. The final decision would be determined whether *X*-likes to drive whether the driving makes tired, whether the cost of taxi is too expensive.

Really analysing to the end all selection possibilities it would be discovered that against the others there would also be advantages and disadvantages, too. The appreciation standpoints can be objectives, and subjectives, too.

Task:

Analyse, that choosing one-one possibility what advantages respectively disadvantages has got against the others.

Let’s take the shopping belonging to the housewife everyday tasks, who after deciding what would she cook for dinner, starts in that shop where she can presumably buy everything according to the necessaries. It is only natural that she grasps automatically the bag needed to shopping, money maybe basket and starts to that vehicle by which she approaches the shopping centre. At shopping she starts almost with logistical punctuality in that sequence with collecting merchandise that by the time she reaches the pay-desk all merchandise should be in her basket what she needs. The rarest case is that she should go back for something and by this to do unnecessary way. The housewife didn’t do else as she unconsciously optimize that way she had to covered to buy the necessary merchandise.

Let’s take a transport example, belonging similarly to among the daily activities surely it can be met a fair number of lorries, camions delivering merchandise on the roads. Let’s suppose that case – taking place also in reality – a camion delivers valuable goods from Budapest to Hamburg and 30 hours are at disposal to cover the distance. Inasmuch the goods arrives late after each hour 1000 EU liability for damages charges the deliverer. After coming to an end of one obligatory resting hour the engine can’t be started because of the failure of starter. According to the service station operating beside the filling station about two hours are needed for repairing. According to practical experiences to tell the time needed to repairs in advance is possible only with great uncertainity as might be ±50% differences but that can also happen the damaged element can’t be repaired. The price of the new starter is 850 EUR the disassembly and the installation 50EUR the repair cost probably 250 EUR. In that case that the express messenger should take reserve starter from Budapest it is out of question because of the long distance (about 800 km). In such or similar cases it has to be considered the cases that can be taken into account and on the basis of justifiable standpoints and after it is needed to choose the most favourable solution.

The consideration standpoints are summed up in Table 14.1.

Table 14.1

Renownable consideration possibilities			
Consideration alternatives	**Costs**	**Increase of delivering time coming about from repair time**	**Emerged risks**
Repair is choosen	250 EUR	2±1 hours	- the repair is not perfect, and the starter breaks down again, - if the repair takes up 3 hours and it still has to be calculated possible traffic jams then the delivery within term is already endangered - the risk thus is rather significant
New starter is installed	900 EUR	1 hour	- the 1 hour deficiency doesn't mean practical risk, - the new starter however increases the safety operation

Considering the cost listed in the Table, the increase of the expected delivery time as well as the possible uncertainities, it is probable that the camion driver will decide beside installing new starter. It is true that to restore the working condition of the camion costs with 650 EUR more but so the delivery of goods is accomplished within date. The customer's confidence wasn't hurt against the deliverer to be basically important is business world. The driver could calmly devote his all attention to the traffic surely the subconscious thought didn't disturb him that operation break down at the starter when could again happen

The short time of repair didn't cause trouble to keep the resting time prescribed, moreover also remained spare time to the shortage of time coming about possible traffic jam.

It can be well perceived on the whole, this multiple analyses of optimizing the comparatively simple delivery process mentioned in the example by which it can be chosen the decision concluded more favourable. The optimization tended finally also to that the customer's confidence didn't damage guaranteeing with this to maintain the long term partner co-operation and the safety of favourable business relation. It can be said thus, though there was no mode to determine the optimum with exact method however thinking over the advantage and disadvantage of versions it was determined by heuristic way to obtain the new starter, so this decision is a near optimum, it can be said a suboptimum.

Let's still take a simple optimizing example touching a totally other problem range namely the environment protection and economy.

Presumably great many people know that structural element in the motorized world which if it is built into the automobiles stops respectively restarts automatically the engine among conditions given. This is very important in the traffic of towns where according to surveys the vehicles stand in 35% of the traffic time. Such built in structural element guarantees 15% fuel savings and reduces CO_2 emission with 15%. Thus more economic and more environment saving vehicle is operated against an earlier construction, namely there is a possibility to operate a more optimal engine.

Countless further examples could be mentioned from our everyday life connecting with the optimal solution based on the common sense. However to determine on scientific basis the optimum and suboptimum of different processes generally tends to the fields of production organization, planning, manufacturing, sale, operation, etc. Thus in the followings it will be presented the analyses connected with also optimization in the mirror of production and operation as well as operation maintenance processes. We aspired to that during discussion of different questions that there shouldn't be need higher level theoretic knowledge – later you will see that this is only partly true -. You can understand the tasks by the logical application of secondary school and university knowledge, however you may often need also the help of connecting disciplines, moreover we endeavour to make also use the possibility of models (verbal, mathematical and figure) offered by special literatures in order to better understanding. Let's sum up shortly without aspiring completeness what should be understood under optimization, optimum, suboptimum ideas and what characteristic possibilities are to determine these, what kind of conclusions can be got to know these.

14.1.1. Questions, tasks:

- Think about that whether it can be adapted for you "the paper, pencil and eraser" as the most important instruments needed for primary representation of intuitive thoughts:
    - If yes, try to describe, to explain your standpoint,
    - If no, motivate and describe your ideas in connection with what do you think better.
- Sum up what research results constituted the scientific basis of production processes.
- Make known the meaning and importance of the cost coverage diagram.

14.2. Optimization, optimum, suboptimum

Optimization, determining the optimum respectively means that analysing activity during which the most favourable result is got by means of summing up the process containing more part-elements. The recognition of optimization thus is based on that logical analysis whether it can be shown their constrasted effects facing one another of different part processes, which bears in itself the possibility to determine the optimum (what can be maximum or minimum) of the process.
Drafted shortly the optimal is as much as to determine the most favourable (the best) in the process given. Looking for the optimal solution includes every human activities should that be the most everyday avtivity or the scientific research, producing different products respectively the operation or maintenance of those.
Let's examine what result can be got in case of different questions. Let's take a figure used to characterize the production efficiency of a plant (almost generally known) showing very clearly the costs, the incomes and the profit needed to produce the product (Figure 14.1).

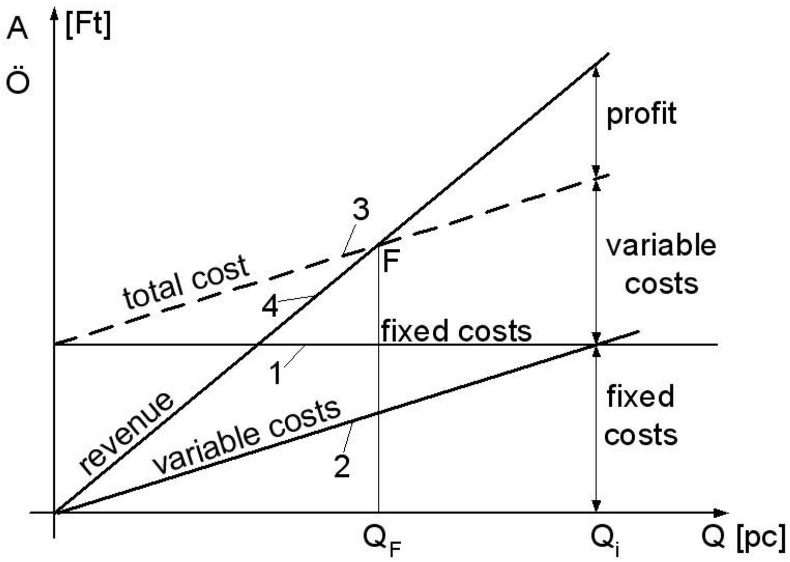


Figure 14.1. Cost coverage diagram.

We indicated the number of pieces of the product produced by the company on the horizontal axis, the total cost...(3) and the revenue (4) on the vertical axis. We indicated in order to the completeness still the two components of the total cost...., namely also the characteristics of the fixed costs (1) and the variable costs (2).

A typical point (F-coverage poin) can be seen on the Figure determined the intersection of the revenue direct line and the total cost direct line. That point has got decisive importance concerning the profitable operation of the company as in case of producing Q_F number of pieces belonging to this intersection the company is neither profitable nor losing. The Figure clearly shows that minimal number of pieces below which must not undertake producing product. The production over the Q_F number of pieces results all the greater profit as greater number of products are produced. There is an opposition to such case of optimization which induces the analyst (it can be said the decision-maker) to set as a goal to fix maximum of profit by utilizing the capacity to be at disposal.
The management knows (if not then they are not suitable for managing position) the behaviour of the market if the number of products produced is increased beyond a certain limit then the revenue decreases grandually in consequence of saturation of respectively oversaturation of the market to such an extent, that the profitable production can become losing (this will be presented later).
Some examples listed so far – to be sufficiently far from each other – show that the optimum search is a process requiring very complex thinking which demands thorough encyclopedic and theoretic knowledges. That what should be understood on optimum Starr [14.1] says "that it has to be found equilibrium among the opposite factors (or aims) of the system. "This has to be well known the whole of the process (question) to be optimized and its component elements, its changes taking place in time, in space, its disturbing effects, etc. The "system view thinking" can be favourable used to the optimum – search of such and similar to this complex processes [14.4] that is shown simplified in Figure 14.2 [14.1].

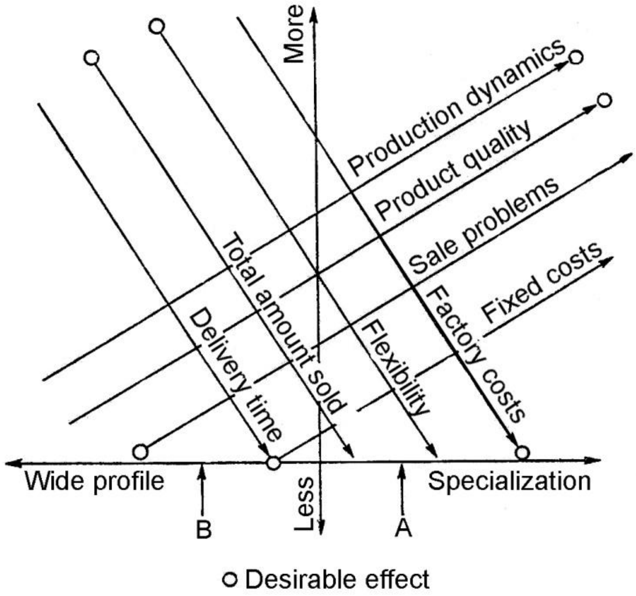


Figure 14.2. Interaction between optimizing and elemental changes.

The figure shows the opposite effects of some general characteristic elements of a producing factory that the “wide profile” (many kind of products), respectively the specialization (homogeneous product) what effects influence on the optimal equilibrium. The change of the production rhythm and of the product quality have got namely favourable course with increasing specialism, but the problems coming about from sale are unfavourable, as well as the increase of fixed costs and deteriorates the flexibility of the company surely the sources of the company special productive forces (single-purpose machines, single-purpose tools, trained workes, etc.) can be changed over with very costly to produce other product. The system-view thinking consequently helps to arrage the processes: what should be considered as a whole, and what as a part, drafting otherwise what is considered as a system and what as a part-system respectively as a sub-system. It has got fundamental importance to draft the main objects first of all and only after should be started to define the system. Churchman presents the character of the question with the following experiment of thought.

“If for example I ask you to describe what is a car. You maybe eliminate at once the thinking and simply list those things which you remember based on your own car: the car wheels, the engine, the form of the car.

You start with that: “Well, the automobile is something which has got four wheels and is driven by an engine.” Then I ask you (trying to join into your mental process) whether it is possible a car with three wheels? You have already seen such car and agree readily to change your definition still thinking not much on that account what does it mean this change of definition. Then – on and on more warlike – I pry into the thing and I ask you, whether is it possible a car with two wheels? You seems a bit confused on this, which refers to that your ideas are not entirely right. Myself continue longer, becoming on and on more unbearable, and I ask the question whether can a car exist without wheel? You get confused still more rather and you already don’t think of the car any longer but of the stupid questioners. Raising the question concerning the car without wheel however is a constructive method to the right view of the system named car. It is possible that just the weels cause the traffic jam and all discomforts of motorization of our age. That car which is capable to hover with some centimeters over the surface of the earth guarantees much more comfortable travel and causes

also less problems at traffic jams, moreover at accidents. Such hovering car can already be accomplished technically in the near future.
The proper method of defining the car to examine at first what it serves, what function has it got and it doesn't starts to list the car components building up its structure. If the sequence of thought is started so with the functions of the car that is to what it serves, then obviously you don't define the car that you start to speak about the four wheels its engine, its dimension and about similars. You start your thoughts that the car serves some people to get from one place to anothere within a certain budget determined in advance. When you started to think then the "definition" of the car presents a new and often totally surprising aspect. This is the system-view approach of the transport with car [14.4].

14.2.1. Questions, tasks:

- Write down what do you understand
    - Optimization, further on
    - Under optimum ideas.
- Write such connections which has got
    - One minimum,
    - One maximum,
    - One local maximum,
    - One local minimum.
- Draft what attitude, readiness and knowledge are needed to plan the optimum respectively an account of a process given,
- Draft those activities by which logical succession it can be reached to determine the optimum of the process.

14.3. Intersection as suboptimum

Let's take a function easy to manage by mathematics for simple survey purpose of the complexity process of the optimization which represents formally clearly different processes which can be optimized (for example inventory – series number, - determining maintenance number discussed later) in order to draw attention to two characteristic points: to the optimum and to the intersection.
The function given should be

$$y = ax + b\frac{1}{x^2}$$

First of all it has to be noticed that the members of the function (ax and $b\frac{1}{x^2}$) reflect contrasted processes with the increasing of the independent variable, that is the first member increases monotonously, the second however decreases monotonously which means that the process suggests optimization possibility.
Let's examine that in case of the change of components (a, b) the value of optimum and the intersection how change as well as its position and what conclusions can be drawn from the basis of results got. This has got importance because it is very frequent that idea that it is sufficient to determine the intersection of the two curves and consider this at the same time optimum, too:

Let's examine three kinds of variations:

No.1. case should be:

$a = 1, 2, \ldots 5$, and

$b = 1$.

No.2. case should be:

$a = 1$, and

$b = 1, 2, \ldots 5$.

No.3. case should be: $ax + b\dfrac{1}{x^2} = 0{,}5x + 4\dfrac{1}{x^2} = y$

At analysing it should be always aspired that the part-processes could be followed more clearly as possible. The above equation offers itself visibly for that purpose to take apart into two parts, which graph can be seen in Figure 14.3.

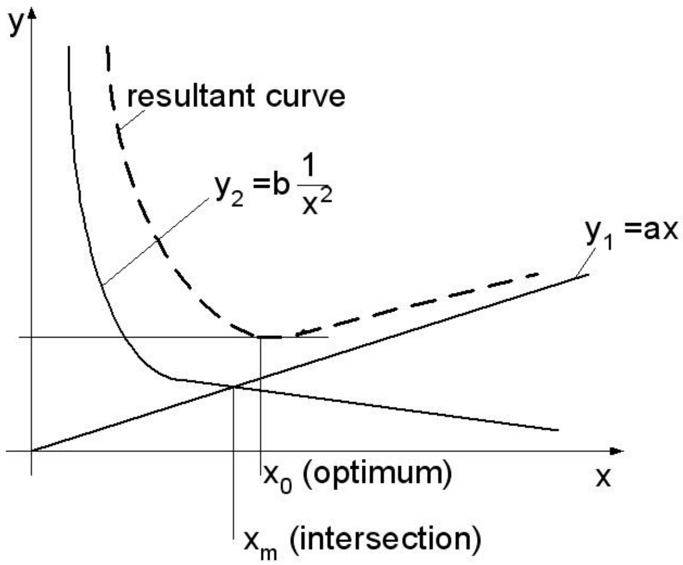


Figure 14.3. Geometric meaning of the optimum and intersection.

$y = y_1 + y_2$ that is

$y_1 = a\,x$

$$y_2 = b\frac{1}{x^2}$$

The condition of the optimum is:

$$y' = \left(ax + b\frac{1}{x^2}\right)' = \left(ax + bx^{-2}\right)' = 0$$

$$a - 2b\frac{1}{x^3} = 0; \quad a = 2b\frac{1}{x^3};$$

$x^3 = 2\dfrac{b}{a}$, thus the place of the optimum: $x_0 = \sqrt[3]{2\dfrac{b}{a}}$

However the condition of the intersection is that the elemental functions should have got common point, namely to be fulfil $y_1 = y_2$

$$ax = b \cdot \frac{1}{x^2}$$

$x^3 = \frac{b}{a}$; thus the intersection of the two functions $x_m = \sqrt[3]{\frac{b}{a}}$

During analysing three things can be interesting:

- the optimum of the function (X_0, Y_0),
- the results of the resultant function (X_m, Y_m) belonging to the intersection of the elemental functions,
- the differences between the coordinates of the intersection and optimum.

It is suitable to the analysis practically to complete the following table (Table 14.2) for purpose of easy survey.

Table 14.2.

a / ***b***	**Optimum** $x_0 = \sqrt[3]{2\frac{b}{a}}$	**Intersection** $x_m = \sqrt[3]{\frac{b}{a}}$	**Differences between the characteristic co-ordinates**	
			$x_0 - x_m$	$y_m - y_0$

A.) Let's complete the Table first according to the "No 1" variation:

a = 0,5, 1, 2, 3, 4, 5, 10.

b = 1

Table 14.3.: Tabular data belonging to *b*=1 value

a	x_{01} $\sqrt[3]{2\frac{1}{a}}$	x_{m1} $\sqrt[3]{\frac{1}{a}}$	**Intersection and optimum co-ordinates**		**Difference between optimum co-ordinates and intersection**	
			y_{m1i}	y_{01i}	$x_{01i} - x_{m1i}$	$y_{m1i} - y_{01i}$
0,5	1,587	1,259	1,259	1,19	0,328	0,069
1	1,259	1,000	2,000	1,889	0,259	0,111
2	1,000	0,793	3,176	3,000	0,207	0,176
3	0,870	0,693	4,161	3,931	0,177	0,23
4	0,793	0,629	5,04	4,762	0,164	0,278
5	0,736	0,584	5,852	5,526	0,152	0,326
10	0,584	0,464	9,284	8,77	0,12	0,514

Ordinate values belonging to b=1 value

$$y_{01.0,5} = 0,5x1,587 + \frac{1}{1,587^2} = 0,793 + 0,397 = 1,19 \qquad y_{m1.0.5} = 0,5x1,259 + \frac{1}{1,259^2} = 0,629 + 0,630 = 1,259$$

$$y_{01.1} = 1x1,259 + \frac{1}{1,259^2} = 1,259 + 0,630 = 1,889; \qquad y_{m1.1} = 1x1 + \frac{1}{1^2} = 2 = 2,00$$

$y_{01.2} = 2x1 + \frac{1}{1^2} = 2 + 1 = 3{,}00$; $y_{m1.2} = 2x0{,}793 + \frac{1}{0{,}793^2} = 1{,}586 + 1{,}59 = 3{,}176$

$y_{01.3} = 3x0{,}87 + \frac{1}{0{,}87^2} = 2{,}61 + 1{,}321 = 3{,}931$ $y_{m1.3} = 3x0{,}693 + \frac{1}{0{,}693^2} = 2{,}079 + 2{,}082 = 4{,}161$

$y_{01.4} = 4x0{,}793 + \frac{1}{0{,}793^2} = 3{,}172 + 1{,}59 = 4{,}762$; $y_{m1.4} = 4x0{,}629 + \frac{1}{0{,}629^2} = 2{,}516 + 2{,}527 = 5{,}04$

$y_{01.5} = 5x0{,}736 + \frac{1}{0{,}736^2} = 0{,}793 + 0{,}397 = 1{,}19$; $y_{m1.5} = 5x0{,}584 + \frac{1}{0{,}584^2} = 2{,}92 + 2{,}932 = 5{,}852$

$y_{01.10} = 10x0{,}584 + \frac{1}{0{,}584^2} = 5{,}84 + 2{,}932 = 8{,}77$; $y_{m1.10} = 10x0{,}464 + \frac{1}{0{,}464^2} = 4{,}64 + 4{,}644 = 9{,}284$

B.) Let's complete the Table according to the "No.2" variation:

a = 1,
b = 0,5, 1, 2, 3, 4, 5, 10.

summing up values belonging to a=1 value

Table 14.4.: Tabular data belonging to a=1 value

b	Optimum $x_{02} = \sqrt[3]{2\frac{b}{1}}$	Intersection $x_{m2} = \sqrt[3]{\frac{b}{1}}$	Intersection and optimum co-ordinates		Difference between intersections and ordinates	
			y_{m2}	y_{02}	$x_{02} - x_{m2}$	$y_{m2} - y_{02}$
0,5	1,00	0,793	1,588	1,500	0,207	0,088
1	1,2599	1,000	2,000	1,889	0,259	0,111
2	1,587	1,260	2,520	2,381	0,327	0,139
3	1,817	1,442	2,884	2,725	0,375	0,159
4	2,00	1,587	3,175	3,00	0,413	0,175
5	2,15	1,71	3,42	3,23	0,44	0,19
10	2,714	2,15	4,31	4,07	0,56	0,24

Ordinate values belonging to a=1

$y_{02.05} = 1 + 0{,}5\frac{1}{1^2} = 1 + 0{,}5 = 1{,}5$; $y_{m2.05} = 0{,}793 + 0{,}5\frac{1}{0{,}793^2} = 0{,}793 + 0{,}795 = 1{,}588$

$y_{02.1} = 1{,}259 + 1\frac{1}{1{,}259^2} = 1{,}259 + 0{,}630 = 1{,}889$; $y_{m2.1} = 1{,}0 + 1\frac{1}{1^2} = 1 + 1 = 2{,}000$

$$y_{02.2} = 1{,}587 + 2\frac{1}{1{,}587^2} = 1{,}587 + 0{,}794 = 2{,}381\,; \qquad y_{m2..2} = 1{,}260 + 2\frac{1}{1{,}259^2} = 1{,}2599 + 1{,}2599 = 2{,}52$$

$$y_{02.3} = 1{,}817 + 3\frac{1}{1{,}817^2} = 1{,}817 + 0{,}908 = 2{,}725\,; \qquad y_{m2..3} = 1{,}442 + 3\frac{1}{1{,}442^2} = 1{,}442 + 1{,}442 = 2{,}884$$

$$y_{02.4} = 2{,}00 + 4\frac{1}{2{,}0^2} = 2{,}0 + 1{,}0 = 3{,}00\,; \qquad y_{m2.4} = 1{,}587 + 4\frac{1}{1{,}587^2} = 1{,}587 + 1{,}588 = 3{,}175$$

$$y_{02.5} = 2{,}15 + 5\frac{1}{2{,}15^2} = 2{,}15 + 1{,}08 = 3{,}23\,; \qquad y_{m2.5} = 1{,}71 + 5\frac{1}{1{,}71^2} = 1{,}71 + 1{,}71 = 3{,}42$$

$$y_{02.10} = 2{,}714 + 10\frac{1}{2{,}714^2} = 2{,}714 + 1{,}357 = 4{,}07\,; \qquad y_{m2.10} = 2{,}15 + 10\frac{1}{2{,}15^2} = 2{,}15 + 2{,}16 = 4{,}31$$

Based on the analysis of the data of tables the following statements can be made:

- The difference between the optimum and the intersection as long as decreases with the "*a*" parameter increase, as far as that it increases with the "*b*" parameter increase, at the same time the differences between the ordinates however increase.
- In the practice in many cases is enough if the intersection is considered also to be the place of the optimum, which naturally is only true with negligence.
- The absolute value of the mistake can be well seen in the table but the importance of the mistake stands out better if the value is indicated in percentage, for example to take the case when *a*=3, *b*=1 and the mistake $x_{01.3} - x_{m1.3} = 0{,}177$ which percentage value, $\varepsilon = \frac{0{,}177}{x_{m1.3}}100 = \frac{0{,}177}{0{,}693}100 = 25{,}54\%$.

This value in the field of technical sciences in the majority of cases is an already inadmissible mistake. Because it has got great importance also to point to the volume of the mistake that could be made at solving one-one problem and if there is an opportunity then to reduce it, too. Surely the savings as against the cost to be needed to indicate the mistake – at the sections of preparing decision of the problem given – can be greater with order magnitude.

The engineer can be restful then if he indicates the mistake or how big is that value at which it is smaller or bigger. Considering the contemplations so far in connection with optimization then it becomes ever clearer that to solve the task given succesfully gives only proper basis of profound knowledge of a row connected science to be very wide – ranging – and to be also deep.

It has to be observed that the manual processing of the example seeming to be worked out in great detail an important element that the readiness tending to recognize the fine details should develop. It is not satisfied that the optimum is somewhere at the surroundings of the two curves' intersection. At the same time it can be said calmly in the case of account of the mistake that the intersection of the two curves presents the suboptimum (the place near to optimum) of the function.

The account, the explanation of the suboptimum can be studied very clearly in Figure 14.4 [14.1], which shows in essence that the temporal change of long-range planning strategies what effect can have onto the cumulative profit.

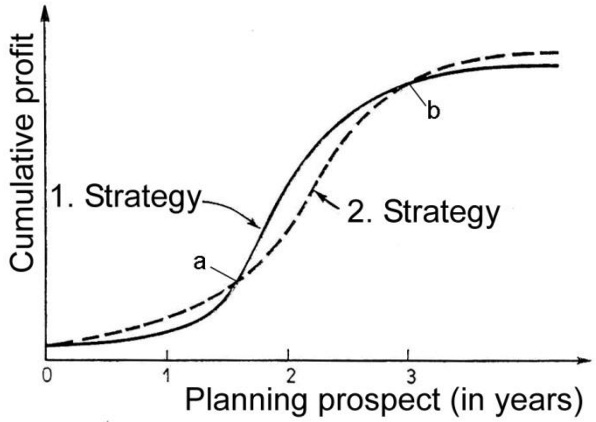


Figure 14.4. Perspective planning and the suboptimum.

The figure shows the expected accumulated profit in case of two kinds strategies in the function of time. It can be seen that the curves intersect each ofher also at two places which means that

- that during the time (this is about 1,5 years) belonging to **a** points the production according to No.2 strategy yields greater profit,
- in the time interval suitable to the **a** and **b** points to apply the No.1 strategy is more profitable,
- the No.2 strategy provides greater profit from the moment to be suitable of **b** intersection.

  It is worth mentioning that the company management is not in easy position concerning the decision making, as the market causes surprises not so rarely and in consequence of that it chooses the more secure, short-distance strategy providing greater profit.

  Let's go back to the finishing of the Figure 14.1 broken off and let's have a look at what happens with the revenue if the produced number of pieces is increased. It has already been referred that in consequence of market effects the revenue (4) increases for some time then it starts decrease and again it intersects the direct line marking the change of the total cost (3) Figure 14.5.

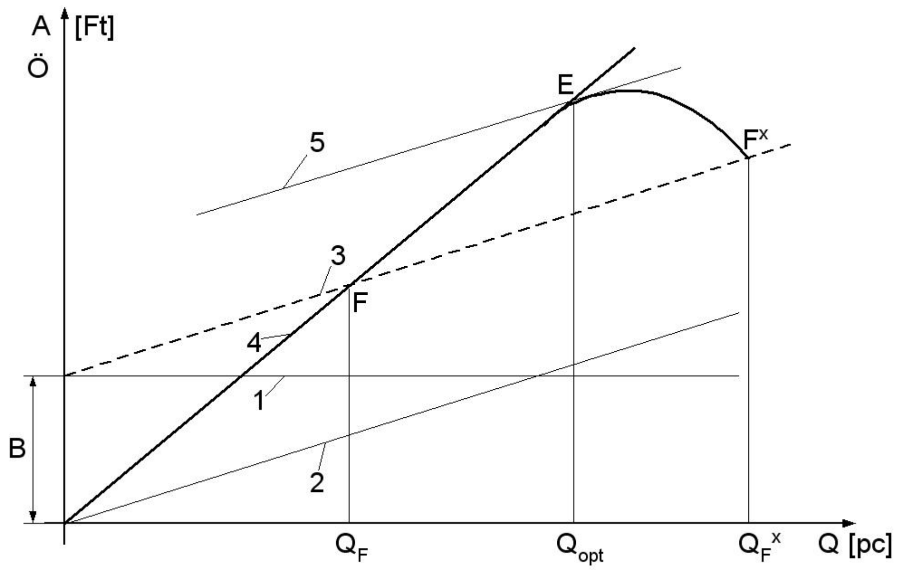


Figure 14.5. Graphic determination of the optimal number of pieces.

