Lecture Notes in Mechanical Engineering

Károly Jármai
Betti Bolló *Editors*

Vehicle and Automotive Engineering

Proceedings of the JK2016, Miskolc, Hungary

Springer

Lecture Notes in Mechanical Engineering

About this Series

Lecture Notes in Mechanical Engineering (LNME) publishes the latest developments in Mechanical Engineering—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNME. Also considered for publication are monographs, contributed volumes and lecture notes of exceptionally high quality and interest. Volumes published in LNME embrace all aspects, subfields and new challenges of mechanical engineering. Topics in the series include:

- Engineering Design
- Machinery and Machine Elements
- Mechanical Structures and Stress Analysis
- Automotive Engineering
- Engine Technology
- Aerospace Technology and Astronautics
- Nanotechnology and Microengineering
- Control, Robotics, Mechatronics
- MEMS
- Theoretical and Applied Mechanics
- Dynamical Systems, Control
- Fluid Mechanics
- Engineering Thermodynamics, Heat and Mass Transfer
- Manufacturing
- Precision Engineering, Instrumentation, Measurement
- Materials Engineering
- Tribology and Surface Technology

More information about this series at http://www.springer.com/series/11236

Károly Jármai · Betti Bolló
Editors

Vehicle and Automotive Engineering

Proceedings of the JK2016, Miskolc, Hungary

Springer

Editors
Károly Jármai
Miskolci Egyetem
University of Miskolc
Miskolc, Egyetemvaros
Hungary

Betti Bolló
Miskolci Egyetem
University of Miskolc
Miskolc, Egyetemvaros
Hungary

ISSN 2195-4356 ISSN 2195-4364 (electronic)
Lecture Notes in Mechanical Engineering
ISBN 978-3-319-51188-7 ISBN 978-3-319-51189-4 (eBook)
DOI 10.1007/978-3-319-51189-4

Library of Congress Control Number: 2017930169

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The production of car and vehicle industry increased greatly in the past decades. People would like to reach the destination as quickly as possible. The quick transportation of persons and goods is more and more important. This is the case in Hungary, where the improvement of the car industry was great in the past decades. Great car producers settled here like Mercedes Benz, Audi, Suzuki, Opel and also small and medium enterprises connected to car element production have developed greatly.

Education has to follow this trend. Vehicle engineering training has a long tradition in Hungary. At the Budapest Technical University and Economics, at the István Széchenyi University in Győr they have a long-term experience in this kind of training. At the University of Miskolc, which is a successor of the Mining and Metallurgical Academy, the first technical higher educational institution on the Earth, founded in 1735, the mechanical engineering training started in 1949. The industrial demand forced the university to start vehicle engineering training also. It was accredited in 2015 and started this year.

The main requirements for cars and car elements are safety, manufacturability and economy. Safety against different loads such as permanent and variable actions is guaranteed by design constraints on stresses, deformations, stability, fatigue, eigenfrequency, while manufacturability is considered by fabrication constraints. The economy is achieved by minimization of the cost.

The main topics of the conference are as follows:

Design: Acoustic investigations, Car electronics, Autonomic vehicles, Fatigue, Industrial applications, Vehicle Powertrain, Modelling and simulation of vehicle informatics and electronic systems, Vehicle navigation, Visual systems of vehicles, Mechatronics, Numerical methods FEM and BEM applications, Vibration and damping, Stability calculations, Structural materials, Structural safety, Structural connections, Analysis and design of structural elements, Design guides, Fracture mechanics, Thin walled structures, Driver assist systems, Hybrid and electric cars.

Fabrication: Forming technologies, Surface protection, Production logistics, Manufacturing technologies, Welding technologies, Heat treatment, Innovative casting technologies, Industrial applications, Maintenance, Environmental protection, Lean technologies, Quality assurance, Gluing technologies, Production, Testing.

Economy: Life cycle assessment, Fabrication costs, Industrial applications, Cost engineering, Structural optimization.

Education: Vehicle engineering training, Dual training, Industrial practice, Training techniques, Training materials.

It is a great pleasure to organize this conference, to give participants an opportunity to show and discuss the new research results in a friendly atmosphere.

The organizers wish all participants successful days to collect new ideas and get new acquaintances.

Miskolc, Egyetemvaros, Hungary
October 2016

Károly Jármai
Betti Bolló

Acknowledgements

The editors would like to acknowledge the co-operation and help of the following organizations

- *Hungarian Vehicle Producers Association (MAJOSZ),*
- *Hungarian Welding Association (MAHEG),*
- *Hungarian Welding Technology and Material Testing Association (MHtE),*
- Foundation for the Development of the Education at the University of Miskolc,
- TÁMOP 4.2.1.B-10/2/KONV-2010-0001 entitled "Increasing the quality of higher education through the development of research—development and innovation program at the University of Miskolc supported by the European Union, co-financed by the European Social Fund."
- *The Mayor and vice mayors of town Miskolc,*
- *Hungarian Trade and Industrial Chamber in Borsod county (BOKIK),*

and last but not least the
University of Miskolc, Hungary, which hosts the conference.

The editors would like to acknowledge the help of the following persons:

László Kota, assistant professor,
Máté Petrik, Ph.D. student,
Éva Engelberth, leader of the dean's office of the faculty of Mechanical Engineering and Informatics.

October 2016

Károly Jármai
Betti Bolló

Contents

About the Editors

Dr. Károly Jármai is Professor at the Faculty of Mechanical Engineering at the University of Miskolc, where he graduated as a mechanical engineer and received his doctorate (dr.univ.) in 1979. He teaches design of steel structures, welded structures, composite structures and optimization in Hungarian and in the English language for foreign students. His research interest includes structural optimization, mathematical programming techniques and expert systems. Dr. Jármai wrote his C.Sc. (Ph.D.) dissertation at the Hungarian Academy of Science in 1988, became a European Engineer (Eur. Ing. FEANI, Paris) in 1990 and did his habilitation (dr.habil.) at Miskolc in 1995. Having successfully defended his doctor of technical science thesis (D.Sc.) in 1995, he subsequently received awards from the Engineering for Peace Foundation in 1997 and a scholarship as Széchenyi professor between the years 1997–2000. He is the co-author (with József Farkas) of four books in English *Analysis and Optimum Design of Metal Structures*, *Economic Design of Metal Structures, Design and optimization of metal structures, Optimum design of steel structure,* and three monographs in Hungarian, and has published over 610 professional papers, lecture notes, textbook chapters and conference papers. He has about 770 independent citations. He is a founding member of ISSMO (International Society for Structural and Multidisciplinary Optimization), a Hungarian delegate, vice chairman of commission XV and a subcommission chairman XV-F of IIW (International Institute of Welding). He has held several leading positions in GTE (Hungarian Scientific Society of Mechanical Engineers) and has been the president of this society at the University of Miskolc since 1991. He was a visiting researcher at Chalmers University of Technology in Sweden in 1991, visiting professor at Osaka University in 1996–1997, at the National University of Singapore in 1998 and at the University of Pretoria several times between 2000–2005.

Dr. Betti Bolló is Associate Professor at the Department of Fluid and Heat Engineering, University of Miskolc, Hungary. She received her M.Sc. degree from the University of Miskolc in Information Engineering (Systems of Power Engineering) in 2003. Her research interests include computational fluid dynamics and internal combustion engines. She wrote her dissertation (Ph.D.) at the Hungarian Academy of Science in 2013. The theme of her dissertation is a numerical investigation of flow past and heat transfer from a heated circular cylinder.

Part I
Design

Investigation of Rolling Element Bearings Using Time Domain Features

Dániel Tóth, Attila Szilágyi and György Takács

Abstract Rolling element bearings can be found widely in domestic and industrial applications. They are important components of most machinery and their working conditions influence the operation of the entire machinery directly. Bearing failures may cause machine breakdown and might even lead to catastrophic failure or even human injuries. In order to prevent unexpected events, bearing failures should be detected as early as possible. Different methods are used for the detection and diagnosis of bearing defects. These techniques can be classified as noise analysis, acoustic measurements, wear debris detection, temperature monitoring, vibration analysis etc. Vibration signals collected from bearings carry detailed information on machine health conditions. This paper deals with a bearing test procedure which based on vibration analysis.

1 Introduction

Vibration monitoring is one of the essential tool that allows to determine the mechanical health of different components in a machine. When the assessment of a ball bearing is performed by vibration analysis, several signal processing techniques can be considered. These techniques can be performed within either the time or the frequency ranges. Among these methods the time domain features are the most appropriate with random signals, where other signal analysis methods are not suitable. These methods facilitate fast data processing and computation. Numerous time domain statistical parameters have been used as trend parameters to detect the

D. Tóth (✉) · A. Szilágyi · G. Takács
University of Miskolc, Miskolc, Hungary
e-mail: toth.daniel@uni-miskolc.hu

A. Szilágyi
e-mail: szilagyi.attila@uni-miskolc.hu

G. Takács
e-mail: takacs.gyorgy@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_1

bearing failures. The most frequently applied stochastic features are the root-mean-square (RMS) value, peak value, skewness, impulse factor, shape factor, clearance factor, crest factor and kurtosis [1, 2].

2 Bearing Test Device

Rolling element bearing condition monitoring can be accomplished by using a test instrument. Such a device is located at University of Miskolc, Department of Machine Tools. The test device is illustrated in Fig. 1.

The equipment is suitable for performing the bearing fatigue and measurement investigations. The particular symbols have the following meanings:

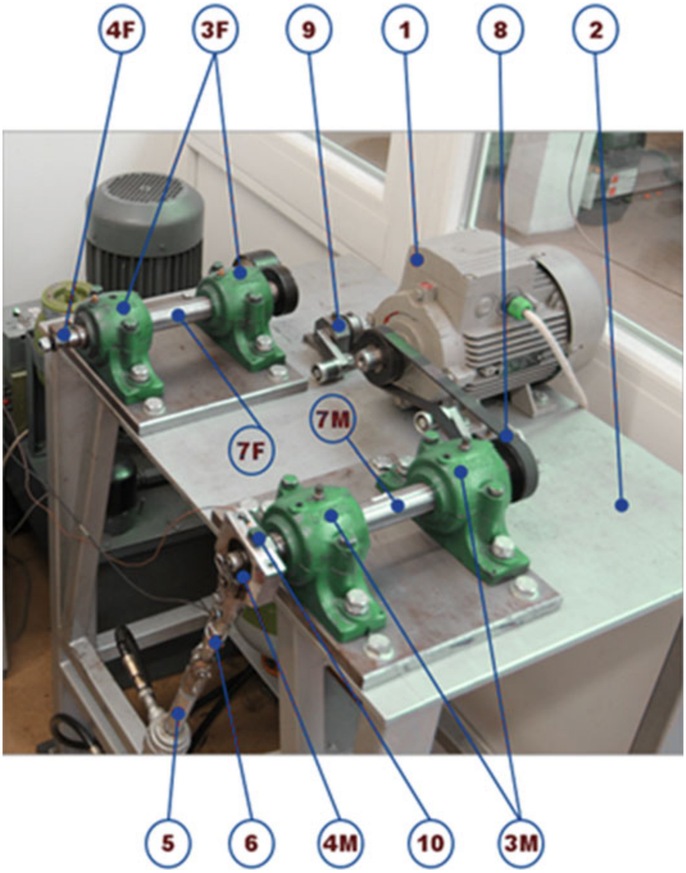


Fig. 1 Experimental test rig

- 1: three-phase motor,
- 2: rigid table,
- 3F: supporting bearings of fatigue side,
- 3M: special supporting plain bearings of measurement side,
- 4F: fatigued bearing position,
- 4M: measured bearing,
- 5: double-acting hydraulic cylinder,
- 6: load cell, the adjustment of hydraulic load,
- 7F: fatigue test shaft,
- 7M: measurement test shaft,
- 8: length ribbed belt,
- 9: belt tensioner,
- 10: piezoelectric vibration accelerometer.

During the measurements the "7M" shaft works at the given rotational speed (1500 $\min^{-1}$), while the "6" hydraulic cylinder exerts artificial load (1 kN) for the "4M" bearing.

3 Description of Investigation

Fundamentally, two proceedings are used for the experimental analysis of rolling element bearings. One method is the fatigue test when the bearings operate until they get permanent damage, and we measure their vibration trends meanwhile. However, the process takes relatively long time, but it can be accelerated with the bearing overload and increased rotational speed. Another technique is the production of one or more artificial failure of the elements of bearings. In this case the vibration signal should be measured and compared to data of faultless bearings. According to the literature [3–5], generally this may use methods such as spark erosion, acid etching, scratching or mechanical indentation. In this research, we used a well reproducible method to create artificial faults. A Rockwell hardness tester applied to make defects to the inner ring of bearings. This method needs a bearing with plastic cage, because it should be disassemble and assemble non-destructively. Figure 2 shows the ball bearing type 6303 which used during experiments.

Fig. 2 The test bearing and the artificial defect on the inner ring

As it was written previously a Rockwell hardness tester is used to cause local defects. The type of this machine is HR—150A. It is suitable for examining the effects of three types of loads. The major load values are 60, 100 and 150 kg. The effects of each loading were examined more than 10 times.

Optical examination can be applied to measure failure size. Polarising microscope is widely used for higher resolution. Carl Zeiss Jenavert polarising microscope is applied to inspect the extent of the defect.

The average extent of the failure is 265 μm in diameter in case of 60 kg load, 411 μm in diameter in case of 100 kg load and 478 μm in diameter in case of 150 kg load. The following illustrations show the effects of different loads (Fig. 3).

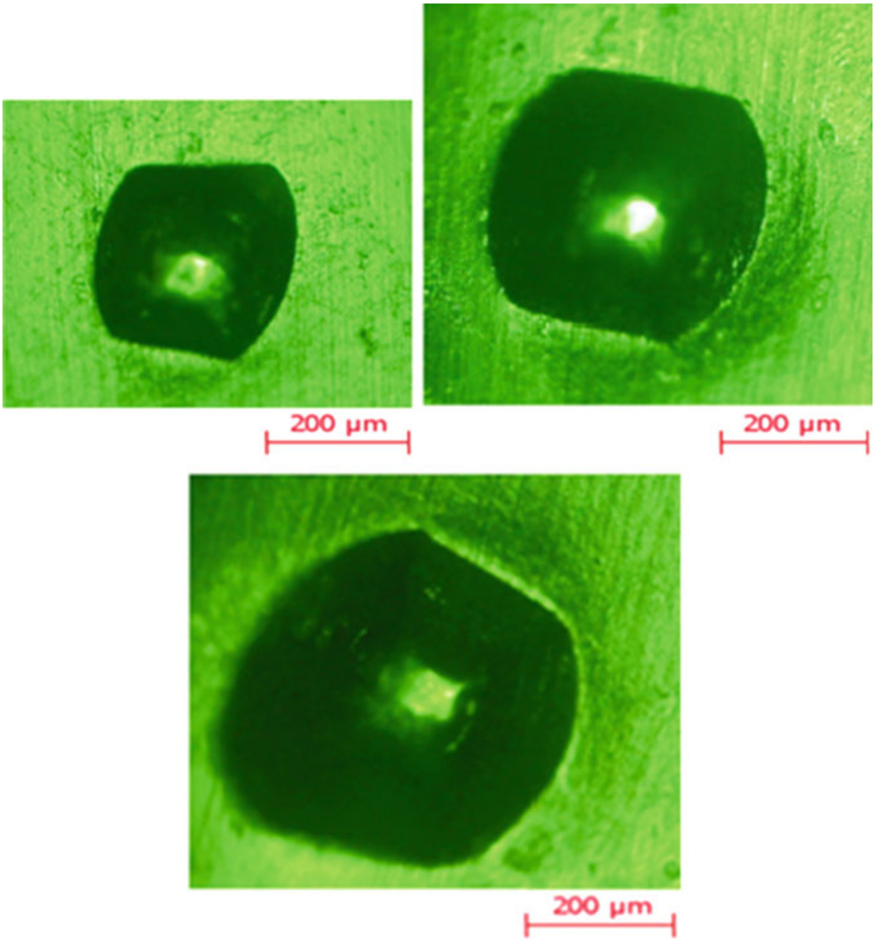


Fig. 3 Inner ring defects in case of 60 kg, 100 kg and 150 kg load (15 times magnification)

4 Analysis of Measurements

During the experiment, first of all the vibration patterns were measured from the examined bearing using piezoelectric vibration accelerometer (the type of it is Kistler 8632C50). After that the artificial defect was created and vibration patterns were measured again. It is followed by time-domain tests during which statistical features have been calculated. These stochastic indexes can be calculated by using the formulas below (Fig. 4).

The measurement cycles are performed at 9.6 kHz sampling frequency. Five vibration samples and 16,384-element samples were taken within each cycle. Statistical features were calculated based on sampled values. These parameters were computed by a program code, which runs in Maple mathematical software. Table 1 contains the statistical parameters in case of 60 kg load.

Table 2 includes the stochastic features in case of 100 kg load.

It is visible that the most of the parameters have doubled under this load. Table 3 contains the statistical parameters in instance of 150 kg load.

It is clearly visible that the statistical parameters of a defective bearing tend to be higher than the values of a normal bearing. The percentage increase is depicted in Fig. 5.

According to the graph it is clear that the Standard deviation, the Peak value and the RMS were the most sensitive to this artificial error. Nevertheless, it is obvious that the Kurtosis and the Skewness also have good correlation.

Fig. 4 Calculation of stochastic features [1]

Peak Value	$\frac{1}{2}\left(max(x_i) - min(x_i)\right)$
Root mean square	$\sqrt{\frac{\sum_{i=1}^{N}(x_i)^2}{N}}$
Crest Factor	$\frac{\text{Peak Value}}{\text{RMS}}$
Skewness	$\frac{\sum_{i=1}^{N}(X_i - \bar{X})^3}{(N-1)S^3}$
Kurtosis	$\frac{\sum_{i=1}^{N}(X_i - \bar{X})^4}{(N-1)S^4}$
Standard deviation	$\sqrt{\frac{\sum_{i=1}^{N}(x_i - x_m)^2}{N-1}}$

Table 1 Statistical parameters in new and damaged conditions (60 kg load)

	Peak value	RMS	Kurtosis	Skewness	Standard deviation		Peak value	RMS	Kurtosis	Skewness	Standard deviation
1	0.5688	0.0761	3.9468	1.0173	0.0461	1	0.6569	0.0792	4.3318	1.0370	0.0480
2	0.6150	0.0813	3.8500	1.1566	0.0485	2	0.6743	0.0831	4.1450	1.1802	0.0491
3	0.5632	0.0784	3.9200	0.9854	0.0463	3	0.6242	0.0790	4.4052	1.0209	0.0475
4	0.6369	0.0759	4.2139	1.0142	0.0449	4	0.6840	0.0784	4.7605	1.0403	0.0467
5	0.6807	0.0730	4.3995	1.0921	0.0442	5	0.7643	0.0802	4.8643	1.3295	0.0480
6	0.5376	0.0746	3.9534	1.0352	0.0523	6	0.6421	0.0755	4.5431	1.0723	0.0542
7	0.6712	0.0865	3.7533	0.8960	0.0508	7	0.7823	0.0921	4.0854	0.9302	0.0546
8	0.5114	0.0864	3.9900	1.0142	0.0441	8	0.6432	0.0887	4.3586	1.0274	0.0452
9	0.5255	0.0976	3.9631	1.0532	0.0553	9	0.6422	0.0992	4.2961	1.0689	0.0568
10	0.5931	0.0867	3.4861	0.8991	0.0521	10	0.7053	0.0888	3.8716	0.9582	0.0529
11	0.6274	0.0974	3.8255	1.0809	0.0580	11	0.7255	0.1032	4.2016	1.0904	0.0583

Table 2 Statistical features in new and damaged conditions (100 kg load)

	Peak value	RMS	Kurtosis	Skewness	Standard deviation		Peak value	RMS	Kurtosis	Skewness	Standard deviation
1	0.6225	0.0827	3.9236	1.1359	0.0477	1	1.1337	0.1482	5.2670	1.6243	0.0906
2	0.6508	0.0776	4.6827	1.1124	0.0471	2	1.1450	0.1425	6.1030	1.4740	0.0884
3	0.7425	0.0806	4.9360	1.1823	0.0495	3	1.3751	0.1461	6.7832	1.7498	0.0920
4	0.6674	0.0701	4.7007	1.3366	0.0435	4	1.1821	0.1260	6.5078	1.8236	0.0778
5	0.6150	0.0813	3.8500	1.2566	0.0485	5	1.0752	0.1487	4.7519	1.6767	0.0861
6	0.6569	0.0779	4.3318	1.0370	0.0463	6	1.1550	0.1385	5.2563	1.3732	0.0820
7	0.5618	0.0628	4.8172	1.2033	0.0385	7	1.0103	0.1120	6.2307	1.6778	0.0728
8	0.5970	0.0819	3.8334	0.9852	0.0488	8	1.0275	0.1456	4.9583	1.3146	0.0873
9	0.6475	0.0757	4.3553	1.0751	0.0459	9	1.1539	0.1358	5.5933	1.4559	0.0848
10	0.6527	0.0763	4.4994	1.0673	0.0455	10	1.1633	0.1360	5.6846	1.4334	0.0832
11	0.6705	0.0782	4.5762	1.2543	0.0484	11	1.1949	0.1366	5.8274	1.7112	0.0915

Table 3 Statistical indexes in new and damaged conditions (150 kg load)

	Peak value	RMS	Kurtosis	Skipness	Standard deviation		Peak value	RMS	Kurtosis	Skewness	Standard deviation
1	0.5824	0.0681	4.0148	0.9524	0.0406	1	1.5039	0.1683	6.0793	1.5524	0.1133
2	0.6386	0.0747	4.2957	1.0603	0.0453	2	1.4868	0.1816	6.5047	1.6997	0.1244
3	0.6777	0.0847	4.3517	1.1184	0.0554	3	1.5195	0.2044	6.5895	1.8040	0.1561
4	0.5527	0.0644	3.7721	1.0339	0.0428	4	1.4658	0.1561	6.0891	1.6297	0.1203
5	0.6292	0.0703	4.9252	1.3477	0.0431	5	1.5429	0.1570	7.3456	2.1513	0.1215
6	0.5875	0.0696	3.8739	0.9273	0.0415	6	1.4876	0.1402	6.1514	1.4710	0.1135
7	0.7160	0.0914	4.4818	1.4629	0.0564	7	1.5337	0.1945	6.5947	2.2913	0.1563
8	0.6309	0.0663	4.4436	1.2635	0.0411	8.	1.4771	0.1474	6.8939	2.0170	0.1122
9	0.6043	0.0839	4.4436	1.1636	0.0588	9	1.3790	0.1957	7.0717	1.8737	0.1592
10	0.5925	0.0825	4.1238	1.0366	0.0487	10	1.3913	0.1752	6.8102	1.6496	0.1334
11	0.6391	0.0855	4.4342	1.1429	0.0518	11	1.4432	0.1944	7.4114	1.8093	0.1425

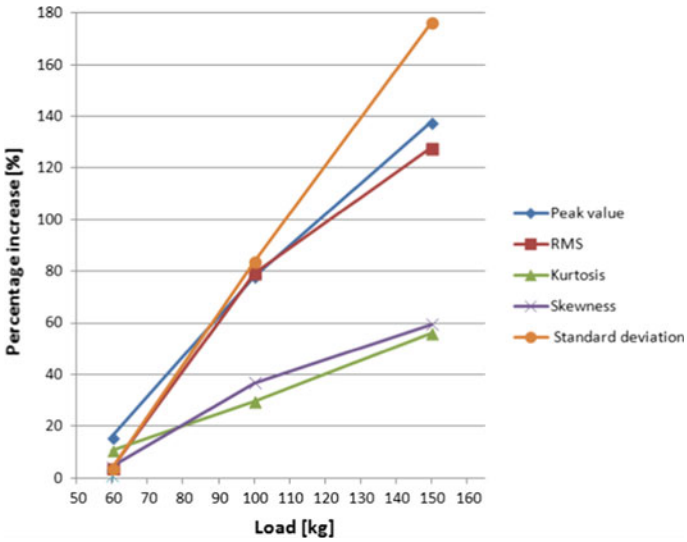


Fig. 5 The percentage increase of loads

5 Conclusion

Trustworthy and accurate measuring methods and devices are inevitable for rotary and bearing condition monitoring. The investigation of vibration signals is a significant technique for monitoring the condition of machine components. Stochastic parameters are widely used as features in failure diagnostics. Present paper shows that the time domain techniques can be effectively used in condition monitoring and fault diagnosis of ball bearings. These methods are reliable tools and they make possible fast data processing.

