INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN
ICED 05 MELBOURNE, AUGUST 15-18, 2005

A METHOD FOR THE CONSIDERATION OF INFLUENCES OF THE SURFACE ROUGHNESS ON THE BEHAVIOUR AND DESIGN OF GLOBAL SYSTEMS

Albert Albers, Lukas Nowicki, Johannes Minx

Keywords: surface roughness, virtual prototyping, product development

1 Introduction

Frictional contacts are a component of almost any technical system. The surface characteristic of the contact partners in this friction contact has a large influence on the dynamic behaviour of the entire system, which affects again the global design. Frequently the design of entire system needs cost-intensive rework respectively the attachment at supplementary systems to remove unwanted characteristics. Because of that the specific system knowledge must be included into the early phases of the product developing process. A virtual test field can be reproduced by the use of recent simulation tools [1].

In the following a method for the generation of surface roughness with desired characteristic for virtual investigational procedure already in the concept phase of the product developing process will be presented. Furthermore, it will be shown how the influence of a friction contact on the system behaviour can be taken into account in a simulation. Surface characteristic

The challenge during the description of the surface characteristic is the allocation of the surface of a numerical value, which describes its substantial characteristics and supplies expressive data concerning its later function behaviour [2]. For the complete description of the surface properties the form of the profile must be considered beside the well-known arithmetic (R_a) and square (R_q) roughness factors or the averaged depth of roughness ($R_{z(DIN)}$). This happens by dint of the linear material ratio curve (Abbott curve), which illustrates the increase of the material load share (M) as a function of the increasing depth of profile (R) of the rough surface. It can be divided into three sections by means of three parameters: reduced peak height (R_{pk}), reduced valley depth (R_{vk}) and core roughness depth (R_k) [3]. The core roughness depth describes the area of the profile which shows the highest increase of the material load share. The reduced peak height gives the shape and the number of the roughness amplitudes. Analogously, the R_{vk} parameter describes the valley section. The sum of these three parameters constitutes the maximum height of the profile (R_t)

Figure 1 shows two surfaces of the same arithmetic mean value as well as profile height but which have different profile characteristics. The upper surface has a relatively big roughness amplitude section and a small valley depth section. For the contact examination this means that the many peaks produce a small nominal contact area, which results in high contact friction and heat. The lower surface, however, has a rather small roughness amplitude section but a bigger valley depth section. Therefore, the nominal contact area is bigger. Also, in lubricated contacts grease pockets can occur in the valley depth section, which, for instance, allows drawing off heat better.

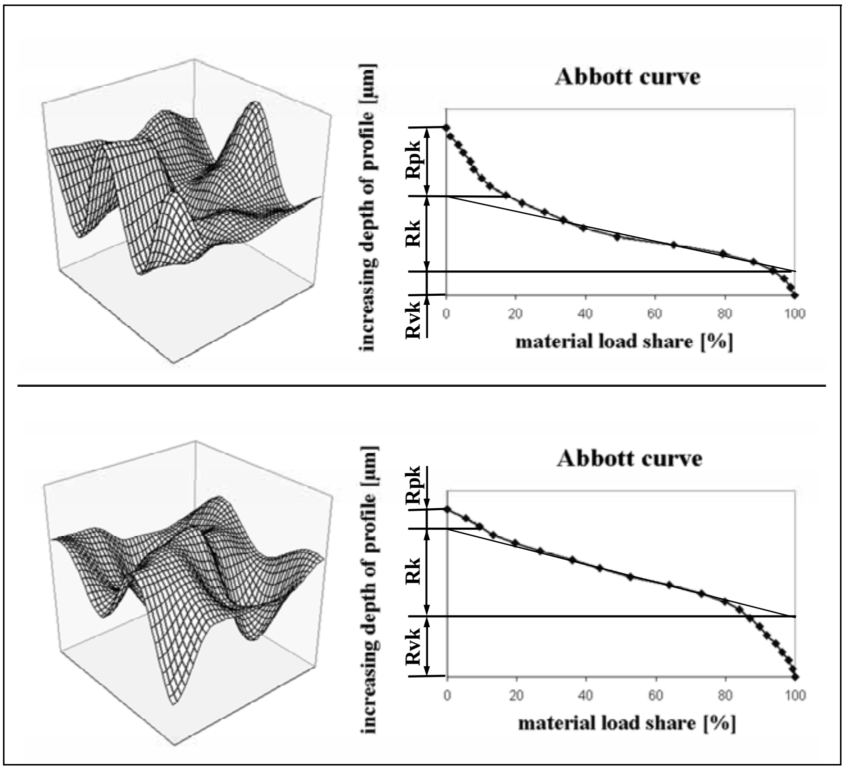

Figure 1. stochastically generated surfaces
upper: R_a=0.4, R_t=2.7, R_{pk}=1.12, R_k=1.23, R_{vk}=0.35
lower: R_a=0.4, R_t=2.7, R_{pk}=0.31, R_k=1.43, R_{vk}=0.96