It is only natural that between the two points (*F* and *Fx*, respectively *QF* and Q_{Fx}) there is one such point where the profit is maximal and accordingly it can be determined that optimal number of product (Q_{opt}), where the profit is maximal. The maximum of the profit can be determined by two kinds of method.

a) The mathematical connections between the costs and the number of pieces is well known.
   That is:
   - Total cost:
     - Generally…$TC=f_{ö}(Q)$
     - In present case…$TC=a\cdot Q+B$ where ***a*** = cost of production of one product, *B*=fixed cost
   - Revenue: …$R=g_A\cdot Q$

The maximal profit is there, where the first derivative of the “*A*” function is equal with the first derivative of the “*Ö*” function.
That-that:

$$g'_A(Q)=(a\cdot Q+B)'=a$$

Solving the equation to *Q* the optimal number of pieces (Q_{opt}) can be got and the difference at this place of the two functions results the obtaniable maximal profit (*P*.)
That-that:

$$P=g_A(Q_{opt})\text{-}(aQ_{opt}+B)$$

In the majority of cases these functions are not at disposal in exact form but with using proper working number measurings which is completed with suitable computer simulation, according to the next “D” point the optimum as suboptimum can be determined with good approximation.

b) It is known (that it can be calculated) the amount of the revenue belonging to certain discreet (determined) quantity (number of pieces) and it is supposed that the fixed costs and the variable costs are linear. (Figure 14.5)
   That is:
   - Total cost:

$$TC=aQ+B$$

- The discrete values of the revenue is shown in Table 14.5

Table 14.5.: Production data

Serial	Number of pieces	Revenue
1	Q_1	R_1
2	Q_2	R_2
3	Q_3	R_3
. . . .	. . . .	. . . .
n	Q_n	R_n

Based on the table marking in turn the value of the ordinate of the revenue beloging to given number of pieces the curve is plotted. The place of the optimal number of pieces marks out the tangent drawn (5) to the revenue curve (4) parallel with the total cost (3) lines.

Naturally there is no need to determine unnecessarily the value pairs to the F^x point, as the decreasing ordinate values already signal the surroundings of the optimum.
It can be said concerning the production processes that their exact description in consequence of their complexity can be accomplished only in details.
Deriving from this at planning different processes the experiences accumulated during practice have got very great importance furthermore the applicability of systems supported with computers wide-spread extensively nowdays.
Come to your mind the "paper, pencil, eraser"(!) as the fixing and planning instrument of engineer's thoughts and only after turn using the method of "cliks".

Task:

Determine the intersection of the $y_1 = \frac{1}{2}x$ and $y_2 = 4\frac{1}{x}$ functions and the optimum of the two functions' resultant. Complete the calculations also with substituting other constants, and draw conclusion from such functions.
Determine further the condition that in which case coincides the optimum with the intersection.
It has to be always before your eyes that the optimization is nothing else then recognizing the complex connection between the system analysis and system synthesis.
Drafting differently to determine the most favourable solution by means of iterative analysis and synthesis of the system connecting elements.

14.3.1. Questions, tasks.

- Draft the content meaning of the suboptimum and knowing this which present time furthermore future time conclusions can be drawn.
- Show the importance of knowing the suboptimum in the process of long-range, strategic planning, furthermore analyse that the company management which quandaries can have concerning the future decisions.

14.4. The idea of the system, its interpretation

It can be met in many nativ and foreign authorts' works with the drafting the definition of the system [for example: 14.2, 14.3, 14.4, 14.5, 14.6, 14.7, 14.8, 14.9, 14.10, 14.11] in which different from one another but essentially similar interpretation of idea of the system are presented.
So for example: Heisenberg [14.5] explains the complex analysis of "the part and the whole" interaction to the interpretability respectively to the recognizability of the system.
Churchman [14.4] refers based on the view of "input-output" model to that "The scientist of the management – theory, probably found very useful to consider the system as totality, its entity of the recognizable things, into that it is given in "resources with different types (people, money, etc.) and from which some kinds of product or service can be got."
Deli-Kocsis-Ladó [14.6] referring to different special literatures drafts the system as summing up in the followings "...on system it is meant the mutual connection of elements (parts) belonging together based some kinds of common criterion."
According to the interpretation of special literatures thus the system is the heap of elements belonging together based on some kind of common criterion, which has got contact with the environment through inputs and outputs, furthermore the various orders and disturbing effect

always affect into it. The so called input – output model (Figure 14.6) name is used to describe the processes in such way. The system changes its condition and as a result of change the modified materials, different products, services etc. or remain in the system or they leave it in output form, that is they come out into the environment.

Two parallel processes can be seen in Figure 14.6 The series of changes of condition happening in the system compose the main process, by means of that the final product appears as output. That can naturally be very diverse (depending on the production system), so for example: in case of refrigerator factory, refrigerator boxes with different sizes and powers, from a shoe factory, shoes with different sizes and styles, from a machine works, different machine – tools (lathe – milling – grinding machines, machining – centres). The auxiliary process is the so called regulating process and is always parallel with the main process. Its meaning is that depending on the efficiency of outputs by changing the inputs the value of certain outputs can be increased or decreased.

Let’s take a bakery as an example, as a system. Let’s start to follow to the end from that the conditions are guaranteed to produce baker’s merchandise, the function of the system which main components are the followings:

- building and the needed premises,
- electric furnaces,
- mixing machines,
- different basic materials having in stock (for example flow products, additives, dairy-products etc.) needed to the routine works,
- skilled workers,
- temporary helps,
- plant management.

The inputs of the system:

- electric energy needed to operate the machines,
- wear and tear coming about operating of machines, equipment,
- used up amount from different basic material,
- skilled workers’ manpower,
- temporary helps’ manpower,
- intellectual contribution (organizing, regulating) of the plant management to make run the system.

The output of the system:

- various sort of breads (kg),
- bun (pc.),
- salted roll (pc.),
- buttered roll (pc.),
- cheesy roll (pc.)
- cake (kg).

The task of the plant management:

a.) to make run the system,
b.) to regulate the system,

It has to be guaranted to make run the system:

- the suitable supply of the outputs level,

- to maintain the running condition of machines, equipment,
- the workers' continuative education,
- the healthy working conditions.

It has to be determined during regulating the system:
- the financial value originating from outputs (income:[Ft]),
- the financial value originating from costs (expediture:[Ft]),
- based on the difference of the two values is constituted the profit, that is the profitability of the company.

The plant management decides on the basis of efficiency indicators on the qualification concerning the whole company, which can be good and further on this level has to be kept, but it can be dissatisfied, too. In this case it has to be analysed in detail the efficiency of different products, based on which the decisions refer to the inputs can concern the quality, quantity, the change of product, but also the market condition alike.
For the reader certainly it was evident unambiguonsly, that the discussions so far concerning the interpretation of the system, tended really to producing companies, and also further on tend to analyse the plant producing processes respectively to optimize the part-processes. That already can also be seen that to direct and regulate the very complex task needs multiple knowledge from the company management and needs time analysis concerning to judge the efficiency of the process. Surely that already could also be seen that a solution seeming optimal can be a suboptimum that can influence the profitability of the company just negative. To make perceptible this let's return to the cost coverage diagram (Figure 14.5) now on the basis of system – view consideration whether it is needed to make certain addition to determine the optimal number of pieces. The answer is because therefore the decision made superficially can just shake the company future, too.
Let's see thus the observations:

a.) so began to maximize the financial characteristics of the production going on in the supposed company, the profit stressed, that it had been defined the numerical data of the products going on in the plant which is therefore a trouble because this thoughtless action can make the company bankrupt,
b.) to the plant supposed the optimal number (Figure 14.5) of pieces determined in the function of number of pieces, inasmuch produces homogeneous product gives the company maximal profit,
c.) inasmuch the plant supposed produces multiple products, then the optimum shown in the figure (Figure 14.5) concerns the maximal profit only one of its product, which means that this is a suboptimum, that doen't mean definitly the company optimum, too.

In this example also appears that the system as defining of the whole and the subsystem (as part) how complicatedly join to each other, respectively in which extent affects the part to the whole.
The statements thus can be summarized in the followings:
- the system is called as the heap of those elements which organizes itself to reach some kind of goal,
- the system can be represented with the so called black box,
- in the system has got contact with its environment by the means of inputs, outputs and other outside (disturbing, - regulating) effects,

- the system produces different goods (products) to the effects of inputs, which leave it as output,
- depending on the complexity of production processes the system dismembers to part systems.
- the inputs of the system are generally materials, machines, instruments, manpower, energy, information etc.,
- the outputs of the system are finished products, semi-finished products, service etc.,
- regulating the system can be accomplished to the input with feed-back of the main characters of the output (Figure 14.6) which can include to the quantity, quality, to the termination of the product, to the production of the product.

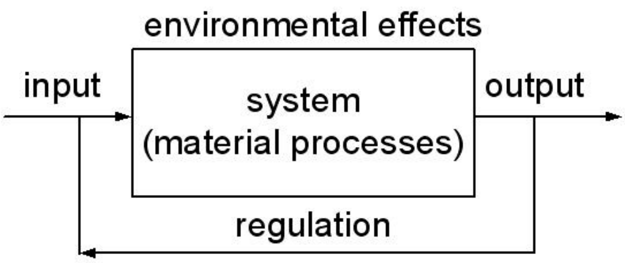


Figure 14.6. Block scheme of material processes.

Task.

Try to define your own group of students on the basis of system-view:

- determine the system itself,
- determine your own group,
- determine the inputs and the outputs,
- determine the elements of regulation,
- determine the disturbing elements,
- evaluate the effectiveness of the system.

14.4.1. Questions, tasks.

- Draft what do you mean under system, try to define the system respectively.
- Analyse the system – ideas defined by different authors, point to what is common in them (if there is such), respectively is there essential contradiction among them and if there is what is that?

14.5. The system theoretic model of the company

Drafting simply the company is a legally independent organization performing direct activity. The owner, the company management respectively determine its activity with that consideration to obtain as greater as possible profit.

The number of workers and also the stock of instruments increase with the increase of production volume of the company. These rate depend naturally on the requirement of resources needed to produce product-, respectively the sort of products. So for example a venture to be organized to assembly certain product (or sort of products) has to start with substantially smaller capital than a venture that wants to establish part production. Surely the part manufacturer needs to start the production starting capital to be greater with scale in contradiction to the contractor undertaking assembly.

It is enough on a large scale if it is compared the suitable machine- and requirements of means of the production profile as in case given there are not significant the differences in premises claim, in assuring social background.

Instruments, tools needed to mounting works:

- universal erecting tool;
    - various fork – box – end wrenches – hexagon spanners (etc)
    - screwdrivers with various dimensions and shapes (flat, star-shaped, etc)
- universal assembly jigs;
    - mounting jigs,
    - disassembling jigs,
    - universal bearing assembling jigs,
- bench,
- vice bench,
- bench drill,
- hand operated assembling press.

The cost of these is some hundredthousand forints.
Machines, instruments needed to part production:

- universal lathe and accesories.
- universal milling machine and accesories.
- sawing machine.
- frame drilling machine.

The cost of these: some million forints.
If it is also the question of esteemed values then the starting capital requirement is significant, in particular if it has to be contracted a loan to give a start to the venture.
It can't be emphasized sufficiently the circumspect drafting of objects, the marketability of the product (that can also be service), if a plant is founded then the arrangement of authorization, if manpower is needed then how can be guaranteed this from the surroundings settlements etc. The company, the venture can be considered as a system which can be dealt with an input-output model (Figure 14.7). I indicated the generally draftable elements of inputs and outputs in the Figure.

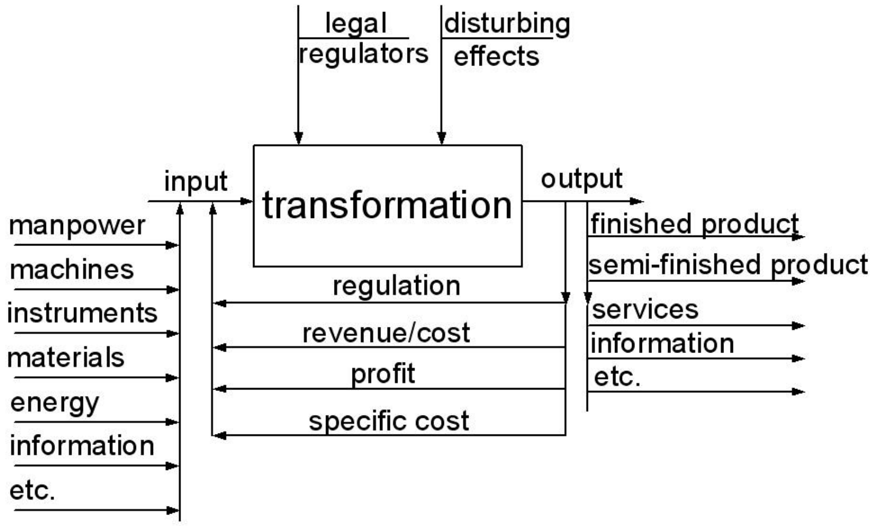


Figure 14.7. More important input-output elements of material processes.

The company (Figure 14.7) as a relatively closed system, has contacts with its environment by means of inputs and outputs, furthermore partly the calculable legal regulators partly the accidental disturbing effects affect the system.
The system changes its condition to the effect of inputs as a result of this change the various products leave the system, that is they come out into the environment.
The system – view characters of the company can be determined according to the followings:

- The company can be considered as a complex, great system as the connection between the elements and those is multifold furthermore their effects into one another can't be described entirely.
- The company is a dynamic system as the time factor fills in basic part in the change of condition of the company which means that certain processes can't be expediently examined statically.
- The company is a self – organizing system, as by the means of changing of external environmental effects is possible to make such measures during which it changes the connection between the company elements so to yield the most effective result.
- The company is a stochastic system because the processes of the change of conditions and the outputs can be made probable (apart from small number and simple cases), that it can be described only with stochastic regularities.
- The company is a hierarchic system as the management performing the direction is multi level and those are in subordinate and superordinate relation.

Multiple determinations models of the system can be found in the special literature however all emphasize the difficulty of these applying in between company conditions. Therefore their importance can be considered first of all as theoretical at the same time their attempt of thoughts are very instructive.
Further on it will be wish to presented such complex mathematical model based on Janik [14.8, 14.12, 14.13] and Zsoldos [14.14, 14.15] works as a case study which has got a model representing a general – and a new subsystem – machine operation – of the company.
The aim of the company is to maximize the profit beside satisfying some kinds of needs. The company subsystems, so the machine operation subsystem have to look for that at what method can it contribute to maximize the profit.
As the task of the machine operation subsystem is to guarantee the conditions main items are the inputs shown in Figure 14.7 of the production the answer is given: the machine operation subsystem can constitute to maximize the profit of the company so, that beside guaranteeing conditions of the production minimizes the costs of the products produced.
As the machine operation is a subsystem of the company is defined on system – theory basis.
The flood of the input – output elements between the environment and system can be considered as the basis as a discription of the company mathematical model, which main components can be the followings (to company pursuing machine industry activities):

- manpower (for example):
    - skilled workers,
    - technicians,
    - foremen,
    - engineers,
    - financial experts,
    - etc.

- machines, instruments:

  - lathes,
  - milling machines,
  - planers,
  - drilling machines,
  - machining centres,
  - etc.

- energy:
  - electric energy,
  - gas,
  - oil,
  - etc.

- various materials:
  - castings,
  - rolled stocks,
  - etc.

- finished products,
- semi – finished products,
- information,
- etc.

The input elements pass on crushing, - casting,- pressing,- cutting,- heat treating, etc, processes within the system and are forming to sale as output – elements.

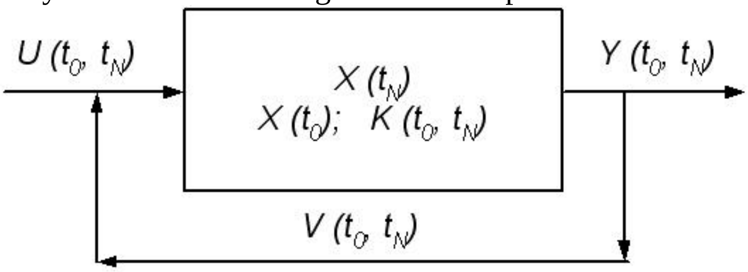


Figure 14.8. System theoretic model of material processes.

Naturally there's no reason to list the number of elements without real given production activity because of this the system (the company) can be defined a material flow but drafting very generally as an energyflow process, which system theory model can be seen in Figure 14.8.

The meaning of certain letters is the following:

$x(t_N)$ = the condition characteristic of the system at given t_N time,

$x(t_0)$ = the condition characteristic of the system in the starting time, in the starting time of the operation of the system,

$U(t_0, t_N)$ = the total inputs flow into the system in the time interval examined,

$Y(t_0, t_N)$ = the total outputs flow out from the system in the time interval examined,

$+K_f(t_0, t_N)$ = the total finished products ready to transportation, produced in the system in the time interval examined, meaning the products and can be marked as source,

$-K_r(t_0, t_N)$ = the total costs needed to operate the system, respectively needed to produce products in the time interval examined which can be interpretated as swallower according to system-theory,
$\Delta V(t_0, t_N)$ = the regulating variable characterizing the process presented which can be financial, but it can have got natural characteristic, too,
t_0 = moment meaning the beginning of the system operation,
t_N = moment meaning the end of the system operation

To describe the operation of the system thus it can be presented a so called energy-flow equilibrium-equation:

$$X(t_N) = X(t_0) + U(t_0, t_N) - Y(t_0, t_N) \pm K_{f,r}(t_0, t_N)$$

and a regulating equilibrium - equation;

$$\Delta V(t_0, t_N) = V_{K_f}(t_0, t_N) - V_{K_r}(t_0, t_N)$$

The member on the right side of the regulating equilibrium - equation depends on the process examined and accordingly the checking and the regulation to appropriate values have to be concerned.

Let's inspect that how the equilibrium - equation can be applied to different processes:

1) The profit is one of the indicators used generally to judge the good operation of a company which is determined at every year's end at balance making.
   The company profit can be determined on the basis of regulating balance – equation:
   a) It has to be calculated by loss that what kind of products and how many pieces were produced by the company in a certain year (for example t_0=from the 1st January 2009 t_N= to 31st December 2009). In order to simplicity let's take 3 types of product, then the yield (that is the source):

$$+K_f = \begin{cases} K_{f1} = 15 \text{ pcs. lathe,} \\ K_{f2} = 20 \text{ pcs milling machine,} \\ K_{f3} = 10 \text{ pcs. drilling machine} \end{cases}$$

   b) It has to be calculated the costs of each product in detail:

$$V_{Kr1}(t_0, t_N) = \begin{matrix} K_{r11} = \text{cost of material} \\ K_{r12} = \text{wage + wage contribution} \\ K_{r13} = \text{cost of energy} \\ K_{r14} = \text{general cost and overheads} \end{matrix}$$

$V_{Kr2}(t_0, t_N)$=
$V_{Kr3}(t_0, t_N)$=

Similarly to the lathe it has to be made account by lots also to the milling machine and also to the drilling machine from the costs.

   c) As the profit of the companies is expressed in monetary terms to establish this it has to be determined the gross production value as well as the value of total costs and the difference between them gives the profit.
      That is:
      - The production value can be got if the number of pieces of the machine – tools is multiplied with the unit price in turn. Let's mark the unit price with the initial letters of the name of machines:
        l = selling price of one piece lathe. Ft/pc,
        m = selling price of one piece milling machine, Ft/pc,

d = selling price of one piece drilling machine Ft/pc.

It can be written:

$$V_{K_f}(t_0,t_N)=l\cdot K_{f1}+m\cdot K_{f2}+d\cdot K_{f3}\quad [\text{Ft}]$$

- The costs can be got if the sum of the money of the machine partial cost is multiplied with number of machine – types manufactured in turn.
- It can be write:

$$V_{K_r}(t_0,t_N)=V_{K_{r1}}(t_0,t_N)\cdot K_{f1}+V_{K_{r2}}(t_0,t_N)\cdot K_{f2}+V_{K_{r3}}(t_0,t_N)\cdot K_{f3}\quad [\text{Ft}]$$

The company profit ($\Delta V(t_0, t_N)$) thus can be determined on the basis of the difference between the sources and swallowers. So the company management can appreciate whether they worked with proper efficiency or not in the year given. Let's however notice that implicitly, it was supposed that the products' sale was without further loss means the exact recording of different elements of the costs.

Task:

Please describe the followings:

- how the material-flow equilibrium-equation looks like, if they fail to sell all products and how it turns into profit.
- what possibilities can you see handling the unsold products.

The company as it pursues multiple activities can't be described with an $X(t_N)$ energy-flowing and a regulating balance-equation but depending on the subsystems furthers are needed. Let's also notice that it doesn't need to represent the balance-equation at every part task in this form surely only the workers would be disturb unnecessarily. The informations required from different levels has to be drafted by simlified way. It is evident that it won't be spoken from inputs to the lathe operator skilled worker in connection with finishing a given workpiece, but it is told that the machining should be carried out with what depth of cut, with what kind of tool in order to satisfy the appropriate quality and level of cost. The regulation thus can include what kind of instruments (for example caliper) should be used to check the part and whether it is enough one final examinations or it is needed to carry out measurings during the process. Instead of energy-flow thus for the accountancy it is important the marking and the size of the material to determine the material cost. Further on the time needed to finish it is necessary to the wage cost. To determine the overheads the marking of the machine – tool is also needed.

The process analyses based on system-theory models thus are reasonable first of all then if great systems are examined and the extent and complexity of the connecting subsystems also justify the application of this method, too.

Let's consider a real company example concerning the machine maintenance subsystem (machine repair, machine maintenance) how can contribute to maximize the company profit.

14.5.1. Interaction of the machine maintenance and company profit

It belongs to the machine-maintenance in important scale guaranteeing in appropriate quality and to date the continuity of the production. To guarantee the good working condition thus has got fundamental importance concerning the profitability of the company. The postponment of the delivery date deriving from a sudden breakdown of a machine also can cause substantial losses for the company, it might jeopardize the company profit, too. Thus in

total to reduce the costs in a given case the optimization it can have an important role maximizing the profit.
The condition-equations (respectively equilibrium-equations):

$$x(t_N) = x(t_0) + U(t_0, t_N) - Y(t_0, t_N) \pm K_{fr}(t_0, t_N)$$

$$\Delta V(t_0, t_N) = V_{K_f}(t_0, t_N) - V_{K_r}(t_0, t_N)$$

It is expedient to start from that aspiration of the company to maximize the profit - $\Delta V(t_{0,}t_N)$. The machine maintenance can contribute to this by minimizing the costs. To decide this question let's examine that theoretically whether it is possible the minimum of the costs on the basis of Figure 14.9 [14.8].

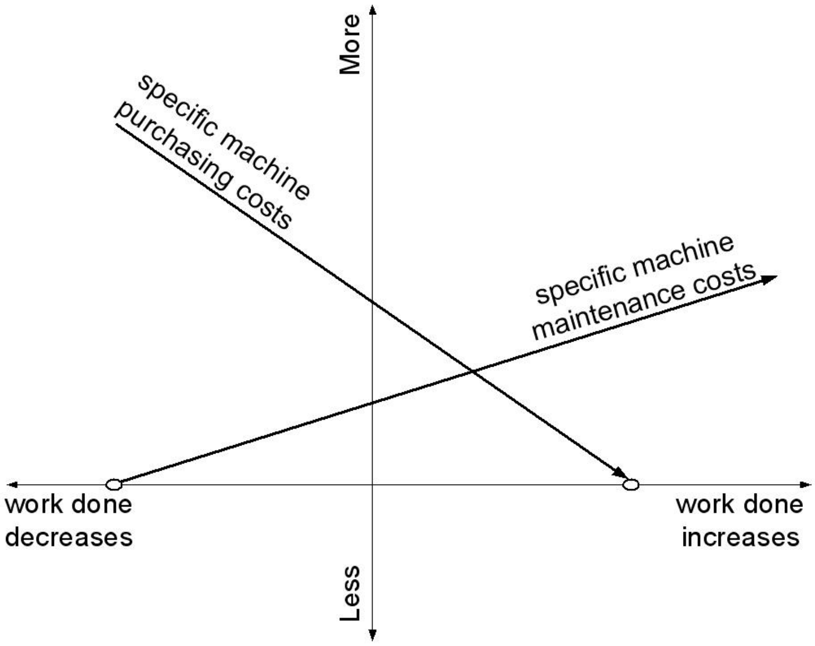


Figure 14.9. Elements with opposite effect suggesting the optimization of machine maintenance.

It can be stated clearly that it can be an optimal usefulness of the machines. I indicated – the amount of the work done – as an independent variable on the horizontal axis, the specific machine purchasing cost and the specific machine maintenance cost on the vertical axis. The tendencies are unambiguous.
Well it is logical if a machine is operated as long as possible the specific investment part is all the smaller. Evidently: if a refrigerator is bought for 100 000 Ft and used for 5 years then the specific yearly cost is 20 000 Ft, but if it used 10 years then the specific yearly cost is only 10000 Ft. That has to be tought at the machine maintenance cost that the machines get into ever worse condition during their operation and so their repair is ever more expensive which means that for example the cost falling to one year also increases. As an independent variable the most suitable has to be always chosen naturally concerning the certain process (for example the time, km, tkm etc.). It is more difficult to determine the cost components. Naturally it can be mentioned some unambiguous components: for example part used during maintenance, energy, lubricant, wage-cost however to take into account the almost hole range

of costs can be determined only knowing of the concrete task and the company as it can be found reference in many special literatures, too [14.16, 14.17, 14.18].
It is still important to determine the object as it determines unambiguously the independent variable and accordingly that to what concerns the optimization. So, for example for a company it can be the goal the optimal service-life to reject a machine which is at the minimum of machine operation costs or it can be the goal guaranteeing the quality that is if the machine can't perform the required production parameters then it is rejected.
It is naturally that the two kinds of objects make necessary the measuring – and evaluation of parameters differring from each another. In the first case the cost oriented parameters come into prominence in the balance-equation, in the second case for example the changing wears, vibrations, the quality can constitute the object of examination.
Let's go back to the statement made previously: the operation maintenance can contribute to maximizing the company profit so that it minimizes the cost, that is:

$$\Delta V(t_0, t_N) \longrightarrow \text{max.}$$

If the minimum of the specific costs is realized,

$$k_r = \frac{V_r(t_0, t_N)}{T} \longrightarrow \text{min.}$$

where
k_r= for example is the specific cost of the machine maintenance [Ft/time] which can be determined expediently also in other measuring unit,
T= is the product of the process respectively a parameter to be characteristics to that [for example: number of pieces, time, km, etc.]
$V_r(t_0,t_N)$= the cost value of the activity carried out (expediently in monetary terms).
The profit of the process ($\Delta V(t_0,t_N)$) is maximal then if the tasks can be carried out with minimal costs accordingly the required conditions. Drafting differently: the profit of the process is maximal there, where cost (k_r) reaches the minimum.
As the components of the equilibrium-equations can't be described exactly in implicit form in consequence of uncertainies and exceptional complexity of the production process, so values measured real to discreet intervals chosen practically of a given machine constitute to determine the optimum of the specific parameter.
Let's look at in general how can be carried out to determine the optimal service-life of a machine given at company level.
Starting bases:

- the aim is to produce a given product at lower cost,
- as the service-life esteemed of a machine-tool (for example: lathe) can be more years (it might also be 10 years) the discreet intervals examined (one-one year), as to an independent variable, the time elapsed from the putting to work the machine is considered,
- the cost elements: for example maintenance, repair, spare-part, energy, wages, wage-contribution, coolant, lubricant, tools, overheads etc. which are determined by the accountancy department at every year's end.
- the costs relating to the year is added at every year's end to the previous years then this value has to be divided with the number of years spent in operation, so the accumulated specific cost [Ft/year] to every year plotted in the function of years can be seen in Figure 14.10.