Acknowledgements This research was supported by the ÚNKP-16-3 New National Excellence Program of the Ministry of Human Capacities.

References

1. Patel J, Patel V, Patel A (2013) Fault diagnostics of rolling bearing based on improve time and frequency domain features using artificial neural networks. IJSRD 1(4)
2. Patidar S, Soni PK (2013) An overview on vibration analysis techniques for the diagnosis of rolling element bearing faults. IJETT 2013

3. Kharche PP, Kshirsagar SV (2014) Review of fault detection in rolling element bearing. IJIRAE 1(5)
4. Patkó Gy, Takács Gy, Demeter P, Barna B, Hegedűs Gy, Barak A, Simon G, Szilágyi A (2010) A process for establishing the remanent lifetime of rolling element bearings. In: XXIV microCAD International Scientific Conference, Miskolc (Hungary), March 2010
5. Howard I (1994) A review of rolling element bearing vibration detection, diagnosis and prognosis. DSTO-RR-0013

Truck Floor Design for Minimum Mass and Cost Using Different Materials

Károly Jármai and József Farkas

Abstract In the chapter the floor structure of a truck produced by a company in Hungary has been investigated. The structure consists of steel members, or extruded Al-alloy longitudinal and cross members as well as a tread deck plate. Using an optimum design process, namely the Hillclimb optimizer, significant mass and cost savings may be achieved by decreasing the deck plate thickness and changing the profile, dimensions and number of cross members. Comparison is made using the combination of the steel and aluminium, or using only steel alone. Design constraints relate to fatigue stress range of welded joints, to local buckling of extruded or normal profiles and to fabrication size limitations. A special loading case is also considered when a wheel is staying on a curb and the floor is distorted.

1 Introduction

There are some trucks for beverage transport, where the truck structure has a steel chassis consisting of two longitudinal beams. The subframe is constructed from two longitudinal beams bolted on steel beams. They can be made from *Al*-alloys, or structural steel. The Al-alloy floor structure has three layers as follows (Fig. 1): cross members welded to subframe, the longitudinal members welded to cross members, tread deck plate distributing the pallet loads. The material of cross members is an *Al*-alloy AlMgSi0.7 according to German standard DIN 1725 [1] of $R_{p.0.2}$ = 215 MPa according to DIN 1748 [2] (international alloy type 6005A). The tread deck plate material is an *Al*-alloy AlMg2.5 (international alloy type 5052). These main structural parts are framed by side rails, which carry the loads from

K. Jármai (✉) · J. Farkas (Deceased)
University of Miskolc, Miskolc, Hungary
e-mail: jarmai@uni-miskolc.hu

J. Farkas
e-mail: altfar@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_2

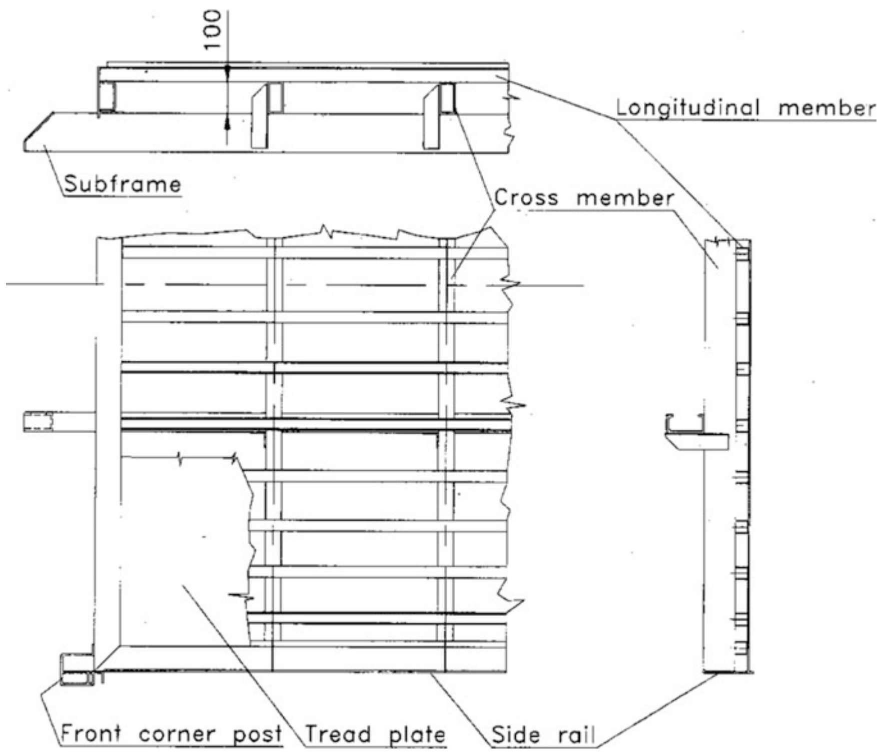


Fig. 1 Truck floor structure

roof, sidewalls and doors. We have made an optimization using aluminium, or normal steel in the floor structure. Due to the fact that the fatigue limit for the steel at Eurocode 3 up to 690 MPa and at IIW recommendation up to 960 MPa does not change, it does not worth to use higher strength steels, only normal structural steel.

Our aim is to decrease the material cost of the floor structure by changing the profile, dimensions and number of cross members, the thickness of deck plate as well as the material grades.

2 Load Cases

2.1 Loads in the Horizontal Floor Position

Two load cases should be considered in the design of cross members as follows: (a) loads due to pallets, roof, door and side walls in the horizontal floor position; (b) the same loading as in (a) but a wheel is staying on a curb, thus, the floor is distorted.

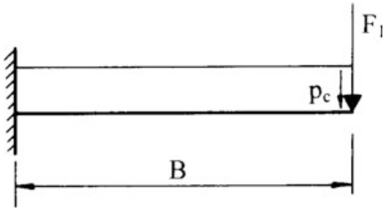


Fig. 2 Loads on the cantilever part of cross members

Loads acting on an outside cross member are as follows:

a corner column		205 N
roof	2060/4	515 N
upper door	1420/2	710 N
front wall	1033/2	516 N
		F_1 = 1946 N

Load from pallets: mass of a pallet is F_p = 8500 N, intensity of the uniformly distributed load is $p = F_p n_p/(BL)$, where the number of pallets placed on the half floor n_p = 5, B and L are the dimensions of a half cantilever floor surface. The uniformly distributed normal load acting on a cross member is $p_c = pL/(n_c{-}1)$, n_c is the number of cross members.

The maximum bending moment in a cross member is (Fig. 2)

$$M_{\max} = \frac{p_c B^2}{2} + F_1 B = \frac{F_p n_p B}{2(n_c - 1)} + F_1 B \tag{1}$$

Calculating with F_p = 8500 N, n_p = 5, B = 720 mm, F_1 = 1946 N one obtains bending moments for different numbers of cross members. This number is limited by the dimension of pallets (800 mm) to $n_{c.min}$ = 10. Since the original number of cross members is 14, we calculate with n_c = 14, 12 and 10. For these values of n_c one obtains

$$M_{14} = 2.578, M_{12} = 2.792 \,\text{and}\, M_{10} = 3.1011\,\text{kNm}.$$

The corresponding shear forces are as follows:

$$Q = F_p n_p/(n_c - 1) + F_1;\; Q_{14} = 5215,\; Q_{12} = 5810 \text{ and } Q_{10} = 6668\,\text{N}.$$

2.2 Loads on the Distorted Floor

Measurements have been carried out on a truck loaded with pallets and with a wheel staying on a curb in a height of 91 mm. The measured deflections have

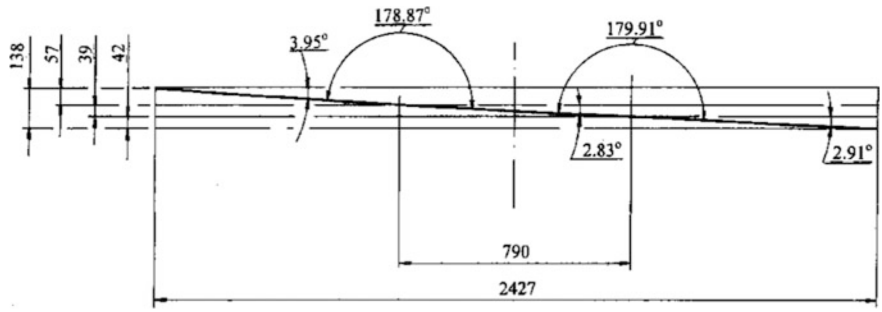


Fig. 3 Measured deflections of a distorted cross member, when a left truck wheel is staying on a curb

shown that the cross members near the wheel being lifted up are loaded by bending as it is seen on Fig. 3. This cross member can be modelled as a cantilever beam of its whole length L_c loaded by a force F corresponding to a deflection w. This deflection can be approximately calculated as $w = 138 - L_c\varphi$, where L_c = 2427 mm, $\varphi(rad) = 2.91°\pi/180° = 0.0508$, thus, w = 15 mm. Furthermore

$$F = \frac{3EI_x w}{L_c^3}; M_{c.\max} = FL_c \tag{2}$$

where E = 7 x 10^4 MPa is the elastic modulus of aluminium, E = 2.1 x 10^5 MPa for steel, I_x is the second moment of area.

3 Geometric Characteristics of Cross Members

The cross-section loaded by bending and shear consists of a cross member and a part of the deck plate (Fig. 4). We calculate an effective width of the deck plate 50t, t is the thickness. In the case of a rectangular hollow section (RHS) the geometric characteristics of this cross section are as follows [3]:

$$A = A_1 + A_2;\ A_1 = 2ht_w + 2bt_f;\ A_2 = 50t^2 \tag{3}$$

$$y_G = \frac{A_1}{A}\left(\frac{h+t}{2} + c\right); \quad y_c = h + c + \frac{t}{2} - y_G \tag{4}$$

$$I_x = \frac{h^3 t_w}{6} + \frac{bt_f h^2}{2} + A_1\left(y_c - \frac{h}{2}\right)^2 + A_2 y_G^2 \tag{5}$$

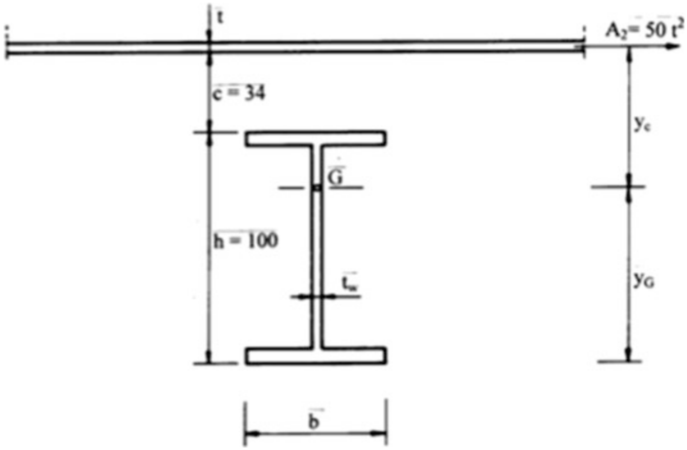


Fig. 4 Cross-sections of cross members

In the case of I-profile (Fig. 4) the characteristics are as follows:

$$A_1 = ht_w + 2bt_f \tag{6}$$

$$I_x = \frac{h^3 t_w}{12} + \frac{bt_f h^2}{2} + A_1\left(y_c - \frac{h}{2}\right)^2 + A_2 y_G^2 \tag{7}$$

In our previous calculations [4] we have made comparisons using the rectangular hollow section, I- and C-profiles. It was found that the best cross section is the I-beam. That is the reason why the I-profile has been chosen.

4 Design Constraints

4.1 Constraints on Fatigue Stress Range for the Horizontal Floor Position

$$\sigma_1 = \frac{M_{\max}}{I_x} y_{\max} \le \frac{\Delta\sigma_N}{\gamma_{Mf}}; \quad y_{\max} = \max(y_G, y_c) \tag{8}$$

$$\tau_1 = \frac{Q}{A_w} \le \frac{\Delta\tau_N}{\gamma_{Mf}}; \tag{9}$$

where $A_w = ht_w$ for I-profile.

Since the cross members are welded to longitudinal subframe beams, they should be designed considering the fatigue of welded joints. According to Hobbache [5] the fatigue stress range for number of cycles 2 x 10^6 in the case of transverse stiffener welded on girder web (detail 512 for structural aluminium alloys) is $\Delta\sigma_C = 28$ MPa. Calculating with a realistic number of cycles $N = 2 \times 10^5$,

$$\log\Delta\sigma_N = \frac{1}{3}\log\frac{2x10^6}{2x10^5} + \log\Delta\sigma_C = 1.78049; \quad \Delta\sigma_N = 60.3\,\text{MPa} \tag{10}$$

For steel $\Delta\sigma_C = 80\,\text{MPa}$ (detail 512 for structural steel, the same as for Al) $\Delta\sigma_N = 172.3\,\text{MPa}$. With a safety factor of 1.25.

For aluminium

$$\frac{\Delta\sigma_N}{\gamma_{Mf}} = \frac{60.3}{1.25} = 48.2\,\text{MPa} \tag{11}$$

For shear it is

$$\Delta\tau_C = 28;\ \Delta\tau_N = 44.3; \frac{\Delta\tau_N}{\gamma_{Mf}} = \frac{44.3}{1.25} = 35.44\,\text{MPa} \tag{12}$$

For steel

$$\frac{\Delta\sigma_N}{\gamma_{Mf}} = \frac{172.3}{1.25} = 137.8\,\text{MPa} \tag{13}$$

For shear it is

$$\Delta\tau_C = 80;\ \Delta\tau_N = 126.8; \frac{\Delta\tau_N}{\gamma_{Mf}} = \frac{126.8}{1.25} = 101.44\,\text{MPa} \tag{14}$$

It should be mentioned that we calculate with the bending moment also from static load F_1 in the fatigue constraint as an approximation in the safe side.

4.2 Constraint on Fatigue Stress Range for the Distorted Floor Position

$$\sigma_2 = \frac{M_{c.\max}}{I_x} y_{\max} = \frac{3Ew}{L^2} y_{\max} \le \frac{\Delta\sigma_{N1}}{\gamma_{Mf}} \tag{15}$$

In the case of distorted floor position the maximum bending moment arises at the end of the cross member, where it is welded to subframe by fillet welds. For this joint, according to [5] (detail No.413) $\Delta\sigma_{C1} = 22$ MPa and a realistic number of cycles $N = 10^5$ it is

$$\frac{\Delta\sigma_{N1}}{\gamma_{Mf}} = \frac{59.7}{1.25} = 47.7\,\text{MPa} \tag{16}$$

4.3 Constraints on Local Buckling of Profiles

Web of I-section (unreinforced)

For aluminium

$$h/t_w \leq 22\varepsilon/g; \tag{17}$$

For steel [6]

$$h/t_w \leq 69\varepsilon/g; \tag{18}$$

$$\begin{aligned} g &= 0.65 + 0.35\tfrac{y_0}{y_c} \quad \text{when} \quad 1 \geq \tfrac{y_0}{y_c} \geq 0 \\ g &= 0.65 + 0.30\tfrac{y_0}{y_c} \quad \text{when} \quad 0 \geq \tfrac{y_0}{y_c} \geq -1 \end{aligned}$$

$$y_0 = y_G - \frac{t}{2} - c \tag{19}$$

$$y_c = h + c + \frac{t}{2} - y_G \tag{20}$$

Flange of I-section (unreinforced)

For aluminium alloy

$$b/t_f \leq 14\varepsilon \tag{21}$$

For steel

$$b/t_f \leq 28\varepsilon \tag{22}$$

For aluminium

$$\varepsilon = \sqrt{\frac{250}{\sigma_{\max}/\gamma_{M1}}}, \text{ for steel } \varepsilon = \sqrt{\frac{235}{\sigma_{\max}/\gamma_{M1}}} \tag{23}$$

4.4 Fabrication Constraints: Size Limitations

Some constant dimensions are prescribed by the original structure as follows:

$$h = 100, \quad c = 34\,\text{mm} \tag{24}$$

The web thickness is limited to

$$t_{w.min} = 3.4\,\text{mm} \tag{25}$$

to guarantee the quality of welding.

The tread plate thickness is limited to

$$t_{min} = 2\,\text{mm} \tag{26}$$

Since the cross members should be welded to side rails, the extruded shapes should not have any reinforcing ribs or bulbs, since they are in the way of welding.

It should be mentioned that the extruded I-profiles with or without reinforcing ribs or bulbs optimized for pure bending have the same minimum cross-section area, thus, the use of ribs or bulbs does not result in mass savings.

5 Optimization Characteristics and Results

The objective function is the cross-sectional area of the cross members and deck plate part (Eq. 3).

The unknown variables are the dimensions of profile flanges b and t_f.

The constraints are as follows: Eqs. 11, 12, 13, 14, 15, 21, 22, 24, 25, 26.

The optimization is performed for *I-profile* and for *three numbers of cross members* n_c = 14, 12 and 10.

Mathematical method: the Rosenbrock's Hillclimb algorithm is used [7].

Results are summarized in Table 1 and 2.

Table 1 Optimum sizes of the frame

	Original aluminium structure	Optimized aluminium	Optimized steel
Structural weight in kg	304.48	172.31	239.7
Structural cost in USD	927.44	550.27	247.40

Table 2 Result of the optimization

Profile		n_c = 14	n_c = 12	n_c = 10
Aluminium I-profile	b	55	60	65
	t_f	7.2	7.2	7.8
	A_1	1332	1404	1536
	m_c kg	104.41	95.18	89.20
	K_T $	927	927	927
	k_T $	0.54	0.63	0.795
	K_c $	304.74	291.78	260.64
Steel I-profile	b	30	30	35
	t_f	2	2	2.1
	A_1	660	660	747
	m_c kg	123.35	105.73	93.28
	K_T $	187	187	187
	k_T $	0.109	0.128	0.153
	K_c $	269.77	252.15	247.4

6 Mass Savings

The mass of the original tread plate of thickness t = 4.5 mm and dimensions 2280 × 6750 mm, taking the density of Al-alloy $\rho = 2.7 \times 10^{-6}$ kg/mm^3, for steel $\rho = 7.85 \times 10^{-6}$ kg/mm^3

$$m_{pl} = \rho t(2280 \times 6750) = 186.98\ \text{kg}.$$

For aluminium

Mass of the optimized Al plate of t = 2.0 mm is $m_{pl.opt}$ = 83.11 kg.

The mass of Al cross members can be calculated as

$$m_c = \rho A_1 n_c L_{cm}$$

where L_{cm} = 2440 mm is the length of a cross member. The calculated masses are shown in Table 2.

The original mass of the tread plate and cross members is

$$m = m_{pl} + m_c = 186.98 + 117.50 = 304.48\ \text{kg}.$$

The mass of the optimum Al solution is m_{min} = 83.11 + 89.20 = 172.31 kg, the mass saving is 132.17 kg for one truck (43%).

For the steel

Mass of the optimized steel plate of t = 2.0 mm is $m_{pl.opt}$ = 146.424 kg.

The mass of steel cross members can be calculated as

$$m_c = \rho A_1 n_c L_{cm}$$

The mass of the steel tread plate and cross members is

$$m = m_{pl} + m_c = 146.42 + 93.28 = 239.7\,\text{kg}.$$

7 Cost Savings

For aluminium

Cost of tread deck plate

London Metal Exchange (LME) price of aluminium [8]	1.559 \$/kg
surcharge	0.9568 \$/kg
total	2.5159 \$/kg

Cost of cross members

Cost of the original plate (t = 4.5) 186.98 × 2.5159	470.44 \$
Cost of the optimized plate (t = 2 mm) 83.11 × 2.5159	209.09 \$

Cost of cross members

LME aluminium total	2.5159 \$/kg
extrusion work upcharge	1.3004 \$/kg
total	k_c = 3.8163 \$/kg

The total cost, including the proportional tool cost can be expressed as

$$K_T = k_c m_c + k_T$$

where

$$k_T = \frac{K_T}{50 n_c L_{cm}}$$

K_T is the tool cost, $50n_cL_{cm}$ is the total length of extruded bars for 50 trucks (one year production).