The existence of both surfaces is justifiable since they are suitable for different applications. The surface with the bigger roughness amplitude section can be used in a friction system like, for example, a continuously variable transmission drive (CVT) where the lubricant is to be displaced as quickly as possible to ensure a loss-free transfer of the transmittable torque. The surface with the bigger valley depth section can be applied in a dynamic system, for instance in a cam drive gear, where in spite of insufficient lubrication an adequately low friction value is demanded.

For further procedures it is necessary to determine quantifiable parameters for the surface characterisation. For the description of the Abbott curve, the material load share (M) is regarded as the function of the depth of profile (R).

$$M = f(R) \tag{1}$$

With their help, the vector $\vec{A}$ of the following kind can be formed:

$$\vec{A} = \begin{Bmatrix} M_0 = M(Rt) \\ M_1 = M(\frac{n-1}{n} * Rt) \\ M_2 = M(\frac{n-2}{n} * Rt) \\ . \\ . \\ M_n = M(0) \end{Bmatrix}, n = 20 \tag{2}$$

For the description of the surface roughness, the root mean square (R_q) is used. However, the procedure introduced here, can also be transferred to the other roughness parameters.

2 Mathematical description of surfaces

A particular demand on the mathematical description method of technical surfaces is its applicability on several manufacturing processes. This requirement excludes the use of polynomials of a constant order since it changes depending on the manufacturing processes. Polynomials describing lapped surfaces would be of a much higher order than polynomials describing lathe surfaces. To solve this task, harmonic functions are used instead of polynomials. The discrete Fourier transform (DFT) allowed the applicability on the most common manufacturing processes.

The Fourier analysis is a mathematical approach to the decomposition of signals into superposed cosine waves of different frequencies and amplitudes. Another special variant of this analysis is the discrete form. This form is used for signals which consist of a number of finite discrete measuring points. As a result, the DFT provides only a finite number of frequencies. If, for instance, a signal of the length L consists of N discrete measured values, it can be divided into a constant component as well as N-1 frequency components, which lie within a range from 1/L to (N-1)/L. Characteristically, the transformed signals are functions of time or space. That is why the unit of the frequency can be [1/second] or [1/unit length]. A Fourier transformed signal has, besides periodicity, another characteristic useful for the result evaluation. Its amplitude spectrum is symmetric. Therefore, for an even number of measured values, only the first N/2 frequencies can be examined. In the case of an uneven number of measured values, the first N/2 + 1 frequencies are important.

For the analysis and synthesis of technical surfaces, a two-dimensional transformation based on the following equations [4] must be carried out:

$$Z(\omega_1,\omega_2) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} z(x,y) * e^{-j*\omega_1*\frac{2\pi}{L}*x} * e^{-j*\omega_2*\frac{2\pi}{L}*y}, \ \omega_1,\omega_2 = 0,1,...,N-1 \tag{3}$$

$$z(x,y) = \sum_{\omega_1=0}^{N-1}\sum_{\omega_2=0}^{N-1} Z(\omega_1,\omega_2) * e^{j*\omega_1*\frac{2\pi}{L}*x} * e^{j*\omega_2*\frac{2\pi}{L}*y}, \ x,y = 0,1,...,N-1 \tag{4}$$

The examined surfaces had the dimensions 500µm x 500µm (sample rate 2 µm) and were manufactured as follows:

- Milling (R_q: 2.83µm, 0.58µm)
- Turning (R_q: 4.25µm, 1.03µm)
- Grinding (R_q: 0.67µm, 0.35µm)
- Lapping (R_q: 0.52µm, 0.21µm)