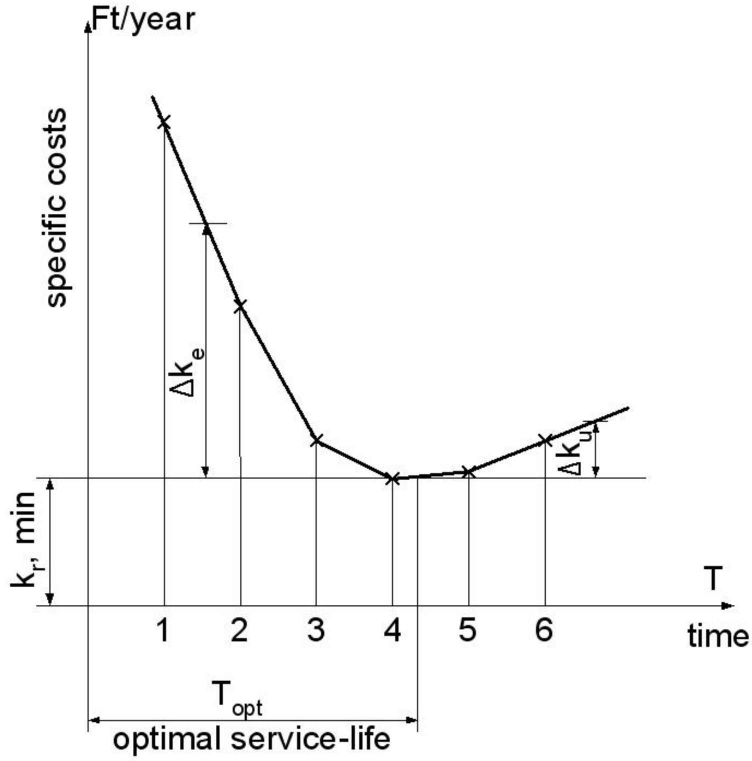


Figure 14.10. Determining the optimal service-life by graphic-analytic method.

Figure 14.10 shows faithfully in its tendency the formation of the expected specific cost (what can be named as operation characteristic) and the determinability of the optimal period of service-life.
This is important for the company management to determine the date of the rejection there are at disposal values based properly by means of their own measurings. It is optimal to reject between the 4th and 5th years concerning the company profit, because both an early rejection (Δk_e) and a later rejection (Δk_μ) cause losses.
The later rejection generally causes smaller profit loss. It is worth to take a note that the machine rejection optimum is a suboptimum, as it produces only a part of the company profit.
The model was tested under operating conditions of powerful tractors and combines with cooperation of larger companies which main results are the followings.

14.5.2. Practical experiences of the model's application in companies

Precedents
To verify the practical applicability of the model such large companies, were looked for to be disposed of several year's data and were ready to take part in the experimental work.
We choose with mutual agreement those machines to be drawn into the examination. These are tractors with various power, combines, ensilage harvesters, in number 60 pieces.
I put down in table all those data composing the basic data of determining the characteristic parameters according to the equilibrium-equations and to indentify of machines belonging to one-one company. It is noteworthy that the not publical data (for example the marking of the names of the companies, the names of the machine manufacturer companies and the types of machines manufactured by them) can be occured in coded form, as naming them can allude to legal questions. Taking this into account I don't make known such data.

Data base and processing

The datum base consists the following main - measurable – data:

- type of machine, its power, its naming,
- the purchase cost of the machine,
- the purchase year of the machine,
- the material cost used to repair the machine,
- the material cost used to the maintenance of the machine,
- the working hours and wage spent on maintenances and repairs,
- the different works done by the machine,
- the amount of fuel, lubricant used and their cost,
- the social security contribution on wages,
- overheads,
- etc.

As a result of arranging and processing the datum [14.14] such parameters, indicators can be determined giving objective possibility to characterize technically and economically the given machine. During the experimental work I concluded from analysing the machine maintenance – and machine operation characteristic curves among parameters determined by means of measuring, grouping of more than 25000 data. Figure 14.11 shows the process of assembling data and processing.

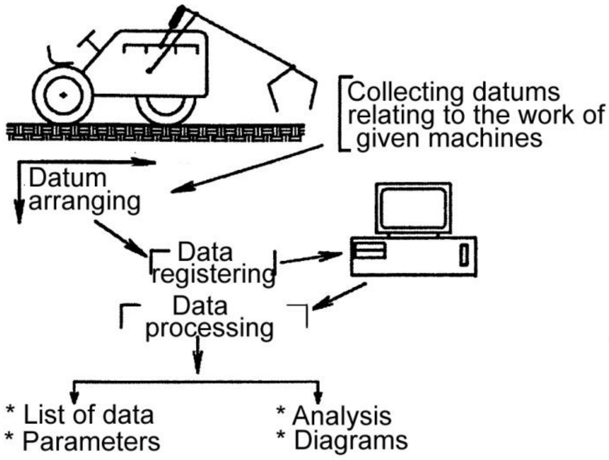


Figure 14.11. Process collecting data and processing.

Results

We indicated partly the specific machine operation costs and partly the specific machine maintenance costs in the function of the work done by them in Figure 14.12, 14.3, 14.14. We indicated the work done in nha (normal hectare) on the horizontal axis. The normal hectare is the unit indicator of different agricultural works (1 nha=25,315 kWh). I indicated the specific machine operation cost (c_{smo}) in Figure 14.12a, 14.13a, and 14.14 and the specific machine maintenance cost (c_{smm}) in Figure 14.12b, and 14.13b on the vertical axis.

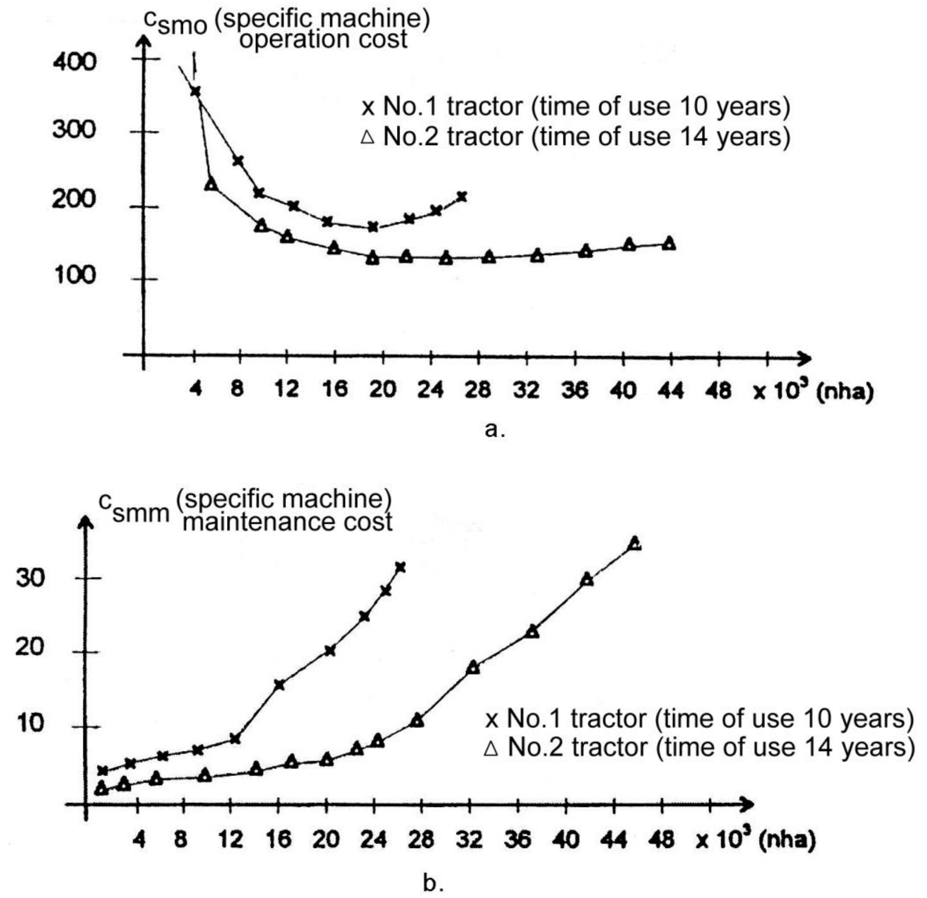


Figure 14.12. Cost characteristic curves of machine maintenance of wheeled tractors.

It can be seen in Figure 14.12 the cost characteristic curves of the special machine operation and the specific machine maintenance of wheeled tractors which verify back faithfully the previous hypothesis made based on Figure 14.9, that the optimal usefulness can exist. In this case as the work done is taken up onto the horizontal axis – as independent variable – the optimal amount of work that can be done is there where there is the minimum of the specific machine operation cost curve. This value at the No. 1. tractor is at 1800 nha, at the No. 2. tractor at 26000 nha respectively. Here it is already worthy of thinking of the company management that how long will they use the given machine, surely the operation beyond the optimum can have got profit reducing effect for the comapany. It doesn't require long reiterated affirmation that evident fact that the company management which will prefer from the two tractors at the next occasion of machine investment.

Let's take one after the other the advantages of the No.2 tractor againts the No.1:

- as the No. 2 tractor produces with less cost the unit product (nha) the company can realize more favourable profit,
- it can do greater amount of work – as with 40% more – within the optimal interval,
- at the optimum surroundings (20000 and 35000 nha) the specific cost doesn't change in significant scale what is very favourable concerning the invessment policy of the company surely the capital can be used more favourable in other field.
- it can be concluded to the reliability of the machines from the steepness of the characteristic curves as the characteristic curves reached too long as the specific

repairing costs (No.2 tractor) also shown in Figure 14.12 give edicence of good construction and manufacturing technology,

- the specific repair cost of the No.1 tractor are significantly greater than the No.2 tractor's and the steep running up of the characteristic curve (Figure 14.12, No.1 tractor) refers to frequent repairs, to second – rate quality of the manufacturing technology.

Let's look at the following examples how the parameters concerning tractors shown are formed in case of combines.

I show the change of the cost characteristic curves of the specific machine operation and the specific machine maintenance of combines with 3 different manufactures in the function of work done (nha) in Figure 14.13.

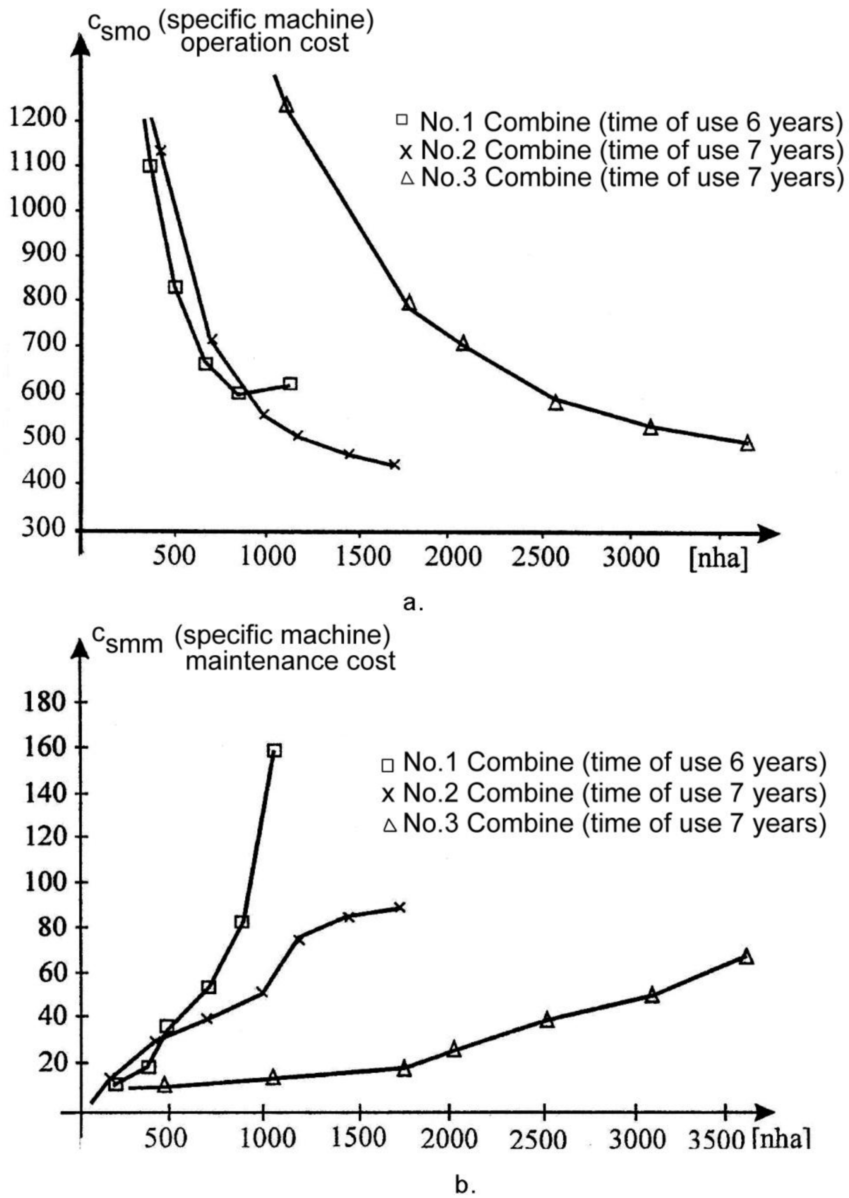


Figure 14.13. Cost characteristic curves of machine maintenance of combines.

It can be seen in Figure 14.13 that the No.1 combine at about 1000 nha already reached the optimum and the company management stopped about after 1200 nha work done as it made the harvesting very costly. That also contributed to keep away from the production that the breakdown of the machine was frequently and as in Figure 14.13.b can be seen its repair cost is also very expensive. The repair cost of the No.1 combine was tenfold comparing to No.3 combine at 1000 nha fulfilment. Despite that the other two combines didn't reach the optimum yet, it can be evaluated that the best selection with very high probability is the No.3 combine for the company. This statement is supported specially well by the specific machine maintenance characteristic curve – it can be seen in Figure 14.13.b – that indicates the company was successful to buy a good quality and reliable combine. That is naturally in case of further purchasings it will invest combine with this type and manufacture.
Machine operation characteristic curves of two caterpillar tractors can be seen in Figure 14.14.

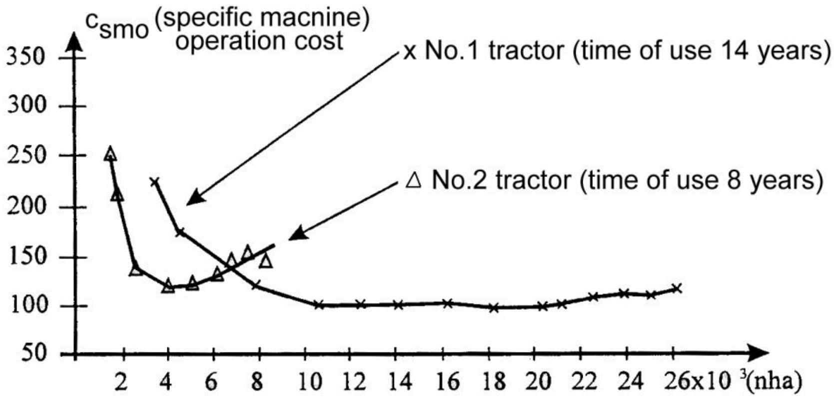


Figure 14.14. Cost characteristic curves of machine operation of caterpillars.

Task.
Try to qualify the tractors based on the analyses so far, value their technical level and using, by which rises greater company profit, if you are the companies manufacturing tractors, what conclusion do you draw on the basis of characteristic curves and what steps would you execute in connection with further manufacturing of No.1 and No.2 tractors.
Further on it will be demonstrated that in case of certain boundary conditions the (Zsoldos-Janik) model what information provides for strategic machine purchasing decisions projected to time horizon.

Transformation operator
It can already be seen the well known machine operation characteristics of the wheeled tractors in Figure 14.15. The power machines with different manufactures and powers are marked with numbers as their concrete marking isn't public.

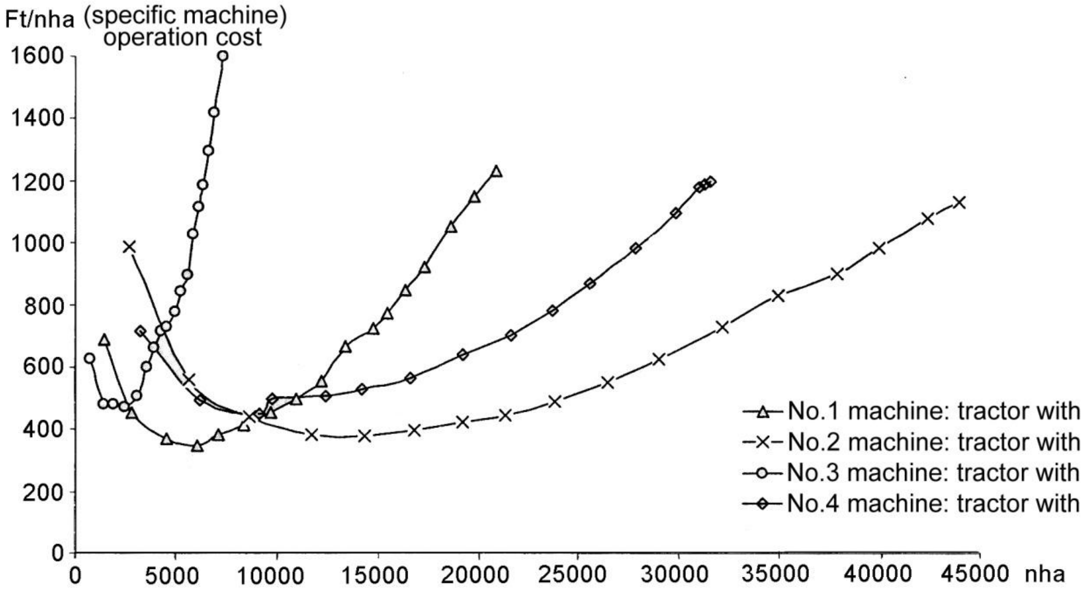


Figure 14.15. Machine operation cost characteristic curves of wheeled tractors.

You can do excellently based on the analyses discussed in detail in precedings that for a given company which power machine is the most advantageous concerning the standpoints of economycalness, the reliability and repair – sensitiveness. We only wish to affix to the analysis that according to my hopes you will also declare the No.2 power machine is the most suitable. The question however can also be put if the power machines perform the same (respectively quasi same) activities then whether the characteristics of the machine change, and if yes, what can be concluded. To decide this we defined a so called transformation operator [14.15],

$$\bar{\Delta} = \begin{bmatrix} \Delta(t_1) \\ \Delta(t_2) \\ \Delta(t_3) \\ . \\ . \\ . \\ \Delta(t_N) \end{bmatrix}$$

using it an activity done by a so called basic machine is with computer simulation process performed with another power machine, which is called virtual machine practically.

The $\Delta(t_i)$- are the multiple factors of the parameters belong to the given year according to content they show that a virtual machine works with what cost level in the year (in the *i*-year) given. In the next step is the same work performed with the transformation operator virtally (with computer simulation) what the chosen machine fulfilled.

As a result objective measuring numbers are calculated that the virtual machine would have worked on what cost level among environmental conditions of the chosen machine.

Comparative diagrams are plotted from the data to be demonstrated by the machine operation characteristics of the basic machine and virtual machine (Figure 14.16).

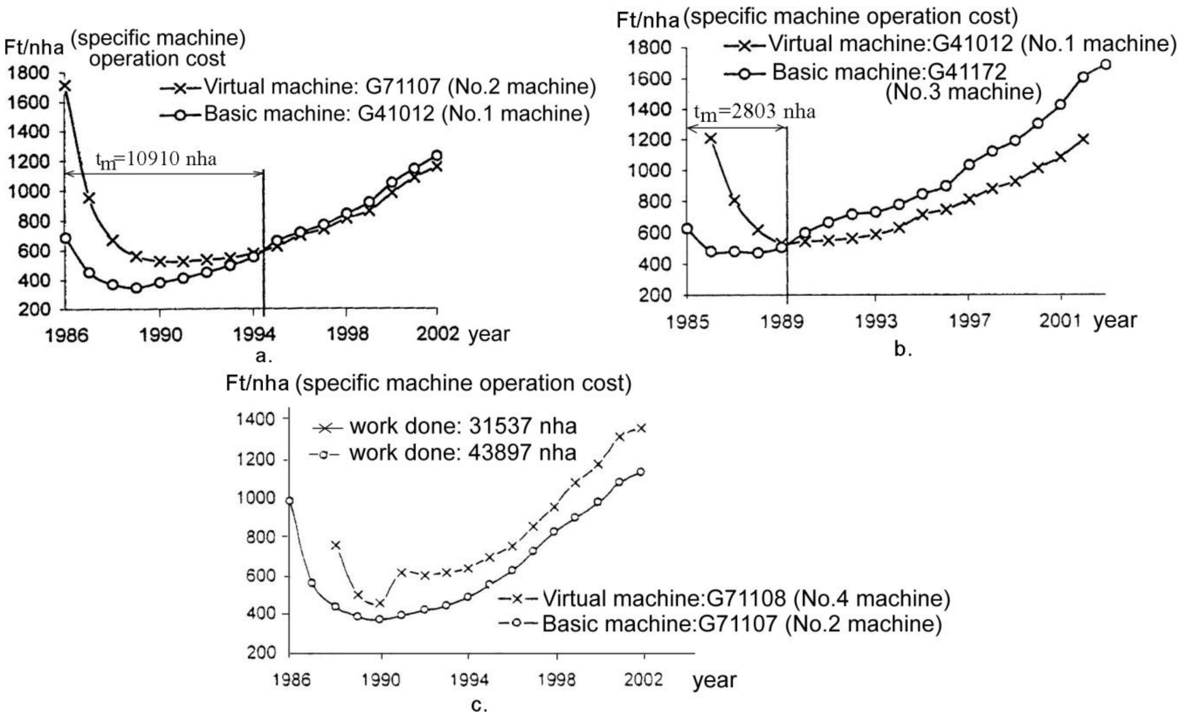


Figure 14.16. Machine operation decision information projected to period.

It is shown in Figure 14.16 that the comparison in pairs of power machines No.1, No.2; No.1, No.3; and No.4, No.2 shown in Figure 14.15 can give possibility to conclude.

Figure 14.16 shows unambiguously that in case of long-range operation concerning the final result there is no difference between the two machines, surely in case of operation beyond the intersection (t_m) of the two operation characteristic curves the costs are nearly the same. That has to induce the investor to consideration that the basic machine needs substantially smaller investment cost than the virtual machine. The suboptimum at the t_m surely calls attention to the importance of long-range strategic decisions. Figure 14.16 shows a typical case. Namely there is a definite difference between the two operation curves before and also after the intersection (t_m). The economic strategy is the following in this case: if the expected time of operation is not longer at t_m (it is near of it respectively) the No.3 machine is chosen thus it can be seen well at the starting costs that the requirement of the investment capital of the No.3 machine is smaller substantially than of the No.1 machine.The savings are substantial what the company can utilize favourably by other profitability investment. However in that case if the expected operation time is much longer than t_m, then the company decides beside investing No.1 machine.

The Figure 14.16. c shows that special case when there is evident totally that the No. 4 virtual machine can't be the strategic alternative of the No.2 basic machine.

14.5.3. Summary

The technical – economical parameters provided by the model worked out by us prove that the Zsoldos – Janik computer machine operation professional system operative application is founded, it provides objective informations for the designers, manufacturers and operators alike of the various of great value machines. That is also naturally that such a system of great value can't be introduced to such level which could make also possible to know the parts, so I

could strive only to present the theoretical structure and the use of the system – theory model. However by the means of analysing real examples I wished to emphasize the importance of the system-view and the system – theory in the optimum searching process.
The following main mathematical connections compose the professional system:
1.) material-flow (energy-flow) balance-equation of the process:

$$x(t_N) = x(t_0) + U(t_0, t_N) - Y(t_0, t_N) \pm K(t_0, t_N)$$

2.) the regulating balance – equation of the process:

$$\Delta V(t_0, t_N) = V_{Kf}(t_0, t_N) - V_{Kr}(t_0, t_N)$$

which can be aim – function practically as the profit of the process is maximal then if the costs – in case of certain boundary conditions – are minimal, that is

$$k_r = \frac{V_{Kr}(t_0, t_N)}{T} \rightarrow min$$

3.) the transformation operator of the process,

$$\bar{\Delta} = \begin{bmatrix} \Delta(t_1) \\ \Delta(t_2) \\ \Delta(t_3) \\ . \\ . \\ . \\ \Delta(t_N) \end{bmatrix}$$

14.5.4. Questions, tasks

- Characterize the company (venture, share company, etc.) according to system – theory,
- Analyse that the aim of a company (venture, etc.) to maximize the profit gets priority in each case or it has to take into consideration other standpoints in certain cases, too.
- Draw sketchy the system-theory input-output model of the company indicates the characteristic factors.
- Write down based on the model the balance – equations to be characteristic to the company and interpret the meaning of certain factors in general.
- Analyse based on the balance-equations that the machine maintenance how can contribute to maximize the company profit.
- Make it known the applicable method – among profitable company conditions– capable to determine the optimal service – life of the machine.
- Analyse the conclusions that can be drawn on the basis of the types of the specific machine operation and machine maintenance characteristics shown in Figure 14.12, 14.13, and 14.14, first of all from the operator's standpoint but also touch upon that what useful information is provided for the designers, the manufacturing companies and also for the merchants.
- Analyse the technical – economical importance of Figure 14.15 and 14.16, evaluate the effect of stategic decisions projected to suboptimum time-horizon to the company profit.

14.6. Organizing mechanical engineering processes

The questions connected with economy – and organization can be answered with the most difficulty in exact respectively in quasi-exact form also in the mechanical engineering as in other production processes. It can be explain first of all that the repeated measurings for examining questions connected to organization would be very costly, on the other hand it is very poorly the assortment of the reliable (mathematical) algorithms capable to apply for very complicated organizational processes. Basically main problem is that these processes can't be described with so called determinant connections and in consequence of this the result can be made only probable. Many disturbing effects can arise between the input and output which can change significantly the final result planned, too. It is a typical organizational question the company cooperating with more transporters having reserves only to limited time guaranteeing the continuity of the production. In case of straight – line production it is enough if only one transporter is late, that the whole production stops at once and the losses can result huge profit loss.
It can be shown clearly with the following simple example the reciprocity of the extreme complex organizational variations of the part production with the economical manufacturing.

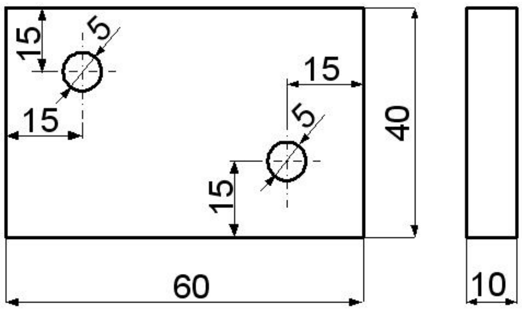


Figure 14.17. Shop drawing for part with hole.

The task is to make two holes with 5 mm diameter into the part shown in Figure 14.17 As an example let's take the following cases:

1.) It has to be made 2 pieces of parts, the limit of accuracy is ± 0,3 mm,
2.) It has to be made 25 pieces of parts, the limit of accuracy is ± 0,1 mm
3.) It has to be made 250 pieces of parts, the limit of accuracy is ± 0,05 mm
4.) It has to be made 10 pieces of parts, the limit of accuracy is ± 0,01 mm

No.1. case:

The so called individual production organizing form characterizes typically to produce one-two products:

- a locksmith gets the task, who draws the places of the holes and marks the centre with centre punch, selects the twist drill with appropriate dimension, if it is needed he grinds it then he makes the hole requested on the drilling machine with suitable dimension. It should be note that this technology can be allowed then if the accuracy is not more exact then ± 0,1 mm. Inasmuch the prescription requires tolerance with hundredth millimeter then the operation should be carried out on high accuracy jig borer in case of one part already .

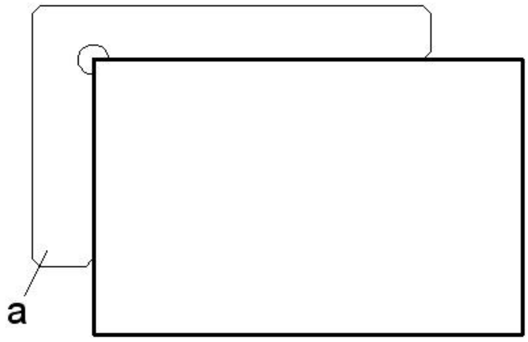


Figure 14.18. Specifying position by stopping.