The results of the calculations are shown in Table 2.

The total cost of the original deck plate and cross members is $K = 470.44 + 457.00 = 927.44$ $

and that of the optimum (10 cross members of I-profile) $K_{min} = 209.09 + 341.18 = 550.27$ $

The cost savings for one *Al* truck is 377.17 $ (39%)

For steel

Cost of tread deck plate

London Metal Exchange (LME) price of steel	0.375 $/kg
surcharge	0.231 $/kg
total	0.606 $/kg

Cost of the optimized steel plate (t = 2 mm) 241.623 × 0.606 146.424 $

Cost of cross members

LME steel	0.606 $/kg
cutting and welding costs	0.475 $/kg
total	k_c = 1.081 $/kg

The total cost, including the proportional tool cost can be expressed as

$$K_T = k_c m_c + k_T$$

where

$$k_T = \frac{K_T}{50n_cL_{cm}}$$

K_T is the tool cost, $50n_cL_{cm}$ is the total length of welded bars for 50 trucks (one year production).

The results of calculations are shown in Table 2.

The total cost of *the Al deck plate* optimum (10 cross members of I-profile) $K_{min} = 209.09 + 341.18 = 550.27$ $

and that of the optimum (10 cross members of welded I-profile) $K_{min} = 146.42 + 100.98 = 247.40$ $

The cost savings for one steel truck is 302.87 $ (55%).

8 Conclusions

In the case of a truck floor welded from *Al* alloy extruded profiles and a deck plate, the systematic optimum design process can result in significant savings in mass and cost compared to the original design. A cross-section is optimized consisting of an extruded cross member and an effective part of the deck plate. The objective function is the cross-sectional area, the design constraints relate to a fatigue stress range of welded joints and local buckling of extruded profiles. Fabrication aspects regarding the size limitations are also considered.

In addition to the loading by pallets in horizontal floor position the case of distorted floor position is also taken into account, when a truck wheel is staying on a curb. The bending moments arising in this position have been calculated on the basis of experimental measurements of deflections.

Optimization shows that the thickness of deck plate can be decreased from 4.5 to 2.0 mm, the original number of cross members can be decreased from 14 to 10, and the original cross member shape (RHS) can be replaced by I- or a C - profile having optimum dimensions. These changes can result in 141 kg mass and 377.17 $ cost savings for a truck structure.

It should be emphasized that, in spite of the torsion of the whole floor in the second loading case, the cross members are loaded by bending, since the torsion is restrained by longitudinal members and by the deck plate. In the case of torsion the RHS profile would be, of course, more advantageous than the open profiles.

Higher tool cost of the RHS for n_c = 12 and 10 is caused by the large width of the profiles, since the height is limited to 100 mm. It can be seen that the higher tool cost does not significantly affect the result.

Using a welded steel deck plate and transversal stiffeners, one can make optimization on the same way. In spite of the mass increment comparing to the aluminium optimum, using steel elements one can reduce the total cost of the structure significantly, with 55%, although for a vehicle the mass is significant in fuel consumption.

Acknowledgements The research was supported by the Hungarian Scientific Research Fund OTKA T 109860 projects and was partially carried out in the framework of the Center of Excellence of Innovative Design and Technologies in Vehicle, Mechanical and Energy Engineering at the University of Miskolc within the EFOP-3.6.1-16-00011 “Younger and Renewing University – Innovative Knowledge City – institutional development of the University of Miskolc aiming at intelligent specialisation” project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund. The second author Prof. József Farkas† has recently passed away. With this paper we remember for his 6 decades long valuable scientific work on steel structure optimization.

References

1. DIN 1725–1983. Aluminiumlegierungen. Knetlegierungen. https://www.beuth.de/en/ standard/din-1725-1/2650507
2. DIN 1748–1981. Strangpressprofile aus Aluminium und Aluminium-Knetlegierungen—Zulässige Abweichungen. www.beuth.de/en/standard/din-1748-4/928185
3. FarkasJ, Jármai K (2013) Optimum design of steel structures, Springer Verlag, Heidelberg 288p. ISBN 978-3-642-36867-7. http://dx.doi.org/10.1007/978-3-642-36868-4
4. Farkas J, Jármai K, Dúl R (2001) Minimum cost design of a truck floor welded from aluminium-alloy profiles, Welding in the World, Pergamon Press, Vol 45, No 9–10, pp 19–22. ISSN 0043-2288
5. Hobbacher A. IIW Recommendations for fatigue design of welded joints and components. IIW-doc. IIW-1823–07, ex XIII-2151r3-07/XV-1254r3-07
6. Eurocode 3. Part 1.1 (2005) Design of steel structures. General rules and rules for buildings. European Committee for Standardization. Brussels
7. Rosenbrock HH (1960) An automatic method for finding the greatest or least value of a function. Comput J 3(3):175–184. doi:10.1093/comjnl/3.3.175
8. London Metal Exchange http://www.lme.com/metals/ Accessed 17 Sept 2016

Theoretical and Parametric Investigation of an Automobile Radiator

Máté Petrik, Gábor Szepesi, Károly Jármai and Betti Bolló

Abstract Automotive radiator is one of the most important devices of the engine cooling system. The function of this equipment is to remove heat from the engine and to keep the engine operating at the most efficient temperature. Nowadays, in the automotive industry, one of the most important project is decreasing the mass. This chapter focuses on calculation and optimization of finned-tube heat exchanger using several methods.

1 Introduction

Automotive radiators are heat exchangers, which used for cooling internal combustion engines. These engines are cooled by circulating engine coolant liquid through the engine block where it is heated, and through this radiator, where it is losing heat to the surrounding air. This coolant circulates from the tubes to the engine block by the coolant-pump. The air forced across the radiator's core, which forced by a fan or by the motion of the vehicle. This air warms up by the coolant, which temperature will be decreased. With the use of the fins, an extended surface is obtainable. This radiator is a surface heat exchanger, so with this extended area the performance will be higher [1].

The cooling system could be divided into two parts: the heat exchanger and the air flow management components (such as pipes, water pump, thermostat, fan). There is cooling performance difference between a car with a 77 kW power and a

M. Petrik (✉) · G. Szepesi · K. Jármai · B. Bolló
University of Miskolc, Miskolc, Hungary
e-mail: vegypet@uni-miskolc.hu

G. Szepesi
e-mail: szepesi@uni-miskolc.hu

K. Jármai
e-mail: jarmai@uni-miskolc.hu

B. Bolló
e-mail: aramzb@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_3

truck with 340 kW. Due to limited space at the front of the engine, the size of the heat exchangers is restricted and cannot be increased. So in case of higher cooing performance, change of the radiator is not enough, investigated of the other parts of the system (fan, fan shroud, coolant pump) is necessary. So, the optimization of an automobile radiator includes the investigation of the single parts and the analysis of interaction between them [2].

According to E. Sany, the performance of the radiator is the function of the inlet temperature and the inlet velocity. Different tube rows and various coolant mass velocities have been analysed in [3]. These results show that the increasing number of tube rows, the performance of the radiator will be increased, but the pressure drop of the device will increase too. The first result is useful, but the free space in front of the device is limited. The second is harmful, because much stronger than it must be used and deposits can be formed.

However, a change of the coolant could increase the performance too. In this chapter, the coolant is water, but in the industry, there are a lot of other coolants are used. Different thermal fluids have been compared in [4]: water, ethylene glycol and propylene glycol aqueous solutions at 30, 40 and 50%. Their results show, that from these the best coolant is the water. In the function of the coolant flow the difference between the first and the second best performance could be approximately 2 kW too. This difference is not negligible. In this chapter, the coolant is water for the best performance.

2 Calculation of the Automotive Radiator—Theorems

As all the other heat exchangers, the performance of the radiator is calculated by the laws of the thermodynamics and criterial equations. The energy from the temperature difference in case of the air and the coolant must be equal to the radiator performance.

These devices are cross-flow structures, where air is the outer and coolant is the inner medium. The tubes were finned to improve the heat transfer. With an increased heat transfer area, the performance will be bigger.

There are several theorems to calculate these finned tubes and shows differences between the theories, which are described below.

2.1 Theorem of Fábry

According to György Fábry, the heat transfer coefficient of finned tubes depends on the outer heat transfer coefficient, the heat transfer areas and the parameters of the fins [5]. The fins disturb the heat convection between the air and the surface. The fin efficiency is the quotient of the disturbed and original heat transfer coefficient.

$$h = h_0\zeta\left(b_b\frac{A_R}{A} + \frac{A - A_R}{A}\right), \tag{1}$$

where

h_0 heat transfer coefficient of the unfinned tubes [W/m^{2}K],
A_R area of the fins [m^2],
A the whole finned area [m^2]

The ζ parameter can be calculated with these formulae:

$$\zeta = 1 - 0.18\left(\frac{l}{t_R}\right)^{0.63}, \tag{2}$$

where

l height of the fin [m],
t_{R} gap between the fins [m]

The b_{b} is the efficiency of the fins, which depends on the χ parameter. Calculation of χ is:

$$\chi = h\sqrt{\frac{2\zeta h_0}{k_R\delta_R}}. \tag{3}$$

The k_R: heat conductivity of the fins material [W/mK] and δ_R: wall thickness of the fins [m].

The overall heat transfer coefficient of the heat exchanger:

$$\frac{1}{U} = \frac{1}{h} + \frac{A}{A_1}\left(\frac{\delta}{k} + \frac{1}{h_1}\right). \tag{4}$$

More fins mean higher heat transfer area, but lower heat transfer coefficient. The h_1 is the heat transfer coefficient inner side of the tube [W/(m^{2}K)] which calculated in the function of the Re-number with these equations:

$$Nu = 3.66 + \frac{0.19\left(\mathrm{Re}\cdot\mathrm{Pr}\frac{d}{L}\right)^{0.8}}{1 + 0.117\left(\mathrm{Re}\cdot\mathrm{Pr}\frac{d}{L}\right)^{0.467}};\ \text{when}\ \mathrm{Re} < 2300 \tag{5}$$

$$Nu = 0.023\left\{1 + \left(\frac{d}{L}\right)^{2/3}\right\}\mathrm{Re}^{0.8}\cdot\mathrm{Pr}^{0.4};\ \text{when}\ \mathrm{Re} \geq 2300 \tag{6}$$

2.2 Theorem of Perry

The heat transfer coefficient between the fins in the function of the face velocity of the air, the outside diameter of the tubes and the center-to-center spacing is calculated by directly according to this method. The heat transfer coefficient [6]:

$$h_f = 5.29 \frac{v_f^{0.6}}{D_0^{0.4}} \left(\frac{p'}{p' - D_0} \right)^{0.6}. \tag{7}$$

The v_f: the face velocity of the air [m/s], D_0: outside parameters of the tubes [m], p': center-to-center spacing [m]

To the fin efficiency, two parameters must be calculated:

$$m = \sqrt{\frac{h_f p_f}{a_f k}}, \tag{8}$$

where p_f: perimeter of the fin [m] and a_f: cross-section area of the fin [m^2].

The Ω must be calculated to determine the heat transfer coefficient:

$$\Omega = \frac{\tanh(mb_f)}{mb_f}, \tag{9}$$

where b_f: height of the fin [m]. The heat transfer coefficient is similar to the previous:

$$\frac{1}{U} = \frac{1}{\Omega h_f} + \frac{A}{A_1} \left(\frac{\delta}{k} + \frac{1}{h_1} \right). \tag{10}$$

The appropriate Nusselt-equations to calculate the inner side heat transfer are:

$$Nu = 3.66 + \frac{0.0668 \cdot \left(\frac{d}{L}\right) \mathrm{Re} \cdot \mathrm{Pr}}{1 + 0.04 \left(\mathrm{Re} \cdot \mathrm{Pr} \cdot \frac{d}{L}\right)^{2/3}}; \text{when } \mathrm{Re} < 2300 \tag{11}$$

$$Nu = \frac{\frac{f}{2} (\mathrm{Re} - 1000)\, \mathrm{Pr}}{1 + 12.7 \left(\frac{f}{2}\right)^{0.5} \left(\mathrm{Pr}^{2/3} - 1\right)}; \text{when } \mathrm{Re} \geq 2300 \tag{12}$$

where f is the friction factor: $f = 0.25 \cdot (0.790 \cdot \ln \mathrm{Re} - 1.64)^{-2}$.

2.3 Cengel's Theorem

This theorem shows a lot of similarities with the previous two theories: the heat transfer of the air side calculated by criterial equations (such as Fábry), but the parameter in the fin-efficiency equations similar to the Perry-theorem. The fin-efficiency as a function of this parameter is shown in a graph [7].

The Nusselt-number equation in case of forced convection over circular cylinders in cross flow is:

$$Nu_k = 0.248 \cdot \mathrm{Re}^{0.612} \cdot \mathrm{Pr}^{1/3}. \tag{13}$$

The hydraulic diameter for the calculation of the Nusselt- and Reynolds-numbers [8]:

$$D_h = \frac{4 \cdot W \cdot A_{free}}{A_{ht}}, \tag{14}$$

where W: the width in the orientation of flow [m], A_{free}: the flow section of the air [m^2] and A_{ht}: the whole heat transfer area of the radiator [m^2]. This type of hydraulic diameter must be used because it is a non-circular cross-section with inner flow breaker parts (the finned tubes). The velocity of the air is equal to the velocity of the automobile.

The parameter for the fin-efficiency:

$$\zeta = \left(l + \frac{1}{2}t\right)\sqrt{\frac{h}{kt}}. \tag{15}$$

The l is the length of the fins [m], and t: wall thickness of the fins [m].

The efficiency depends on this ζ value. An approximation curve was used in the calculations.

The heat transfer coefficient calculated by:

$$\frac{1}{U} = \frac{1}{\zeta h_f} + \frac{A}{A_1}\left(\frac{\delta}{k} + \frac{1}{h_1}\right). \tag{16}$$

The criterial equations for the inner heat transfer are:

$$Nu = 3.66 + \frac{0.065 \cdot \left(\frac{D}{L}\right)\mathrm{Re} \cdot \mathrm{Pr}}{1 + 0.04\left(\mathrm{Re} \cdot \mathrm{Pr} \cdot \frac{D}{L}\right)^{2/3}}; \text{when } \mathrm{Re} < 2300 \tag{17}$$

$$Nu = 0.125 \cdot \mathrm{f} \cdot \mathrm{Re} \cdot \mathrm{Pr}^{2/3}; \text{when } \mathrm{Re} \geq 23000, \tag{18}$$

$$f = (0.790 \cdot \ln \mathrm{Re} - 1.64)^{-2}. \tag{19}$$

3 Comparisons of the Theorem's Results

The previous methods were used to calculate a commercially available cross-flow car cooler. Parameters are shown in Table 1.

The results in Table 2, in case of steel structural material, are as follows:

Table 2 shows, that the calculation methods bring the same results. As a real radiator is calculated, the outlet coolant temperature is a technical data. Increased the performance this temperature will be decreased, which is the main goal of the following calculations. When the fins are made of aluminium, the results with the three theorems are according to Table 3.

There is a four-time ratio between the heat conductivities of the steel and the aluminium. However, in case of the Fábry- and Perry-theorem the performance does not significantly changed. The Cengel-method is sensitive to the change of the material, but the others are not. The fin efficiency has the most considerable difference: in the first case is changed between 0.2 and 0.99, but in the other case changed in a much smaller interval. If the fin material is copper it causes the largest heat conductivity, which is about 386 W/(mK) according to Table 4.

Table 1 Basic data and used symbols of a commercially available cooler

Length of the radiator	L	378 mm
Height of the radiator (length of tubes)	H	378 mm
Width of the radiator	W	34 mm
Outer diameter of the tubes	D_{out}	8 mm
Wall thickness of tube	t_{tube}	1 mm
Numbers of tubes	n_{tube}	20 pieces $\times$ 2 lines
Heat conduction of the tubes	λ_{tube}	50 W/mK
Wall thickness of fins	t_{fin}	0.1 mm
Length of the fins	l_{fin}	18.9 mm
Spacing between the fins	s_{fin}	1.3 mm
Numbers of fins	n_{fin}	270 pieces/tube
Heat conductivity of the fins	λ_{fin}	50 W/mK
Inlet air temperature	T_{airin}	20 °C
Inlet coolant temperature	T_{coolin}	90 °C
Face velocity of air	v_{air}	70 km/h

Table 2 Results in case of steel structural material

Theorem	Performance (W)	Outlet coolant temperature (°C)	Outlet air temperature (°C)
Cengel	11,538	61.58	26.8
Fábry	11,155	62.52	26.57
Perry	11,294	62.18	26.66

Table 3 Results in case of aluminium structural material

Theorem	Performance (W)	Outlet coolant temperature (°C)	Outlet air temperature (°C)
Cengel	12,451	59.33	27.78
Fábry	11,264	62.26	27.04
Perry	11,537	61.59	27.21

Table 4 Results in case of cooper structural material

Theorem	Performance (W)	Outlet coolant temperature (°C)	Outlet air temperature (°C)
Cengel	12,652	58.84	27.9
Fábry	11,279	62.22	27.05
Perry	11,548	61.56	27.21

Table 5 Results in case of three structural materials

Fin material	Performance (W)	Mass (kg)	Performance/mass (W/kg)
Steel	11,538	5.83	1940.27
Aluminium	12,451	2.00	6212.58
Copper	12,652	6.65	1902.17

Based on these tables can be obviously determined that better material heat conductivity and fin efficiency cause higher performance. The sensitivity analysis is calculated by the Cengel-method and the results shown below.

4 Sensitivity Analysis, Optimization

The heat exchanger performance depends on the structural material. However, these materials have different density and mass. When the geometry is given (such in the previous calculations), this performance and mass ratio are comparable as it is visible in Table 5.

The steel is the less expensive, but has the lowest heat conductivity, so it causes a large equipment. At a constant performance, the smallest device made of copper, but it has a much higher price. The values of the aluminium are between the values of steel and copper, but this ratio is three-times higher. So consider the price of the radiator, the aluminium is the optimal choice.

In [3], they got similar results: the copper is the best choice, but the aluminium has an approximately 4–5% less performance than the copper.

From this point, this performance/mass ratio change will be investigated in the function of the change of the parameters. These parameters are the number of tubes, the size of the tubes, number of fins and the width of the cooler.

4.1 Preliminary Calculations

In this calculation, the local optimum value is investigated in the function of fin-parameter, tube size and numbers of tube-rows. The tube sizes are standard values (Ø6 × 1 mm, Ø8 × 1 mm and Ø10 × 1 mm), and the rows are 1, 2 or 3. The spacing between the fins are changed in the function of the fin-parameter.

As it can be seen in Fig. 1, when the number of rows increasing, the performance and performance/ratio values are increasing too. There are local maximum values of the fin-parameter of 250–300 values. So the increasing the fin parameter the performance will be increased, but with it, the mass will more increasing, so their ratio will be decreasing. However, in case of 1 row the radiator cannot satisfy the required performance.

In Fig. 2, in case of Ø8 × 1 mm tubes, the performance is increasing, but the mass is also increasing too, so the investigated performance/mass values are smaller, then the first case. The Ø10 × 1 mm tubes are also investigated, but this decreasing is observed. The decrease in the performance is because of the increasing of the coolant flow area. The mass velocity of the coolant is constant, but the free flow area is growing (the diameter and the number of tubes rises). So the velocity of the water is reduced. This lower velocity causes lower Reynolds number, which will affect lower heat transfer coefficient. The local maximum point is shown Fig. 3 in the function of the fin-parameter.

From the calculations, this maximum point is at the 272 1/m fin-parameter. In this case there are 102 fins on the tube, and the spacing between them is 3.64 mm at this width. Any of the parameters are changed, this local maximum point will vary.

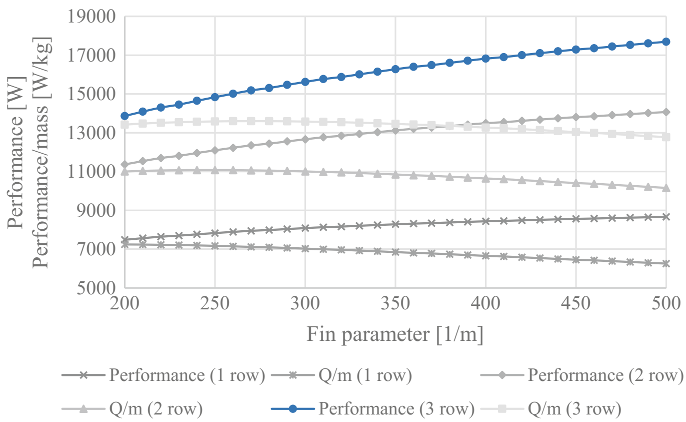


Fig. 1 Performance of radiator from Ø6 × 1 mm tubes

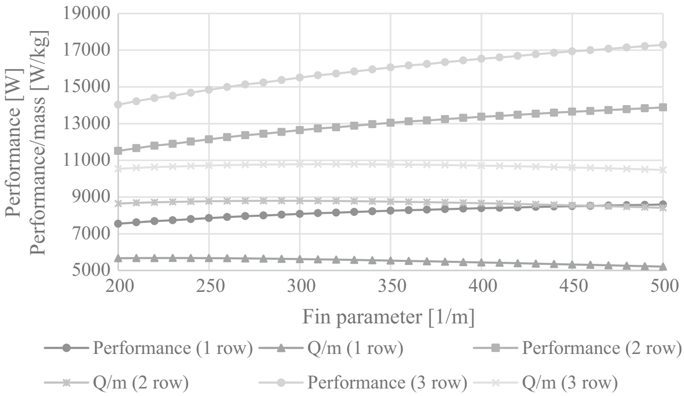


Fig. 2 Performance of radiator from Ø8 × 1 mm tubes

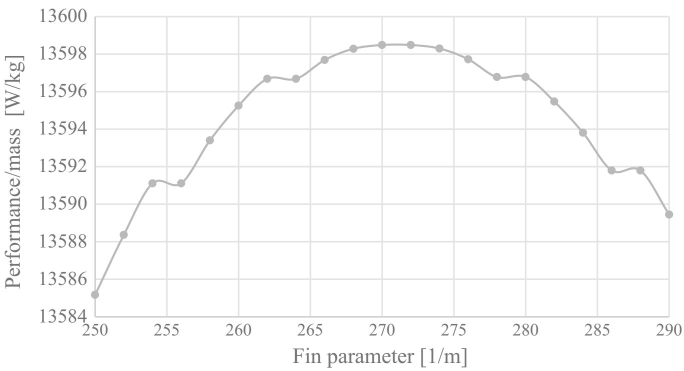


Fig. 3 Local maximum point in case of 3 rows and Ø6 × 1 mm tubes

4.2 *Effect of the Radiator Width*

The width is changed between 34 and 56 mm. Figure 4 shows the results:

Figure 4 shows that a wider radiator is better, but the optimum point move to the smaller fin-parameters. So the free space in the car defines the maximum built-in width. This free space is a basic planning data. This chapter shows a modified geometry where larger performance can be reached with lower air velocity in Table 6.