For each surface, the frequency spectrum with its corresponding amplitudes, the Abbott curve as well as the R_q-value were determined. The examination focused on how many of the approximately 31000 amplitudes play a significant role. For this, in case 1 all amplitudes were counted which were more than 15% of the highest value. In other cases the bound was reduced to 10% (case 2) and to 5% (case 3). Figure 2 shows the analysis results of two surfaces: turning (R_q = 1.03µm) and lapping (R_q = 0.21µm) as well as their dominant amplitudes of case 1. Both examples reflect very well the trend of the whole examination. All surfaces can be described by a manageable amount of frequencies. All significant amplitudes were in the first five columns and rows of the Fourier matrix. Manufacturing processes with geometric definite cutting edge showed less dominant amplitudes than manufacturing processes with geometric indefinite cutting edge. The number of amplitudes depended very much on the roughness value. With decreasing R_q-values, more high-frequency components occurred. In case 1 the number of relevant amplitudes depending on the manufacturing process was between 11 – 28 amplitudes. In case 2 this range increased to 17-54 amplitudes and in case 3 to 32-198 amplitudes. In order to examine in how far it is allowed to neglect the remaining frequency components, a reverse transform with solely the dominant amplitudes was carried out in all three cases and the reversely transformed surfaces were compared with the original versions. Already in case 1, the reduced surfaces showed characteristics similar to the measured surfaces. This allowed the conclusion that surfaces with desired characteristics can be generated with a small number of amplitudes.

3 Evolutionary Algorithms

For generating the technical surface with the given characteristic, an optimization algorithm was chosen, which is based on the biological process of evolution. This optimization algorithm was developed in Germany at the beginning of the seventies [5, 6] and is nowadays a established as standard tools for optimization. It is used for problems with a huge area of design and the is highly nonlinear with lots of local extremes. Searching for the optimum is done by generating multiple, stochastically distributed points in the design area – called individuals - an weighting them with respect to the value of the objective function. Based on the individuals with the best fitness, a new group – called generation – is created. This process is repeated iteratively until an optimum is reached.

Based on equations (2) and (4), the following objective function is defined:

$$f(z, \vec{A}_{Pattern}) = \vec{A}(z) \bullet \vec{A}_{Pattern} \tag{5}$$

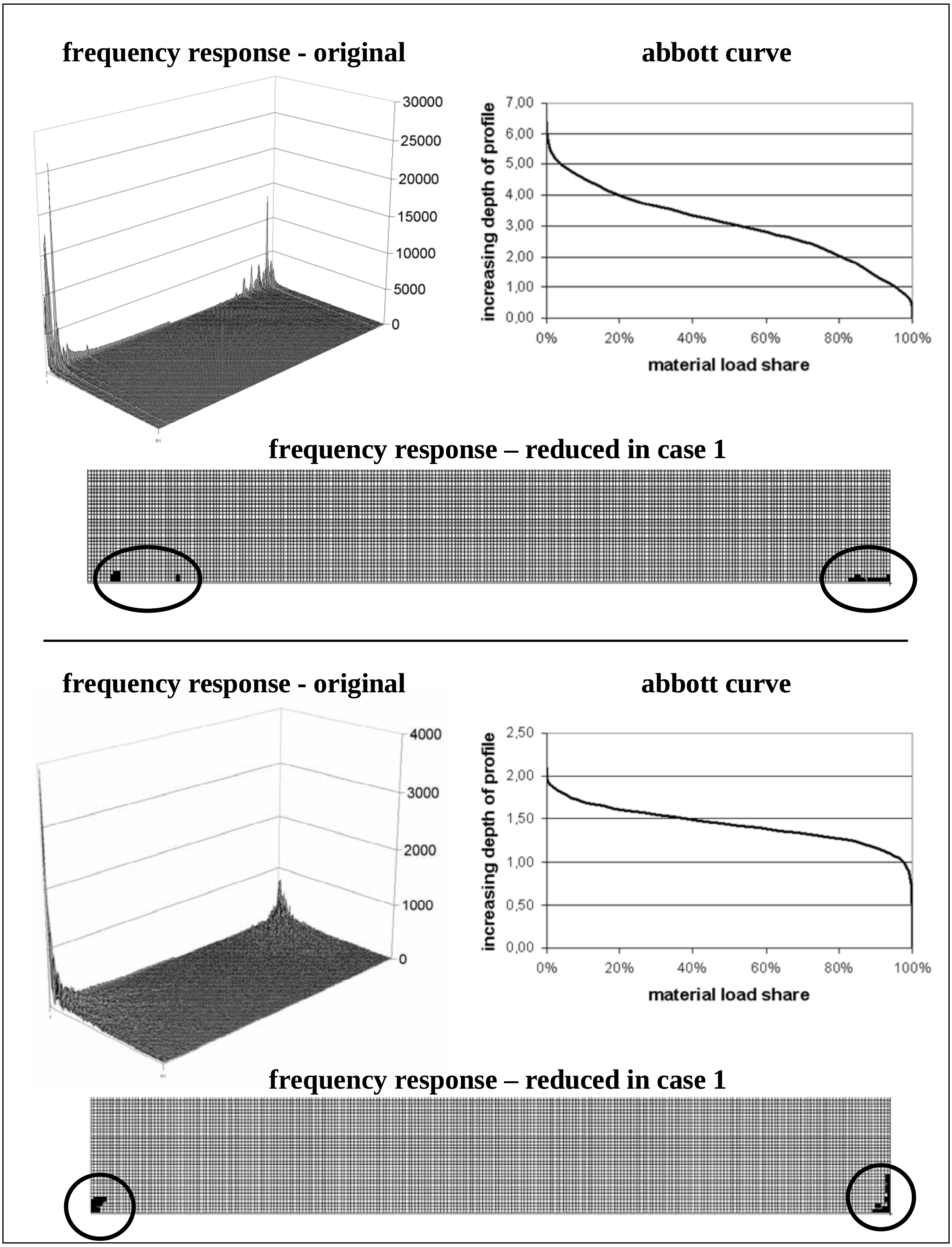

Figure 2. frequency spectrum and abbott curve of measured surfaces
top: turning Rq=1.03µm , number of dominant amplitudes: 18
bottom: lapping Rq=0.21µm, number of dominant amplitudes: 27

Based on that the following formulation for the optimization problem is defined:

$$\begin{aligned} &\text{Maximize} \quad f\,(\,z\,,\,\vec{A}_{Pattern}\,) \\ &\text{Subject to} \quad h(z) = R_{q\_constraint} \end{aligned} \tag{6}$$

This means, that the scalar product of the vector $\vec{A}$ for the generated surface and vector $\vec{A}$ for the measured surface should be maximized with the restriction, that value for $R_{q\_constraint}$ must correspond to the given value. Figure 3 shows the process for the optimization schematically.