No.2. case:

Really the ± 0,1 mm accuracy can be accomplished still with very careful work by the method made known in No.1 case, however in such a case still arises the question of economicalness in the technologist and the making time of one part can be reduced significantly by simple stop-block (*a*) shown in Figure 14.18. Not taking place the time needed to mark out and to orient into to be co-linear the tool centre print with the centre of the hole. The location of the part is altogether some seconds surely the part has to be pressed only by hard to the stop-block (*a*) fixed onto the table of the drilling machine. The location is guaranteed by the table of the drilling machine and the *a* stop-block, the pressing force is provided by the worker's handpower.

Task.

Determine that the number of pieces at which manufacturing more parts make possible that the No.2. method can be carried out with less specific (auxiliary time needed to produce one part) process of time than with No.1. method. The cutting time (main operation time) furthermore the time needed to clamp – and to sharpen is considered the same in both cases, so with these elements of time it is needn't to calculate. The data needed for calculations are the followings.

No.1 method:

- preparation of instruments needed to mark out (scriber, centre punch, caliper, surface preparation to drafting) 30 min/25 pcs.
- marking out and punching, 5 min/pc.
- to orient the centre of holes of the workpiece to the tool, 2 min/pc.

No.2 method:

- it is supposed that 100 min. is needed to manufacture the stop-block,
- 20 min. is needed to orient the machine table and the stop-block together to the tool and to fix the stop-block to the table,
- making impact of the workpiece to the stop-block to drill the first hole (0,5 min.), turning to drill the second hole 0,5 min./pc.

No.3. case:

In case of 250 pcs.of parts and the required accuracy already suitable jig is needed for manufacturing. The so called drilling jig can be used economically for such tasks, which by means of their constructions guarantee always the identical places of the workpiece, so the jig bushing to be in the jig determines the place of the hole.

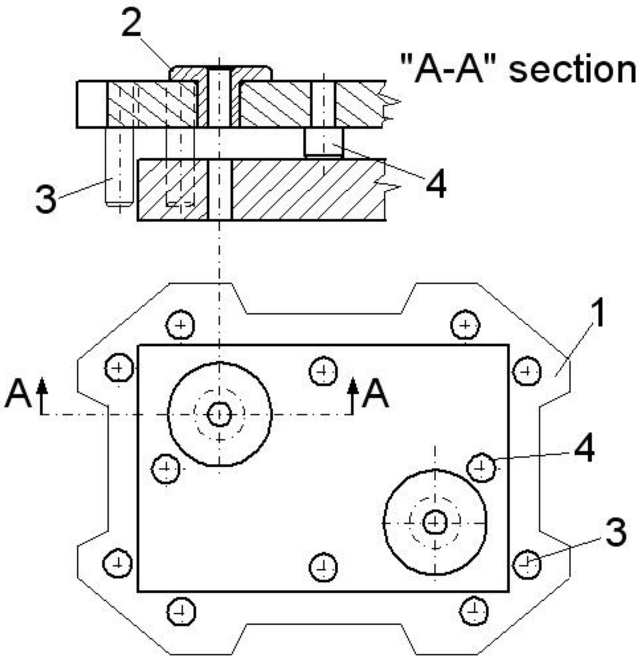


Figure 14.19. Specifying position by plate jig.

Figure 14.19 shows one possible variation of the task's solution which substance is that all elements needed are built into the plate jig (1):

- the jig bushes (2) guarantee the placing of the tool,
- 8 pcs. of limiter pins (3) guarantee the placing of the part,
- elements (4) are needed to press the part onto the table.

Such and similar to this plate-jig solutions (inasmuch the accuracy requirements can still be kept) are favourable because of the shorter auxiliary times. Drilling the holes altogether consists to place the workpiece onto the table of the drilling machine then the plate-jig and the operation can be made. The pressing of the workpiece onto the table is made by hand-power with the elements (4) fixed into the plate-jig.

No.4. case:

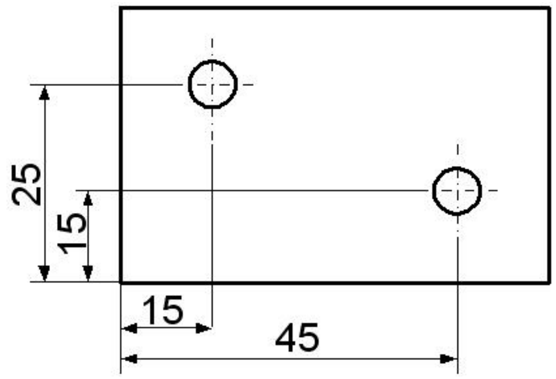


Figure 14.20. Specifying hole position by surface locating element.

In case of close-limit accuracy (0,01 mm or 0,001 mm) requirements the jig borers are used very favourably whether piece production or serial production comes up. The position of the

holes as compared to each other is determined to the basis which is the perpendicular planes of the workpiece.
The special built in measuring systems in the machine – tool guarantee the necessary accuracy.

- It is worth to consider that the solution of such a task seeming to simple what kind of knowledges, requires too. Surely is has to be known the mechanical engineering technology, the machine – tools, the jig clamping workpieces, and jigs guiding tools to choose the optimal the quasi-optimal solution respectively.
- As the organizing processes of mechanical engineering are extremely ramifying they include multiple questions because of these further I shall deal with such parts of questions which understanding don't require (highlevel theoretic) complementary knowledges. As you will see a part of the models have got verbal characters to be formed based on for long years rather on the practical experiences of decades.
- For those inquirers who would like to get deeper knowledges in the topic of planning manufacturing processes, I suggest them those Hungarian authors' works [14.14, 14.20] which knowledge are indispensable for engineers working in the field of production, research and education.

14.6.1. Organizational forms, models of mechanical engineering

Under the organizational models of mechanical engineering – in a narrow sense – are understood really the arrangement related to each other of those machines, instruments and work sites (plant arrangement) where it can be accomplished the manufacturing of the part, the part unit, the machine given. Literature [14.21, 14.22, 14.23, 14.24] are at disposal for the inquirers in the special literature.
It is very important to consider the followings during planning the arrangement:

- to develop the possible most favourable machine, operation, line of operations (minimizing production area),
- high production volume (favourable income),
- low costs, (maximal profit),
- customers' best service, (favourable market stability),
- to satisfy the company workers' demands (insignificant work force migration).

The aims listed are important alike in case of planning new plant or developing existing plant.
The various organizing types can be put into three groups according nowdays to the classical arrangement – naturally it can also be said existing standard models -.

1. **Defined work site**. Its typical character is that the product (the object of the work) remains on the spot and the workers go from one spot to another taking with themselves tools, instruments and maybe parts needed. Its advantage is first of all in case of heavy products occupying big space and as a consequence of these are movable with difficulty. The arrangement can accomodate to the changing requirements, it doesn't demand costly planning, it is characteristic piece – and small series production. It is suitable alike to produce new products and to maintain or repair the used ones. Its main fields of application railway – and tram carriages, buses, agricultural power – and machines, road machinery. It is characteristic the defined work site organization first of all to work sites formed to service and maintain lorries.
2. **Arrangement according to technology**. Its characteristic is that workshops, respectively plants are formed from work sites needed uniform technology. So for example foundry, smithery, welding, heat treating, cutting workshops can be formed,

because of this it is called workshop piece production respectively. It advantage is not to be sensitive to the changing of variety of products furthermore to the operation sequence, to the lack of materials, to the workers' absence, the machine breakdown doesn't endanger the production continuity, it is characteristic to piece-and small series production. This form of organization is characteristic to producing new products and to see to the maintenance and service workshops connected together with car-saloons organized to service passanger cars. Certain processes can be reconciled only partly because of this the transit time is relatively long.

3. **Arrangement with objective (product) according to production respectively**. It is typical that they produce one product or type of product. The product (part) according to the operation sequence gets from one work site to another. The so called grouped or continuous production is formed from the arrangement of work sites (machines). Its advantage is that the route of material handling and the amount stored decrease substantially, the quality level can be increased favourably by the means of higher level specialization (it is economical in case of due great number of pieces). Its characteristic is the mass production furthermore the work is done on the basis of manufacturing documentation dissolved into elemental working processes. Its important characteristic is that in greater part single-purpose machines are applied with using jigs mainly and tooling – up. It isn't a simple task to organize (to synchronize) the operations succesively but together with this the transit times can be formed favourably.

   In order to understand better the mentioned things before let's have a look how can be made clearer the forms of organization with exemplary figures.

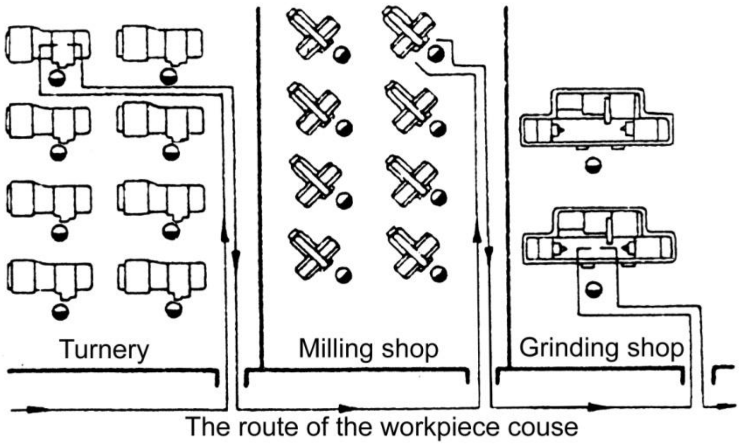


Figure 14.21. Organization according to workshop (piece-production).

Basic characteristic is the organization concerning technology that the work sites are formed by a way suitable to accomplish works with similar characteristic as a result of which it can be talk about turnery, milling shop, grinding shop (Figure 14.21) etc. in general it can be talk about workshop principle arrangement. That means that the workshops executing the same technology form workshops, respectively plants. Depending on the multiplicity of machines, instruments, operations needed to producing or re-producing (repair) product it could be listed wide scales of workshops: for example turnery, milling shop, grinding shop, welding shop, plating shop, body ironer shop, diagnostic shop, disassembling shop, part cleaning shop, assembly shop, quality control shop, etc. The product to be produced is finished by means of more workshops common work, the workshops do one-one technological phase of the product and in consequence of this the workpiece can also be got back just several times to the same workshop. In order to reduce the routes determining the operation sequence of technology has got great importance.

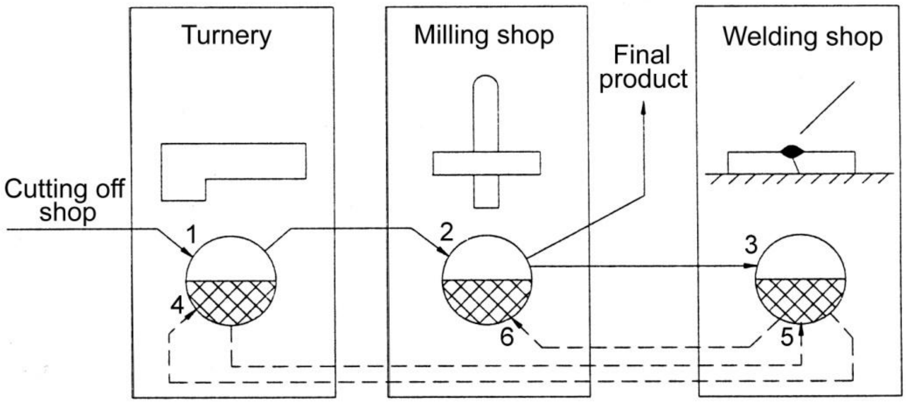


Figure 14.22. Transport route deriving from operation sequence.

The Figure 14.22 shows well that the product what long route has to do there and back if the technologic instructions prescribe the operation sequence:

1. Turning,
2. Milling,
3. Welding,
4. Turning,
5. Welding,
6. Milling

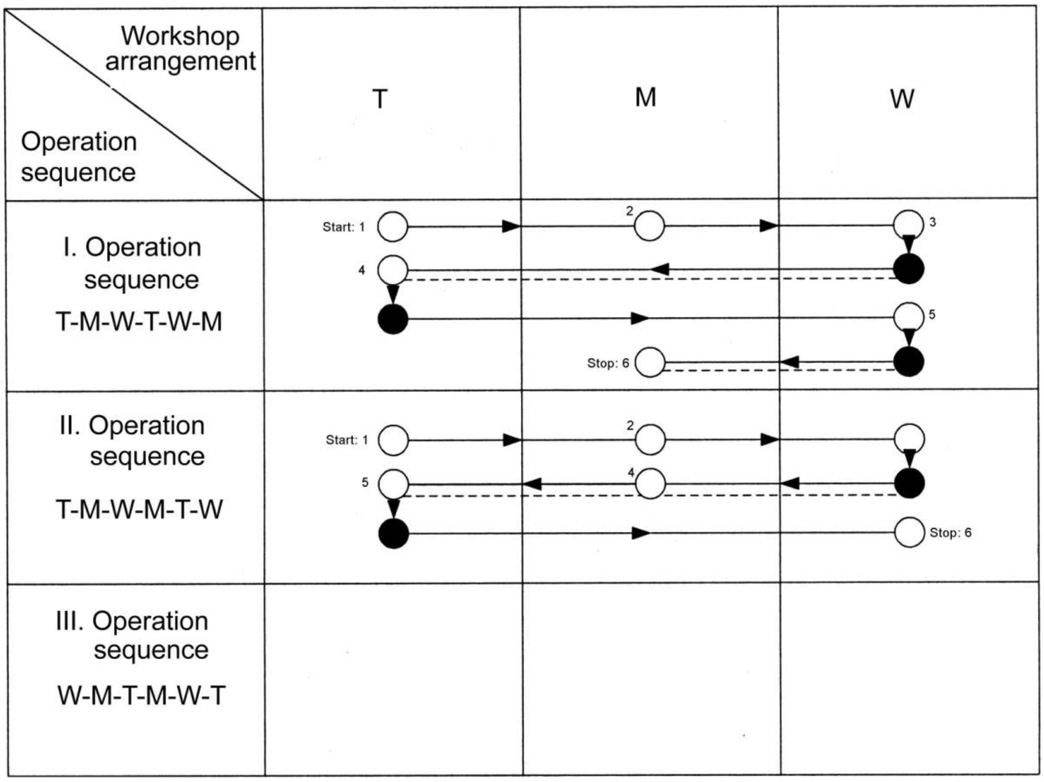


Figure 14.23. Reciprocating transport route in case of interchangeable operation.

The example is shown for using the so called back-route diagram that can be applied then if succession of certain operations can be inverted. This is important therefore because with such simple process figure it can be determined the minimal transport route of the part which has got reducing effect to the costs naturally. That is the product can be produced with less cost of production.

Task.
Make the back-route diagram according to the III. operation sequence, then decide which operation sequence provides the shortest transit time.

This example shows immediately also to one of the disadvantages of the workshop production as the transit time of the product is very long. Its further disadvantage is that the connection of the workshops is irregular between each other, but within the workshop there is no regular work connection between certain work sites yet, what also touches disadvantageously the utilization of machines and this appears naturally in the higher costs. Its advantage is that the machines arranged according to workshop principle can be well-arranged, the technology can be checked, properly it isn't sensitive to the change of profile, it can switch over producing new product flexibly as the universality is characteristic to the resources (machines and skilled workers). The organization according to workshop principle satisfies typically the conditions of the production (mechanical engineering) of individual products.
It is within the scope of **organization with objective standpoint** on collective respectively continuons manufacturing system.
Organization with collective standpoint. The form of collective organization in its original interpretation placed the machines, work sites suitable of the technological sequence spatially near from time to time in one workshop. Later on it was developed of the form of collective organization a version oriented to one part at which it is needn't to change the arrangement of machines, that is the workshop principle arrangement can be kept.
The grouping with objective standpoint of the work sites is expedient in that case if it is mentioned on a product returning into the production with due frequency with great amount relatively. In such case the work sites are organized to one place according to the requirement of technical operations (for example turning, milling, drilling, welding) namely suitably of the operations listed: 2 pcs. of lathes, 1 pc. of milling machine, 1 pc. of drilling machine, one welding work-bench.

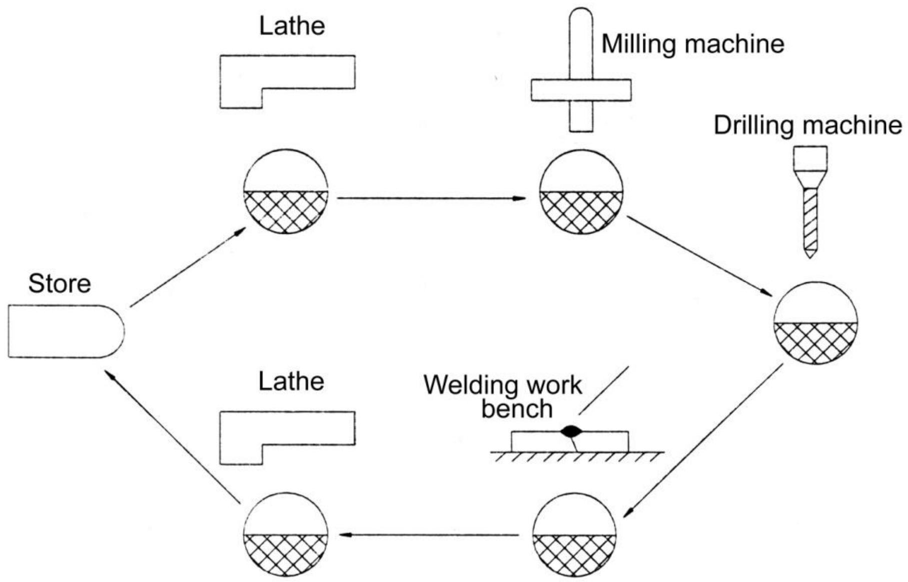


Figure 14.24.Work-site arrangement suitable for operation sequence (grouped organization).

Figure 14.24 shows the sketchy arrangement. There is a possibility to such arrangement in the case of workshop manufacturing if the amount of certain products makes this reasonable. The manufacturing groups formed so are often organized as independent workshop.
Instead of collective production still the cyclic production is also used the arrangement itself is called producing cycle. It is characteristic of the collective (cyclic) manufacturing system that substantial part of the operative equipment are universal in order to accomodate well to the change of product ensuing in turn. Its advantage is still the routes are shortened significantly (what is shown well at the comparison of Figure 14.22. and 14.24.) in consequence of the production equipment drawn together into small place spatialy, the costs become favourable by the means of good utilizability in the consequence of economical use of jigs and the tooling-up and the large number of pieces.
It should be noted that the cycles can't be burden in one-hundred percent in majority of cases with products planned and because of it has to be calculated at any time with operations aiming the utilization $\left(\sum t_n\right)$ and the operations coming in the cycle $\left(\sum t_c\right)$.
The number of machines (work sites) needed to complete certain operation (M) is given by the requirement of time for operation $\left(\sum t_r\right)$ and the quotient of the avaliable productive time (T_M) per machine:

$$M = \frac{\sum t_r}{T_M}$$

Three parameters are used to judge the group economically:

- the average burden (utilization) of the group,
- the degree of closed condition related to the product,
- the degree of closed condition related to equipment.

The average burden of the cycle (a_b) is the sum of the operation time of all products machined in the cycle ($\sum t_{in}$) divided with the productive available time of the machines pulled into the group $\left(T \cdot \sum M_T\right)$

$$a_b = \frac{\sum t_{in}}{T \cdot \sum M_T}$$

The burden of the cycle is all the better as much better approximates this value to the one. The degree of closed condition (*DCC*) means that what proportion of part of the total operation time ($\sum t_{in}$) of products counted to the group goes on within the cycle, what can also be drafted that the operations coming out $\left(\sum t_c\right)$ in what extent loose the closed condition.

$$DCC_t = \frac{\sum t_{in} - \sum t_c}{\sum t_{in}}$$

The degree of closed condition related to producing equipment (DCC_{pe}) expresses that the time of operation going on there of the products counted to the group $\left(\sum t_{in} - \sum t_c\right)$ amounts what proportion of part of the total operation time going on in the producing group, that is the operations coming into the cycle $\left(\sum t_e\right)$ loosen the closed condition in what extent.

$$DCC_{pe} = \frac{\sum t_{in} - \sum t_c}{\sum t_{in} - \sum t_c + \sum t_e}$$

Basically it can be said that the time (transient) to be spent producing product with grouped respectively with cycle production organization is wished to reduce. Similarly it makes possible reducing the transient time the grouping based on certain standpoints of the parts to be machined.

The essence of the grouping is that in case of products manufactured in piece or some number of pieces one-one part of the parts on the basis of certain similarity signs can be listed to one group and so the production can also be organized as whether small or large – series in the organizational form of the workshop production.

Great advantage is that preparing the production becomes simple, the transient time is more favourable, uniform technology can be worked out for the group. The grouping of parts is possible according to different standpoints. Characteristic groups can be formed according to the followings:

- requiring same machineries,
- workable with the same technological operations,
- to be manufactured with the same operation time,
- having the same shape (for example shafts, discs),
- parts making from the same material.

It has to be mentioned definitely the continually wider spread of the computer added designing and of the intelligent integrated manufacturing systems from which it can be got information in the various special literatures (for example [14.19], [14.20])

Organizing continuous production system. The work sites are placed relating to each other according to the arrangement of the objective principle of work sites. The workpieces get from one work site to another according to the operation sequence. Every operation belongs to other and other work sites.

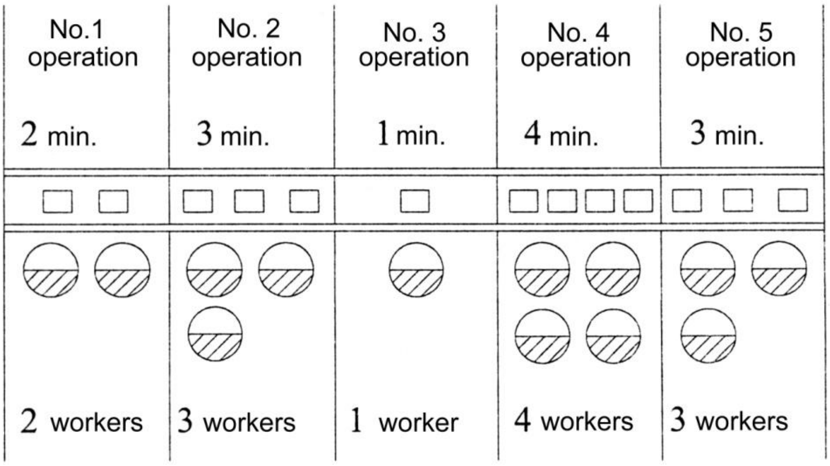


Figure 14.25. Straight-line (continuous) production.

Figure 25 shows clearly as an example that in case of one straight-line production supposing 1 min. straight-line passing time (cycle-time) how many workers has to put into work in certain operation places.

It is only then economic organizing continuous production if its heap of a sort is very high. Its advantage is the transit time of complicated products can be short extremely the production area and the available time of equipment can be utilized favourably, the production can be well directed, the production management can be simplyfied. Synchronizing of operation times to be made on different work sites means the greatest difficully. The synchronizability of the work sites can be said ideal then if it is succeeded to determine operation times needing equal intensity in every work site section. This is a very hard work needs a lot of time technological operation analyses, which is refunded at very great number of pieces only. In case of straight – line production special single – purpose machines are put into production generally trained workers do the majority of the work. However workers with very high – level professional qualification are needed to machine maintenance activity in order to prevent failures quickly and workmanlikely as stopping any operation place – for the time of preventing the failure – calls forth the stop on the total straight-line production what results significant increase of cost.

Let's see the three production organizing models mentioned that which are those main steps that can be the characteristic centres in planning process concerning preparing onto a new task (for example to producing some kinds of product).

14.6.2. General model to plan a product to the production process.

Organizing a production process based on certain precedents (market-research, marketing, social requirement, tender etc.) requirement arises to produce some kind of product. The precedents can present themselves in two forms:

- what has to be produced,
- what is the number of pieces,
- at which date,

the planning can concerned:

a.) planning of the producing – system which can tend towards the product, the process, the equipment, the workpower in totality towards the whole plant or company activity.

b.) regulating the production processes to directing production programs, to regulating the stocks, - the products, - the productivity.

Thus basically it has to be distinguished sharply from each other the planning of the production process and the regulation of production process despite that the planning and regulation act almost in unity at any time in the field of production. The scope of the management is determined first of all that whether the question is from developing existing plant, process or greenfield investment. In case of development great restiction appears already if it has to be made strategic decision and so the changes will be accomplished within the framework of tactical decision first of all.
In case of all new investments however the strategic decisions get important role. The multiplicity of tactical decisions has to be made in the practical situations comparatively with not great energy devotion at the same time the strategic decisions require greater preparations, more significant energy devotions. The regulation is connented to the process of execution, the operations and to the evalution. The complexity of this question is shown clearly in Figure 14.26 [14.8] that shows the qualitative changing of financial characteristics belonging to main centres of total cycle of some kinds of product.

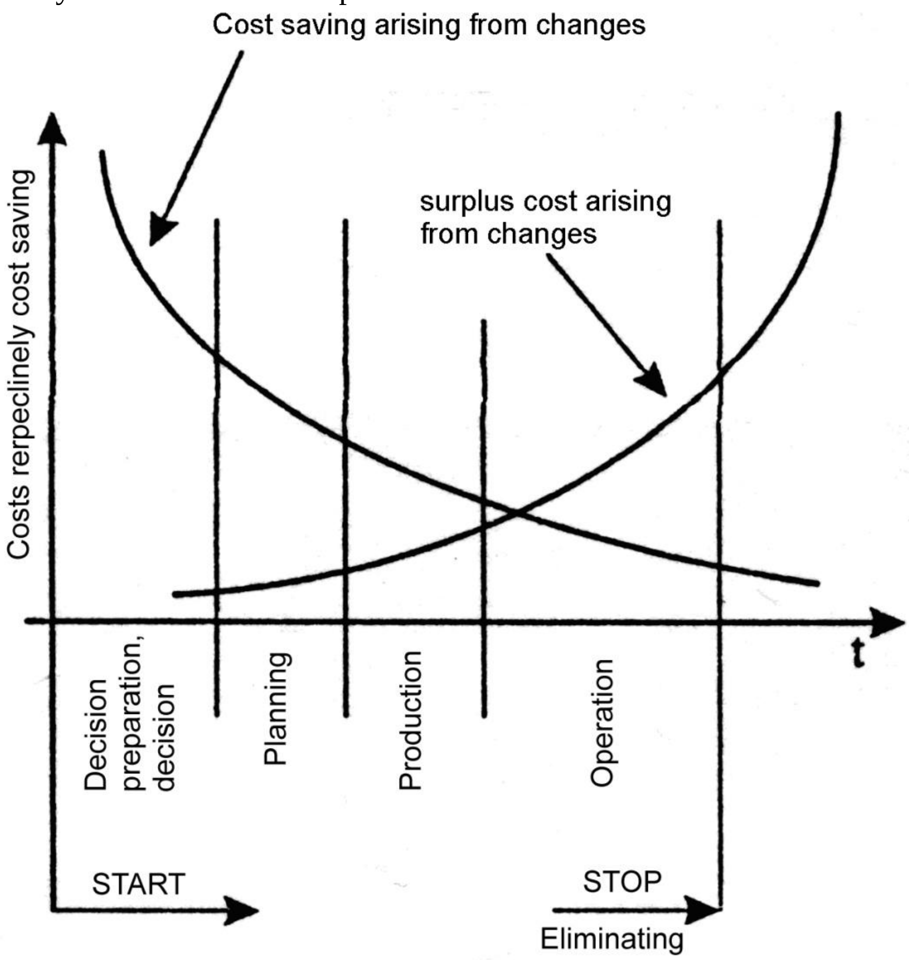


Figure 14.26. Financial factors belonging to main characteristic centres of machine life-cycle.