Whit this geometry, according to Cengel-method calculation the requested performance (approx. 12 kW) can be reached with 16 m/s air velocity, but the original heat exchanger needs more air velocity to produce this performance.

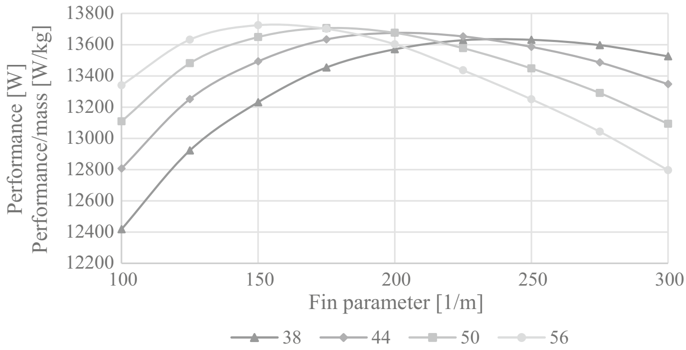


Fig. 4 Effect of the width to the radiator performance

Table 6 Original and the modified heat exchanger geometry

	Original	Optimized
Width	34 mm	56 mm
Tube size	Ø8 × 1	Ø6 × 1
Number of tubes	2 × 20 pieces	3 × 25 pieces
Spacing between fins	1.3 mm	6.77 mm
Length of fins	18.9 mm	15.12 mm
Number of fins	270 pieces	56 pieces
Mass	1.68 kg	1.10 kg

5 Summary

This chapter point out this area is still an interesting topic in the field of heat exchangers. Several calculation methods were introduced to calculate the performance of the car radiators which is a typical cross-flow air-liquid finned heat exchanger. As this chapter showed, a lot of parameters can affect the performance, but the main parameters are the radiator width and the fin parameter. Two calculations presented where the performance change can be calculated in the function of these parameters. A new construction is also demonstrated where lower air speed is enough to reach the acceptable performance. It would be very important in case of electrical cars where the noise is critical.

Acknowledgements The research was supported by the Hungarian Scientific Research Fund OTKA T 109860 project and was partially carried out in the framework of the Center of Excellence of Innovative Vehicle, Mechanical and Energy Engineering Design and Technologies at the University of Miskolc.

References

1. Carl M, Guy D, Leyendecker B, Miller A, Fan X (2012) The theoretical and experimental investigation of the heat transfer process of an automobile radiator. In: 2012 ASEE gulf southwest annual conference, April 4–6, El Paso, Texas, vol 1 (128), pp 1–12
2. Esmaeili Sany AR, Saidi MH, Neyestani J (2010) Experimental prediction of Nusselt number and coolant heat transfer coefficient in compact heat exchanger performed with ε-NTU method. J Engine Res 18:62–70
3. Charyulu DG, Singh G, Sharma JK (1999) Performance evaluation of a radiator in a diesel engine—a case study. Appl Therm Eng 19(6):625–639
4. Oliet C, Oliva A, Castro J, Pérez-Segarra CD (2007) Parametric studies on automotive radiators. Appl Therm Eng 27(11):2033–2043
5. Fonyó Z, Fábry G (1998) Vegyipari művelettani alapismeretek. Nemzeti Tankönyvkiadó
6. Green DW, Perry RH (2008) Perry's chemical engineers' handbook
7. Cengel YA (2003) Heat transfer—a practical approach, 2nd edn.
8. Amrutkar PS, Patil SR, Shilwant SC (2013) Automotive radiator—design and experimental validation 3(4):1–10

Past and Present: Teaching and Research in Vehicle Engines at the University of Miskolc

Szilárd Szabó, Péter Bencs and Sándor Tollár

Abstract The Department of Fluid and Heat Engineering, established as the Department of Machine Operation in 1952, hosts research and teaching in engines for vehicles. In the first two decades after the foundation of the department, besides engines in vehicles, steam engines still played an important role as they were used in locomotives. At that time the department owned several dynamometers and besides teaching about engines, the department carried out research and development for the national production of engines. After a decade of transition, in the second half of the 1990s the production of internal combustion engines started to flourish again. This was manifested first in teaching, but from 2010 in research as well. The replacement of the old laboratory equipment with modern measuring instruments and engines and the construction of a full departmental laboratory for engine diagnostics contributed to this. All of these factors explain why the Department of Fluid and Heat Engineering is in charge of teaching internal combustion engines in the new (starting from 2016) Vehicle Engineering B.Sc programme.

1 Introduction

The Department of Machine Operation was established in the Faculty of Mechanical Engineering at the Technical University of Heavy Industry by Decree 25 of 1951. The first head of department was Professor Alajos Lancsarics (1903–1963) (Fig. 1).

S. Szabó (✉) · P. Bencs · S. Tollár
Institute of Energy Engineering and Chemical Machinery, Department of Fluid and Heat Engineering, University of Miskolc, Miskolc, Hungary
e-mail: aram2xsz@uni-miskolc.hu

P. Bencs
e-mail: arambp@uni-miskolc.hu

S. Tollár
e-mail: aramts@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_4

Fig. 1 Alajos Lancsarics, founding professor of the department

Since its foundation, the department has hosted research and teaching in engines for vehicles. After its educational and research profile became established, the name of the department was changed to the Department of Fluid and Heat Engineering by Decree 52341/1965, MM on 15 March 1965.

The organisational structure of the Faculty of Mechanical Engineering and Informatics was modified by the formation of institutes in the autumn of 2013, leading to the establishment of the Institute of Energy Engineering and Chemical Machinery. Two previously existing departments joined to form the institute, and our activities go on within the Institute.

2 In the Beginning

Alajos Lancsarics was legendary for his enthusiasm for teaching machinery, and especially heat engines. As the vice rector for financial affairs, it was his task to set up and expand the equipment available for workshops and laboratories, which were quite rudimentary in the beginning. He managed to provide the necessary background needed for teaching about internal combustion engines, despite the numerous difficulties faced in the 1950s.

Heat engines were a major part of the content taught at that time, with a focus on internal combustion engines and also steam engines (see Fig. 2), as their role in locomotives was still important at that time.

It was mainly the tireless work and wide professional knowledge of Alajos Lacsarics that contributed to the development of teaching materials for the newly formed department (Fig. 3). He was enthralled by motorization, and in his opinion a degree in mechanical engineering could not be granted to someone lacking a thorough knowledge of these machines. In the laboratory classes for engines, he arranged trips for pairs of students between Miskolc and Hatvan in a steam locomotive, or organised a special train to take all of the mechanical engineering students on a visit to a power plant, using the Campus-Tiszapalkonya line. As early as the 1950s, students were learning how to drive motorcycles (Fig. 4), automobiles, tractors and combines.

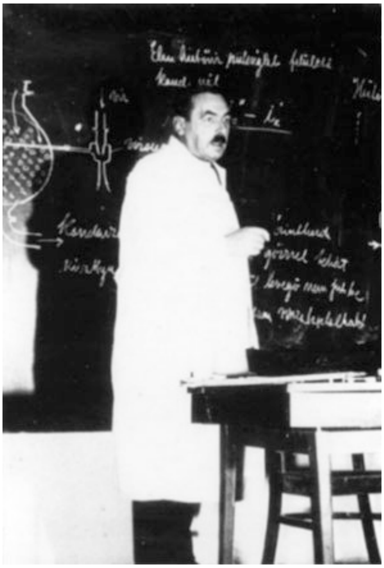


Fig. 2 Prof. Lancsarics lecturing on heat engines in the 1950s

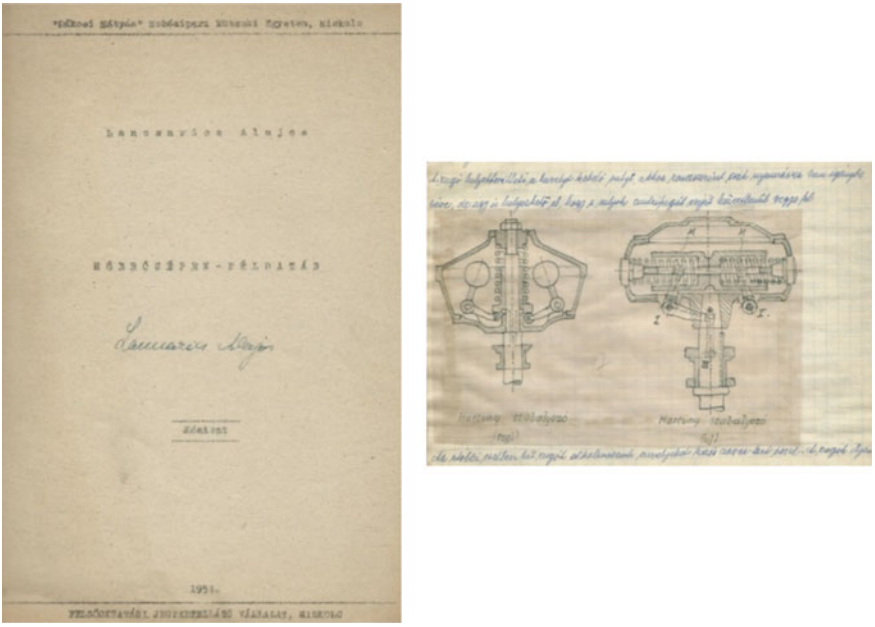


Fig. 3 Collection of problems for heat engines (*left*); hand-drawn lecture notes of a control unit and a regulator, by A. Lancsarics (*right*)

Fig. 4 Students learning how to drive a motorcycle on campus

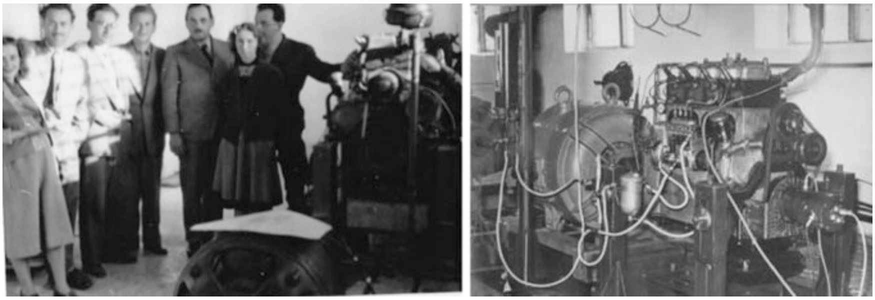


Fig. 5 Motor testbeds in the 1950s

In the first two decades, the department laboratories were located in the make-shift premises earlier inhabited by the prisoners who built the first buildings of the university. In addition to housing various types of vehicles, there was also an engine test lab with a testbed. That the members of the department had their photograph taken with this equipment (Fig. 5) testifies to their strategic importance.

During those first two decades, the department was involved not only in teaching but also in R&D on behalf of Hungarian motor manufacturers. These activities also supported the research progress made by lecturers. An example is the 1966 thesis for the university doctor's degree written by György Vida on the factors determining heat transfer in the cylinders of a diesel engine.

From the mid-1960s the teaching and research profile of the department altered somewhat, under the leadership of Tibor Czibere, with hydraulic machines gaining more emphasis among the turbomachines. Naturally internal combustion engines remained a focus of interest, as shown by the fact that half of the space in the modern laboratory (established in 1969) was given over to the engine testbeds, while the other half was used for the teaching of and research on hydraulic machines (Figs. 6 and 7).

In the 1990s, most of the motor manufacturing in Hungary took place on a licensing contract basis, and the amount of R&D was quite limited. During this period the main focus for engines was in teaching, while hydraulic turbines were a topic not only in the classroom but also for research.

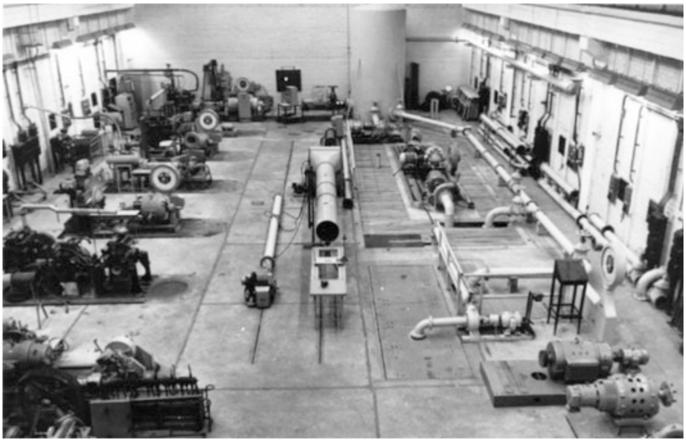


Fig. 6 The new department laboratory (1969), with testbeds on the *left* and underground water reservoirs and hydraulic machines on the *right*

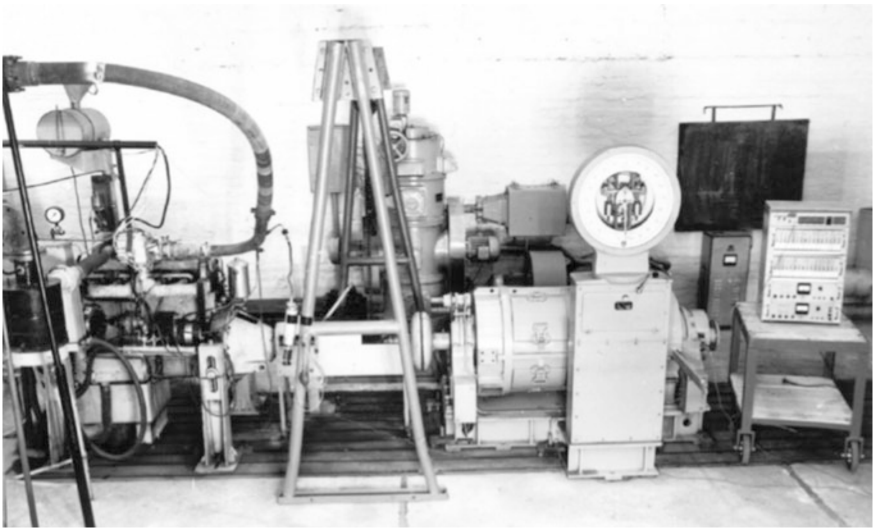


Fig. 7 A modern testbed with dynamometer (1969)

With the settlement of multinational companies in Hungary in the new millennium, the production of internal combustion engines started to flourish in Hungary. At the same time gas engines and gas turbines came to the front in energy production. That is why the demand for engineers specialising in these fields has increased again. Awareness of these trends motivated the department to revise the curriculum and update the equipment needed for teaching and research. Within the subjects of Turbomachines, Power Stations and Systems of Power Engineering within the B.Sc programme in Energy Engineering special attention was given to heat engines, steam turbines, gas turbines and gas engines. In the area of internal combustion engines educational equipment was first procured: a 90 hp 1.4 Opel Astra G petrol engine of 1997, a 2.0 HDI (High Pressure Direct Injection) Peugeot diesel engine of 2002, vibration diagnostic equipment called EASY-VIBER, a Testo type exhaust gas analyser and Würth type automotive diagnostic equipment. Using this equipment, new courses were launched such as Internal Combustion Engines and Engine Diagnostics.

As the next step, in 2010 an engine test stand for the engine braking of single cylinder petrol and diesel engines produced by GUNT was purchased (see Fig. 8). The test stand was produced as equipment for engineering education but can also be used for doing some simple research. For example, using this equipment it became possible to investigate fuel consumption [1] or the use of additives for reducing fuel consumption. Using the apparatus, students investigated the effect of auxiliary devices within complex assignment and final thesis projects.

Fig. 8 An engine test lab system with a drive motor and dynamometer produced by GUNT (2010)

3 Nowadays

The largest development took place in 2012. The research and teaching laboratory was established in two rooms of the departmental laboratory for engine diagnostics. The establishment of the laboratory was mainly supported by a grant for infrastructural development. The engineering and technological design and the implementation of the project were carried out by Energotest Ltd. [2]. The name of the laboratory became the Lancsarics Engine Test Lab after the founder of the department.

The engine test room was established using existing machines. The 2.0 TD Common-rail diesel engine used for the first experiments was awarded to the University of Miskolc by Audi Hungaria Motor Ltd. The tested engine (Fig. 9) was placed in a soundproofed room separated from a control and teaching cabin. Soundproof fire doors between the two rooms ensure visibility and accessibility.

The fixing of the engine to the base of the machine and positioning it to the dynamometer is ensured by specially developed engine holder palettes. The

Fig. 9 Interior of the engine test laboratory

connection between the engine and dynamometer for the transfer of torque and speed (of revolution) is ensured by a Cardan shaft with a very flexible rubber element for absorbing torsional vibrations and equipped with fitting disks. The 8000 rpm dynamometer (maximum power 250 kW, maximum torque 1200 Nm) is water-cooled eddy current dynamometer with impulse modulation control electronics. A 44 kW induction motor with frequency converter speed control is attached to the engine brake side of the powertrain system, constituting a compound braking unit developed by Energotest. The controllable induction motor enhances the dynamics of the system substantially and also the effectiveness of the dynamic tests. Also, the induction motor is capable of driving the test engine, hence the engine operation can be avoided and tests can be carried out using an electric motor to rotate the engine, ensuring the investigation of engine friction.

The environmental and technological boundary conditions are provided by a preparatory and serving measuring system developed and built by Energotest with the following elements:

- A preparatory system for liquid coolant that provides cooling of the test motor using a water–water heat exchanger, built-in pipe system, a pump whose parameters match those of the system, and control elements.
- A preparatory system for diesel fuel provides the fuel needed to drive the test motor. The cooling of the fuel is carried out with a fuel–water heat exchanger.
- A preparatory system for the air intake of the motor provides the proper filtering of the air needed for motor operation, with adjustable de-pressurization and temperature control in the temperature range of 15–40 °C.
- An exhaust system removes the emitted fumes with corrosion-free piping and a compensator. A test section allows emission measurements.
- A ventilation system for the room provides continuous ventilation of the lab, with ventilators removing and introducing air, air ducts, fume hood, filtering system and rain proofing elements.
- An accelerator unit ensures the control of modern E-gas systems as well as conventional bowden cable systems.
- A CAN bus data acquisition system can be configured flexibly and the potentially extendable version allows the measurement of motor and environmental parameters up to 20 signals. It can also control the preparatory system and that of the test engine. The whole system is computer controlled with Hungarian language Energopower engine test bed software installed in an industrial computer built into the rack-type control desk.

The up-to-date engine test laboratory is suitable for investigating energy processes taking place in internal combustion engines and for carrying out diagnostic and emissions tests on diesel engines for passenger cars and light commercial vehicles, in use today or for future development, as well as for educational activities related to the practical operation of motors.

With a dynamometer, during motor operation, it is possible to determine charge characteristics, investigate part loads and test cycles defined by the user, measure fuel consumption, determine specific fuel consumption curves, and investigate the effect of changes in environmental parameters on performance. In addition, with the environmental system the effects of various types of motor oil and fuel can be investigated, and harmful emissions can be analysed using an AVL emission meter of "0" accuracy class.

Being able to rotate the engine by an electrical motor enables us to investigate the friction at different speeds, and to carry out tests as used by manufacturers for engine development. This is a fundamental measuring technique for the downsizing experiments expected these days.

Since being put into operation the engine test bed has been engaged in three large research projects when long tests for different purposes have been carried out on diesel engines:

- 100-h test to check the laboratory operation at different engine operating conditions
- 800-h test of deposition and soot deposition
- 500-h test of motor wear and oil consumption.

All of these factors explain why the Department of Fluid and Heat Engineering is in charge of teaching about internal combustion engines in the Vehicle Engineering B.Sc programme being launched in 2016. This includes subjects such as Internal Combustion Engines and Motor Vehicle Engine Diagnostics.

Research within the department on two doctoral-level research topics is currently underway: one deals with experimental and numerical investigation of the exhaust system and the other with the development of an alternative valve.

In addition to measuring devices and apparatus, the software package ANSYS Fluent and AVL motor diagnostics software are available for numerical simulation.

Further development of the engine test laboratory is planned in order to accommodate the latest models of engines, and we would like to expand the research topics by purchasing equipment for vibration diagnostics in order to investigate mechanical effects of vibration.

Acknowledgements The research was partially carried out in the framework of the Center of Excellence of Innovative Design and Technologies in Vehicle, Mechanical and Energy Engineering at the University of Miskolc.

The described article was carried out as part of the EFOP-3.6.1-16-00011 "Younger and Renewing University – Innovative Knowledge City – institutional development of the University of Miskolc aiming at intelligent specialisation" project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund.

References

1. Tollár S, Mátrai Z (2012) Investigation of the effect of different diesel fuels on operating parameters of an engine. GÉP 62(9):57–60
2. Szilágyi G (2012) New dual-function dynamic engine test lab at the University of Miskolc. GÉP 62(9):5–6

Alternating Current Hydraulic Drive the Possibility of Applying in the Automotive Industry

Tamás Fekete

Abstract For the drive technology tasks are used different system drives. One of these drives is the well-known direct-current hydraulic drive (DCH). The hydraulic drives can be classified among a new version of the new hydraulic drives: the alternating-current hydraulic drive (ACH). The spread of the drive innovation problems hinder, in turn it has many favourable features, which influence positively the transmission properties of the drive. With the development of this type of drive is dealt at the Department of Machine Tools the leadership of Dr. János Lukács about 40 years. The subject in the field of several doctoral dissertations, patents, scientific articles and student papers have been written. I got into the research work within the framework of doctoral training. My task was within the topic of alternating-current hydraulic drives: the transmission properties of the synchronous alternating-current hydraulic drive. I would like to introduce on the Department of Machine Tools that I built help of alternating-current hydraulic drive, this drive is advantageous and disadvantageous features, knowing that they these transmission properties can be used advantageously machines of the automotive industry.