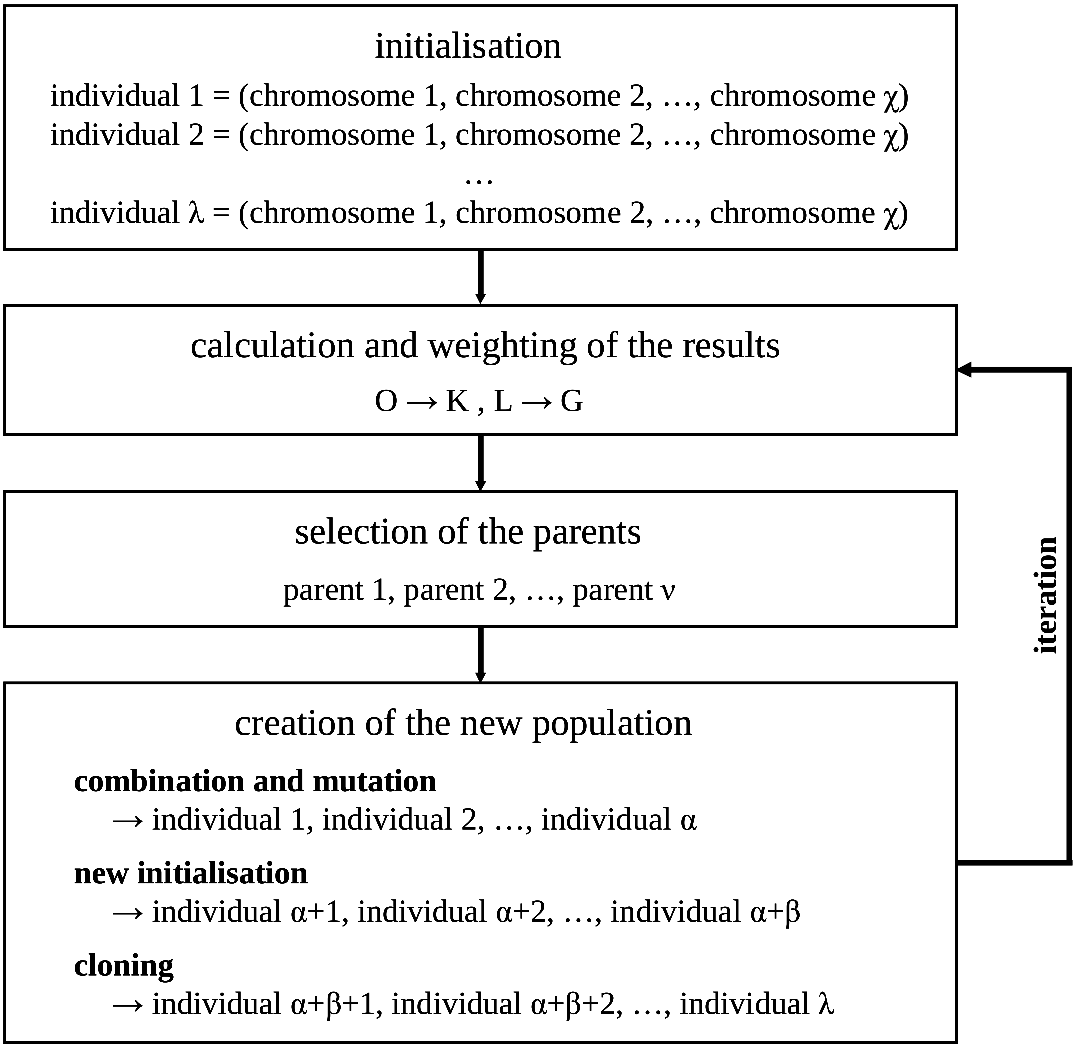

Figure 3. flow chart of the optimization process

In the first step an initial population is created. This population consists of a number of λ individuals which all have χ chromosomes. The value for each chromosome is generated randomly following an equal distribution. In the following step for each individual the value of the objective function and the equality constraint is calculated. The calculated values of the objective function are sorted in descending order.

$$O = \begin{pmatrix} f_1(z_1, \vec{A}_{Pattern}) & h_1(z_1) \\ f_2(z_2, \vec{A}_{Pattern}) & h_2(z_2) \\ . & . \\ . & . \\ f_\lambda(z_\lambda, \vec{A}_{Pattern}) & h_\lambda(z_\lambda) \end{pmatrix}, \tag{7}$$

$$f_1(z_1, \vec{A}_{Pattern}) > f_2(z_2, \vec{A}_{Pattern}) > ... > f_\lambda(z_\lambda, \vec{A}_{Pattern})$$

For the weighting of the objective function and constraints, two weighting functions are established,

$$K = \begin{pmatrix} k_1(\lambda, i) \\ k_2(\lambda, i) \\ . \\ . \\ k_\lambda(\lambda, i) \end{pmatrix}, \; k_i = \lambda + 1 - i, \; i = 1,2,...,\lambda \tag{8}$$

$$L = \begin{pmatrix} l_1(z_1, R_{q\_constraint}) \\ l_2(z_2, R_{q\_constraint}) \\ . \\ . \\ l_\lambda(z_\lambda, R_{q\_constraint}) \end{pmatrix}, \; l_i(z_i, R_{q\_constraint}) = \frac{1}{\left(\left|h_i(z_i) - R_{q\_constraint}\right| - 0.12\right)^4}, \; i = 1,2,...,\lambda \tag{9}$$

which were joint to a over all weighting function.

$$G = \begin{pmatrix} g_1(k_1, l_1) \\ g_2(k_2, l_2) \\ . \\ . \\ g_\lambda(k_\lambda, l_\lambda) \end{pmatrix}, \; g_i = k_i * l_i, \; i = 1,2,...,\lambda \tag{10}$$

The first ν individuals with the highest weighting factor were elected as parents. All other individuals are discarded. With help from the parents the next generation of individuals are build. The following strategies are used to generate the children:

- combination and mutation
  for each chromosome of a new individual two parents are chosen randomly. The children earn the mean value of the parents chromosome. According to this strategy, α new children are generated. These children can underlie a mutation which means, that there a small modifications for the inherited properties. For each chromosome a mean value μ and the

standard deviation σ is calculated from all parents. This statistic is adjusted in each step of the iteration. In case of a mutation, the chromosome gets a random value in the range of $\mu \pm \sigma$. The probability that a mutation of a chromosome occurs is $1/\chi$.

- new initiation
  to avoid the algorithm form sticking in a local optimum, which would not lead to satisfying results, a number of β new children is generated randomly.
- cloning
  to prevent the risk of deterioration, all ν parents are cloned for the next generation.

For all λ new individuals the process of calculating the objective function and the constraint is done again. This loop is continued as long as none of the stop condition is not satisfied. For the considered task, a maximum number of iteration is given.