Based on Figure 14.26 it is very instructive to think on the qualitative changes of financial characteristics belonging to characteristic centres of the main process lasting from the creation of the machine to the machine rejection life-cycle of the machine. The main activities belonging to the life-cycle of the machine can be found on the horizontal axis in chronological sequence, the costs on the vertical axis however. It can be read that the activities following one another in chronogonically what effect have got on the product cost. The two curves – still if qualitive only, too – show well that what competent cost saving possibilities respectively changing costs belong to certain activities. During activities decision preparation and decision extremely great savings can be reached with small changing cost. The changing costs in case of activities following one another are ever greater and at the same time the cost saving possibilities are ever smaller. The changing costs in case of machine maintenance activities needed in operation period are the greatest. This also means at the same time that in case of right economical regulation the companies why reject the machines become obsolate morally and technically rather and don't repair however. This also means that it has to be dealt with the quality improvement in the periods of the market research, the aim-planning and decision preparation, the planning, the designing and manufacturing (thus the determining of the machine quality and reliability) of the machine life-cycle- as in the further period of the machine life-cycle (operation, maintenance) there is real possibility to guarantee the quality and reliability of the machine bought respectively to guarantee as much as less degree deterioration. By means of complying with the workman-like operation specifications to guarantee the quality can be done the most in the period of the machine life-cycle-as the practice shows that, too.

However it is fact that the machine has to be also rejected after certain working hours against to the workmanlike operation and maintenace, as it was presented in the chapter 14.5 based on concrete operative measurings.

Three large groups can be made from the life time of the machine:

- **Reuse**: those parts can be listed here which can be built back into their original place without changing, they can be utilized is some kind of form on other field or they are built back into their original place as renewed parts, or entirely new other product is produced with some kind of technology.
- **Destruction**: the non utilizable parts or other wastes (for example used oil, plastic parts, textiles etc.) are burnt with observance of regulations of environmental protection.
- **Long-range storage**: those dangerous wastes which destruction are not possible because of technical reasons or is not economical has to be stored with observance of regulations of strict environmental protection.

It is important to mention that from the standpoint of the production safety it has got decisive importance to comply with the strict regulations of service, running maintenance of machines, equipment.

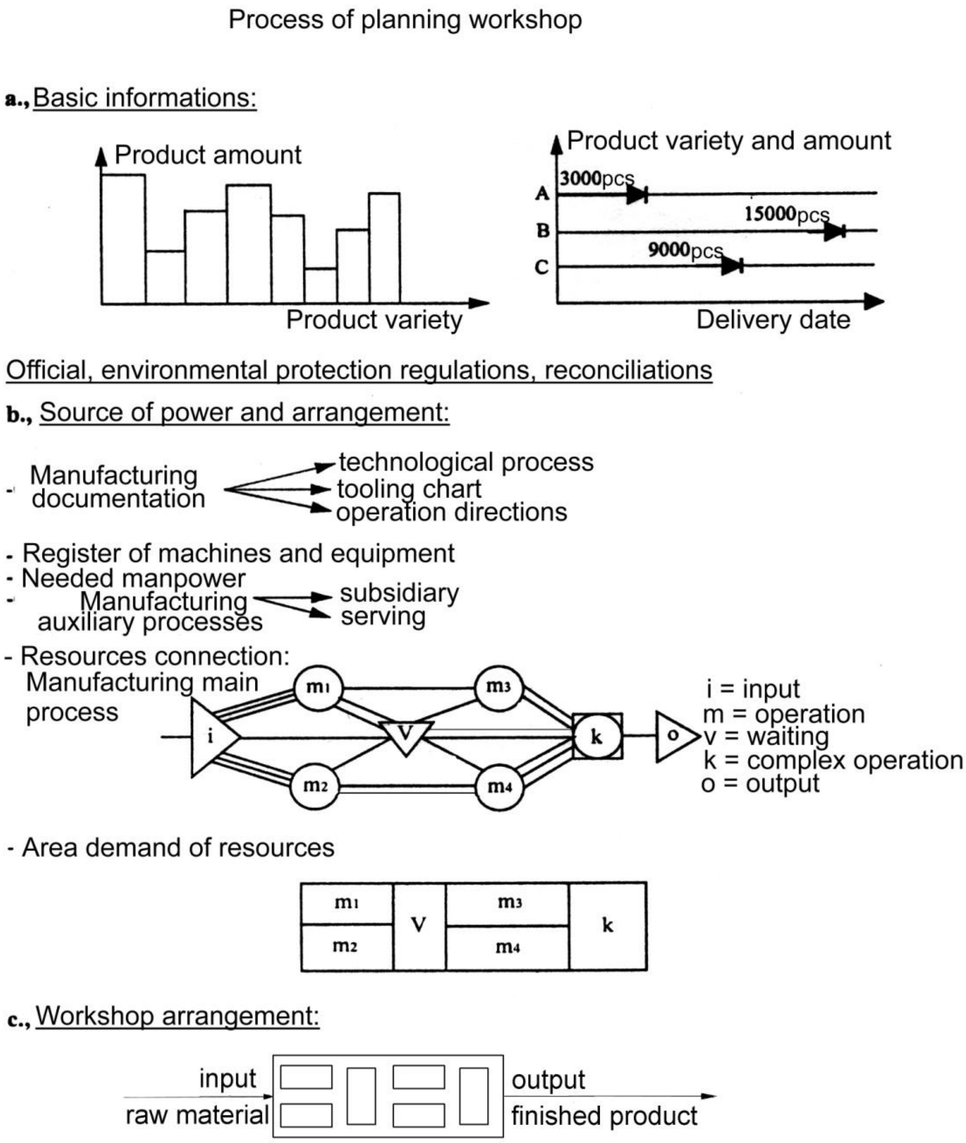


Figure 14.27. Workshop planning process.

It can be followed with attention the process of locating new plant in Figure 14.27 indicating the consecutive sequence of centres stressed (*a, b, c*). This logical process is also true to developing new plant with that limited provision that it already has to be taken account the existing establishments, machines, equipment, human resources.

The main phases of planning are the followings:

a.) Basic informations:
- What has to be manufactured,
- In what quantity,
- To which date, in which delivery timing.

b.) Parameters respectively data to be determined based on starting data:
- The manufacturing process, operations,
- Deadlines,
- Equipment, machines, instruments,
- Workpower,

- Manufacturing auxiliary processes (serving-, subsidiary processes),
- Connecting system of certain activities,
- Area demand of certain machines, activities.

c.) Detailed disposition drawing from workshop arrangement in accordance with space requirement of certain machines, operations taking into account the guarantee of the routes of material flow.

Regulating characteristics: The process-regulating system of production includes the following main elements:

- stockpiling,
- quality securing,
- production programs,
- costs.

It is important to take into account the regulations concerning the national – and international standards, the regulations of the authorities and the local authority.
The importance of standards and the regulations of local anthority appear in the phase of decision preparation as the quality securing and the environment protection have got significant role in the sphere of plant, industrial activities.
I sum up the more important characteristics of the process' main phases in the followings. The basic informations include the what the how many and the deadline. In the framework of what has to be manufactured is usual to give the name of product, its mark, that is its some kind of specification in order to identify the possibilities. Further information can be given with the complete drawing documentation (assembly and detail drawings, part drawings), power-, weight- or geometric data.
The quantity means that number of pieces what the customer determines in the contract in every case.
The deadline (times) has also to be fixed in the contract in order to avoid later debates.
It can already be judged from the construction and volume of the product to be produced those, technological processes by means of the product is finished. So it can also be measured that whether it has to be developed piece-, serial – or mass production organizing form furthermore the by–products of different technologies mean what environmental damages and what measures has to be executed to moderate or to stop these in the planning phase. It is expedient to show on clearly arranged columnar – and linear diagrams (Figure 14.27 the **a** part) the volume and finishing date of the varieties of the product.
In case of development at determining the resources demand it has to started from the company profile first of all as it has got determining role basically that the company has got what specialization. In case of a company with diverse profiles the universal equipment count first of all while in case of a pure homogeneous production profile rather the single-purpose machines are more dominant. The technological documentations have got basic importance naturally to determining the human and mechanical resources containing the informations needed. The machines needed can be determined if the total time of rate comes forward in year's level requiring identical technology is divided with the real annual working hour capacity determined according to the appopriate shift-number of one machine. These calculations have to be made to all fields of speciality (turning, milling, grinding, loksmith, welding etc.). Thus the need of machines in case of a technology given:

$$M = \frac{T}{A_{tM} \cdot n_s}$$

where:

M = the number of machines needed,
A_{tM} = the real annual basic time of one machine concerning to one shift,
$T = \sum \frac{t_p \cdot N}{60}$ the summed up piece period needed to yearly machining,
t_p= the time needed to machine one piece (min.),
N = yearly number of parts machinable on the machine with same type and dimension,
n_s = number of shifts.

The calculated value can be a floating number, which is rounding to integer value, of course. The values got are rounded upward generally. Suppose that the periodic checking, inspections as well as the running repairs are made on days to be kept as a holiday or during free shifts.
At determing the manpower also has to be set out from the yearly requirement of work arising on certain work sites. The number of machine workers with various disciplines and categories as well as the number of skilled workers working on the other work sites (welding, locksmith, assembling, quenching, etc.) can be determined per work site. The technological documentations (tooling chart, operation direction, etc.) contain, the man-hour demand arising in the period of plan.
Thus the manpower of work site given:

$$W_p = \frac{T}{W_{yat} \cdot ws}$$

where:
W_p = the workpower needed,
T = the man-hour demand of a work site given,
ws = the number of work site looked after by one worker (for example one worker handles more machines),
W_{yat} = one worker's real yearly basic working time, which is determined from the yearly basic time (365 days) the holidays, the days to be kept as a holiday, the holiday with pay, the calculated days coming about from illness are subtracted so it can be calculated with yearly 2000 hours.
A_{tM} = the holidays, the days to be kept as a holiday are subtracted from the yearly basic time (365 days) and counting with 1 shift and 8 hours 2040 hours (rounded monthly 170 hours) basic time can be planned.

The manufacturing main process can be sketched from the technological documentation (Figure 14.27.b part).
So called auxiliary processes complete the manufacturing main process which belong to the sphere partly serving – partly subsidiary activities.
The processes serving the production are those activities which built – into the main process guarantee its self- support character. Such characteristic serving process is the transport the material handling respectively. The subsidiary processes help the manufacturing those activities which are accomplished settling independently from the main process and only the result of those operation in connected with the main process.
The guarantee of tools, the jigs for workpieces and for tools, the different gauges are the typical instruments of subsidiary processes.
The area demand of the machines can be determined partly with help of indexes based on practical experiences partly concerning to a machine given in concrete form. The ground plot enclosing the machine can be considered as a start in latter case the demand of place of

wardrobe storing jigs, tools belonging to the machine, the safety route marking the dangerous approach of the machine in case of need, the range of work needed to serve the machine and the route suitable of material handling.
Based on all these it can be done the ground plot – linear – of the workshop arrangement (Figure 14.27.c part) on which the machines, different work sites, transport routes, etc. are shown true–to–scale.
Making such true – to- scale disposition drawing requires a lot of work naturally. The production system (as a system) has to be taken to such part processes (subsystems) to which are available experimental, theoretical and models added by computer. Further on it will be shown some such cost saving method models in which cases the logical step of optimizing can be followed well and they are clearly arranged. However never lose sight that the optimum of part processes is suboptimum first of all and the evaluation of its practical effect has got on the whole process requires separate examination.

14.6.3. Cost efficient optimizing methods, models

14.6.3.1 Organizational forms of transit times

The chronological cadencing of production determines fundamentally the realization of planned output of product. Transit time is referred to the time spent from the taking in hand of the raw material till the finishing. I would draw attention here to the so called Gantt diagrams used very widely to the cadencing of production tasks in other words the lined schedule. The reticular programming methods however spread ever greater domain in the latter decades, which make use of the assistance of computer programs for more complicated organizational tasks.
The largeness of the transit time is given by the totality of different activities arising during manufacturing and other requirement of time. The largness of the transit time consits of four kinds of times spent:

- T_c : technological time of cycle containing those times when the product forming is made according to instruction.
- T_{bop} : time between operations to be needed: to transport the workpiece from operation to operation, to checking the quality, it has to be waited until certain number of pieces come together at one work site and only after is passed on to the next work site, etc.
- T_{np} : the time of natural processes, ageing of metals, heat treatment of metals, painting, drying.
- T_i : time of interruptions which can be happened from the pauses connecting with the system of working hours (for example days off, lunch break, interruption in case of intermittent manufacturing), pauses between work shifts.

The whole transit time is:

$$T_w = T_c + T_{bop} + T_{np} + T_i$$

Among the whole transit time of the manufacturing to determine the technological time of cycle (T_c) means the most important case. Its length is determined by the time of rate, the power of the machine to be at disposal, the technological parameters prescribed to the operation. The simplest case of chronological programming of the production process can be studied at determining the manufacturing process of one part. During manufacturing the part goes from one work site to another according to the technological sequence. Figure 14.28 shows the manufacturing process in case of one part.

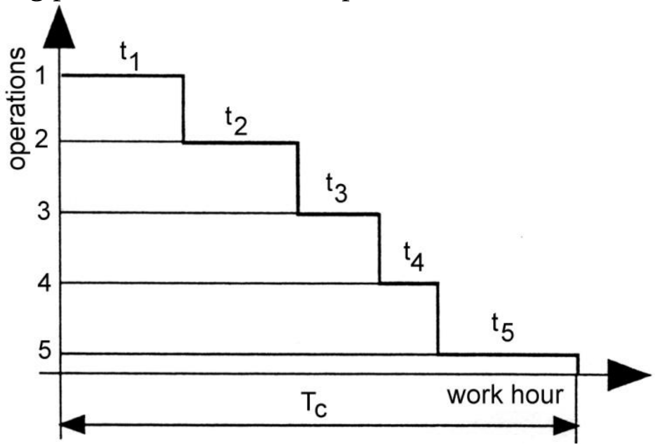


Figure 14.28. Graphical presentation of manufacturing transit time.

The technological time of cycle (T_c) is equal with the mathematical sum of the operation times occuring in the process.

$$T_c = t_1 + t_2 + t_3 + t_4 + t_5$$

it can be written in general:

$$T_c = \sum_{j=1}^{n} t_j$$

where:

t_j = means the operation time given,
n = the number of operations,

The problem is more complicated if the serial–production time of cycle has to be determined. Examining two – operational series of parts the following variations can arise: the two operations are equal with each another ($t_1=t_2$), the first operation is shorter than the second ($t_1<t_2$) the second operation is shorter than the first ($t_1>t_2$). The length of the transit time of the three cases taking into account that two series is examined and within one-one series the number of parts is ten can be seen in Figure 14.29.

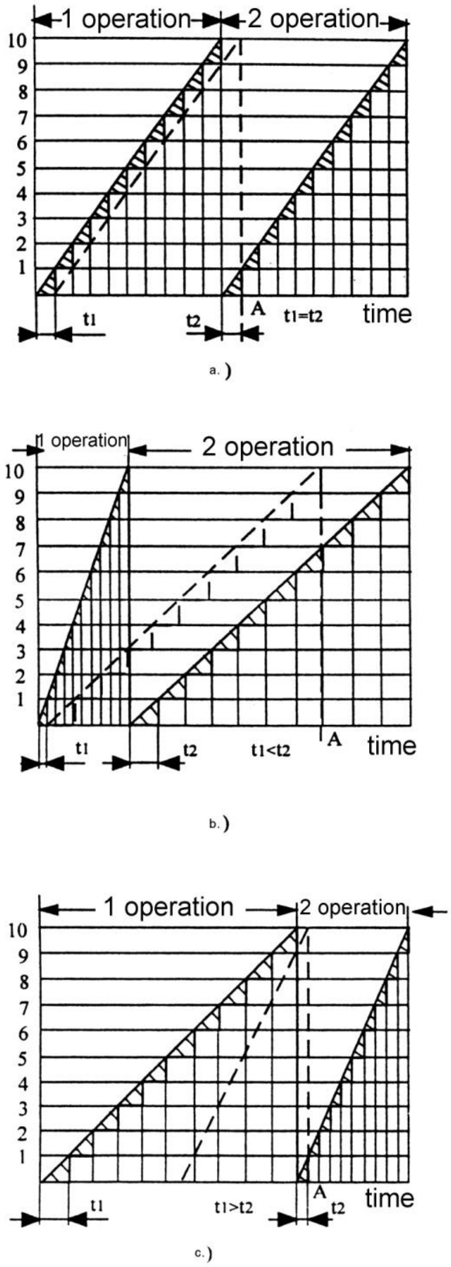


Figure 14.29. Variation possibilities to organize the transit time in case of serial-production.

The operation time is indicated on the horizontal axis, the number of parts on the vertical axis. One-one triangle corresponds to one-one part on the figure. In order to better comparison it is supposed that the total time (t_1+t_2) of the two operations are the same in all three cases.
It can be seen from the figure if the parts are passed on successively in series then the transit time is the same in all three cases.

$$T_c = \sum_{j=1}^{s} t_{1j} + \sum_{j=1}^{s} t_{2j}$$

where:

s = the number of parts in the series.

The transit time will be different if the parts are passed on from one work site to another. The relation of the two parts to each other determines the period of cycle what the dotted line shows. In the first case (Figure 14.29.a) when it is supposed that first and second operation time are equal, the time of cycle takes shape on very simply. Surely after the finishing of the first operation the part can go to the next operation, and with the other parts similary continuously as it is counted with technological time only, waiting times don't arise in not any relation. It can be write thus:

$$T_c = st_1 + t_2 = (s+1) \cdot t$$

In the second case (Figure 14.29.b) when ($t_1 < t_2$) the second operation can be started immediately as it is longer than the first so with waiting in work site hasn't to be counted in case of further workpieces. The waiting time of the part however increases on and on as the second operation time is longer than the one.
The time of cycle:

$$T_c = t_1 + st_2$$

In the third case (Figure 14.29.c) it is not expedient the immediate passing on of the parts from the first operation to the second because the second work site should wait for the next workpiece in every case. One-one waiting time between operations: $t_1 = t_2 = w$, which means, that the work site is unutilized. This waiting time has to be eliminated so that the second work site starts only then, when it can only finish continuously the machining of the series.
In case of such organization the waiting time of the machine doing the second operation:

$$W = st_1 + t_2 - st_2 = s \cdot (t_1 - t_2) + t_2$$

This getting free capacity can be used to carry out other task. It can be established examining the possible variations of the transit time of series of part that three fundamental cases can be distinguished from each other. There are the followings in turn:

- following each other, continuous passing on,
- parallel passing on,
- overlapped (mixed) passing on.

It is characteristic to the movement following each other of the series of part that the operations following each other can start only if the previous operation is already finished on all members of the series given. (Figure 14.30)

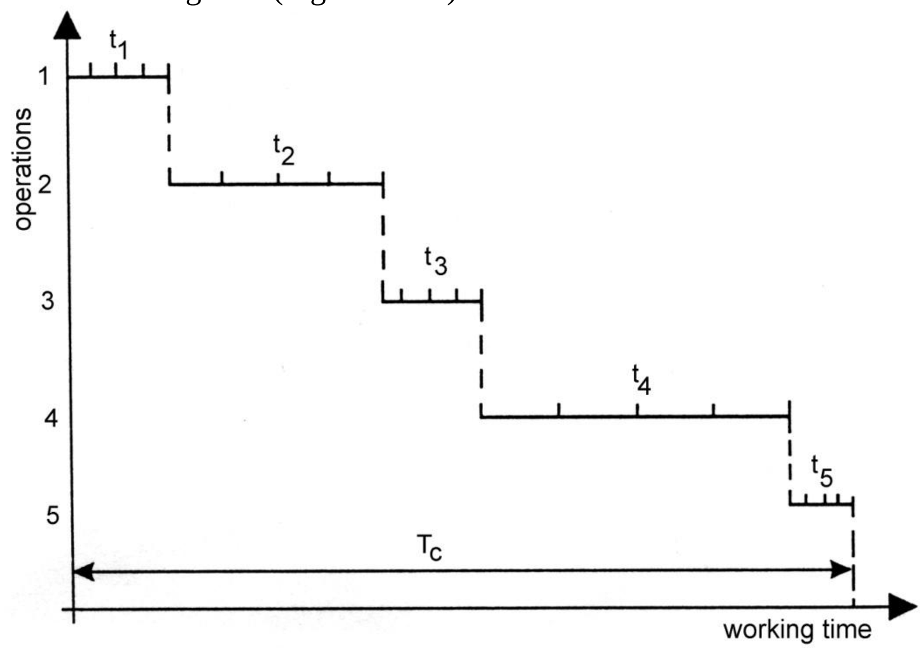


Figure 14.30. Formation the time in case of successive transfer of part-series.

Certain operation times ($t_1, t_2, t_3, \ldots. t_j \ldots t_n$) of process repeat so many times as many part (s) are in the series. The transit time of the operation:

$$T_c = st_1 + st_2 + st_3 + \ldots\ldots + st_j + \ldots\ldots + st_n = s\sum_{j=1}^{n} t_j$$

This menthod of organization is very simple it makes possible the continuous work within given series but its very great disadvantage is that the transit time is very long, the waiting time of parts are very long the stockpiling costs are the greatest in this case. At the same time it can be organized well, the process can be checked well.

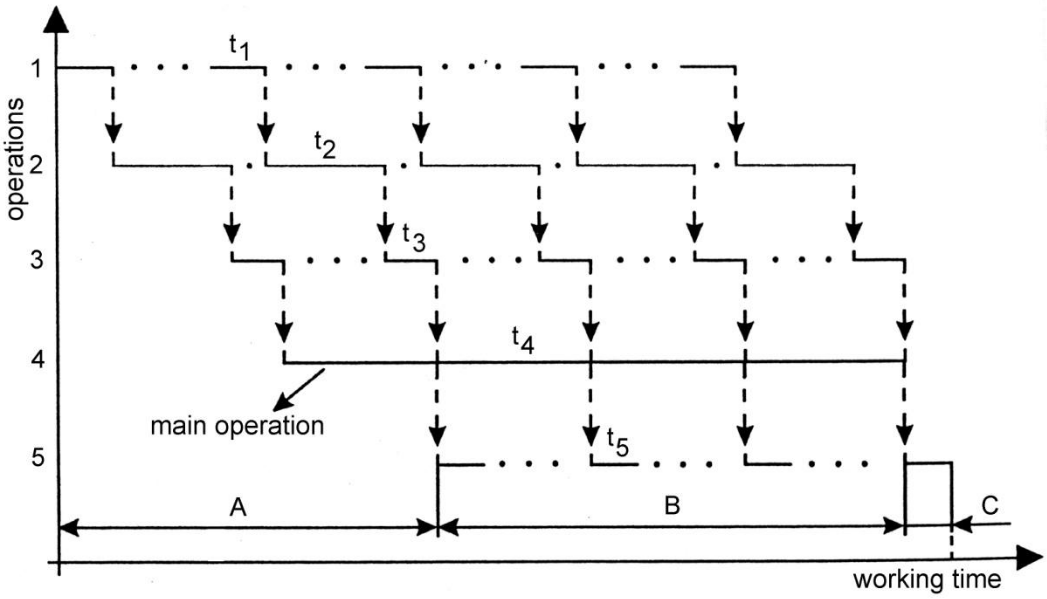


Figure 14.31. Formation of transit time in case of parallel moving of part-series.

It is characteristic to the parallel moving of the series of part that one-one series of part should get without difficulty after starting from the first operation place without waiting to the next operation place following that as it can be seen in Figure 14.31.
The time of cycle:

$$T_c = A + B + C,$$

$$A + C = t_1 + t_2 + t_3 + \ldots.. + t_n = \sum_{j=1}^{n} t_j$$

$$B = (S - 1) \cdot t_{main}$$

It can be seen from the figure that the main operation (t_{main}) has the greatest influence on the time of cycle. Similary it is essential the each other relation of the main time with the other operation times.
Namely this determines the waiting times of certain work sites the utilizability of the capacity respectively. The dotted places show well the length of waiting times to be between operations. The transit time can be reduced further if more work sites are assigned to finish the main operations. Naturally in such a case the next longest operation (in this case t_2) takes over the role of the main operation in every case.
The disadvantage of this menthod of organization is that the waiting time of the work sites working with shorter operation times can be very long comparing to the main operation.

However its advantage is the very short transit time what is favourable concerning the requirement of customers because in case of market relations one of the important conditions of winning the work is the pledge of the deadline. At the same time except the place of the main operation the others are forced to wait what results loss for the company. Thus the offered price has to be determined so that undertaking the task should be profitable despite the loss arising from the waiting time.
It can also be seen here that optimization also needs the manysided analysis of the given task from that standpoint, that a part–optimum namely suboptimum should be the optimum for the whole company, too.
The advantage of the two kinds of organization methods (next and parallel) introduced unites the mixed (overlapped) moving of the series from one work site to another. The characteristic of this organization is that still before the machining of one series would be finished at given operation place already starts the machining of the workpieces finished on the next work site. (Figure 14.32).

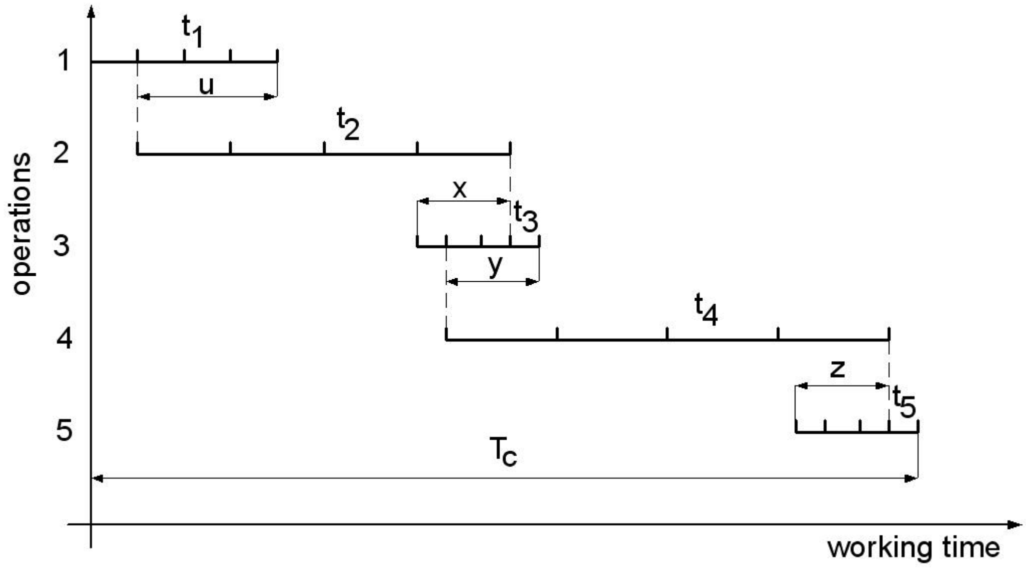


Figure 14.32. Formation of transit time in case of lapped (mixed) moving of part-series.

It can be reached with such method of organization that the work sites can be burden continuously at the expense of that, the transit time is longer with something than at the paracell organization form.
Its main characteristics are the followings:

- the series is passed on from one work site to another in details,
- on certain members of the series more operation are made at the same time,
- the starting moment of the operation next to the row is determined so that the operation should be continuous.

The transit time (T_c) can be determined according to the following method:

$$T_c = s\sum_{j=1}^{n} t_j - [u + x + y + z]$$

where:

$u = (s\text{-}1)\ t_1$
$x = (s\text{-}1)\ t_3$
$y = (s\text{-}1)\ t_3$
$z = (s\text{-}1)\ t_5$

That determines the starting time of certain series that from the two neighbouring operation times whether the next is longer or shorter. For example the No.1 operation time is shorter than the No.2. That means as the first member of the series at the No.1 place is ready it can be transferred to the No.2 work site as t_2 is longer than t_1. Inasmuch the next operation is shorter comparing with the previous (for example $t_3<t_2$) then the series is started so in the third operation place that from the end of the No.2 operation $(n\text{-}1)t_3$ time (Figure 14.35) is measured back. So it can be determined exactly on certain operation places the starting of the series at given operation place.

Task.

Prepare in order to exercise the true-to-scale figures of the transit times in case of parallel or mixed organization based on the data given below:

- the number of parts to be in the series: s = 5 pcs.
- the operation times following each other: t_1 = 4 min.
  t_2 = 8 min.
  t_3 = 3 min.
  t_4 = 5 min.
  t_5 = 2 min.
- determine the length of the transit time,
- count the loss of the waiting time arising from forced standstill,
- summarize the advantage and disadvantage of certain organizational form,
- try to establish another connection of determining the T_c technological transit time and check whether you have got similar result.

During solving the next example you will see evidently that solving a task occuring in the practice generally can be solved by multiple iterative trial only.

The task is, it has to be manufactured multiple – stepped shaft with complicated surface which one of its end has got gear form. From this shaft Q=2500 pieces has to be manufactured with 20 months deadline. The request of the customer is that the delivery should be carried out in 10-10 monthly cadencing in 2 parts.