Keywords Synchronous alternating-current hydraulic drive · Hydromotor · Hydrogenerator

1 Introduction

By all of the vehicles, used in other areas of technology, and by all of the implement (not knitted track travelling) the fundamental problem is the design of the controllability and the planning of his drive is necessary to the operating of his ensuring movement equipment. The implementing of the drives may be using mechanical, electrical, pneumatic and hydraulic energy is. The drives are from among operated by technology listed introduced hydraulic drives I would like to introduce a drive,

T. Fekete (✉)
University of Miskolc, Miskolc, Hungary
e-mail: tamas.fekete@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_5

which is based on the positive properties of drive technology in the automotive industry, movement tasks of mobile appliances can be used.

According to the electrical analogy of the hydraulic drives can be divided into two main groups: direct-current hydraulic drives (DCH) and alternating-current hydraulic drives (ACH) [1].

By direct-current hydraulic drives, the energy transfer fluid (for example oil) between pump and motor(s) does unconfined movement. In case of direct-current hydraulic drives, the exotherm is substantial because of the hydraulic resistance. This exotherm must be drained out of the system. The draining of generating heat can carry out with the help of a heat exchanger or big size holder.

By the alternating-current hydraulic drive the oil in the phasepipes performs small amplitude alternating movement, therefore smaller in the phase space the thermogenesis.

In phase transmission are oil by alternating-current hydraulic drive makes an alternative motion. Therefore, temperature dissipation is lesser, because the parabolic speed distribution in phase transmission can't develop for reciprocating motion. The alternating-current drive can be divided into two groups: synchronous and non synchronous.

The ACH drives have two types. The alternating current synchronous drive (S-ACH in Fig. 1), the alternating current nonsynchronous drive (ACHM) (and the alternating current asynchronous drive (A-ACH)). The ACH drives have three main units: the alternating current hydrogenerator (ACG), the alternating current hydromotor (ACM) and the phase space (phasepipes) [2].

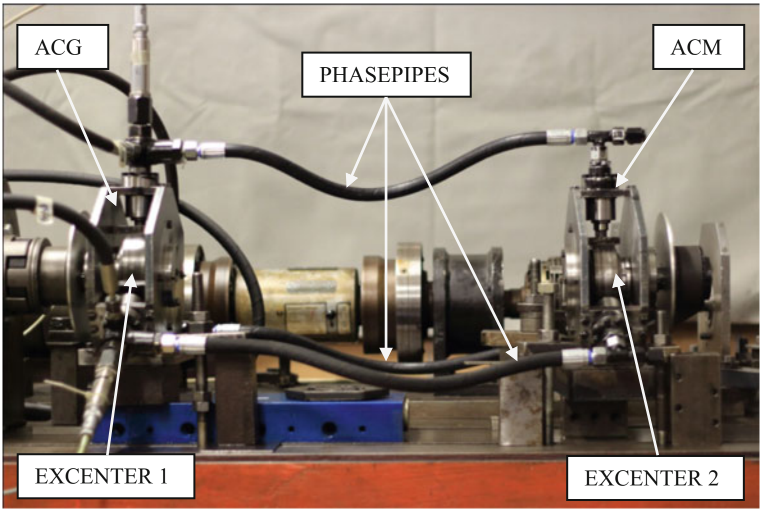


Fig. 1 The experimental drive made by me of the synchronous alternating-current hydraulic drive (S-ACH)

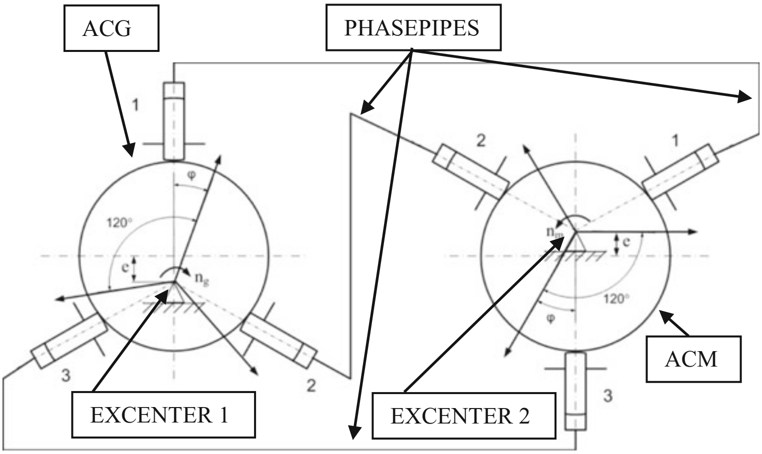


Fig. 2 Schematic outline of the synchronous alternating-current hydraulic drive (S-ACH)

The ACG of the pilot model has an eccentric exciting element. The shaft of the ACG is driven by a direct current hydro motor. By applying a direct current hydro motor simply can be enable the frequency controlling by the setting of the driving flow. We can also set the amplitude of the phase-flow with a double eccentric wheel when the drive is standing [3].

The experimental drive made based on the schematic outline (Fig. 2).

2 The Synchronous Alternating-Current Hydraulic Drive (S-ACH)

Contrast to direct-current drive, one advantage of alternating-current hydraulic drive is that, big size holder doesn't need. A central hydrogenerator (Fig. 3) generates the pulsating liquid flow (ACG), and propulsion of hydromotors (ACM) (Fig. 4) on given equipment with pulsating liquid flow can be feasible separately or in combination.

The hydrogenerator driven takes place by an external drive, which can be by an electric motor, gasoline or diesel internal combustion engine or a hydromotor. In my case the drive takes place by direct-current hydromotor. The direct-current hydromotor drive is shown in the Fig. 3.

The generator is an energy source in the chain of the energy transfer. The main parts of the generator are: the phase cylinder, the implied phase pistons and the motion mechanism of the phase pistons. These all cause the generation of the hydraulic sections, which determine the liquid flow. In case of synchronous drive case, the design of both the ACG and ACM is the same. In the Fig. 4 we can see the hydromotor with the load module.

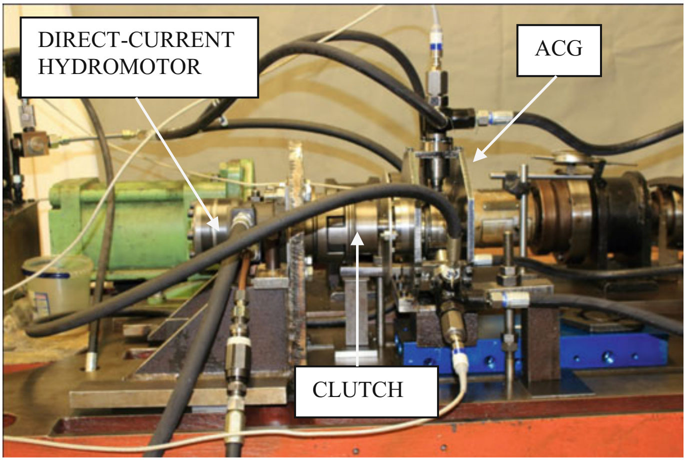


Fig. 3 The hydrogenerator of the synchronous alternating-current hydraulic drive (ACG) with the direct-current hidromotor

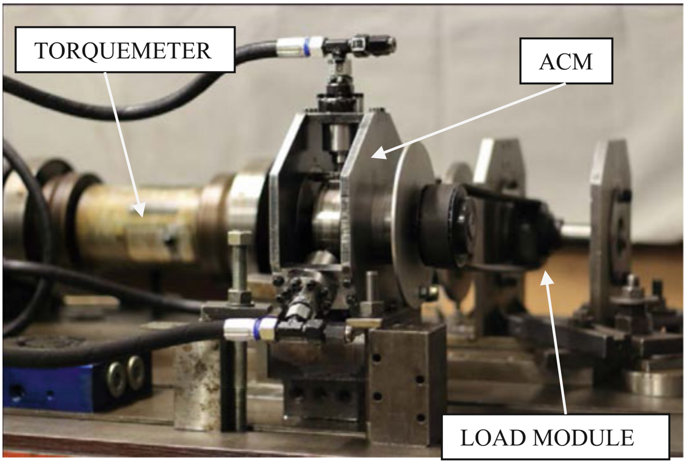


Fig. 4 The hydromotor of the synchronous alternating-current hydraulic drive (ACM) with the load module

From knowledge of the theorem of movement of excenter we can determine by the movement of the pistons. Examine the movement of excenter and that is the impact of the movement of the pistons. Based on the Fig. 5 we can determine that the value of ρ is differentially associated changing of the angle φ.

The theorem of movement of the alternating-current hydraulic drive it can be calculated as follows:

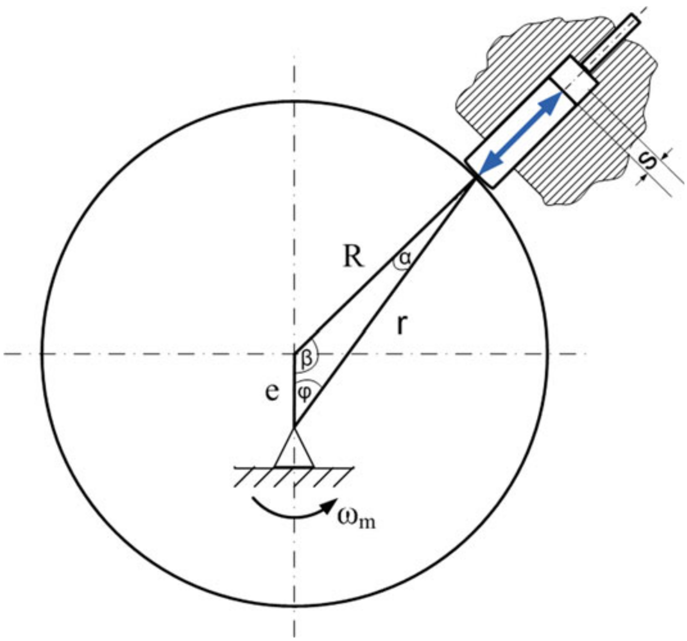


Fig. 5 The effect of the movement of the excenter of the piston displacement

$$r = R + e \cdot cos\varphi, \tag{1}$$

where r—a selected point of the excenter of movement distance from the pivot point [mm], R—radius of the excenter [mm], e—the eccentricity [mm], φ—[°].

In case of 120° offsets, why the pistons in case of three-phase spaced relative to one another 120°:

$$\left.\begin{aligned} r_1(\varphi) &= R + e\ \cos\varphi \\ r_2(\varphi) &= R + e\ \cos\left(\varphi + \tfrac{2\pi}{3}\right) \\ r_3(\varphi) &= R + e\ \cos\left(\varphi + \tfrac{4\pi}{3}\right) \end{aligned}\right\}. \tag{2}$$

So the displacement of the pistons is:

$$\left.\begin{aligned} s_1 &= e[1 - \cos(\omega t)] \\ s_2 &= e\left[1 - \cos\left(\omega t + \tfrac{2\pi}{3}\right)\right] \\ s_3 &= e\left[1 - \cos\left(\omega t + \tfrac{4\pi}{3}\right)\right] \end{aligned}\right\} \tag{3}$$

where s—displacement of piston [mm].

The alternating-current hydraulic drive reach at creep loads. These loads are ohmic inductive and capacitive hydraulic resistances. I disregard for these simplicity trial mode.

The characteristic curve of the liquid flow, generated by ACG, is all most same like sine function. The $\tilde{Q}_i(i = 1, 2, 3)$ phase liquid flow can be calculated with the following equation:

$$Q_i = Q_0 \cdot \sin\left(\omega t + k \cdot \frac{2\pi}{3}\right), \tag{4}$$

The absolute value of the liquid flow isc

$$Q_0 = A \cdot e \cdot \omega_g \cdot, \tag{5}$$

where A—surface of the piston [m^2], e—the eccentricity [mm], ω_{g}_angular velocity of the generator [1/s].

The law of motions of the pistons in the three hydraulic cylinders is

$$\left.\begin{aligned} x_{g_1} &= \mathrm{r}\ \cos \omega t \\ x_{g_2} &= \mathrm{r}\ \cos\left(\omega t - \tfrac{2\pi}{3}\right) \\ x_{g_3} &= \mathrm{r}\ \cos\left(\omega t - \tfrac{4\pi}{3}\right) \end{aligned}\right\}, \tag{6}$$

The surface of the foreheads of the pistons is the same.

The

$$\phi_{h_1} = L_{h_1} Q_1 \tag{7}$$

where ϕ_h—the hydraulic flux, L_h—the hydraulic induction factor, Q—volume flow rate or liquid flow, and

$$\phi_{h_1} = \frac{m}{A^2}\frac{dx_{gl}}{dt}, \tag{8}$$

from these relations

$$L_h = \frac{m}{A^2}. \tag{9}$$

With substitution Q the liquid flows are calculatedly:

$$\left.\begin{aligned} Q_1 &= Ar\ \sin(\omega t) \\ Q_2 &= -Ar\ \sin\left(\omega t - \tfrac{2\pi}{3}\right) \\ Q_3 &= -Ar\ \sin\left(\omega t - \tfrac{4\pi}{3}\right) \end{aligned}\right\}. \tag{10}$$

The summing of the three liquid flows the following amount:

$$Q_1 + Q_2 + Q_3 = 0 \tag{11}$$

is got [4, 5].

The functioning of the hydraulic generator starts the alternating current of liquid flows in the phase pipe, and these liquid flows forced the excenter disc of the motor to rotate.

The essential condition for operating the system is that summarise of the three liquid flows must be zero ($\sum Q = 0$). If the condition does not come true, pressure increase will appear in the phase pipe.

Each phase space of the capacitive pressure changes can be determined:

$$\begin{aligned} \Delta p_1 &= -\frac{1}{C_H}\int \omega_g a \ \sin\left(\omega_g t + \right) dt \\ \Delta p_2 &= -\frac{1}{C_H}\int \omega_g a \ \sin\left(\omega_g t + \ -\frac{2\pi}{3}\right) dt \\ \Delta p_3 &= -\frac{1}{C_H}\int \omega_g a \ \sin\left(\omega_g t + \ -\frac{4\pi}{3}\right) dt \end{aligned} \tag{12}$$

The phase printing may not be less than zero value (the phase space may must be tensioned, such as with pump). The phase pressures and knowledge of the torque arm of the actuation torque can be calculated:

$$M_h = Ap_{c1}k_1 + Ap_{c2}k_2 + Ap_{c3}k_3. \tag{13}$$

3 Measurement

We conducted the experimental measurements with flexible phasepipes, which we can see in Fig. 6. Based on the measuring arrangement the advantage of the flexible phasepipes is that the drive can be transmitted in any way, disadvantage that the capacitive character of the system is increased, so that the driving power is reduced.

The target of the measurement, we got to know to thoroughly the transmission properties of the synchronous alternating-current hydraulic drive (S-ACH), context of theoretical torque and speed (rotation) characteristics.

We can be determined with the measurement configurations main characteristics of the drive, the rotation speed of the hydromotor and load capacity, and their context using the appropriate measuring instruments.

Under the given circumstances must be examined which drive should choose to purpose (flexible or rigid phasepipe). In this paper, we try to give an answer to this question and we would like to demonstrate its characteristics and measurement form and method, wherewith parameters can be controlled. For this purpose, with measurements we would like to demonstrate the synchronous drive, it's a implemented method and characterized by state parameters (Fig. 7).

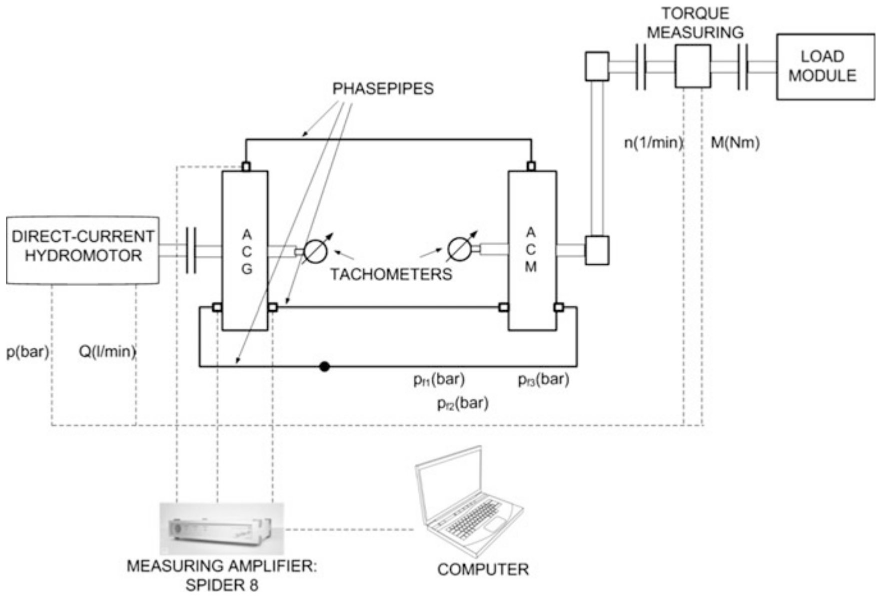


Fig. 6 The measuring arrangement

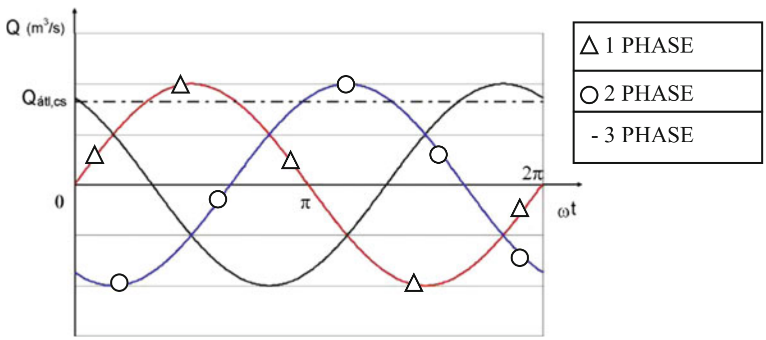


Fig. 7 The characteristics of the liquid flows in the phasepipes

4 Conclusion

From the present paper, it can be deduced, that the alternating-current hydraulic drive is suitable automotive industry drive technology tasks can be used. This technique gives the development opportunity, that the generator speed adjustability of modern electrical engineering can be realized. Later, it should have a more thorough examination of the effects of heat. Later, it should have a more thorough examination of the heat effect.

Third consideration is the possibility of the incorporation of other alternating-current hydraulic drive, e.g.: a hydrotransformer. We would with this hydrotransformator increase the distance between the hydromotor and the hydrogenerator, but at the same time the temperature can be reduced. We could be achieved with the hydrotransformer using that a small generator a bigger size hydraulic motor (and vice versa) to operate.

References

1. Fekete T (2014) The alternating current synchronous hydraulic drive. Annals of faculty of engineering Hunedoara, Int J Eng 12:235–238
2. Lukács J, Erdélyi J (2005) Operation and construction issues of the AC hydro generator phase pistons. Pneumatika, hidraulika, hajtástechnika, automatizálás. Vol. IX. pp. 60–63 (in Hungarian)
3. Erdélyi J, Fekete T, Lukács J (2008) The constructional and operational characteristics of contraction cylinder. Pneumatika, hidraulika, hajtástechnika, automatizálás Vol. XII. pp. 3–5 (in Hungarian)
4. Pattantyús ÁG (1951) Practical Fluid mechanics. Budapest, Tankönyvkiadó (in Hungarian)
5. Ponomarjov Sz D (1966) Strength calculations in mechanical engineering. Vibrations SHOCKS. Budapest, Műszaki Könyvkiadó (in Hungarian)

Comparative Destructive and Non-Destructive Residual Stress Measuring Methods for Steering Rack Bar Semi-Product

József Majtényi, Viktor Kárpáti, Márton Benke and Valéria Mertinger

Abstract It is well known that residual stress is introduced within solid materials during many types of processing methods, including heat treatments, machining, grinding, casting, etc. The type and magnitude of the formed stress state can be various depending on the type and conditions of the treatment and the geometry of the sample. The presence of residual stress can either be harmful or useful. If undesired residual stress is arisen within a machine component during its manufacturing steps, it can lead to deformation. Since the geometry of automotive components must be kept strictly within tolerances, more and more attention is given to the importance of residual stress in the field of the automotive industry. Many methods have been developed to characterise the residual stress. In the present study, the results of a destructive and non-destructive residual stress measuring methods for induction hardened automotive rack bar semi-products were compared. The theory of non-destructive method associated with the distortion of the crystal lattice, from which, residual stress can be quantitatively calculated, while the destructive method that we have developed specially to test this product at the Lech-Stahl Veredelung GmbH is very fast, suitable for qualification, and easily inserted into the manufacturing line.

J. Majtényi (✉) · V. Kárpáti · M. Benke · V. Mertinger
Institute of Physical Metallurgy, Metalforming and Nanotechnology,
University of Miskolc, Miskolc, Hungary
e-mail: jozsef.majtenyi@lech-stahlveredelung.de

V. Kárpáti
e-mail: cirkonium2000@gmail.com

M. Benke
e-mail: fembenke@uni-miskolc.hu

V. Mertinger
e-mail: femvali@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_6

1 Introduction

Residual stress is arisen in solid state materials due to various types of manufacturing methods such as heat treatments, plastic deformation, machining, grinding, casting, etc. The type (tensile/compressive) and magnitude of the stress state can be extremely various depending on the characteristics of the processing method, conditions, the geometry and material properties of the component. The occurrence of residual stress can either be harmful or useful [1]. The formed residual stress within a component can harmfully affect its geometry during many types of manufacturing steps, therefore, more and more attention is payed to the phenomenon of residual stress in the field of the automotive industry. Nowadays, many methods are available to describe the residual stress state of a component. The majority of these examination methods are based on the principle that mechanical stress is always balanced within a component. However, if material is removed from the component, the original stress state will undergo a change which results in the loss of the previous geometry, that is, the component undergoes macroscopic deformation. Despite these macroscopic deformation-based examinations are capable for qualification or even classification, none of them provide accurate stress values. Diffraction based techniques are based on the strain of the crystal lattice, from which, residual stress can be calculated using mechanical constants of the examined material [1, 2].