4 Results of the surface generation

Two optimization tasks were set up for the manufacturing process "turning" in order to verify the method. For both optimizations the abbot curve from figure 2 was used for reference. The constraint in optimization 1 was a value of 1.5µm for R_q and a value of 2.5µm for R_q in optimization 2. The following optimization parameters were used:

- Optimization 1
  number of individual $\lambda = 61$
  number of parents $\nu = 7$
  number of combined children $\alpha = 49$
  number of new children $\beta = 5$

- Optimization 2
  number of individual $\lambda = 60$
  number of parents $\nu = 5$
  number of combined children $\alpha = 35$
  number of new children $\beta = 20$

Figure 4 shows the results for both of the optimizations. You can see the characteristic of the abbot curve of the reference surface (bottom) as well as the characteristic of the abbot curve received from optimization 1 and optimization 2. Both optimization results show the character of the chosen manufacturing process. The scalar product of the vector $\vec{A}$ for the first result with vector $\vec{A}$ for the reference surface yields to a value of 0.9999 which means that the characteristic of both abbot curves is very similar. The equality constraint could be preserved very good with an value of 1.5005µm for R_q. The quality of the second result is less accurate than the first one but still good enough. The scalar product of the vectors is 0.9947 and the value of R_q is 2.4959µm.

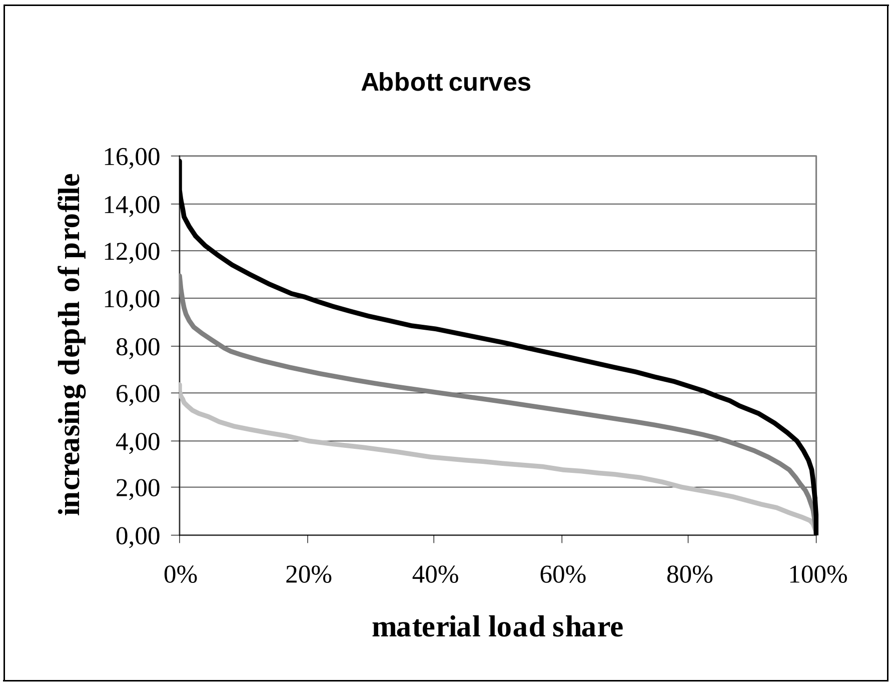

Figure 4. Abbott curves
bottom: pattern curve
center: first optimized surface Rq = 1.5µm
top: second optimized surface Rq = 2.49µm

Regarding the stochastically generated surfaces in figure 5, it is obvious, one can see, that their topography shows the characteristic of a turned surface. For this kind of manufacturing process it is typical, that the surfaces shows grooves in manufacturing direction in regular distance. So the roughness value measured in manufacturing direction is significantly higher the measured perpendicular.

Different from the real surface the generated surfaces seem to be more smooth at the grooves and peaks. This is due to the reduction of frequency numbers in surface generation. Considering more frequencies in the generation process of the surfaces would lead to a better conformance.