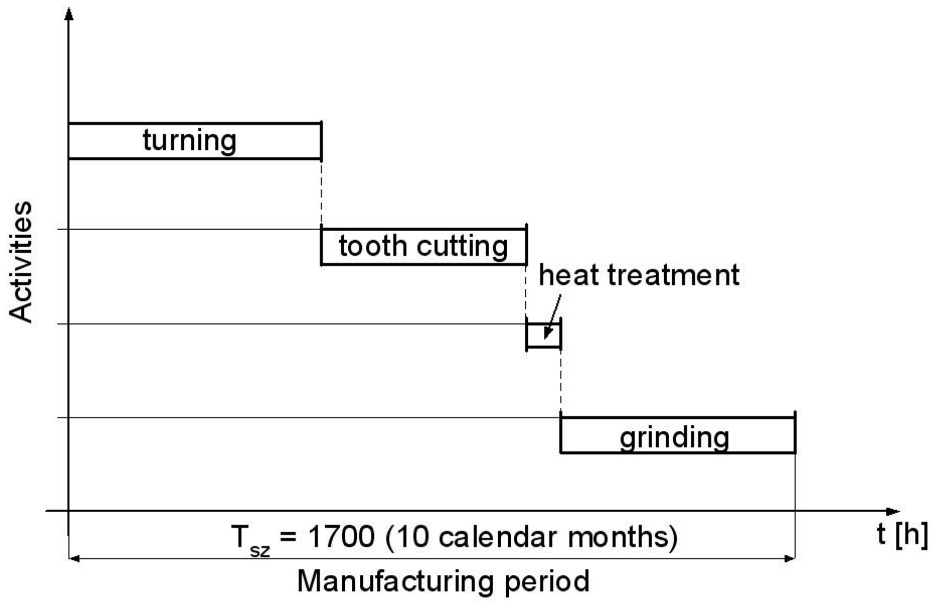


Figure 14.33. Activity plan.

The first and most important task of the company is to make properly detailed technological plan (tooling chart, operation directions) in order to perform the task undertaken. The activity plan (Figure 14.33) shows clearly the succession of the production process. We indicated the time on the horizontal axis and on the vertical axis the activities. Based on the tooling chart – containing in detail the technological data concerning the manufacturing (for example type of the machine-tool, its power, the operation sequence of manufacturing, the time needed to different operations, etc.) – it can be determined the number and composition of the machine – tools needed as well as the worker requirement according to profession. Let's accept the following data based on the tooling chart:

- first operation is turning:
  - operation time, t_{pT} = 90 min/piece
- second operation is tooth cutting:
  - operation time, t_{ptc} = 60 min/pc.
- third operation is heat treatment:
  - total operation time, T_{pht} = 80 hours
- fourth operation is grinding,
  - operation time, t_{pg} = 60 min/pc.

The total time needed to manufacture one run (1250 pieces), it should be marked this with T_1 (as the first variation, it will be seen that still more will also be needed surely the possible best is looked for, that is the production organizing process is optimized):

$$T_1 = T_t + T_{tc} + T_{ht} + T_g = 1250 \cdot 1{,}5 + 1250 \cdot 1 + 80 + 1250 \cdot 1 = 4455 \ hours$$

The number of machines, work sites respectively are: M_1=4, the workers' number in case of one shift is also 4. Taking into account the 10 monthly delivery dadline as the run (Q=1250 pcs) time limit of manufacturing, it is evident that this variation of organization is not suitable to fulfil the task undertaken. That is the using one machine for one operation with that obligation that when the machining of the last piece of the run has been finished on one-one operation place it can get only to the next operation place doesn't give acceptable solution.

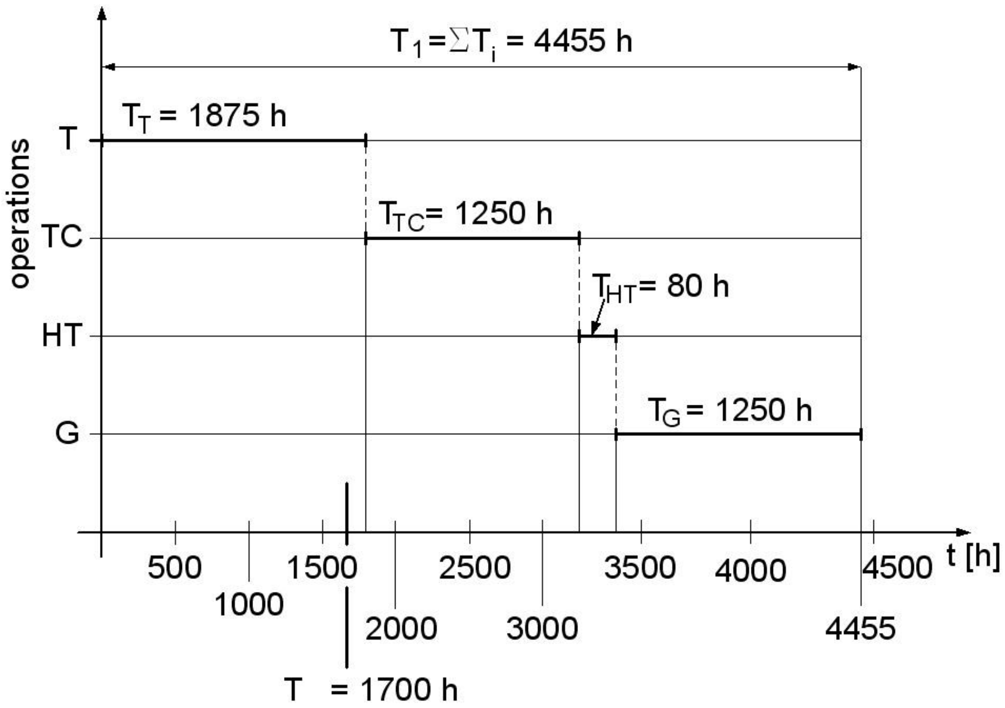


Figure 14.34. Transit time in case of No.1 variation.

Figure 14.34 shows well, what was told. The working time is indicated on the horizontal axis, on the vertical axis the operations (turning, tooth cutting, heat treatment and grinding). It can be seen that only the time needed for turning exceeds the delivery time limit with 175 hours.
Let's have a look on the second case when the task is accomplished by more machines and the delivery deadline would be considered as basic time, that is the possibility of accomplishment of this would constitute the number of machines needed. According to this variation the number of machines to one shift is:

$$M_2 = \sum \frac{t_{pt} \cdot Q}{W_{yat}} = \frac{1{,}5 \cdot 1250 + 1 \cdot 1250 + 80 + 1 \cdot 1250}{1700} = 2{,}62$$

The 2,62 number of machines is not possible naturally, so it has to be rounded to integer upwards or downwards. The rounding downwards can compose the object of consideration if it is only with one-two tenth over the whole and there is real basis of that with better organization, with greater power the task undertaken can be accomplished to deadline.
In present case the upwards rounding is unambiguous, that is let's take the number of machines to 3 and let's calculate the value of the transit time (T_2).
In case of three machines the cutting capacity to be at disposal to one day: 3·8=24 hours, which means that the manufacturing time of the run decreases to one third at certain operation places.
That is:

$$T_2 = \frac{1875}{3} + \frac{1250}{3} + \frac{80}{3} + \frac{1250}{3} = \frac{4475}{3} = 1485 \ hours$$

This variation gives already more favourable result, surely it is near enough to the delivery cycle of the run (T_d=1700 hours). However let's try to take nearer the transit time to the delivery cycle. Let's look at that case if not three machines is used at the grinding operation but only two and this should be the third variation (T_3). That is:

$$T_3 = \frac{1875}{3} + \frac{1250}{3} + \frac{80}{3} + \frac{1250}{2} \cong 1693 \, hours$$

this transit time is equal almost ideally with the delivery time (T_0=1700 hours).
It can be considered thus as optimum.
Thus it can be seen that optimizing such simple task leads to succesful solution after several trials, too. The method of solving such characteristic tasks is based on that the optimum searched is placed between a so called lower limit and an upper limit. Succesful solution can be expected after multiple trial or approximation. The practical experience naturally reduces significantly the time needed to the solution.

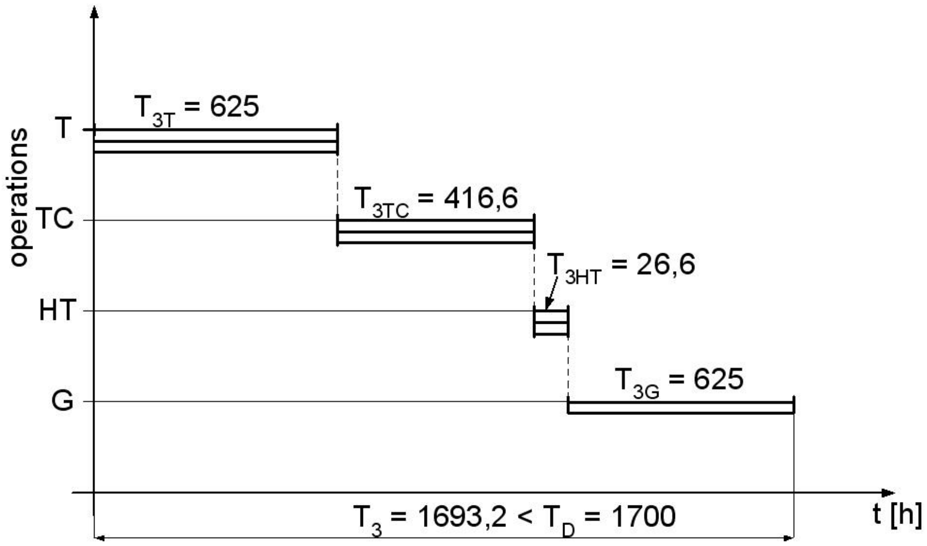


Figure 14.35. Optimal transit time according to third variation.

The operation sequence suitable to optimal time can be seen in Figure 14.35 according to the third variation. The lines belonging to one-to-one operation symbolize the machines working parallel in the same time. It is to be noted that the solution of such task needs some minutes altogether for a technologist working in the practice. Let's notice that only two standpoints were emphasized of the optimizing possibilities namely to observe the delivery deadline (10 months) and the one shift with that purpose first of all, the thinking and logical process of optimization should be followed clearly. If it is also wish to take into consideration the costs what is done by company is a case given, then the task becomes more complicated in an increased degree and instead of as such "paper, pencil" technics it is expedient to appeal computing technique accessory collection. The so called computer simulation models can be used in particular to solve such tasks. However it has to be emphasized that the great systems referred is possible to use in effect if the basis of optimazation has been attained.

Task.

Try to work out such form of organization to the example shown previously, that:

the transit time (T_c) should be shorter without increasing the number of machines.

14.6.3.2 Economic series number

The basic question of the series – production is that in case of producing different products is expedient organizing series whether how many pieces it should contain considering the economic standpoints.

This question is raised because during finishing one after the other parts made of different materials and having various shapes with machine – tool (work site) given for example the so called preparational costs can be reduced by increasing the number of pieces, but at the same time the costs coming about from binding working capital and storing increase. These factors depending on the possibility of optimazation to determine the economic series number (Figure 14.36)

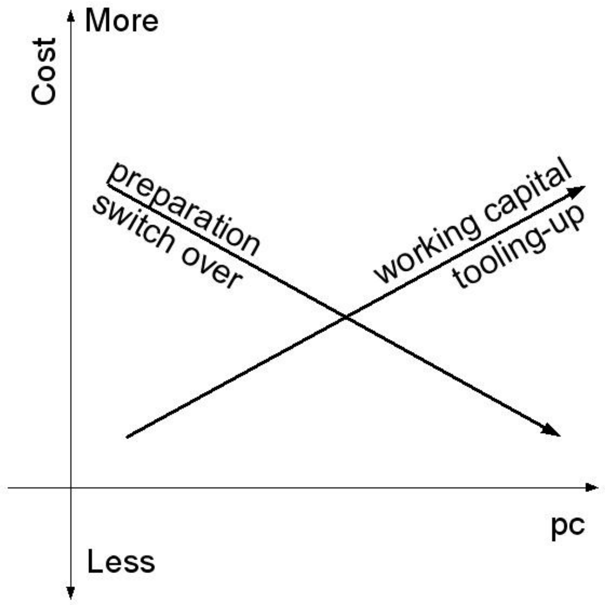


Figure 14.36. Elements affecting opposite to the optimal series number.

The number of pieces of the series can be optimized by the means of making numerical the factors to affect contrary with each other. (Figure 14.37)

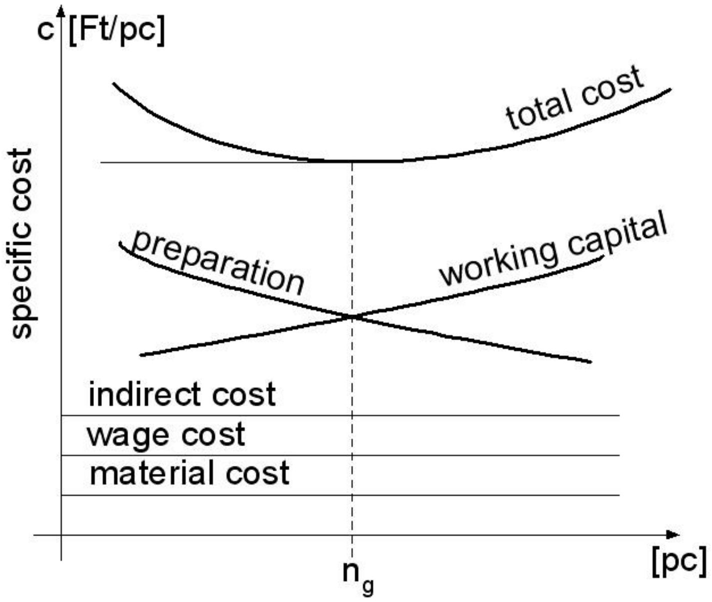


Figure 14.37. The effect of main cost elements of series number to the optimal series number.

The special literature is [14.25] connects to Kurt Andler's name the foundation of the theory concerning the taking to pieces of series and this calculation process will be followed, too. First of all let's try to draw up that what is the main object of examining the series – production and to reach the aim which components has to be taken into accunt. As different products are produced to sell them, naturally it happens, that there is the most favourable income from selling the product if the company can produce it with the possible least cost, which can be realized by the means of minimizing the producing cost of the series. Naturally the expected market reception requires important surveying furthermore the obtainable manufacturing cost-level as these can be determined from the standpoint of obtainable profit.

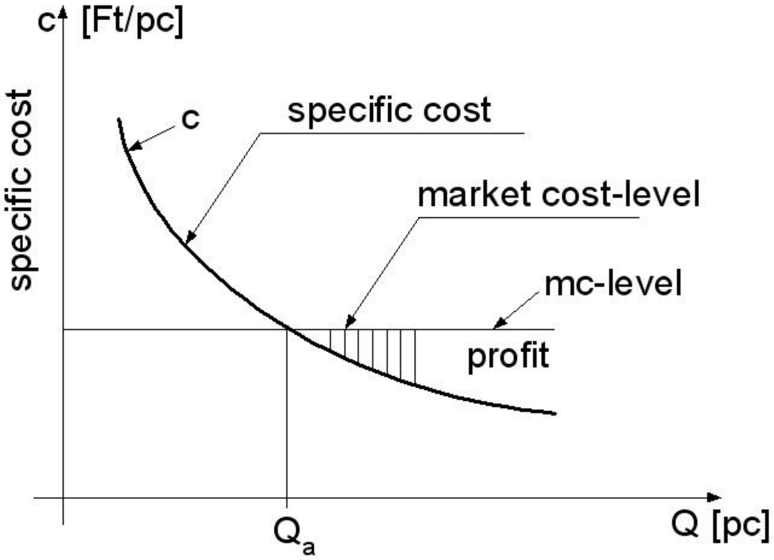


Figure 14.38. Critical number of pieces (Q_a).

Figure 14.38 shows that critical number of pieces (Q_a) – to be determined by the intersection of the specific cost curve (*c*) and the direct line (*mc-L*) marking the market cost–level–under which the manufacturing would be a deficit and over which thus profitable. Thus the main components of the costs are:

- cost of market research,
- planning cost,
- planning cost of manufacturing documents,
- planning cost of resources (human- and machinery),
- operation cost of manufacturing (workers' hourly wage, machine operation hourly costs),
- jig planning and manufacturing cost, tooling-up costs,
- costs of series' switch-over,
- stockpiling costs,
- overheads (heating, lightning, maintaining),
- contributions (social security contribution of wages),
- general costs (non effective workers' wages),
- amortization costs,
- other costs.

It is to be noted, that there can be different deviations and arrangements in the statements of companies, accordingly the items listed by me refer only to diversity of costs.
Thus the aim is to determine the economical series number (n_e). To starting let's write down the manufacturing cost M_c of the series (n_e).

$$M_c = A \cdot n_e + B$$

where:

A= part of the constant cost, indipendent from the series number regarded to the product unit (for example material, wage, machine hour, amortization, indirect, etc.)

B= preparation and finishing costs to be in connection with the starting – or restarting of the series (for example setting, adjustment, cutting tools -, measuring instruments, etc.)

The specific cost regarded to one part (*k*)

$$k = \frac{M_c}{n} = A + \frac{B}{n}$$

n – is the prevailing series number

Let's write down the specific cost to the unit element of two different series (n and n_1) as well as the difference between them.

That is:

$$k = A + \frac{B}{n},$$

$$k_1 = A + \frac{B}{n_1}$$

and let's constitute the difference of the two costs

$$k - k_1 = (A + \frac{B}{n}) - (A + \frac{B}{n_1}) = B \cdot (\frac{1}{n} - \frac{1}{n_1})$$

Let's mark with Q_m the number of parts manufactured yearly and with *Z* the producing cost, then it can be written that the decrease of producing costs (instead of n supposing n_1 series):

$$Z - Z_1 = Q \cdot B \cdot (\frac{1}{n} - \frac{1}{n_1})$$

The criterion of the going over to greater series can be that the Z-Z_1 savings is greater than the volume of the working capital's binding increment belonging to n_1 series.

The financial loss arising from the binding of working capital can be described with the following equation:

$$V = F \cdot P_a \cdot p \cdot r$$

where:

V= loss from binding working capital (expressed in money),

F= the circulating fund needed to manufacturing (expressed in money),

P_a= binding time – periodicity – (year),

p= proportion factor of the binding loss (year),

$r = \frac{Q}{n}$ number of series manufactured yearly,

The binding time: $P_a = \frac{n}{Q}$

The circulating fund needed to the manufacturing, supposing the equal use,

$$F = \frac{A \cdot n + B}{2}$$

Thus the loss,

$$V = \frac{A \cdot n + B}{2} \cdot \frac{n}{Q} \cdot \frac{Q}{n} \cdot p = \frac{A \cdot n + B}{2} \cdot p$$

Difference of the n and n_1 series manufacturing loss:

$$V_1 - V = \frac{A \cdot n_1 + B}{2} \cdot p - \frac{A \cdot n + B}{2} \cdot p = \frac{A \cdot n_1}{2} \cdot p + \frac{B}{2} \cdot p - (\frac{A \cdot n}{2} \cdot p + \frac{B}{2} \cdot p)$$

After completing the reduction:

$$V_1 - V = \frac{A \cdot p}{2} \cdot (n_1 - n)$$

Reviving the criteriums put forward previously that the cost savings (Z-Z_1) of the number of pieces to be manufactured yearly (Q) has to be greater than the loss of binding working capital belonging to the whole number of pieces (Q) it can be written:

$$Z - Z_1 \geq V_1 - V$$

Substituting the left – and right ride expressions it can be written:

$$Q \cdot B \cdot (\frac{1}{n} - \frac{1}{n_1}) \geq \frac{A \cdot p}{2} \cdot (n_1 - n)$$

Arranging the equation, it can be written:

$$\frac{Q \cdot B}{n \cdot n_1} \geq \frac{A \cdot p}{2}$$

The solve such type tasks it can be used advantagely the switch over to the limit based onto small (elemental) difference changes, that is n_1= n+Δn substitution – with certain neglect – the inequality to ***n*** can be solved. That is:

$$lim \frac{Q \cdot B}{(n + \Delta n) \cdot n} \geq \frac{A \cdot p}{2} \quad \text{respectively} \quad \frac{Q \cdot B}{n^2} = \frac{A \cdot p}{2}$$

expressing – *n* – the economical series number of pieces can be got:

$$n_e = \sqrt{2 \cdot \frac{Q \cdot B}{A \cdot p}} \quad \text{(piece)}$$

The series number needed to economical manufacturing of Q number of pieces thus it can be determined very simply if there are at disposal the value of factors to be in the formula (A, B, p) by means of established calculations.

Supposing in order to simplicity according to the following data:

Q=10000 pcs.

B=50000 Ft

A=10000 Ft

p=0,1

the economic series is:

$$n_e = \sqrt{2 \cdot \frac{10000 \cdot 50000}{10000 \cdot 0,1}} = 1000\,pcs.$$

Task.

Point to that in the $n_1 = n + \Delta n$ relation it was rightful the certain neglect and what is this.

Héberger and Cserhalmi [14.25] make known several calculating model to be suitable determining the optimal series number. Among them the formula worked out by Demin $n = 5{,}77\sqrt{\frac{QB}{A}}$ deserves attention because it is very simple and can be handled easily. Authors analyzing the models made known by them, they make known those standpoints by which certain modifications – to be suitable for the plant given – can be done expediently. Some such stressed standpoints as an example:

- the economic series number determined according to different models has to be modified often in accordance with the technological organizational, economic peculiarity of the plant given, however the effect of this is very small,
- the plant administration can be simplified by certain modifications,
- it is expedient to determine the transit time of the series number in accordance with the service-life between two renewals of the jigs and tools,
- the area of the store can be a limited factor by the upper limit of the series number,
- it has to be checked the harmony the period time of the series number and of the circulation speed of the working capital.

Try to think over and sum up that the solution of the mathematical connection got to determine the optimal series number is extremely simple in knowing the factor (it is a secondary school task). At the same time to determine these a wide knowledge of economic ,-mechanical engineering technological – and production organizational disciples are needed. Don’t forget that the base needed of the optimizing conditions is the readiness of recognizing the problem and solving the problem, to which is accompanied with a very innovative attitude and a very various lexical knowledge, furthermore readiness to integrate the part and the whole.

14.6.3.3 Optimal stock volume

The stockpiling – the reservation can be taken in the daily act in the most direct form. It is natural that different amount is bought from various foodstuffs depending on how long they can be stored or the fresh goods how it is liked. Buns are bought only that pieces to be enough for supper and breakfast. Flour is bought so that it should be enough till one –two months, too. It is natural if something runs out people go to the shop and buy what is missing. People go to the shop with that consciousness that it has got the product what is needed. As many people think so the manager has to store adequate great amount of merchandise to be able to satisfy the customers completely. This is very important from the customer’s satisfaction standpoint.

It is also an important condition of utilizing economically the productive working resources, that the supply of materials should be guaranteed continuously. Surely the interruption of supply of material means the stoppage of the process (production) what can cause extraordinary damages for the company. It is well known the organization of work of the large car factories that they count one – two hours reserves what requires very strict discipline from the transporters considering the fulfillment of deadlines. The reservation is an important question concerning the machine operation standpoint, too.

The reservation makes significantly easier the planning of the machine running maintenance activity and everyday practice. For example it is “easier” significantly to maintain, to repair a vehicle if from them or at least from certain structural units for example from cylinder head, dynamo there are reserves. At the same time the machine operation of such vehicles can also be planned more reliable namely the production program to be accomplished by them can be fulfilled more safely. The statements would give reason that the companies should reserve parts and machines in”unlimited” numbers. Naturally that fact delimits to this that reducing the costs of production interruption gives reason to the increase of the number of reserve machines, the costprice and maintaining price give reason to reducing the reserves. The optimal reserve has to be searched evidently between the two extremity (infinite many reserves, zero reserve) arising from this. The problem of stockpiling appears in the fields of the manufacturing, commerce and of maintenance is similar. The importance of stockpiling is in that the compulsion intervals of differenct activities can be stopped, it can be kept on some kind of optimal level.

The basic problem of the stockpiling is really at any time that what how many should be reserved and where. It can also be seen if the number of storing places are reduced then the stockpile volume can also be reduced relatively but the transport routes and the service time can increase. Contrary to this if the number of storing places are increased then the stockpiles also increase but it can be quicker to satisfy the claims.

As the stockpiling also means cost so its financial limits as a restricting conditions has to be taken into account first of all. Other restricting conditions can be for example the mutual connection among different products or the weight or the operation length of time.

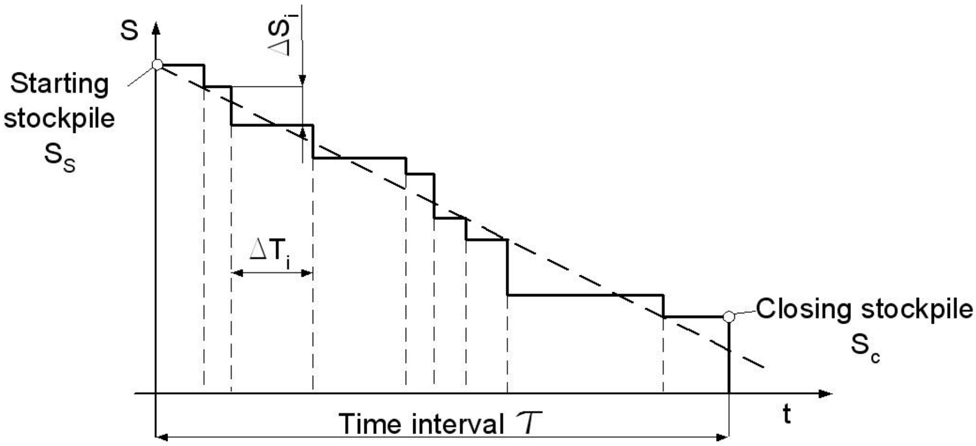


Figure 14.39. Basic scheme of stockpiling.

The problem of stockpiling can be studied clearly in Figure 10.39 The change of the stockpile (s) can be seen in the function of the time (t), furthermore it is indicated the starting stockpile (S_s) belonging to t=0 value and the closing stockpile (S_c) belonging to $t = \tau$ value. The τ means that interval to which belongs S-piece stockpile and as a safety reserve the N_c – closing stockpile. As the need is random like therefore different losses ΔS_i belonging to different Δt_i. The sections can be replaced with a continuous direct line which means neglect but at the same time it guarantees certain calculating possibility. Supposing that the time between posting the order and receiving the goods is zero it can be distinguished two basic methods of the stockpiling:

- periodic method (Figure 14.40)
- damping method (Figure 14.41)

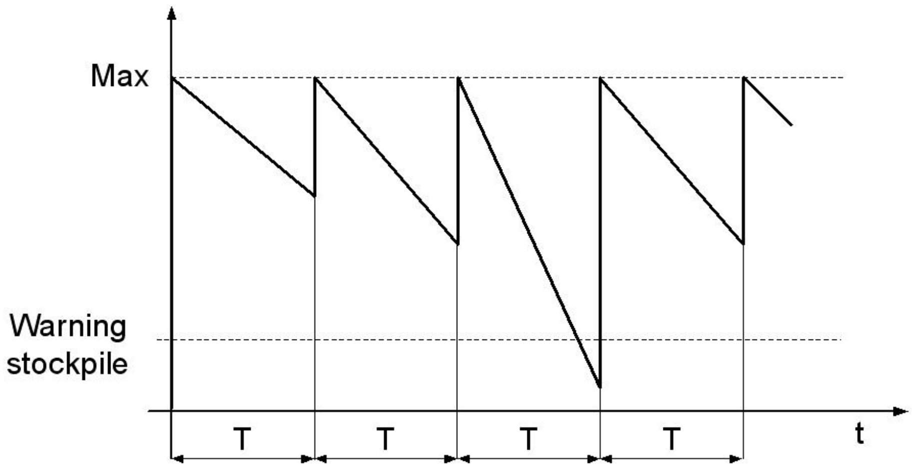


Figure 14.40. Periodic stockpiling.