In the present manuscript, the residual stress states of automotive steering rack bar semi-products produced by Lech-Stahl Veredelung GmbH were investigated. The main processing steps of the steering rack bars are induction hardening, annealing, burnishing and polishing in a continuous processing line. It is known that the developed stress state resulting from induction hardening can be various depending on the conditions of cooling and component geometry [3–5]. Furthermore, if subsequent processing steps follow, the final stress state of the component can be much more complex. This study focuses on the stress states formed after quenching, tempering and, in addition, to achieve higher stress values, stress was intentionally introduced into the steering rack bars through direct water cooling after stress relaxation heat treatments. The effect of quenching temperature on forming stress states was also examined. The stress states were characterised using non-destructive (sampling-free) X-ray diffraction and macroscopic deformation. The different examinations in principle were carried out on the same steering rack bars for better comparison. The correlation was searched for between the results of the two methods.

2 Experimental

Steering rack bar semi-product rods with diameter (D) of 26 mm and length of 500 mm made of 37CrS4 type steel were processed in a continuous manufacturing line at Lech-Stahl Veredelung GmbH, in Oberndorf am Lech, Germany. The first main manufacturing step was heat treated which three applied quenching temperatures, being 860, 880 and 900 °C, while annealing temperature was constant, namely 720 °C. Afterwards, the rods were subjected to machining and polishing, which were followed by stress relaxation heat treatment and final polishing. For some rods, water cooling was applied after stress relaxation to induce residual stress within the semi-products. The residual stress state of the rods was characterised with non-destructive (sampling-free) X-ray diffraction method at the Institute of Physical Metallurgy, Metalforming and Nanotechnology of the University of Miskolc, Hungary with a Stresstech X stress 3000 G3R centreless X-ray diffractometer. The residual stress was calculated from the shift of the {211} reflection of ferrite using Cr Kα radiation, Young's modulus (E) of 210000 MPa and Poisson's ratio of 0.29. The reflections were recorded in ψ geometry and $\pm$ 3 tilting positions in the −45°/45° tilting range [6]. Stress was measured in 5 points on each of the eight generatrix of the rods Fig. 1. For one rod, stress was measured in a point series 20 mm from each other on the eight generatrix. The stress was measured with less scatter (error) than 50 MPa in every measurement point, therefore scatter ranges are not marked in the figures.

If the true cause of bending is searched for, the difference between the stress values of the opposing generatrix must be considered. This is true for pure bending of a rod, where a tensile and a compressed genetratrix exists on the opposing sides of the rod. If an additional homogeneous stress is applied to the whole rod, for example, compressive stress, the stress value of the tensile generatrix decreases while the stress of the compressed generatrix increases. Therefore, to measure the true tendency of bending, stress asymmetry (σ^*) must be calculated as the absolute value of the difference of stress values along the two opposing generatrix of the rods. An example for the formula of stress asymmetry for generatrix "A" and "E" is given in Eq. 1 [7].

$$\delta^*_{AE} = |\delta_A - \delta_E| \tag{1}$$

During macroscopic deformation examinations, the eccentricity measured 20 mm from the end of the rods. Afterwards, a 180 mm long section with a depth of 0.3xD was machined from the rods 20 mm from the end of the rods on the generatrix of the highest measured eccentricity Fig. 2. After machining, eccentricity was again measured and the difference (Δ) of eccentricity before and after machining was calculated to characterise macroscopic deformation.

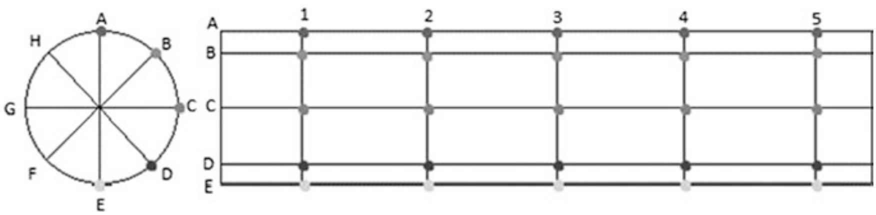


Fig. 1 Residual stress measurement mesh of the examined rods

3 Results and Discussion

3.1 Residual Stress States

Figure 3 shows the measured residual stress values of the rods after quenching from (a) 880 °C and (b) 900 °C, prior to annealing. Average stress values are inserted. The stress values are low, varying between −100 and 100 MPa. The average stress values are ~40 MPa. No differences can be observed between the stress states after quenching from 880 and 900 °C.

Figure 4 shows the measured residual stress values of the rods after annealing. Three quenching temperatures were used: (a) 860 °C, (b) 880 °C and (b) 900 °C.

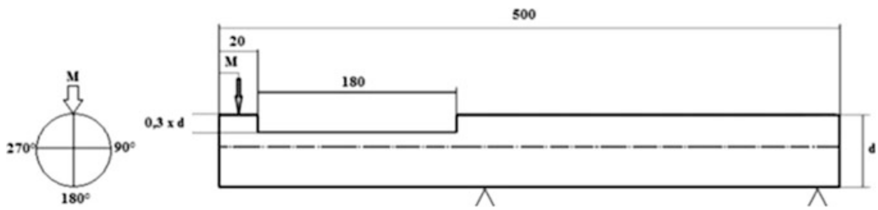
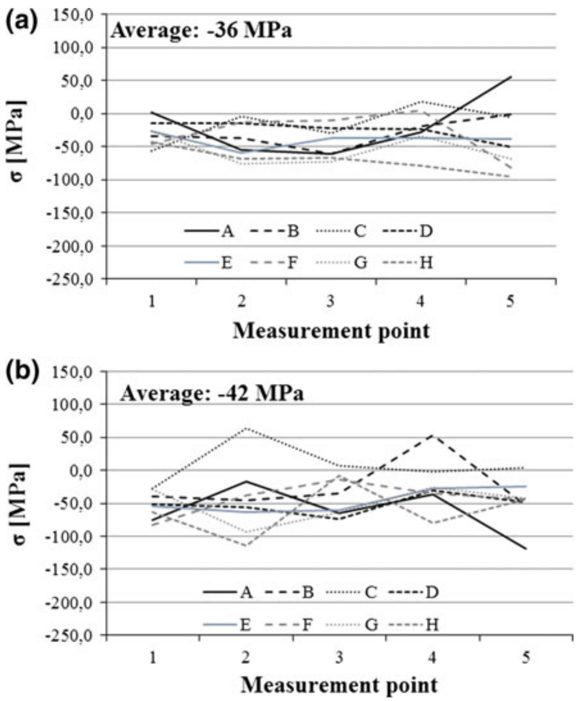


Fig. 2 Macroscopic deformation measurement performed by Lehl-Stahl Veredelung GmbH

Fig. 3 Residual stress distributions of the examined rack bar rods quenched from **a** 880 °C **b** 900 °C



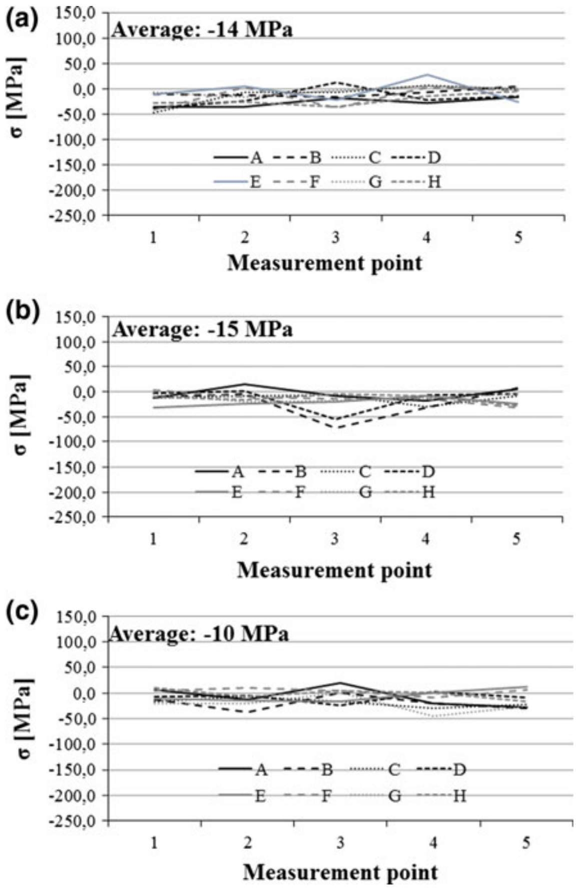


Fig. 4 Residual stress distributions of the examined rack bar rods after annealing. Quenching temperatures are (a) 860 °C (b) 880 °C (c) 900 °C

Average stress values are inserted. The rods are stress-free, the measured values typically vary between −50 and 50 MPa. The average stress values are ∼14 MPa. Again, no differences were found between the stress states using different quenching temperatures.

Figure 5 summarizes the measured stress values of the rods subjected to water cooling from the stress relaxation temperature. Again, three quenching temperatures were applied: (a) 860 °C, (b) 880 °C and (b) 900 °C. Average stress values are inserted. For all three rods, notable compressive stress was induced due to the water cooling. The average stress values are around −350 MPa. The notable role of the quenching temperature was not found once again.

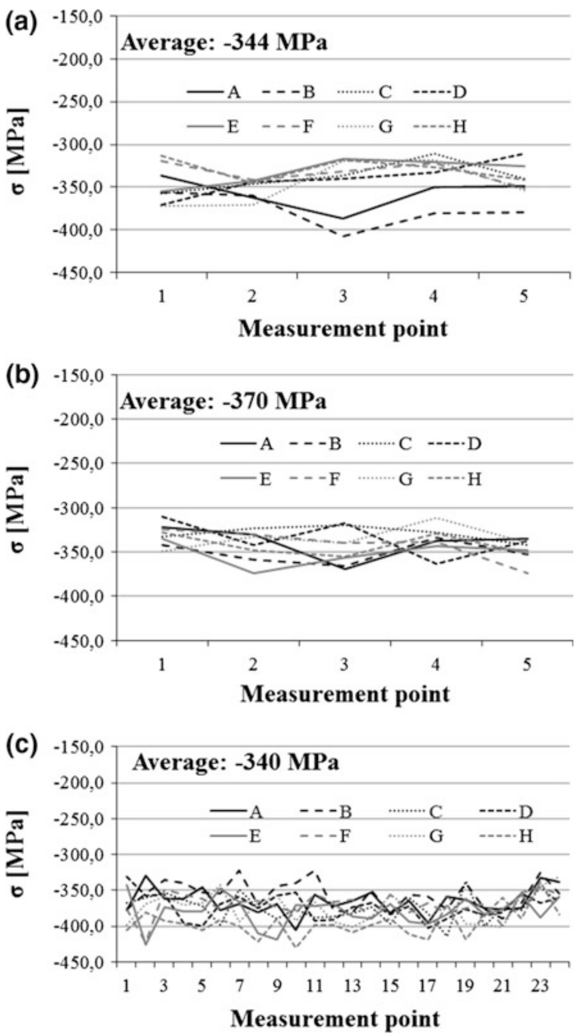
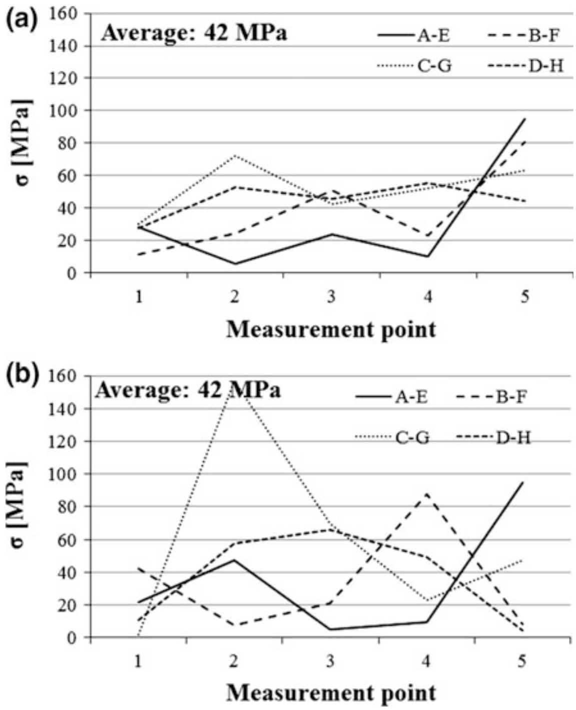


Fig. 5 Residual stress distributions of the examined rack bar rods with intentionally induced stress. Quenching temperatures are (a) 860 °C (b) 880 °C (c) 900 °C

3.2 Stress Asymmetry

Figure 6 shows the stress asymmetry of the rods quenched from (a) 880 °C and (b) 900 °C. The average values of stress asymmetry are inserted. In Fig. 6 b, there is a high stress asymmetry value, being above 140 MPa. However, all other values and the average values are low, being ~40 MPa. Evaluating a detailed stress and stress asymmetry mapping connected with macroscopic deformation examinations in Ref. [7] revealed that solitary, outlier stress and/or stress asymmetry values are

Fig. 6 Stress asymmetry distributions of the examined rack bar rods quenched from (a) 880 °C (b) 900 °C



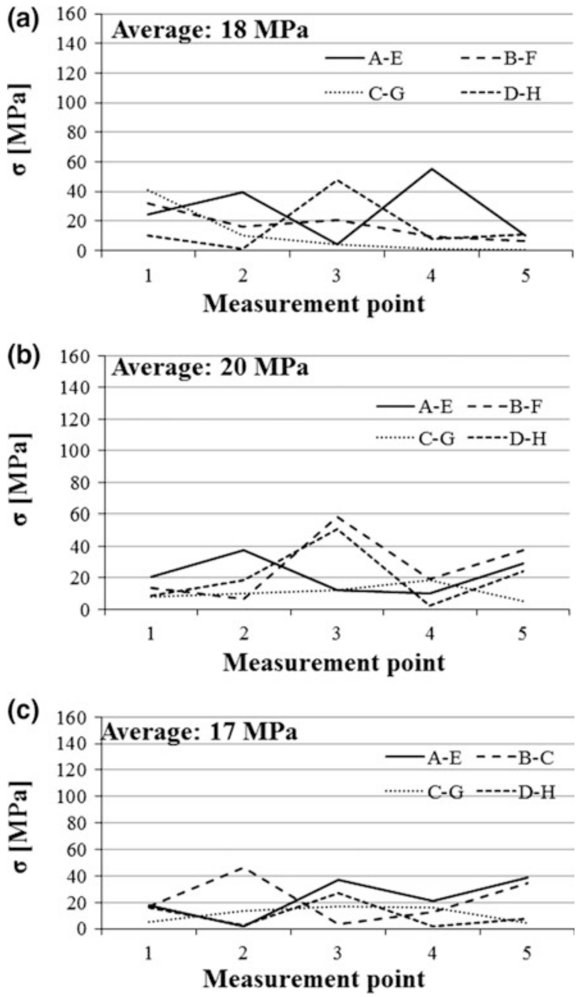
not responsible for the bending tendency of the rods. Based on that conclusion, the sole high stress asymmetry value found here is considered as an outlier data.

Figure 7 shows the stress asymmetry of the annealed rods with (a) 860 °C (b) 880 °C and (c) 900 °C quenching temperatures. The average values of stress asymmetry are inserted. The measured stress asymmetry values and average stress asymmetry values are low, being around 20 MPa.

Figure 8 shows the stress asymmetry of the rods induced during direct water cooling from the stress relaxation temperature. For all three rods, the stress asymmetry values are low, the average values are $\sim$25 MPa.

Table 1 summarizes the measured eccentricity values of the rods before and after machining and the calculated Δ values. The average stress values are also included. The states of the rods are marked as follows: quenched: Q, quenched + tempered: QT, intentionally stressed after stress relaxation: IS. It can be seen that in the quenched and annealed states, lower stress values, stress asymmetry values and Δ were measured. However, when the rods were subjected to direct water cooling after stress relaxation, lower stress asymmetry, but high stress and Δ values were measured. In this study, the high Δ values, that is, bending tendency of the rods cannot be associated with high Δ values, as was seen in our previous study [7]. On the other hand, the strong bending tendency is clearly associated with higher stress

Fig. 7 Stress asymmetry distributions of the examined rack bar rods after annealing. Quenching temperatures are (a) 860 °C (b) 880 °C (c) 900 °C



values within the rods. This correlation can easily be understood if one takes into account that during macroscopic deformation examination, the material removal is asymmetric. If low stresses are present within the rod, removing material from one part of the rod will not affect the stress state notably and the rod will not tend to bend. On the other hand, if the rod has a symmetric, high stress state, the asymmetric material removal destroys the stress balance and the rod bends. This can also easily occur during the tooth machining of the steering rack bar. The applied quenching temperatures had no effect on the stress states and bending tendency of the rods.

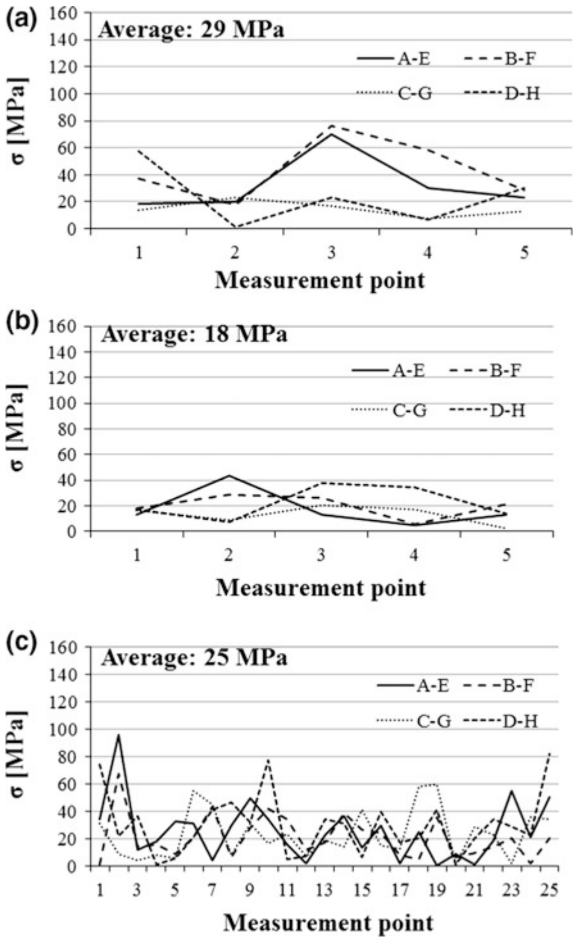


Fig. 8 Stress asymmetry distributions of the examined rack bar rods with intentionally induced stress. Quenching temperatures are (a) 860 °C (b) 880 °C (c) 900 °C

Table 1 Measured eccentricity Δ, average stress values and average stress asymmetry values of the examined rods. quenched: Q, quenched + tempered: QT, intentionally stressed: IS

State	T_Q (°C)	Eccentricity prior to machining (mm)	Eccentricity after machining (mm)	Δ (mm)	Average stress (MPa)	Average stress asymmetry (MPa)
Q	880	1.040	1.525	0.485	−36	42
Q	900	0.250	0.510	0.260	−42	42
QT	860	0.520	0.330	0.190	−14	18
QT	880	0.440	0.515	0.075	−15	20
QT	900	0.350	0.335	0.015	−10	17
IS	860	0.045	−2.720	2.765	−344	29
IS	900	0.085	−3.075	3.160	−370	18
IS	900	0.060	−2.760	2.820	−340	25

4 Summary

In the present manuscript, the correlation was searched for between the stress states and bending tendency of steering rack bar semi-products. The steering rack bars were examined in quenched, tempered, and intentionally stressed states. The effect of quenching temperature was also investigated. The stress states were examined using non-destructive (sampling-free) X-ray diffraction method, while bending tendency was examined with macroscopic deformation, during which, material was removed from one side of the rods. It was found that symmetric, weak stress states formed after quenching and tempering and no bending were measured. However, when the rods were intentionally stressed by direct water cooling after stress relaxation heat treatment, symmetric, but large (~ -350 MPa) stresses formed, which was associated with a strong bending tendency. The results clearly showed that in a high stress state, even if the stress state is symmetric, asymmetric material removal leads to a strong bending tendency. The examined quenching temperatures (860, 880, 900 °C) had no effect on the results.

Acknowledgements The described article was carried out as part of the EFOP-3.6.1-16-00011 "Younger and Renewing University—Innovative Knowledge City—institutional development of the University of Miskolc aiming at intelligent specialisation" project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund.

References

1. Totten G, Howes M, Inoue T (2002) Handbook of residual stress and deformation of steel. ASM International, Ohio
2. Krawitz AD (2001) Introduction to diffraction in materials science and engineering, Columbia
3. Krauss G (1980) Principles of heat treatment of steel. ASM ASM 15:240–245
4. Richard E (2006) Haimbaugh: Induction heat treating. The Materials Information Society, Ohio
5. Rudnev V (2003) Handbook of induction heating, Libary of Congress Cataloging—in-Publication Data, USA
6. Xstress 3000 G3/G3R system Operating instructions and instrument documents, 2012
7. Majtenyi J, Benke M, Mertinger V, Kazinczi T. The effect of quenching temperature and polishing force on the residual stress and deformation of rack bar semi-products, Mat Sci Forum, in press

Dynamical Modelling of Vehicle's Maneuvering

Ákos Cservenák and Tamás Szabó

Abstract This chapter deals with a dynamical modelling of autonomous vehicle, including different manoeuvrings. Solution of the state-space form was performed using the Runge-Kutta method. A special purpose program has been written in the Scilab software system. Manoeuvrings of different cornering and overtaking are simulated.

Keywords Mechatronics · Autonomous vehicle · Runge-Kutta method · Cornering · Mobile robot

1 Introduction

Nowadays prevention of accidents is very important in modern vehicles. An autonomous vehicle can sense faster the environment, than a human driver. Furthermore, the modern vehicle can prevent more accidents, than an older one [1–4]. The research and education of such problems inevitable are important. The vehicle engineering education has been started in this semester at the University of Miskolc. The preparation of the lecture notes is in progress. This paper is a contribution to this task.

The autonomous vehicle can also be considered as a mobile robot. The motion along a straight line can be easily performed, but the vehicle often has to move along arch, when cornering or overtaking are taking place. However, the motion along arch may have a more critical situation. When the vehicle moves along an arch with high speed, the vehicle might suffer instability which is called under steering or over steering behaviour [5].