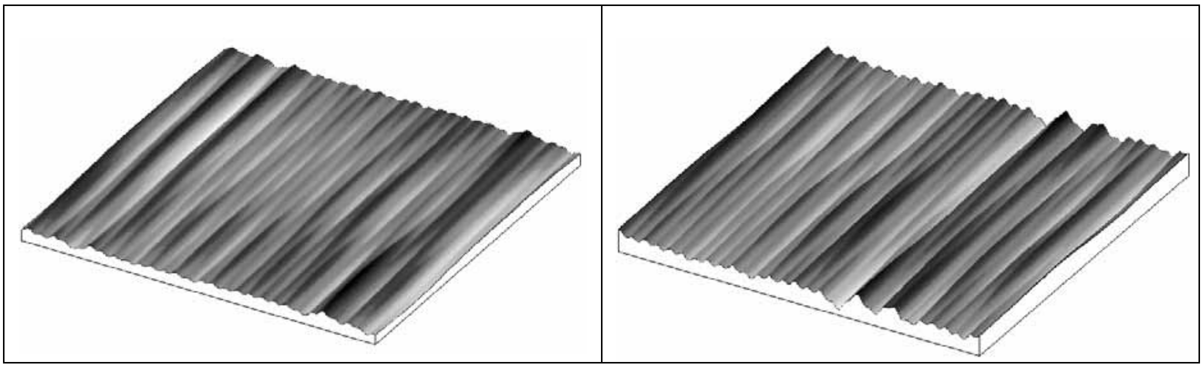

Figure 5. stochastic generated surfaces
left: R_q = 1.5µm
right: R_q = 2.49µm

5 Influence of the friction contact on the system behaviour

The dynamic behaviour of a technical system is constantly in an interaction with the mechanisms of the friction contact. On the one hand the friction value of the friction contact depends on the boundary conditions (forces, moments, speed, acceleration) of the system. On the other hand friction influences the system behaviour like, for instance, frictional vibration in clutch systems. In order to examine this fact thoroughly by means of the simulation method, it is necessary to apply the analysis of both, the Multi-Body-System (MBS) and the Finite-Element-Model (FEM) at the same time. To describe the situation in the friction contact the FEM is used. The MBS, however, describes the system behaviour. Because of the interaction between the system and the contact this procedure is an iterative process. The MBS simulation can be carried out in real time, whereas the contact analysis requires long calculation times because of the many degrees of freedom. It is impossible to link both simulation methods. To solve this task a Lookup-Table, which has been successfully employed in other industrial areas, is used. This solution consists of two steps:

- parameter study of the friction contact
- system analysis

5.1 Parameter study of the friction contact

The aim of the first step is to construct a n-dimensional matrix by means of a parameter study which establishes the friction value in view of relevant parameters. Figure 6 shows this process.

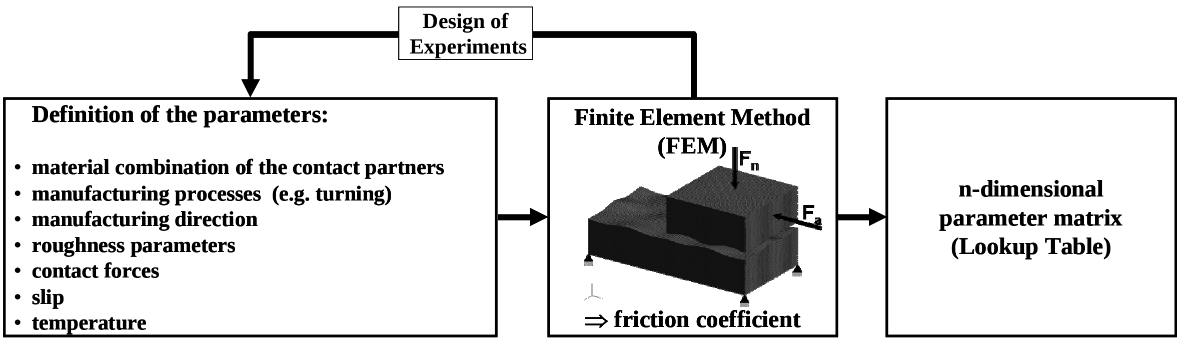

Figure 6. building of a n-dimensional parameter matrix

First, the artificially generated surfaces are imported into a fully parameterized contact model. The transient model consists of two three-dimensional rough solid bodies, which are in contact with each other. An elastic-plastic material behaviour is used as a basis. The nodes of the lower contact partner are fixed underneath its surface. A normal force F_n and an axial force F_a is exerted on the upper contact partner. Generally, it is possible to replace the axial force by a give move in axial direction. To determine the friction value an energy balance is drawn. On the assumption that the friction losses are entirely produced by the deformation process (W_{def}), which can be determined by the energy of all finite elements, the friction value (μ) can be calculated as the function of deformation process, friction path (s) and normal force (F_N).

$$\mu = \frac{W_{def}}{F_N \cdot s} \tag{11}$$

The next step is to choose an amount of parameters which have crucial influence on the situation in the contact. Here, two groups of parameters are distinguished. The first group contains characteristic values, which result from the choice of material and the production:

- mechanical and thermal properties of the materials
- surface profile
- roughness value

The second group contains parameters which directly depend on the dynamic behaviour of the system:

- contact forces
- slip between the contact partners
- temperature

By applying design of experiment (DOE), parameter combinations are established which allow optimal statements and at the same time keep the computational effort relatively down.