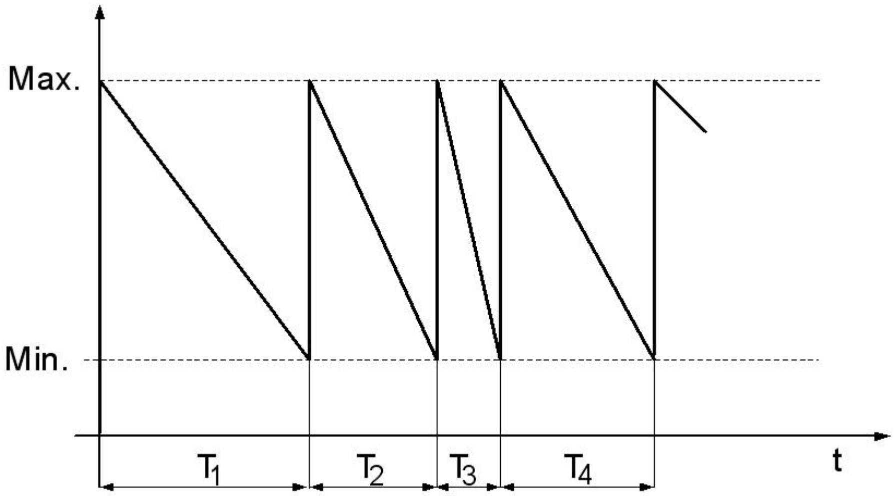


Figure 14.41. Damping stockpiling.

It is characteristic to the periodic method, that a T – time interval is determined in advance at the end of which to replace the stockpiles runs out comes in turn. The disadvantage of this method is that it can happen the stockpiles run out, its advantage is that it is automatic.
It is characteristic of the damping method that replacing the stockpile is constant contrary to this the time interval (T_i) is changing. There is no risk running out of the stockpile but it is more difficult also more costly to operate the system.

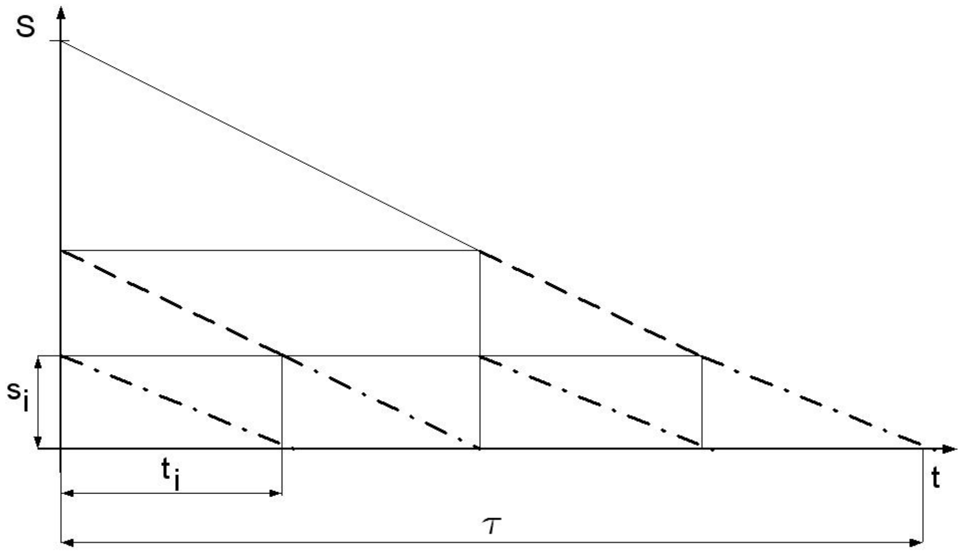


Figure 14.42. Geometrical scheme of stockpile reducing.

The next consideration can be made to determine the economic (optimal) volume. Let's suppose that the stockpile falling at τ time interval is bought at the same time (Figure 14.42). In this case the total cost (C_t) is made up from the charge (C) and from the storing cost (C_s).
That is: $C_t = C + C_s$,
The charge is determined by the transporter, the storing cost depending on customer's possibility can be determined with the following relation:

$C_s = c_{sd} \cdot N \cdot \tau \cdot \frac{1}{2}$; where: c_{sd} – is the storing cost of one piece of product for one day.

This approximation means the highest cost, as $t=0$ moment it has to be mobilized the charge and the storing cost of N – number of pieces, too. The figure shows how is needed to think concerning the cost reduction. The halving method offers itself well writing down the single costs to n_i -:

$$C_{si} = c_{ni} + c_{sd} \cdot \frac{n_i \cdot t_i}{2}$$

where: c_{ni} – means the cost of single order (n_i) and the second member of the equation's right means the storing cost of the n_i – volume. This sum naturally is only the N – part of that cost as if whole would be bough at the same time. Let's write down the total cost to τ – time interval now (Figure 14.43):

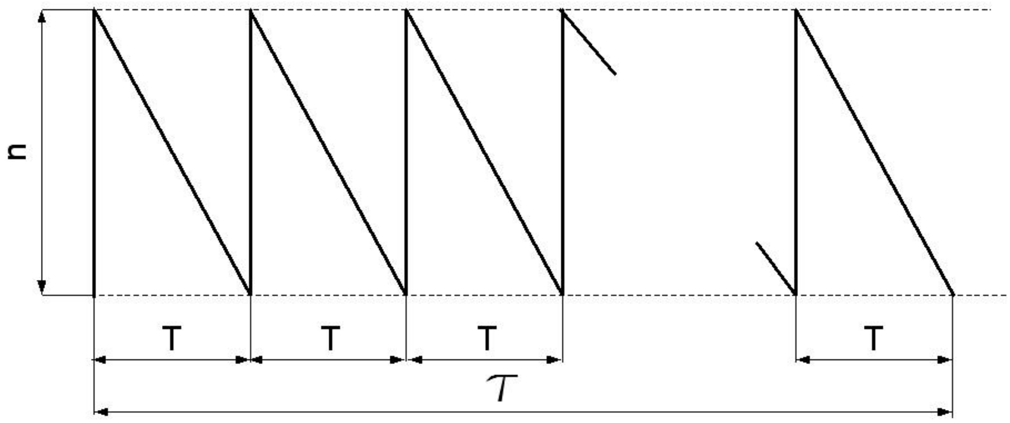


Figure 14.43. Geometrical scheme of optimal stockpile.

$$C_t = c_n \cdot \frac{N}{n} + c_{sd} \cdot \frac{n \cdot T}{2} \cdot \frac{N}{n}$$

The right side members are multiplied with $\frac{N}{n}$ as the n single stockpile of supply has to be bought and stockpiled $\frac{N}{n}$ times. Let's notice that $\frac{N}{n}$ can be replaced with $\frac{\tau}{T}$ and let's write down the total cost:

$$C_t = c_n \cdot \frac{N}{n} + c_{sd} \cdot \frac{n \cdot T}{2} \cdot \frac{\tau}{T}$$

doing the reduction

$$C_t = c_n \cdot \frac{N}{n} + c_{sd} \cdot \frac{n \cdot \tau}{2}$$

If it is supposed, this function can be optimized. Surely the members to be on the right side of the equation have got opposite effect with the – change of – *n*, the first number has got reducing – the second member has got increasing tendency, thus optimum exists.
Determining the first derivative of this cost function according to "*n*" and making equal with zero the economical series number (n_e) can be got:

$$C'_t = -c_n \cdot \frac{N}{n^2} + c_{sd} \cdot \frac{\tau}{2} = 0$$

$$n_e = \sqrt{2 \cdot \frac{c_n \cdot N}{c_{sd} \cdot \tau}}$$

Task.

Think over on that at what mode could you determine the optimum (quasi-optimum respectively) then if c_n is not independent from the number of pieces of the volume (n) and motivate why is logical that the cost of single ordering depends on the number of pieces.

14.6.3.4 Hot and cold redundancy systems

The hot and cold redundancy contain the reserve elements built into different machines. The word redundant means surplus its technical meaning covers better the reserve idea, as certain machines, equipment are designed so that at breakdown the hot – or cold built in redundant (reserve) element should take over the function of the part (unit) broken down. Instead of the hot redundancy and cold redundancy rather the hot – reserve and cold – reserve name spread in Hungarian practice, in accordance with this I also use so. It is important to remark that the reserves are examined as non repairable element.

a.) Hot – reserve without repair

The hot – reserve is characterized that the given element in Figure 14.44 can be for example the No.2, at breakdown the function is taken over by the other elements by the means of their parallel connection.

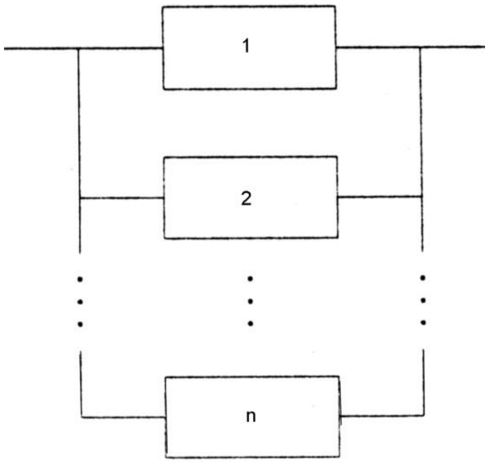


Figure 14.44. "Hot" connection redundancy.

Such machine system is shown on the figure, which one basic element is completed with n-1 reserve element in order to increase the reliability (*R(t)*) of the system. Thus let's have a look how changes that the reliability of the No. *N-1* reserve system. The following reliability, breakdown distribution functions respectively

$$R_1(t), R_2(t), \ldots, R_n(t)$$

should mark the reliability of certain elements.

$$Q_1(t), Q_2(t), \ldots, Q_n(t), \quad [\text{where } Q_k(t) = 1 - R_k(t)]$$

however should mark their unreliability.

The reliability of the system is marked by the $R(t)$ its unreliability is marked by the $Q(t)$ breakdown distribution functions. It results from the determination of hot – reserve that the breakdown of the system ensues in that moment when its last element gets unserviceable. Therefore to that the machine system should be unserviceable during *t* time. It is needed that during this time all or – *n* – number element should be unserviceable. The breakdown of certain elements are independent from each other, in accordance with the multiplication proposition of probabilities:

$$Q_{Mn}(t) = Q_1(t), Q_2(t), \ldots, Q_n(t).$$

If this inequality is transcribed to the reliabilities, the

$$R_{M\,n} = 1 - [1 - R_1(t)] \cdot [1 - R_2(t)], \ldots [1 - R_n(t)]$$

connection can be got.

In this case if it is supposed that the reliability of the elements is identical, that is

$$R_1(t) = R_2(t) = ... = R_n(t) = R(t)$$

then the inscription of the unreliability becomes simple significantly,

$$Q_M(t) = Q^n(t)$$

and the reliability in accordance with this:

$$R_M(t) = 1 - [1 - R(t)]^n$$

The unreliability equation as a result of its simplicity provides favourable possibility to determine the number of elements to this contrary to determine the unreliability respectively

- if it is given the reliability of the elements $R(t)$ and it has to be determined the number of those reserve elements by which the $Q_M(t)$ unreliability doesn't exceed its given value, then from

$$Q^n(t) \le Q_M$$

inequality

$$n = \frac{log\,Q_M}{log\,Q(t)}$$

- if the number of elements is considered as given and it has to be determined that what should be the unreliability of certain elements, then

$$Q(t) = \sqrt[n]{Q_M(t)}$$

connection can be used.

Task.

Calculate the reliability of such machine group consisting of fans which has got one basic element and two reserve elements. The reliability of the base element as well as reserve element should be 0,5 furthermore calculate that how many reserve elements are needed to a system operating with $R(t)$=0,89 reliability, if the unreliability of the system elements is $Q(t)$=0,57

b.) Cold – reserve without repair

The elements of the cold – reserve don't operate till the breakdown of the basic elements, thus they can't break down till replacing the basic elements, and they don't influence the reliability of the system consisting basic elements. It is supposed in addition of the mentioned that the time needed to change the elements broken down with new ones is zero practically and inasmuch there is also a need for an equipment with change-over switch its operation is absolute reliable. Cold – reserve is used often in the mechanical engineering. For example the number of those buses having in reserve within the bus company is qualified as cold –reserve, which are put into traffic if one of them broke down. The reserve electric aggregate of a incubator plant or power machines, cargo vehicles reserved for harvesting can serve similar purpose.

Similarly to the hot – reserve let's examine the operation of such a machine – group to be cold – reserved which consists of one basic element and a n-1 reserve element (Figure 14.45). $R_k(t)$ should mark the reliability of the k – element in the reserving sequence, its uncrealibility thus the $Q_k(t)$.

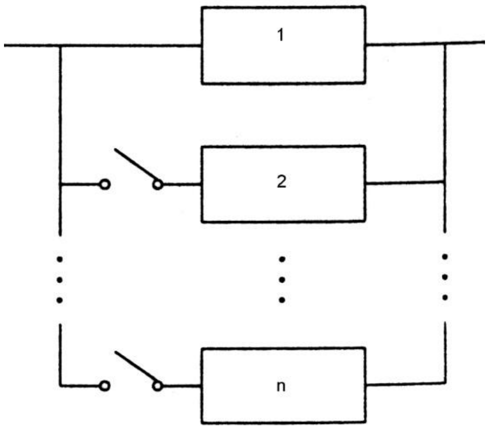


Figure 14.45. “Cold” connection redundancy.

After the basic element operated till τ_1 time, it brakes down and the first reserve element replaces it, which operates till τ_2 time, etc. The last reserve element brakes down after τ_n time operation and together with it the whole machine group brakes down. In such way the machine group to be cold – reserved T_n operation time (service – life) is:

$$T_n = \tau_1 + \tau_2 + ... + \tau_n .$$

The t_k quantity are independent from each other, and

$$P\{\tau_k < t\} = Q_k(t).$$

$Q_{Hn}(t)$ should mark the unreliability of the machine group to be cold – reserved. The $Q_{Cn}(t)$ function as the n – distribution of the sum of the independent member can be determined from the following approximate formula,

$$Q_{Cn}(t) \cong \frac{Q_1(t),Q_2(t),...,Q_n(t)}{n!}$$

On the other hand the unreliability formula to the hot – reserve case,

$$Q_{Hn}(t) = Q_1(t),Q_2(t),...,Q_n(t)$$

Forming the quotient of the two kind of reserve functions:

$$\frac{Q_{Hn}(t)}{Q_c(t)} = n!$$

which shows that the unreliability reduces significantly at passing over to cold – reserve.

Task.

Determine the reliability and the unreliability based on the data given at hot – reserving, compare the values and draw conclusions.

14.6.3.5 Optimal expert demand

Determining the expert demand to be engaged with machine maintenance causes difficulty in those cases, when it has to be guaranteed the needed staff preventing, respectively preceding repair (maintenance) tasks of the machines working in the production. In the majority of cases the maintaining teams are formed based on experimental causes and so perpetual debate goes on whether the staff practice as profession with repair is much few. Namely if the staff is few then the loss arising from the standstill of the machine broken down can be great maybe greater with scales than one or two maintenance man’s wage.

The machine broken down can cause two kinds of loss for the company:

- the one the reconditioning (repairing) cost,
- the other however the production loss due to the machine standstill which can be considered as cost likewise.

Both costs depend on the maintenance men' number substantially. (Figure 14.46)

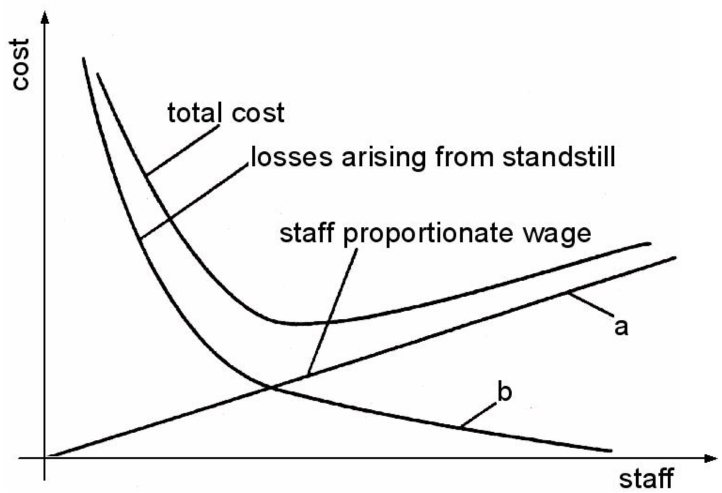


Figure 14.46. Optimal expert demand.

It can be seen if the staff is increased then the wages increase linearly but at the same time the loss arising from standstill reduces. The minimum of the resultant of the two curves gives the optimal staff. This system can be applied successfully where the losses can be determined appropriate exactness and the losses can be expressed in forint. Inasmuch the optimum is determined to skilled workers both with special knowledge and with universal qualification, then that which solution results the less maintenance total cost by the means of compering the optimums.

Practical experiences show in case of universally trained maintenance men the minimum of the total cost is more favourable. This can be explained with that, the standstills reduce significantly in this case due to the quicker fault prevention and doing maintenance operations.

14.6.4. Questions, tasks.

- Summarize what is the interaction between the number of piece of the given product and its prescribed tolerance and between the manufacturing process chosen.
- Summarize the characteristic forms of manufacturing organization, characterize the criterions of application of those!
- What is the importance the mutual interchangeability of operations?
- Make known the characteristics the advantages, the disadvantages of the grouped manufacturing. Draw comparison between the two kind of groupings (machine grouping – cycle – parts grouping respectively).
- Determine the more important characteristics advantages, disadvantages of the mass production (continuous production). What causes one of the greatest problems during planning the continuous production?

- Try to analyse for example in case of a given machine-tool the possibility of the savings appearing in the main centres of the total cycle of life (Figure 14.26). Think over that what faults can be commited theoretically in certain centres, what activities are needed to correct these and what financial a “cousequences” have they got. What is the consequence if the machine – tool still has been manufactured underestimating the largeness of the real fault. How would you decide on the future fate of the machine manufactured: would you sell, reject, or sell as iron scrap, etc.? Make risk analysis (you can built upon also to the Table 14.1 or if you don’t spare the “pains” it is worth and instructive looking after in the special literatures).
- Present a picture to be well arranged and to be hold up to the elements of main activities of the greenfield investment planning process of a plant needed producing given product (products). Make illustrative figures to certain activities.
- Make known the organization form of different production transit times, give reason that which form of organization can be used expediently among in what conditions of production organization.
- Prove – on the basis of logic – that the series optimal number of pieces can be determined in the possession of due reliable plant data in case of manufacturing identical parts in great number of pieces (it is useful to show certain processes graphically in figure).
- Give reason why it is important to reserve less product concerning both the manufacturer and the commercial firms.
- Make known the characteristic reserving models, those advantages and disadvantages.
- Deduct the mathematical connection serving to determine the optimal stockpile.
- Analyse the importance of the hot – and cold – reserve systems without repair, furthermore give reason which you consider more favourable and why. Make known based on what considerations – and plant data could you determine the number of optimal maintenance men?

14.7. Complementary examples to the subject

No.1 task:
Let’s determine that number of pieces (*n*) at which manufacturing more parts already it takes less specific time (auxiliary time needed to produce a part) to produce with the No.2 method than with No.1 method. (chapter14.6). The cutting time (main operation time) furthermore the time of tool-grinding and clamping is the same in both cases and so it doesn’t need to calculate with these time elements.
To solve the task the comparison of specific time requirements is considered as basis practically.
The specific time elements according to the No. 1 method:

- instruments needed to tracing: tracing point, centre punch, caliper, surface preparation to tracing,
  30 min./25 pcs ≈ 1,2 min/pc……………………………………………1,2 min./pc.
- tracing and centre punching……………………………………………5,0 min./pc.
- setting the centre of holes of the workpiece to the tool………………...2,0 min./pc.

The specific time requirement (needed to finish one part)
according to No. 1 method…………………………………………………...t_1 = 8,2 min/pc.

The specific time elements according to No. 2 method:

- time needed to fabricate the stop-block…………………………………100 min.
- common setting the stop-block and machine-tool table to the tool, the stop-block fixing to the table………………..……..20 min.
- the pressing of the workpiece to drill the first hole (0,5 min), turning to drill the second hole 0,5 min/pc……………………..………..1 min./pc.

Marking with t_2 the specific time requirement according to No.2 method it can be written:

$$t_2 = \frac{100 + 20}{n} + 1$$

The limit number of pieces looked for can be determined by the equality of the specific costs calculated with two kinds of method. It can be written thus,

$$t_2 = t_1$$

that is,

$$\frac{100 + 20}{n} + 1 = 8{,}2$$

The number of pieces looked for from the relation,

$$n = \frac{120}{7{,}2} \cong 17$$

The limit number of pieces rounded n = 17 pcs. which means that in case of producing 17 pcs. parts the No. 2 method not only is more economical but also results more accurate machining that the No. 1 method. This number of series belongs to the small series category as with simple jig using in case of small number of pieces the piece of time decreased and at the same time the accuracy increased.

No. 2 task:

Let's determine the utilization indicators to be characteristic to the cycle in case of grouping (chapter 14.6.1, Figure 14.24) organizational form.

The production capacity characteristics of the cycle:

2 pcs. lathes, yearly capacity (2040 hours/pc.) 4080 hours…………….…….L_{yc} = 4080 hours,
1 pc. milling machine, yearly capacity…………………………………….M_{yc} = 2040 hours,
1 pc. drilling machine, yearly capacity…………………………………….D_{yc} = 2040 hours,
1 pc. automatic welding equipment, yearly capacity……..…………………W_{yc} = 2040 hours.

The yearly cutting and welding working hour requirements planned to the cycle according to profession are the followings:

- Turning……………………………………………………….……T_i = 4200 hours,
- Milling………..…………………………………………………..M_i = 1800 hours,
- Drilling……………………………………………………………D_t = 1200 hours,
- Welding……………………………………………………….……W_i = 750 hours.

The yearly working hour requirement planned to the cycle:

$$\sum t_{in} = T_{wh} + M_{wh} + D_{wh} + W_{wh} = 7950 \ hours$$

Let's notice that the following organizational possibilities has to be taken into account in order to accomplish the tasks economically (which is characteristic in the reality generally):

- The capacity of the 2 pcs. lathes is 4080 hours, the cutting requirement is 4200 hours however. Thus it has to be guaranteed the outsourcing 120 hours capacity, that is………………………………………….. t_c = 120 hours.

- There is capacity excess in the three other work sites, that the working hour requirement of the operations entering in the cycle can burden economically, that is:

k_{eM} = 240 hours,
k_{eD} = 840 hours,
k_{eW} = 1290 hours.

Based on the data listed the indicators of the characteristic burdening respectively of utilization of the cycle (chapter 14.6.1) can be determined:

- The average burden of the cycle: $a_b = 0{,}78$

$$a_b = \frac{\sum t_{in}}{T\sum M_T} = \frac{7950}{10200} \cong 0{,}78$$

- The degree of closed condition concerning the product of the cycle: $DCC_t = 0{,}98$

$$DCC_t = \frac{\sum t_{in} - \sum t_c}{\sum t_{in}} = \frac{7950 - 120}{7950} \cong 0{,}98$$

- The degree of closed condition concerning the producing equipment of the cycle:

$$DCC_{pe} \frac{\sum t_{in} - \sum_{tc}}{\sum t_{in} - \sum t_c + \sum t_e} = \frac{7950 - 120}{7950 - 120 + 2370} \cong 0{,}77$$ DCC_{pe} = 0,77

The global indicators can be accepted in reality. However it is expedient to carry out the examinations of the part elements in both cases. It is obvious in present case the working hour requirement of the drilling – and welding operations don't reach neither together yet the yearly capacity of one machine (the yearly working hour capacity to be due to one worker), because of this it is worth to consider that the welding skilled worker should look after the two work sites. This connection is still logical because the drilling operation can be trained within short time generally.

No. 3 task.

Let's determine the economical number of series based on the formula worked out by Demin (chapter 14.6.3.2) in the knowledge of the following datums:

- Number of pieces to be manufactured………………………………Q = 10000 pcs.
- Indipendent costs from the number of series (for example material,- wage,- direct costs, overhead – etc.)………..A = 10000 Ft
- Preparation – and finishing costs connected with the starting – or with restarting of the series (for example setting, adjustment, cutting tools, measuring instruments, etc.)…………………………………………………...B = 50000 Ft

Substituting the data into Demin's formula:

$$n = 5{,}77\sqrt{\frac{QB}{A}} = 5{,}77\sqrt{\frac{10000 \times 50000}{10000}} = 1290\ pcs.$$

The economical number of series:……………………….…………………n_e = 1290 pcs.

If it is compared with Andler's formula (chapter 14.6.3.2) it can be seen in determined value that it can be got in order of magnitude nearly the same value.

No. 4 task.

The formula to determine the optimal stock volume (chapter 14.6.3.3) can be handled very simply and logically from mathematical standpoint, knowing the constants the optimum can be determined almost in an instant. However in the practice as it can be seen in the followings that those tasks are not near so simple as there is no general that the cost of lot can be considered independent from the volume (number of pieces) of the lot (c_n).

The optimal stock volume can be determined thus based on the following starting data.

- Customer needs 120 000 pcs. parts yearly.
- The cost of parts are piece by piece: Cp = 1000 Ft/pc.
- The storing cost of the parts piece by piece for one day C_{sd} = 10 Ft/pc.

The customer can do the following calculations based on the available data for the sake of that the purchasing price and the storing cost requires what total cost.

No.1. variation

The part volume for the whole year is bought at the same time, that is………N = 120 000 pcs.

The purchasing price of one piece……………………………………………c_p = 1000 Ft/pc.

In Figure 14.47 enclosed, supposing uniform decrease, the use of part is shown in the function of time (days).

Figure 14.47. Purchasing parts in one lot.

The purchasing price of the parts: $P_{p1} = C_p N_1 = 1000 \times 120000 = 120000000 \quad Ft$

The storing cost of the parts, thus: $C_{s1} = C_{sd} \frac{1}{2} N_i T_1 = 10 \times \frac{1}{2} 120000 \times 365 = 219000000 \quad Ft$

The total cost, thus: $C_{t1} = P_{p1} + C_{s1} = 120000000 + 219000000 = 339000000 \quad Ft$

No. 2. variation

The customer thinks so (rightly) that he/she doesn't order the whole volume (N) at the same time only its half, the other half in the second half-year.

The part decrease is shown in Figure (Figure 14.48) in the function of the (days), with the help of data given the purchasing and storing costs then the total cost can be determined.

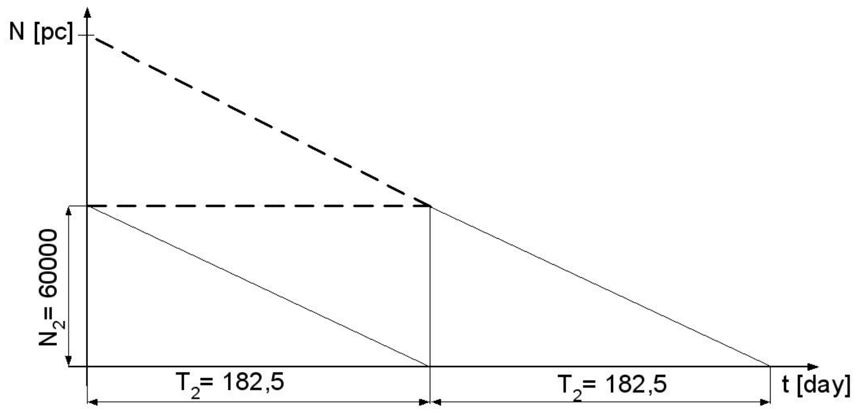


Figure 14.48. Purchasing parts in two lots.

The purchasing price of the parts for the whole year:

$$P_{p2} = 2 \times N_2 C_p = 2 \times 60000 \times 1000$$

$$P_{p2} = 120000000 \ \ Ft$$

The storing cost of the parts (Figure 14.48):

$$C_{s2} = C_{sd} \frac{N_2 T_2}{2} 2 = 10 \frac{60000 \times 182{,}5}{2} 2 = 109500000 \ \ Ft$$

The yearly total cost:

$$C_{t2} = P_{p2} + C_{s2} = 120000000 + 109500000 = 229500000 \ \ Ft$$

The savings arising from the storing cost:

$$C_{s1} - C_{s2} = 219000000 - 109500000 = 109500000 \ \ Ft$$

It is only natural that the customer halves the single purchasing volume and so the storing cost still will be less.

No. 3. variation

It is shown in Figure 14.49 that the part decrease how takes shape in case of quarterly delivery.