The linear single track model is used to describe motions of the vehicle, which are detailed in [5]. The solution of a differential equation system is carried out by

Á. Cservenák (✉) · T. Szabó
Robert Bosch Department of Mechatronics, University of Miskolc,
Miskolc, Hungary
e-mail: cservenak.akos@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_7

Runge-Kutta numerical method written in the Scilab software system. The program can compute vehicle lateral slip angle, relative and absolute yaw angle, yaw rate and the track of vehicle's center of gravity in time.

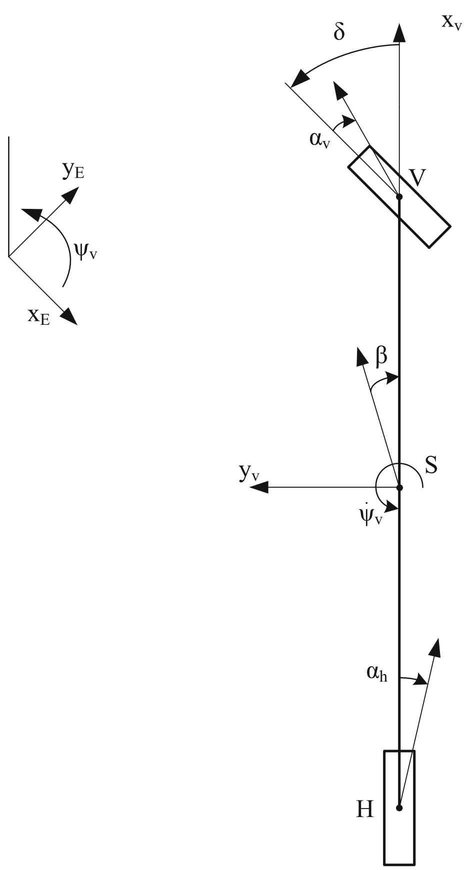
Examples of different manoeuvrings, including lane changing, overtaking, cornering in different cases will be shown for a vehicle [6] at constant velocity. These examples are very useful in engineering education.

2 Dynamics of Vehicle's Movement

A single track model is shown in Fig. 1, which can describe the dynamics of a four-wheel vehicle. The notations are given as follows:

- α_v and α_h: lateral slip angles of front and rear tire, respectively,
- β: vehicle slip angle,
- δ: steering angle,
- ψ_V: vehicle yaw angle,

Fig. 1 Linear single track vehicle model with required angles



- $\dot{\psi}_V$: vehicle yaw rate,
- V point: the center of front wheel,
- H point: the center of rear wheel,
- S point: the center of gravity
- x_v and y_v coordinate axes attached to the vehicle
- x_E and y_E coordinate axes of the reference system

In the ideal case non-holonomic constraints would determine the motion. However, our model is more realistic, the velocity of the center of a wheel usually does not lie in its plane. This is due to the fact that the friction is considered and lateral slip always experienced in case of cornering. Therefore, this model can also be applied to determine critical speed in different situations. To answer such question it is necessary to solve the dynamics of the vehicle.

3 Equations of Dynamical Model

The formulation of the equations of the dynamical model can be found, e.g., in [5]. A short summary will be given in the sequel.

The self-steering gradient "*EG*" is an essential quantity to characterize the behaviour of a moving vehicle. Its value remains constant in our investigations during motion. It depends on the mass m, the wheelbase l, the distance between front wheel and center of gravity l_v, the distance between rear wheel and center of gravity l_h and the cornering stiffnesses $c_{\alpha,h}$ and $c_{\alpha,v}$. Its value can be determined by the following formula:

$$EG = \frac{m}{l} \cdot \left(\frac{l_h \cdot c_{\alpha,h} - l_v \cdot c_{\alpha,v}}{c_{\alpha,v} \cdot c_{\alpha,v}} \right). \tag{1}$$

When the value of "*EG*" is lower or higher than zero, the behavior of vehicle is called over steering or under steering, respectively.

For a given path the cornering radii are known at each point. To follow this path the steering angle is needed, which can be determined by the following formula:

$$\delta = \frac{1}{\rho} \cdot \left(l + EG \cdot v^2 \right), \tag{2}$$

where v is the velocity of the centre of gravity. Otherwise, if the steering angle is given, one can determine the cornering radius as follows:

$$\rho = \frac{l + EG \cdot v^2}{\delta}. \tag{3}$$

The equation of the motion is given in state-space form:

$$\dot{\underline{x}} = \underline{\underline{A}} \cdot \underline{x} + \underline{b} \cdot u, \tag{4}$$

where the "$\underline{\underline{A}}$" matrix is the system matrix

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 1 & 0 & 0 \end{bmatrix}, \tag{5}$$

and "a_{11}", "a_{12}", "a_{21}" and "a_{22}" are given as:

$$a_{11} = -\frac{1}{v} \cdot \frac{c_{\alpha,v} \cdot l_v^2 + c_{\alpha,h} \cdot l_h^2}{\theta}, \tag{6}$$

$$a_{12} = -\frac{c_{\alpha,v} \cdot l_v - c_{\alpha,h} \cdot l_h}{\theta}, \tag{7}$$

$$a_{21} = -1 - \frac{1}{v^2} \cdot \frac{c_{\alpha,v} \cdot l_v - c_{\alpha,h} \cdot l_h}{m}, \tag{8}$$

$$a_{22} = -\frac{1}{v} \cdot \frac{c_{\alpha,v} + c_{\alpha,h}}{m}. \tag{9}$$

θ is the moment of inertia and the state vector is

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \dot{\psi}_v \\ \beta \\ \psi_v \end{bmatrix}, \tag{10}$$

the dot denotes the derivation with respect to time, $\underline{b}$ is the control vector

$$\underline{b} = \begin{bmatrix} \frac{c_{\alpha,v} \cdot l_v}{\theta} \\ \frac{1}{v} \cdot \frac{c_{\alpha,v}}{m} \\ 0 \end{bmatrix}, \tag{11}$$

finally, the system input is the steering angle, i.e., $u = \delta$.

4 Simulation Results

In this Section the dynamical simulations of a Suzuki car will be carried out. The vehicle dimensions [6] and the rest of the estimated parameters are listed in Table 1.

Three different manoeuvres, i.e., an overtaking, a turning to left direction have been analysed in the Scilab software system.

Table 1 The parameters of the vehicle

Name of dimension	Dimension	Quantity	Unit
Mass of vehicle	m	775	kg
Moment of inertia	θ	2000	$kg \cdot m^2$
Front length	l_v	1.032	m
Rear length	l_h	1.343	m
Wheel base	l	2.375	m
Front cornering stiffness	$c_{\alpha,v}$	60000	$\frac{N}{rad}$
Rear cornering stiffness	$c_{\alpha,h}$	45000	$\frac{N}{rad}$

This car is regarded as a supermini car since its weight is relatively low and very popular to use within cities. In the following examples the maximum velocity of the car will be 70 km/h.

4.1 Overtaking

Firstly, an overtaking of another vehicle has been computed at constant velocity v = 50 km/h. The manoeuvre of the vehicle is controlled by given steering angle. The steering angle $\delta = 0$ along a straight line (before $t_1 = 3s$; between t_2 and t_3; and after t_4). During the first phase of overtaking the steering angle $\delta = \delta_0 \cdot sin(\omega \cdot (t - t_1))$ is changing between t_1 and t_2, where $\delta_0 = 2°$ and $\omega = 2.45 \frac{rad}{s}$, while in the second phase the steering angle $\delta = -\delta_0 \cdot sin(\omega \cdot (t - t_3))$ is changing between t_3 and t_4. Runge-Kutta method is used for the solution of the differential equation system.

The results, i.e., yaw angle, vehicle yaw rate, vehicle lateral slip angle and the steering angle as input are shown in Fig. 2.

The track of the vehicle and the curve of the center of gravity are also obtained in awareness of the velocity and the time function of absolute yaw angle. The track is illustrated in Fig. 3, while the curve of the performed path is in Fig. 4.

4.2 Turning in Left Direction

A turning of vehicle in left direction has been described in the sequel. The velocity is kept constant, i.e. v = 15 km/h. In this example a car will take a left turn. The cornering radius is prescribed $\rho = 4.5\,m$, and the car runs along an arch of a quarter of a circle. While getting close to the corner ($t_1 < 2s$) and leaving ($t > t_2 \sim 3.7s$) it the steering angle should be zero to drive the car along straight lines. To start the cornering the steering angle is changed abruptly to a constant value according to

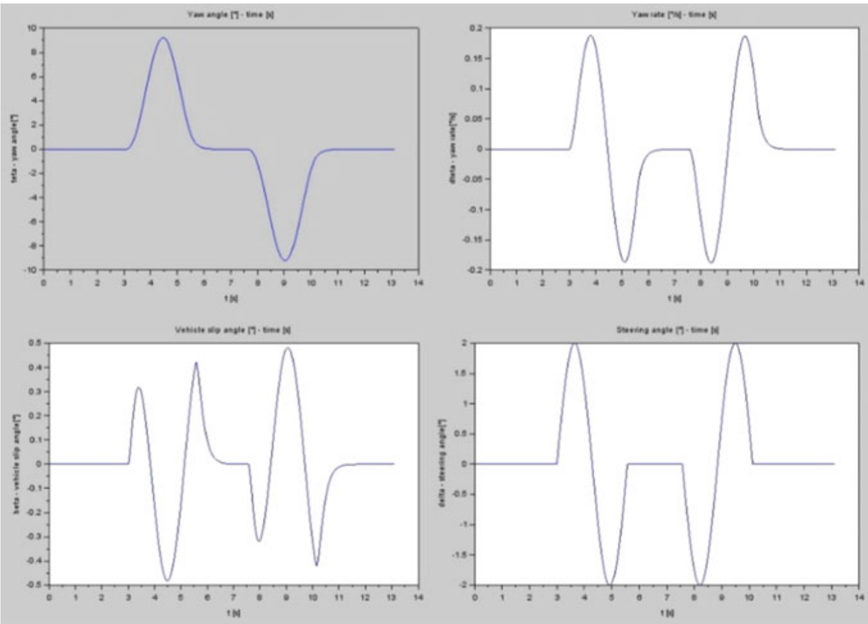


Fig. 2 The slip-, the yaw-, the steering angles and yaw rate of movement during overtaking

Fig. 3 The track of vehicle during overtaking

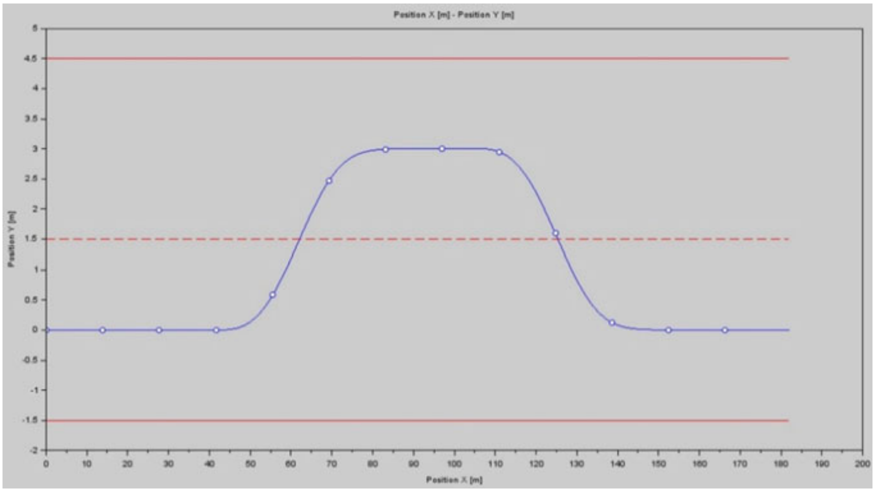


Fig. 4 The curve of path during overtaking

(2) and the time of manoeuvre is calculated with the help of the velocity and the arc length. Then at the end of the cornering the steering angle is changed back to zero.

The results for the vehicle yaw angle, the vehicle yaw rate, the vehicle lateral slip angle and the steering angle are shown in Fig. 5. One can see that the trajectory of the steering control has got sharp changes, however the response curves of the vehicle, i.e., the vehicle yaw rate and the vehicle lateral slip angle are smoother. The track of the vehicle and the curve of path are illustrated in Fig. 6.

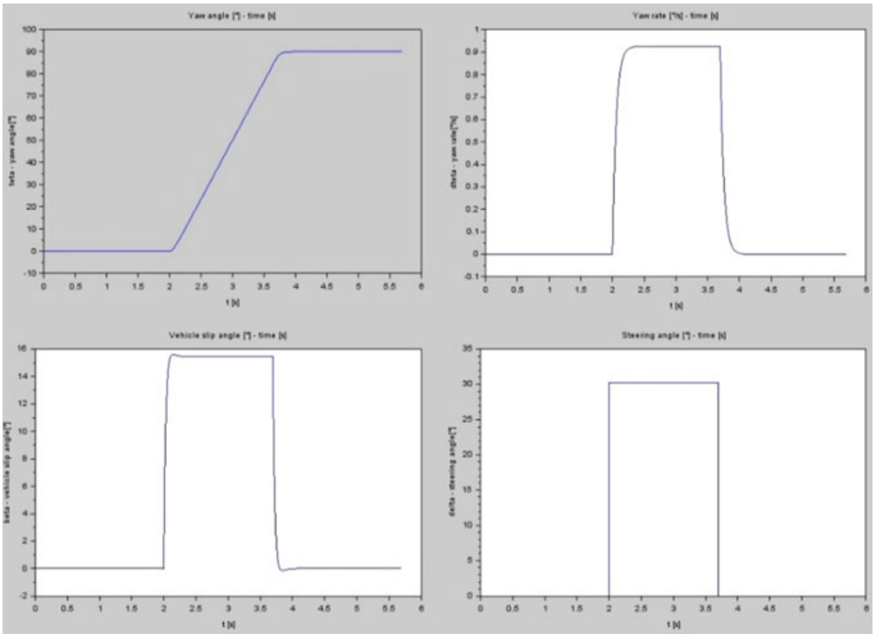


Fig. 5 The slip-, the yaw-, the steering angles and yaw rate of movement during *left* turning

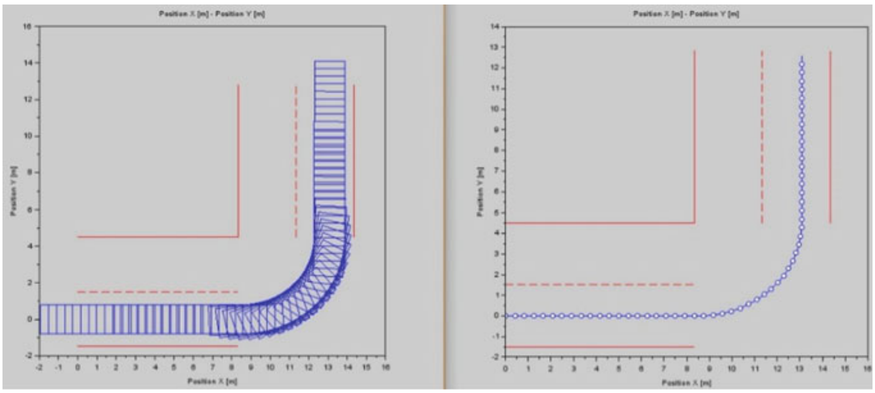


Fig. 6 The track of vehicle and curve of path during *left* turning

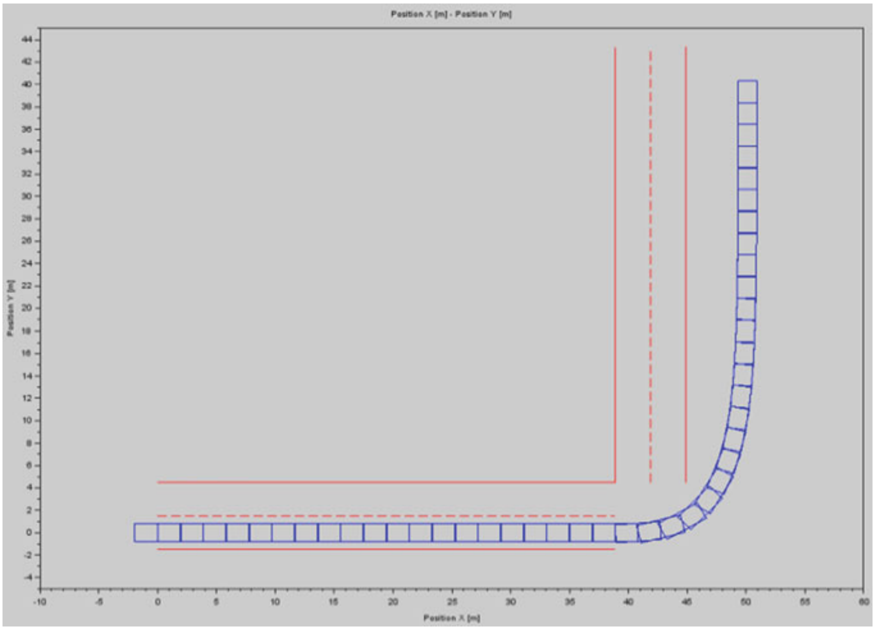


Fig. 7 The track of the vehicle turning at 70 km/h velocity

Repeating the computation at velocity 70 km/h but with the same steering control approach the result is shown in Fig. 7. It can be seen that the vehicle cannot keep the desired lane, which may lead to accident. It shows, the advantage of a simulation program one can perform numerical experiments to determine the appropriate steering controls for different velocities, which can be implemented into an autonomous vehicle or mobile robot.

5 Conclusions

Dynamical modelling of a linear single track model has been analyzed numerically in this paper. The vehicle can be controlled by prescribing the steering angle or the cornering radius. The solution of the model provides the vehicle yaw angle, the vehicle yaw rate, the vehicle lateral slip angle, the steering angle and the path.

Two manoeuvring problems have been examined. In the first one steering control was applied for an overtaking at velocity 50 km/h. In the second problem the cornering radius was given and the turnings have been performed at 15 and 70 km/h. In the latter case instability of the manoeuvre was experienced.

The two examples demonstrate the efficiency of the applied single track model.

Acknowledgements This research was carried out in the framework of the Center of Excellence of Mechatronics and Logistics at University of Miskolc.

References

1. VDA, Verband der Automobilindustrie e.V (2015) Automation from driver assistance systems to automated driving
2. Meiyuan Zhao (2015) Advanced driver assistant system, threats, requirements, security solutions, Intel Labs
3. Gietelink O, Ploeg J, De Schutter B, Verhaegen M (2006) Development of advanced driver assistance systems with vehicle hardware-in-the-loop simulations. Veh Syst Dyn 44 (7):569–590
4. Golias J, Yannis G, Antoniou C (2002) Classification of driver-assistance systems according to their impact on road safety and traffic efficiency. Transp Rev 22(2):179–196
5. Schramm D, Hiller M, Bardini R (2014) Vehicle dynamics, modeling and simulation, Springer. ISBN 3-5403- 6044-1
6. Suzuki Swift 1.0 GLX technical specifications, http://www.carfolio.com/ specifications/models/car/?car = 85745

Developing a Rotary Internal Combustion Engine Characterised by High Speed Operation

László Dudás

Abstract The paper deals with the development of a new internal combustion engine having a rotary piston. The introduction presents the short history of the evolution of the rotary combustion engines. The main part of the paper introduces the new patented internal combustion engine structure that has three rotational parts only: the rotary piston—rotor—, the rotary housing—rotary chamber—and the synchronizing gears. After the description of the structure, the work of the engine is discussed, compared to the usual internal combustion engines and the advantages and disadvantages are analyzed. Those properties that make possible the high speed operation are emphasized. Besides the constructional characteristics some manufacturing tasks are also presented, especially the very important precision finishing manufacturing of the working surfaces of the rotor and the rotary chamber.

1 Introduction

Though the development of automobile engines focuses on the electric motors special areas may use the new, efficient versions of the conventional internal combustion engines. Such area may be the motors of light sport airplanes and large drones for carrying humans, where the use of the patented construction of the author's rotary engine is promising. Using the new industrial ceramics as motor material adumbrates high efficiency of engines working quasi adiabatic manner at high temperature. The paper introduces a new rotary internal combustion engine construction that characterized by extremely high rotational speed ability owing to its simplicity and frictionless working.

There is a long evolutionary history of rotary type engines that resulted in the new motor presented in this paper. Two parallel ways of the inventions led to the

L. Dudás (✉)
University of Miskolc, Miskolc, Hungary
e-mail: iitdl@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_8

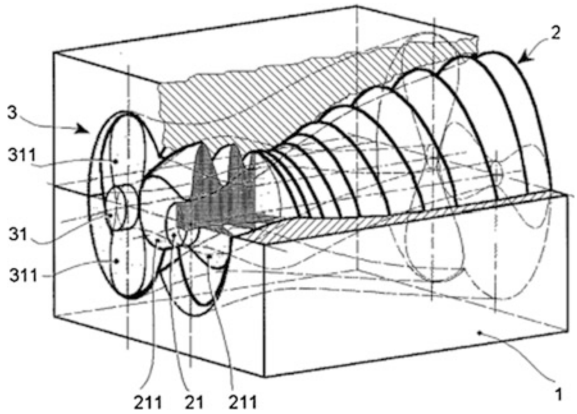


Fig. 1 Perna's invention is a modified screw compressor that works as an engine [11]

mentioned invention. The one is the screw compressor line and the other is the progressive cavity pump line. The *screw compressor line* started with the patent of Krigar in Germany [1] and continued with different innovative types having changing pitch, improved profiles and angle variation between the axes of the rotors [2–7]. In the last years appeared Perna's construction that can be used as internal combustion engine as it has compressing and expanding sections as shown in Fig. 1.

The *progressive cavity pump* line started with the patented construction of Moineau in United States in 1932 [8]. His construction characterized by a rotor that moved with a planetary motion in the static chamber substituting the rotations of the rotor and the chamber. The special helicoid surfaces of the rotor and chamber formed closed cavities that moved along the axis conveying fluid. Other patented versions apply fixed axes for the rotor and the rotary chamber with eccentricity [9, 10]. The version that is close to the discussed motor invention has two regions of the threads with different pitches. This construction can perform compression or expansion, depending on the rotational direction [9]. Figure 2 shows an example.