5.2 System analysis

The results of the first step are a basis for the MBS. Before the system analysis, the material properties of the contact partners as well as the topography of their contact surfaces must be chosen. In the course of the subsequent analysis, the MBS makes a statement about the forces, moments, speed and acceleration occuring in the system.On the basis of these boundary conditions the friction value is calculated by using the parameter matrix and is later passed on to the Multi-Body-analysis (figure 7).

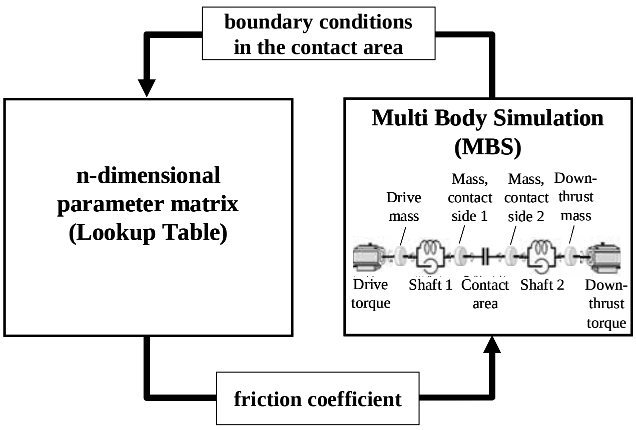

Figure 7. sequence of the system analysis

6 Conclusions and Future Research

The presented work has shown a methodical process which allows to generate technical surfaces with desired characteristics. Characterization of the surfaces is done by the abbot curve and the roughness parameters. For mathematical description the Fourier transformation is used. The generation of the surfaces is done by an evolutionary optimizer. Furthermore, it has been shown how the influence of a friction contact on the system behaviour can be taken into account in a simulation. The methods of the FEM and MBS were used fort this. To link these methods a Lookup-Table was applied.

The presented examples of optimized surfaces have shown that the application of this method is successful. Further work has to be done to determine which and how many frequencies have to be taken into account as design variables to generate surfaces which are even more realistic. An increasing number of design variables has an influence on the convergence behaviour of the optimization process and the computational effort which has to be done. Therefore it has to be investigated how many frequencies of the Fourier spectrum can be taken into account without loosing the efficiency of the method. At IPEK, the method for taking into account the influence of surface topography on the system behaviour is applied in the power train area. Currently, the parameter matrix is filled with data from experiments as well as from simulations. In future, using these methods, statements about design requirements (mass, inertia, stiffness) of a technical system with regard to a desired dynamic behaviour can be made in a very early phase of the product developing process.

References

[1] Albers, A.; Ott, S.; Nowicki, L., „Ganzheitliche Untersuchungsmethode von friktionskontaktinduzierten Schwingungen in Antrieben“ VDI-Berichte Nr. 1786, Kupplungen und Kupplungssysteme, VDI-Verlag, Düsseldorf, 2003, page177-192.

[2] Rometsch, R, Letzner, R-D, „Rauheitsmessung, Theorie und Praxis“, Hommelwerke GmbH, Schnurr Druck, 1993.

[3] DIN EN ISO 13565, „Oberflächenbeschaffenheit: Tastschnittverfahren“, Part 1 till Part 3. Hrsg, Deutsches Institut für Normung. Berlin, Köln: Beuth-Verlag, 1997.

[4] Sundararajan, D, „The Discrete Fourier Transform“, World Scientific Publishing Co. Pte. Ltd., Singapore, New Jersey, London, Hong Kong, 2001.

[5] Kursawe, F, Schwefel, H-P, „Optimierung mit Evolutionären Algorithmen“, Automatisierungstechnische Praxis, vol. 39 nr. 9, 1997, page 10-17.

[6] Rechenberg, I, „Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution“, Frommann-Holzboog, Stuttgart,1973

[7] Rienäcker, A, „Instationäre Elastohydrodynamik von Gleitlagern mit rauen Oberflächen und inverse Bestimmung der Warmkonturen“, Diss. RWTH Aachen, 1995

o. Prof. Dr.-Ing. Dr. h.c. Albert Albers
University of Karlsruhe (TH)
IPEK, Institute of Product Development
Kaiserstr. 12
76131 Karlsruhe, Germany
Tel.: +49 721 608 2371
Fax: +49 721 608 6051
albers@ipek.uni-karlsruhe.de