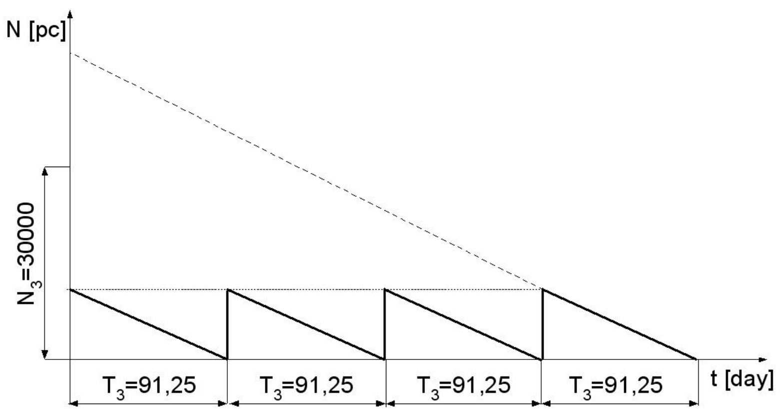


Figure 14.49. Purchasing parts in four lots.

The purchasing price of the parts corresponds to the values of previous variations naturally as the price didn't change.

The purchasing price of the parts for the whole year:

$$P_{p3} = 4 \times C_p N_3 = 4 \times 1000 \times 30000 = 120000000 \;\; Ft$$

The storing cost of parts:

$$C_{s3} = C_{sd} \frac{N_3 T_3}{2} 4 = 10 \frac{30000 \times 91{,}25}{2} 4 = 54750000 \;\; Ft$$

The yearly total cost :

$$C_{t3} = P_{p3} + C_{s3} = 120000000 + 54750000 = 174750000 \;\; Ft$$

Comparing the storing costs of No.1 and No.3 variations it can be seen, that the part delivery belonging to the No. 3 variation results more that 160 million forints savings for the customer company.

By reducing the delivery volumes in the same time the storing cost still can be reduced further, but that is also natural that the collateral costs emerging at the delivery company (switching over costs from one series to another, delivery – and other costs) increase the costs needed to produce parts. So over certain limit the transporter won't undertake already the increase of the yearly number of orders with the same conditions. Let's look at the variation when the transporter fixes higher price to the product depending on the part deliveries.

No. 5 task

Starting data to calculate the various variations (I suggest that always should be made drawings, figures to solve the tasks).

Yearly ordering volume: N = 120 000 pcs.

The storing cost of the product piece for by piece one day: C_{sd} = 10 Ft/pc.

Ordering variations and cost indicators belonging to those:

1. **Order for one year in one lot:**
   - number of pieces………………………………………...……..N_1 = 120000 pcs.
   - piece price………………………………………………………P_p = 1000 Ft/pc.

   The cost of the ordering lot for one year:

   $$C_{o1} = P_{p1} N_1 = 1000 \times 120000 = 120000000 \;\; Ft$$

   The storing cost of the ordering lot of one year:

   $$C_{s1} = C_{sd1} \frac{N_1 T_1}{2} = 10 \frac{120000 \times 365}{2} = 219000000 \;\; Ft$$

   The cost of the yearly volume in case of ordering in one lot:

   $$C_{t1} = C_{o1} + C_{s1} = 120000000 + 219000000 = 339000000 \;\; Ft$$

   It is still the same naturally in this case as in the previous example.
2. **Order for one year in two lots:**
   - number of pieces (number of pieces of one lot)…………………...N_2 = 60000 pcs.
   - piece price………………………………………………………P_{p2} = 1020 Ft/pc.

The cost of the ordering lots for one year:

$$C_{o2} = 2 P_{p2} N_2 = 2 \times 1020 \times 60000 = 122400000 \;\; Ft$$

The storing cost of the ordering lots for one year:

$$C_{s2} = 2 C_{sd2} \frac{N_2 T_2}{2} = 2 \times 10 \frac{60000 \times 182{,}5}{2} = 109500000 \;\; Ft$$

The cost of the yearly volume in case of ordering in two lots:

$$C_{t2} = C_{o2} + C_{s2} = 122400000 + 109500000 = 231900000 \ \ Ft$$

3. **Order for one year in three lots:**
   - number of pieces (number of pieces of one lot)……………….…....*N₃* = 40000 pcs.
   - piece price……………………………………………………….*P*$_{p3}$ = 1050 Ft/pc.

The cost of ordering lots for one year:

$$C_{o3} = 3P_{p3}N_3 = 3\times1050\times40000 = 126000000 \ \ Ft$$

The storing cost of the ordering lots for one year:

$$C_{s3} = 3C_{sd3}\frac{N_3T_3}{2} = 3\times10\frac{40000\times121{,}666}{2} = 72999600 \ \ Ft$$

The cost of the yearly volume in case of ordering in three lots:

$$C_{t3} = C_{o3} + C_{s3} = 126000000 + 72999600 = 198999600 \ \ Ft$$

4. **Order for one year in four lots:**
   - number of pieces (number of pieces of one lot)……………......N_4 = 30000 pcs.
   - piece price……………………………………………………..P_{p4} = 1300 Ft/pc.

The cost of ordering lots for one year:

$$C_{o4} = P_{p4}N_4 = 4\times1300\times30000 = 156000000 \quad Ft$$

The storing cost of the ordering lots for one year:

$$C_{s4} = 4C_{sd4}\frac{N_4T_4}{2} = 4\times10\frac{30000\times91{,}25}{2} = 54750000 \ \ Ft$$

The cost of the yearly volume in case of ordering in three lots:

$$C_{t4} = C_{o4} + C_{s4} = 156000000 + 54750000 = 210750000 \ \ Ft$$

Comparing the total costs it can be established that among the possibilities offered by the transporter the optimal solution for the customer is the order in three lots yearly. It can be made based on the No. 4 and No.5 examples the following statements:

- The calculations to be carried out in such simple way can be used very advantageously in those cases when offer is asked from more transporter and it has to be decided which is the most favourable.
- It can be sensible unanimously that what great importance has got the phase of decision proceedings (as in all other cases; too). For example chapter 14.6.2 Figure 14.26) concerning maximizing the profit by choosing the economical production process. Surely the detailed working out of the No4. and No5. tasks that maximum some days work (to search companies, solicitation for price offers, correspondence, processing offers come in) can result might 100 000 000 Ft savings for the customer, too.

No. 6 task.

The Janik – Zsoldos system theory model was presented in the chapter 14.5 which offers manysided solutions for qualifying complicated equipment technically-economically used in the practice, which main indicators: machine operation characteristics, machine maintenance characteristics, simulation characteristics for strategic decisions etc.

We wish to show within the framework of this task on the other hand – with thought experiment character – that analysing the optimizing process of the whole operation cycle of the machines with many elements (for example public road vehicles, building industry machines,

agricultural machines, etc.) what technical -, economical – maybe social questions are raised to be worth to consider.
As a starting basis the cost function [14.11] of maintenance in the long period is considered which takes into the sale of used machine and the repair costs. The following formula gives the function connection of the factors listed:

$$\Gamma(t) = A_0 - A_0\varphi(t) + \psi(t)$$

where: A_0 = is the purchase price of the equipment,,
$\varphi(t)$ = selling factor $\varphi(0) = 1$ and $\varphi(t)$ decreases monotonusly (Figure 14.50),
$\Psi(t)$ = accumulated maintenance cost $\Psi(0) = 0$ and $\Psi(t)$ increases monotonously (Figure 14.51)

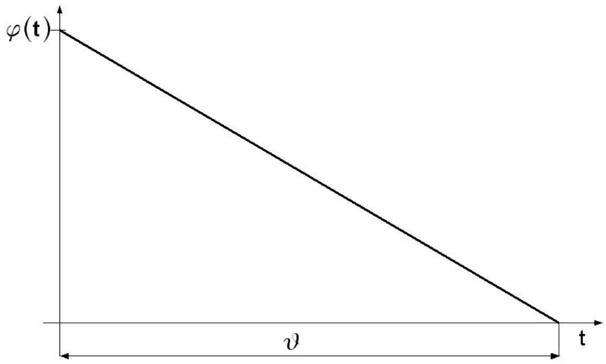


Figure 14.50. The change of selling factor.

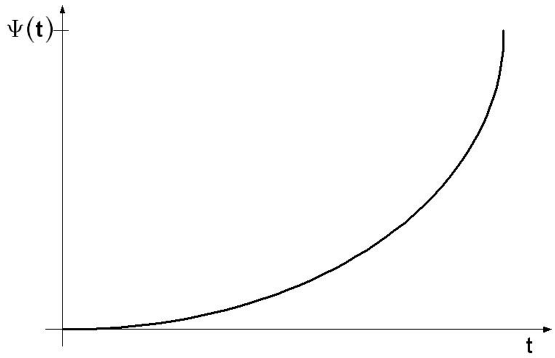


Figure 14.51. The change of maintenance factor.

As the model gives the maintenance cost for long period the optimal service – life can be determined by minimizing the specific cost.

$$\gamma(t) = \frac{\Gamma(t)}{t} = \frac{1}{t}\left[A_0 - A_0\varphi(t) + \psi(t)\right]$$

As the $\varphi(t)$ and $\Psi(t)$ are given in implicit form, the derivative of the specific function can be marked out only:

$$\gamma'(t) = \left[\frac{\Gamma(t)}{t}\right]' = \frac{t\Gamma'(t) - \Gamma(t)}{t^2} = 0$$

That fact limits the practical applicability of the relation that to determine the $\varphi(t)$ and $\Psi(t)$ factors there are at disposal altogether as much information what it can be determined from the boundary conditions. So it can't be transformed directly to characterizing the operation cycle of machines operating in different fields of production at the same time the question is that by the means of our hypothesis what results can be got in that cases if different values are given to $\varphi(t)$ and $\Psi(t)$.
It is unambiguous entirely that infinite possibilities are as opportunity – theoretically – to choose the two parameters, however considering the manageability of the function three characteristic operation cycle curves can be generated to which is worth to follow with attention.
First case:
$\varphi(t)$ and $\Psi(t)$ should be linear, for example:

$$\varphi(t) = 1 - \frac{t}{\vartheta} \text{ and } \Psi(t) = kt$$

where: t = is the operation time,

ϑ és k = are constants.

Substituting the data taken up and constituting the specific function the following result can be got:

$$\Gamma(t) = A_0 - A_0(1 - \frac{t}{\vartheta}) + kt$$

$$\gamma(t) = \frac{A_0}{t} - \frac{A}{t} + \frac{A}{\vartheta} + k$$

$$\gamma(t) = \frac{A}{\vartheta} + k$$

Figure 14.52 shows if there would be such machine then its specific cost always would be the same independently from the time.

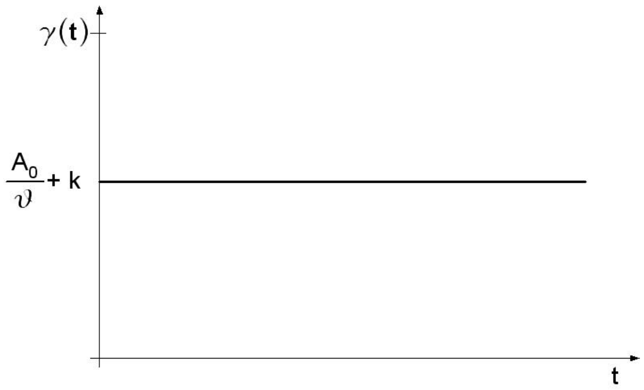


Figure 14.52. The change of the specific machine maintenance cost.

Let's look at that whether it is possible according to present knowledges such a machine in the reality and this hypothesis whether can be proved as a conceptual experiment:

Based on the Chapter 14.5.1, Figure 14.9. the followings can be said:

- Both $\varphi(t)$ and $\Psi(t)$ satisfy that practical experience formally that $\varphi(t)$ has got decreasing $\Psi(t)$ has got increasing tendency in the function of the operation time.
- However contradictions seem in connection with (ϑ) constant to be in $\varphi(t)$ factor, which Figure 14.50 shows clearly. The formal interpretation range of the function is from t=0 to t=∞. This also means that (ϑ) also increases (Figure 14.50) in accordance with t – value which is in fundamental contradiction with that suppose that (ϑ) is constant.
- In order to simple handling of the $\Gamma(t)$ function linear connection is supposed between the cost of machine maintenance and operation time. The practice and the special literature as well as the characteristics of machine maintenance made known on the chapter 14.5.2 prove that the costs of machine maintenance increase in an increased degree in the function of the operation time.

After all summing up the process of the first phase of thought experiment it can be got that conclusion it is needed to submit the results got based on the hypothesis to thorough checking, first of all – if there is a possibility – with the practice but besides with raising our questions whether it is possible, whether it is acceptable, whether it reflects the reality.

This dialogue has to be continued until acceptable answer can be got whether the result obtains proof or rejection.

Second case:

$\varphi(t)$ and $\Psi(t)$ should be non linear, for example:

$\varphi(t) = e^{-\lambda t}$ and $\psi(t) = k_0 \cdot (e^{\mu t} - 1)$

where: t = is the operation time

λ, μ, k_0 = constants

Substituting the data taken up and forming the specific function the following result can be got:

Running maintenance function for long period.

$$\Gamma(t) = A_0 - A_0 e^{-\lambda t} + k_0 (e^{\mu t} - 1)$$

Specific running maintenance function.

$$\gamma(t) = \frac{1}{t}[A_0 - A_0 e^{-\lambda t} + k_0 e^{\mu t} - k_0]$$

We are interested in the specific running maintenance function what from has got. It can be established that in the $\gamma(t)$ 0 and $+\infty$ interval continuous curves and can be generated to be seen in Figure 14.53 depending on the parameters.

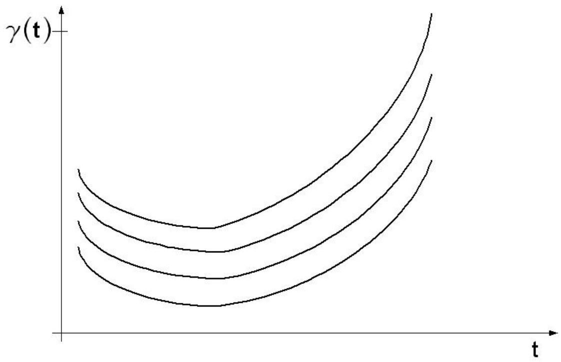


Figure 14.53. Specific machine operation curves.

The $\gamma(t)$ value of the function at t=0 locus can't be determined directly as both the numerator and the denominator have got 0 value, that is the connection is indeterminated,

$$\gamma_{t \to 0}(t) = \frac{o}{0}$$

However there is no cause to despair if the boundary conditions are known, that is the constants, surely by substituting t – values the specific curve can be plotted.

Let's suppose that the constants are the followings:

A_0 = 100 000 EUR

λ = 0,5

μ = 0,8

k_0 = 10 000 EUR

Let's write down the $\gamma(t)$ in transposed form:

$$\gamma(t) = \frac{1}{t}\left[A_0 - k_0 + k_0 \cdot e^{\mu t} - \frac{A_0}{e^{\lambda t}}\right]$$

The following Table is expedient to use to determine the discrete data for plotting the function looked for.

Table 14.5.

t	+ A_0-k_0	λt	$e^{\lambda t}$	- $A_0 \frac{1}{e^{\lambda t}}$	μt	$e^{\mu t}$	+ $k_0 e^{\mu t}$	$\gamma(t)$
0,0001	90000	0,00005	1,00005	99995	0,00008	1,00008	10000,8	58000
0,001	90000	0,0005	1,0005…	99950…	0,0008	1,0008…	10008,…	57988
0,01	90000	0,005	1,0050	99502	0,008	1,00803	10080,3	57800
0,1	90000	0,05	1,0511	95138	0,08	1,08328	10832,8	56940
0,5	90000	0,25	1,2840	77881	0,4	1,4918	14918	54074
1,0	90000	0,5	1,6487	60653	0,8	2,2255	22255	51602
2,0	90000	1,0	2,7182	36789	1,6	4,95303	49530,3	51370
3,0	90000	1,5	4,4816	22313	2,4	11,02317	110231,7	59306
4,0	90000	2,0	7,3890	13533,5	3,2	24,53253	245325,3	82698

Based on the data of the Table the specific costs can be presented in the function of the time of use (Figure 14.54). It can be established from the Table 14.5 that the optimal time of use: $t_{opt.}$ = 2 years.

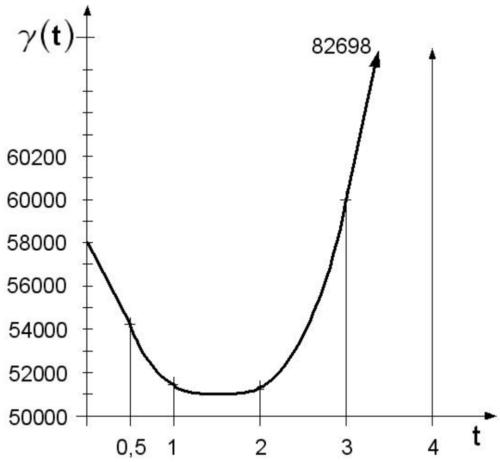


Figure 14.54. Machine operation characteristic curve drawable on the basis of table data.

Based on the Figure it can be seen that the optimum is between 1 – and 2, thus according to the Table 14.5 the 2 years as suboptimum can be make better if for example the specific cost is also determined in more discrete moments (t=1,2;1,4;1,6 and 1,8). This calculation is left to you, but let's have a look what help gives the mathematics to such case when a function has got indeterminate form, that is in this case the value is

$$\underset{t \to 0}{\gamma(t)} = \frac{0}{0}.$$

Let's write down again the formula of the specific cost:

$$\gamma(t) = \frac{1}{t}[A_0 - A_0 e^{-\lambda t} + k_0(e^{\mu t} - 1)]$$

By substituting t=0, the result is that the numerator and denominator is zero, too.

$$\gamma(t)_{t\to 0} = \frac{A_0 - A_0 e^{-\lambda t} + k_0(e^{\mu t} - 1)}{t} = \frac{0}{0}$$

The value of the function (t=0) can be determined by using the L'Hospital formula.
Namely

$$\gamma(t)_{t\to 0} = \frac{f(t)}{g(t)} = \frac{f'(t)}{g'(t)} = \frac{\lambda A_0 e^{-\lambda 0} + \mu k_0 e^{\mu 0}}{1}$$

That is

$$\gamma(t)_{t=0} = \lambda A_0 + \mu k_0$$

Let's substitute the constants:

$$\gamma(t=0) = 0{,}5 \times 100000 + 0{,}8 \times 10000$$

$$\gamma(t=0) = 58000 \ \ EUR / year$$

This value agrees with the tabular value determined with attempting repeatedly. If the Figure 14.54 is thus compared with characteristics of specific running maintenance got by real operative application of the Janik's machine running maintenance model discussed in details in 14.5 chapter, it can be said that by substituting of factors $\varphi(t)$ and $\Psi(t)$ choosing arbitrarily – the function given – reflects theoretically really one of the typical technical – economical parameters of the machines operated within working conditions.
A wide range of knowledge in different disciplines is needed to make optimization. Surely to answer such "simple" question as that the specific cost function ($\gamma(t)$) at t=0 locus has got what value in case of knowing the L'Hospital – formula needs one –two minutes only.

Third case:
Should be $\varphi(t)$ exponential (decreasing monotonously) and $\Psi(t)$ linear increasing monotonously:

$$\varphi(t) = e^{-\lambda t}; \ \psi(t) = kt$$

Substituting the factors taken down and forming the specific function the following result can be got.
The running maintenance function for long period.

$$\Gamma(t) = A_0 - A_0 e^{-\lambda t} + kt$$

The specific running maintenance function.

$$\gamma(t) = \frac{1}{t}\left[A_0 - A_0 e^{-\lambda t} + kt\right]$$

The specific cost function (Figure 14.55) at t=0 locus can be determined by the help of the L'Hospital formula.
That is:

$$\lim_{t\to 0}\gamma(t) = \lim_{t\to 0}\frac{f(t)}{g(t)} = \lim_{t\to 0}\frac{f'(t)}{g'(t)} = \frac{\lambda A_0 e^{-\lambda 0} + k}{1}$$

That specific cost function, thus,

$$\gamma(t)_{t=0} = \lambda A_0 + k$$

Let's have a look the function value at $t = \infty$ locus.

$$\lim_{t\to\infty}\gamma(t)=\lim_{t\to\infty}\frac{f(t)}{g(t)}=\lim_{t\to\infty}\frac{f'(t)}{g'(t)}=\frac{\lambda A_0 e^{-\lambda 0}+k}{1}$$

The specific cost function (Figure 14.55) thus, $\gamma(t)_{t\to\infty}=k$

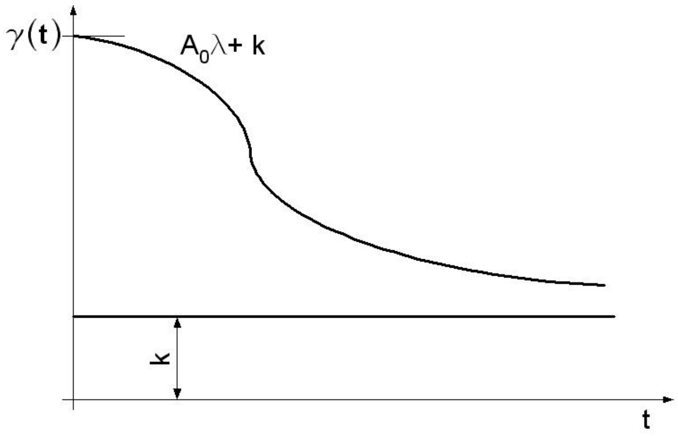


Figure 14.55. Machine operation characteristic curve of an equipment with "infinite" service-life.

After these it is instructive imagine on that a machine with the characteristic of running maintenance shown in Figure 14.55 what would mean practically.

- the energies devoting to exploit then further processing of raw materials moderate significantly, the environment burdening would reduce,
- it would be minimal entirely the maintenance costs of parts made of special, quasi infinite service–life material, surely it can be supposed, that the economical long service-life is in synchron with the insignificant maintenance, however it can be supposed that to manufacture such machine would be extremely expensive and the market couldn't pay for it.
- what I mentioned as facts are not sure of being advantageouses concerning the rate of employment, too. Surely it can be supposed rightly that substantional manpower capacity would also get free both in the fields of exploitation and of processing further of service, which can cause social and of national economical problems if it exceeds a certain volume.

This example also shows well that optimization requires wide-spread interdisciple knowledges. The bases of these knowledges can be got learning the university subjects on appropriate level. Establish natural thinking attitude that recognizing the interdisciple connections between different subjects how could you utilize in optimizing processes.
It is possible to reach if during learning different subjects there is an aspiration to learn such knowledge aggregation which makes possible laying the foundation the talent of recognizing the problem and solving it.

14.8. Literature:

[14.1] Starr, M. K.: Production Management Systems and Synthesis. Prentice-Hall, Inc. Englewood Clifs, New Jersey, USA, 1972 (in Hungarian: Rendszerelméletű termelésvezetés, termelésszervezés. Közgazdasági és Jogi Könyvkiadó, Budapest, 1974.)

[14.2] Erdélyi, F.: Some technological relations of the global crisis. (in Hungarian: A globális válság néhány technológiai vonatkozása. Gépgyártás, XLIX. Évfolyam, 2009. 3. szám. 13-21. old.)

[14.3] Ackoff, R. L.: A concept of Corporate Planning. John Wiley et. Sons, Inc. 1970 (in Hungarian: Operációkutatás és vállalati tervezés. Közgazdasági és Jogi Könyvkiadó. Budapest, 1974.)

[14.4] Churchman, C. W.: The systems approach. Dell Publishing Co., Inc. 1968 (in Hungarian: Rendszerelmélet. Statisztikai Kiadó Vállalat, Budapest, 1974.)

[14.5] Heisenberg, W.: Physics and beyond. Harper and Row Publishers, Inc. 1971 (in Hungarian: A rész és az egész, beszélgetések az atomfizikáról. Gondolat, Budapest, 1974.)

[14.6] Deli, L., Kocsis, J., Ladó, L.: Economic calculations based on system theory (in Hungarian: Rendszerelméleten alapuló gazdaságossági számítások. Műszaki Könyvkiadó, Budapest, 1975.)

[14.7] Horváth, M., Somló, J.: Optimizing cutting machinings and their adaptive control (in Hungarian: Forgácsoló megmunkálások optimálása és adaptív irányítása. Műszaki Kiadó, Budapest, 1979.)

[14.8] Janik, J.: System-view planning of agricultural machine maintenance. HAS doctoral dissertation, Budapest, 1980. (Library of the Hungarian Academy of Sciences) (in Hungarian: Mezőgazdasági gépfenntartás rendszerszemléletű tervezése. MTA doktori értekezés. Budapest, 1980.)

[14.9] Janik, J.: System-view control of machine maintenance of agricultural companies. (in Hungarian: Mezőgazdasági vállalatok gépfenntartásának rendszerszemléletű irányítása. Akadémiai Kiadó, Budapest, 1979.)

[14.10] Jándy, G.: System analysis and control. (in Hungarian: Rendszerelemzés és irányítás. Statisztikai Kiadó Vállalat, Budapest, 1975.)

[14.11] Kaufmann, A.: Méthodes et modéles de la recherche opérationelle. Dunod Paris 1962. (in Hungarian: Az operációkutatás módszerei és modelljei. Műszaki Könyvkiadó, Budapest, 1977.)

[14.12] Janik, J., Szíjjártó, O.: Complex Evaluation System of Agricultural Machines. Agricultural Research, December 1993. Vol. 2, No.4, pp. 4-6.

[14.13] Janik, J., Vermes, P.: Machine using time – optimal methods or the optimum of the methods. (in Hungarian: Géphasználati idő – optimális módszerek vagy a módszerek optimuma. Gépgyártástechnológia, 2000. jún. 6. 36-42 old.)

[14.14] Janik, J., Zsoldos, I.: Complex qualification system of machines. (in Hungarian: Gépek komplex minősítési rendszere. Gépgyártás, 2001. április. 6-8. old.)

[14.15] Zsoldos, I., Janik, J.: Virtual machine utilization characteristics. (in Hungarian: Virtuális géphasznosulási jellemzők. Járművek és mobilgépek, Gödöllő, ISBN 978-963-269-227-2, 2011. 89-92. old.)

[14.16] Derman, C.: On Sequential Decisions an Marcov Chains. Man. Sci., Vol.9., No.1., Október 1962. pp. 16-24.

[14.17] Bode, B.: Die Nutzungsdauer – eine ökonomische Kategorie. Fertigungstechnik und Betrieb, 1962. No. 8. sz. pp. 553-555.

[14.18] Chapman, H.: Cost Reduction Trough Improved Maintenance Operations. Techniques of Plant Engineering and Maintenance. Clapp and Poliak, New York, Vol. XV. 1964. pp. 17-26.

[14.19] Dudás, I.: Mechanical Engineering Technology. (in Hungarian: Gépgyártástechnológia. Miskolci Egyetem, 2001.)

[14.20] M. Horváth, S. Markos (editors): Mechanical Engineering Technology. (in Hungarian: Horváth, M., Markos, S. (szerkesztők):Gépgyártástechnológia. Műegyetem Kiadó, 2002.)
[14.21] Böhme, C., Borghardt, H. J., Kirberg, A.: Informationsbuch für technologen metallverarbeitende industrie. VEB Verlag Tachnik, Berlin, 1964. (in Hungarian: Gyártástervezési kézikönyv. Műszaki Könyvkiadó, Budapest, 1967.)
[14.22] Kocsis, J.: Organizing production systems. (in Hungarian: Gyártási rendszerek szervezése. Műszaki Könyvkiadó, Budapest, 1972.)
[14.23] Maynard, H. B.: Industrial Engineering Handbook. Mc Graw-Hill Book Company, New York, 1956. (in Hungarian: Gazdasági mérnöki kézikönyv. Műszaki Könyvkiadó, Budapest, 1977.)
[14.24] J. Janik (editor and author): Machine maintenance.(in Hungarian: Janik, J. (szerkesztő és szerző) Gépfenntartás I. II. Dunaújvárosi Főiskola, Dunaújváros, 2009.)
[14.25] Andler, K: Rationalisiernung der Fabrikation und optimale Losgrösse. München, Oldenburg, 1929. (in Hungarian: Héberger, K., Cserhalmi, Gy.: A gazdaságos sorozatnagyság számítási módszerei és bibliográfiája. Tankönyvkiadó, Budapest, 1965.)