Among the above-mentioned constructions the suggestion of Perna as the member of the screw compressor type family is suitable for internal combustion engine [11]. Similarly, applying compressing and expanding sections for the progressive cavity type machines a motor can be constructed as shown in Fig. 3. This patented new motor [12] is the subject of the following discussions. The construction is characterized by continuously changing pitch and an elliptical rotor section curve. In the following the next sections will describe the construction, the advantageous and disadvantageous features. Among the advantages the properties that make especially suitable the construction for working with high rotary speed, will be emphasized. The perfect work of the construction depends on the small gaps between the rotor and the rotary chamber, so a section will analyze the grinding possibility of the rotor.

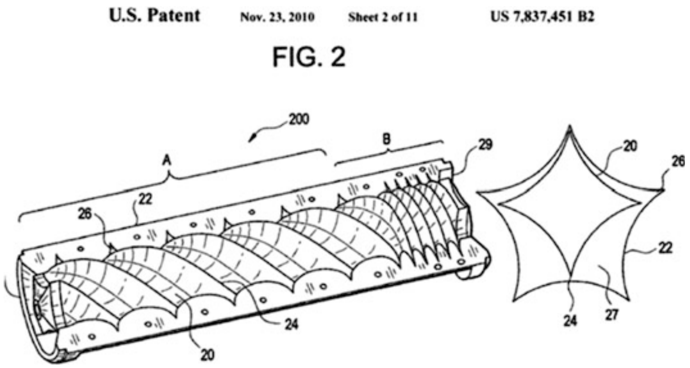


Fig. 2 Murrow–Giffin type turbine [9]

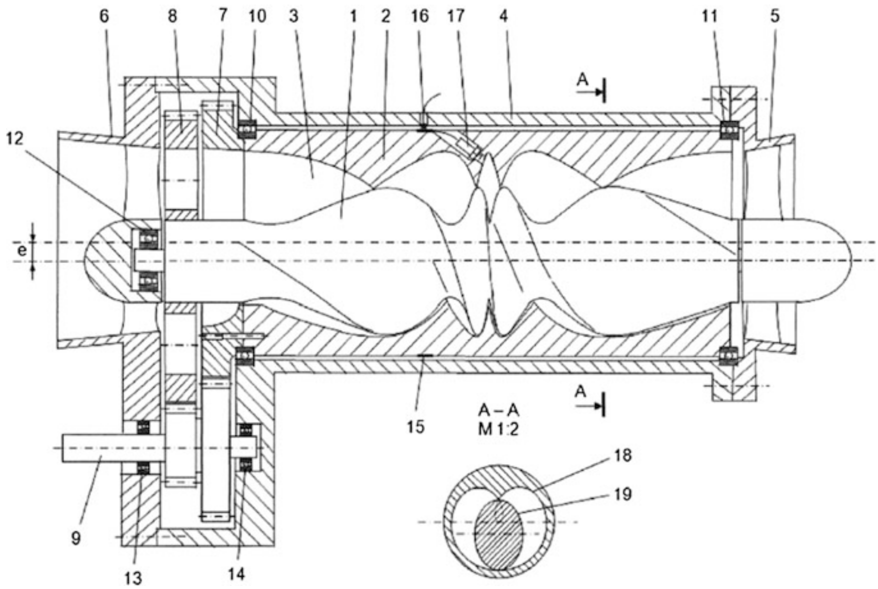


Fig. 3 The patented engine construction [12]

2 The Suggested New Internal Combustion Engine

The new engine practically is a spatial Wankel engine [13]. According to the original idea of Wankel, the triangle-like rotor and the chamber also rotated around fixed axes. This construction was modified later by Paschke [14] to provide an easy input of fuel-air mixture and exhausting of the hot expanded gas. Paschke dismissed some disadvantage of the motor, but its advantage, the clear rotary motions, was lost. Our patented engine construction unites the Wankel principle of clear rotational motion with the easy sucking in the fuel and the easy exhausting, as can be followed in Fig. 3.

2.1 The Build-up of the Motor

This motor is one of the simplest internal combustion engine constructions in the world. It has only three working parts: the rotor 1, the rotary chamber 2 and the synchronizing gear 9, similarly to the original Wankel-idea. The rotor and the rotary chamber have helical working surfaces that close cavities. These surfaces are conjugate surfaces in the relative motion. A very small gap between the surfaces provides the frictionless working. There is an e eccentricity between the two axes. The intersection area along the axes is constant, equal to the difference in the area of close curves 18 and 19. The cavities are larger at the two ends of the engine and become smaller in the middle section because of the continuously changing pitch. There are synchronizing gears 7, 8 and 9 to provide the 1:2 ratio between the rotor and rotary chamber. A spark plug 17 is applied in the small cavity section.

2.2 The Working of the Motor

When the rotor and the rotary chamber rotate the closed cavities move from the left intake side to the middle of the motor. The opened cavity at the left side of the motor sucks in the fuel-air mixture. Then the cavity closes and becomes smaller and smaller and compresses the mixture. In the middle the spark plug starts it up. The hot gas expands the cavity and such a manner rotates the rotor and the rotary chamber. Finally the cavity opens at the right hand side of the motor and the expanded gas effuses from the cavity and has been removed totally by the moving walls of the cavity. The gears synchronize the rotation of the chamber and the rotor and the axle 9 serves as the output shaft. This shaft can also be used for starting the motor.

3 Positive and Negative Properties of the Engine

In this section the advantageous properties and the disadvantages will be discussed, in many times in comparison to the well-known Otto motor.

The main advantage of this motor is its simplicity. It results in many consecutive advantages.

3.1 Cost

The simple construction results in low production and maintenance cost, long lifespan and increased reliability. The motor works without valves and valve synchronizing mechanisms, has less parts, less mass, less volume requirement, needs less material than the usual Otto motors. The knowledge of setting the suitable synchronizing of the valves for Otto motors is large, many books deal with it in detailed form. The cooling of the exhaust valves is especially hard problem in Otto motors. The calculation of the perfect length of the exhaust pipes and position of the inlet and exhaust manifolds is a sort of art.

3.2 Efficiency

In this motor there is no waste energy because of valves, springs, alternating mass of piston, friction between cylinders and pistons and between cylinders and piston rings, between valves at valve-seats, between the valve stems and valve controlling mechanisms. The frictionless working of the rotor without any sliding seal in the chamber allows very high rotational speed in opposite to the Wankel engine too. The high rotational speed means high power. The function curve of the compression and expansion can be set freely choosing proper pitch-function along the length of the motor. There is no possibility for interfusion of the content of the different cavities, the cavity contents are separated in every moment of the working, the gas moves linearly with the cavities. The required turbulence of the fuel vapor to achieve speedy combustion can be generated in the suitable form of the intake slot. The air pollution can be limited owing to prolonged and repeated sparking. There is no need for exhauster and sound damper as the pressure of the exhaust gas can be as low as in the environment. Without water cooling the temperature can be 25–35% higher [15, 16]. Using industrial ceramics for the rotor and the rotary chamber the cooling can be avoided and very efficient quasi adiabatic working cycle can be realized owing to the extreme high combustion temperature, see Table 1. The combustion temperature can be higher—1350 °C—than the melting point of the metals. Using glass ceramics having low thermal conductivity the quasi adiabatic process can be achieved, a low heat rejection engine can be made and the

Table 1 Engine material properties [17]

Material	Thermal conductivity ($Wm^{-1}K^{-1}$)	Max design temperature (°C)
Metallic	54	1000
High performance ceramics	43	1350
Glass ceramics	4	1000

high temperature results in higher efficiency and less not burnt material in the exhaust gas.

3.3 Noise

There is no noise resultant from valves and valve controlling mechanisms and wearing in valves. There is no vibration, the motor rotary parts can be balanced statically and dynamically. This is the main dynamical advantage, the rotary parts simple rotate. The form of the combustion chamber decreases the effect of the periodicity of explosions, the axial forces have a small acceleration effect at the beginning of the combustion, the tangential forces that cause the rotation act through longer space and time interval, causing smoother operation.

3.4 Properties Resulting in High Speed Operation

Owing to frictionless working of rotor and rotary chamber the speed is not limited by the friction, lubrication. There are no alternating, accelerated parts in the motor. Moreover, choosing suitable function for pitches in the intake section a continuous, constant intake speed can be achieved. As the red line [18], so the physical RPM-limiting factor is very high for this construction, it can easily outperform the 5,000–7,000 RPM of gasoline cars or even the 9,400 RPM of Mazda RX-8 Wankel engine. It seems to be not a problem to exceed the very high 19,000 RPM redline value of race motorcycle engines or 20,000 RPM of modern formula one cars. The rotation speed of internal combustion engines increases when the size decreases. The top speed of the smallest reciprocating engine of the world is 30,000 RPM [19]. Such small size engine applies special fuel and ignition by compression. For the suggested motor construction the RPM is limited by the rolling bearings or by the centrifugal forces arising in the rotary chamber. Gears can work up to 40,000 RPM [20].

3.5 Disadvantages

Unfortunately, there are some disadvantages:

- Complex surface of the rotor and rotary chamber need special technology
- The small gap between rotor and rotary chamber needs special attention in the timing of the explosion
- The inertia of the oil sealing in the small gap between chamber and rotor surfaces decreases the efficiency

- Higher noise caused by rolling bearings opposite to sliding bearings of Otto motors.

To avoid friction a small gap is required between the rotary chamber and the rotor, similar to screw compressors. To achieve good sealing a very small gap is essential, so special surface finishing method needed. This technology will be presented and analyzed in Sect. 4. Unfortunately, this technology increases the cost of manufacturing.

4 Manufacturing Considerations

The manufacturing considerations of this new motor type were presented in [21] in detailed form. The manufacturing of the parts of the motor is easy except the rotor and the rotary chamber. These parts have special helical surfaces having changing pitch, so the usual methods applied for producing threads and worms can't be applied. The chamber is produced with uniting two halves of it. The inner helicoids of the chamber machined in the halves by milling and the grinding is also possible with the small diameter disk form grinding wheel when the two halves are assembled. But the efficiency of such a process is very low.

In case of rotor the outer helical surface can be manufactured more easily. The machining can be imagined using a Gellért-type polygon turning machine tool, or milling is also possible with disk form tools. The finishing needs grinding and for this the process and special grinding machine and grinding wheel that was invented by the author can be used. This latter technology will be discussed in the following subsection.

4.1 Analyzing the Grinding Possibilities

This section deals with the analysis of grinding of the rotor. The need of grinding is justified by the required small gaps to achieve perfect sealing between the cavities. The problem is similar to the sealing of screw compressor rotors [22, 23]. As the rotor has a special helicoidal surface with changing pitch, conventional grinding using a surface of revolution shape grinding wheel is impossible. The problem is similar to grinding of tapered worms or hourglass worms. This problem is solved by the author who proposed a new grinding machine structure and special grinding wheel [24]. The patented grinding machine shown in Fig. 4 is suitable for grinding quasi-helical surfaces, e.g. tapered and hourglass worms. The novelty is that the machine applies a 1:1 rotation ratio between the workpiece and the grinding wheel, so the two surfaces are conjugate surfaces. For the generation of the grinding wheel surface an enveloping by the rotor surface method was applied. Such process can be performed using the Surface Constructor software application developed by the

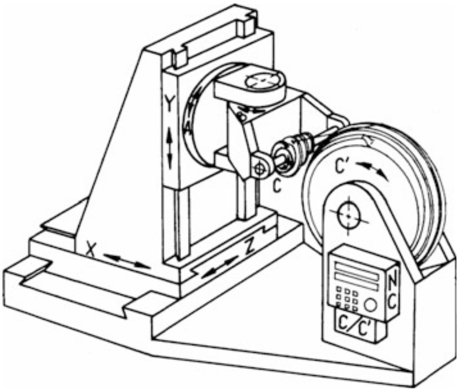


Fig. 4 Grinding machine for grinding quasi-helical surfaces

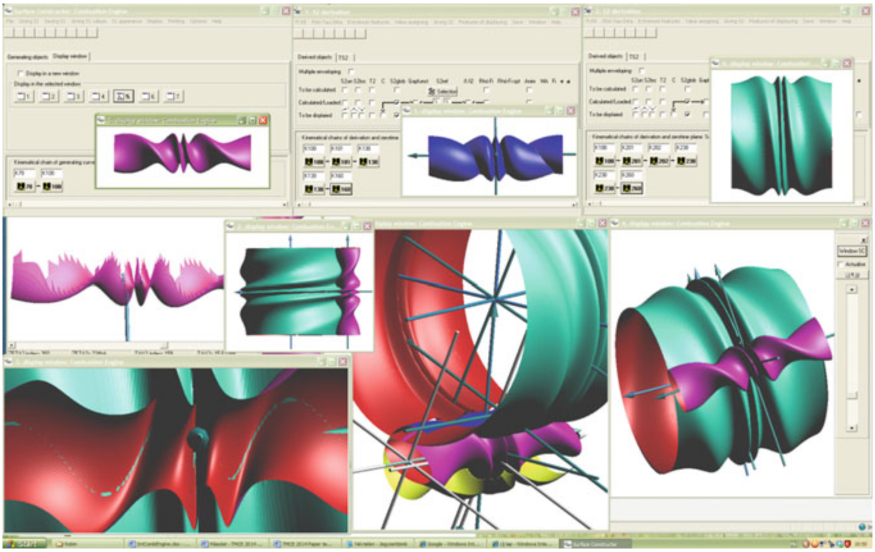


Fig. 5 Designing and checking the surfaces in the Surface Constructor

author. If the grinding wheel can be generated without undercutting, i.e. with clear enveloping, then the rotor can be grinded exactly with the generated grinding wheel. In the followings this generating process and the checking against undercut will be presented.

Figure 5 shows the Surface Constructor software application. The software was applied to create the rotor surface, for generation of the inner surface of the rotary chamber and for generation of the grinding wheel surface for the rotor grinding. The upper left window shows the rotor, the upper middle window shows the chamber

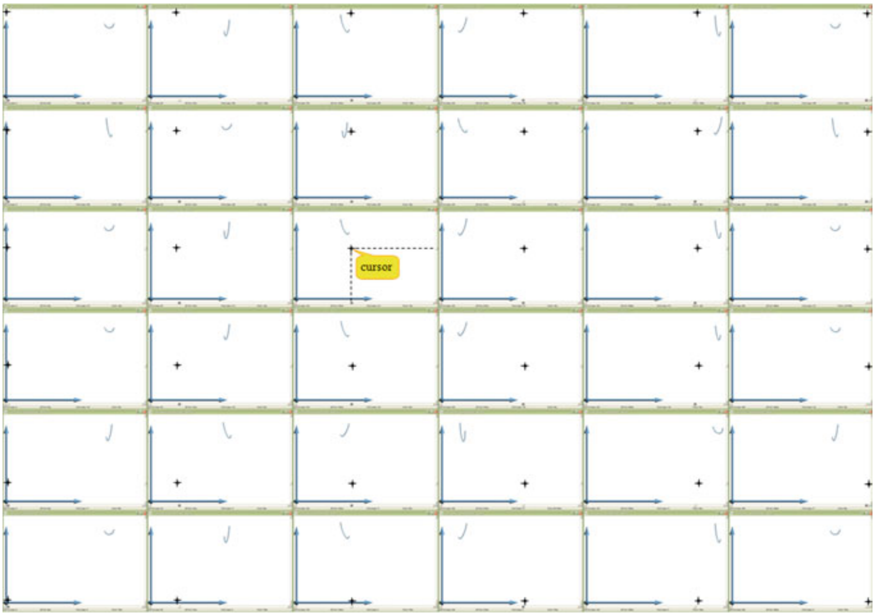


Fig. 6 Checking against undercut using Fi2 functions

surface and the upper right window shows the grinding wheel surface. The lower windows show the grinding wheel and the rotor in contact, and some other windows show the undercut analysis. These results appeared after inputting the required expressions and data (after entering the Fi2 grinding machine kinematics and setting the required Zeta2, Rho2 and Tau2 expressions and parameters for the generating process). To avoid undercut it was important to determine the right tilting angle for the rotary axis A of the grinding machine. Earlier experiences adumbrated that for the modelling of tapered or hourglass worm the best grinding tilting angle is equal to the least pitch angle. This value was 5° for the helicoids of the rotor.

The checking against undercutting was accomplished using the special visualization capability of Surface Constructor. In a dedicated window we can scan the whole Tau2, Zeta2 parameter domain of the $F2_2$ grinding wheel surface and the window for every point draws the Rho2 = Rho2(Fi2) function. Such a function can reveal all types of local undercut and global cut with its shape. The scientific background that includes the theoretical basis of the Surface Constructor can be found in [25]. If the Rho2 = Rho2(Fi2) function shows inflection or a local maximum, then the inspected point of $F2_2$ is a local undercut point. The perfect grinding wheel points are characterized by a smooth valley form. If all the $F2_2$ points in the analyzed Tau2, Zeta2 parameter domain show perfect shape, then the surface has not undercuts or edges. The mapping window acts as a scanning window at the same time, so gives an easy way to scan all the points of the $F2_2$ grinding wheel surface. By scanning the points the problematic locations can be detected. The checking against undercut proved that even a 0° tilting angle results in perfect grinding surface. Figure 6 shows some points with the Rho2 = Rho2(Fi2) function curves demonstrating the method. The grinding wheel can be produced for exact grinding of the rotor surface, but making such a wheel is a complicated, expensive process, thus it is suggested in case of mass production of rotors.

5 Summary

The paper introduced a new, patented rotary internal combustion engine. The advantages and disadvantages were discussed and the suitability for high speed operation was justified. Finally the finishing operation of the rotor using a special form of grinding was presented. In this process the special capability of Surface Constructor for detection undercuts was applied. In the future the analysis of the perfect pitch functions and the proper sparking will be analyzed. The mathematical modelling of the pressures and forces will be the following task to determine the moment and the power of the motor.

Acknowledgements This research was partially carried out in the framework of the Center of Excellence of Mechatronics and Logistics at the University of Miskolc. The financial support is acknowledged.

References

1. Cicerale FH (2009) Kenne Bell Supercharger Ford Mustang Install—Superbad. Mustang360°, article published 1 March 2009. http://www.mustangandfords.com/how-to/engine/mmfp-0903-2004-ford-mustang-kenne-bell-supercharger-install/. Accessed 1 Sept 2016
2. Abe YT (1996) Fluid rotary apparatus having tapered rotors. United States patent US 5533887 A, 1996
3. Becher U (2003) Twin screw rotors and displacement machines containing the same. United States patent US 20030152475 A1 2003
4. Jacobsson BO, Gabelli A (1990) Screw compressor. European patent EP0376373 A1, 1990
5. Kapp B. Improvements in screw pumps. Patent GB696732 (A), 1953
6. Martin H, Sachse R (2006) Rotors for a screw compressor. Patent EP 1722104 A2, 2006
7. North MH (2001) Screw vacuum pump. United States patent US 20010024620 A1, 2001
8. Moineau LRJ (1932) Gear Mechanism. United States patent US 1892217 A, 1932
9. Murrow KD, Giffin RG (2009) Axial flow positive displacement turbine. United States patent 2009/0226336 A1, 2009
10. Wiedenhoefer JFA (2009) Non-contact seal for positive displacement capture device. United States patent US 20090220369 A1, 2009
11. Perna V (2003) Equipment with mutually interacting spiral teeth. United States patent. US 2003001267 A1, 2003
12. Dudás L (2015) Rotary piston internal combustion engine. Hungarian patent HU230082, 2015

13. NSU Wankel rotary engines and cars. http://www.tripod.com/rotary/pg05.htm. Accessed 2 Sept 2016
14. Der Wankelmotor und sein Erfinder. http://www.nsuro80.ch/das-auto/der-motor/. Accessed 10 Sept 2016
15. 2012 Ceramic Rotary Engines, Inc: Ceramic Rotary Engine, http://www.youtube.com/watch?v=YnfFEf7wyww, Elérve: 10 Sept 2016
16. Stosic N, Smith IK, Kovacevic A (2003) Opportunities for innovation with screw compressors. In: Proceedings of the institution of mechanical engineers, Part E: J Pro Mechanical Eng 217(3):157–170. http://www.staff.city.ac.uk/~ra601/oportsvi.pdf. Accessed 23 Aug 2016
17. Salazar F (1998) Internal combustion engines, Department of Aerospace and Mechanical Engineering. University of Notre Dame, Notre Dame, IN 46556, 30 April 1998
18. Wikipedia. "Redline", http://en.wikipedia.org/wiki/Redline. Accessed Sept. 21, 2016
19. Ronney PD (2016) Hydrocarbon-fueled internal combustion engines: "the worst form of vehicle propulsion… except for all the other forms", Online publication: http://ronney.usc.edu/whyicengines/WhyICEngines.pdf. Accessed 11 Sept 2016
20. Triveni Engineering & Industries Ltd. product catalogue: High Speed Gears and Gearboxes. (2016) Online publication: http://www.trivenigroup.com/download/gearbox-catalogue.pdf. Accessed 11 Sept 2016
21. Dudás L, Biró M, Novák LL (2016) Construction modeling and manufacturing analysis of a new rotary combustion engine. In: INES 2016 20th Jubilee IEEE international conference on intelligent engineering systems, June 30-July 2, 2016, Budapest, Hungary pp 89–94
22. Holmes CS (2004) Inspection of screw compressor rotors for the prediction of performance, reliability, and noise. In: International Compressor Engineering Conference, Paper 1692. http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=2691&context=icec. Accessed 3 Sept 2016
23. Kacalak W, Majewski M, Budiak Z (2015) Worm gear drives with adjustable backlash. J Mech Rob 8(1):7–17
24. Dudás L (1992) Grinding machine, for grinding non-surface of revolution surfaces, especially conical and globoid worms. Hungarian patent HU P9003803, 1992
25. Dudás L (2010) Modelling and simulation of a novel worm gear drive having point-like contact. In: Horváth I, Mandorli F, Rusák Z (Eds.) Proceedings of TMCE 2010 Symposium, Ancona, pp 685–698

Simulation Methods in the Vehicle Noise, Vibration and Harshness (NVH)

Károly Jálics

Abstract The chapter introduces the simulation methods (MBS, FEM, SEA) which are generally used in the vehicle NVH. Alongside also hybrid methods will be introduced. An overview will be given about the usage of the methods depending on the frequency range for the simulation and prediction of the NVH behaviour of full vehicle and its components.

1 Introduction

The basis of the simulation methods was laid down already in the sixties of the past century (e.g. Finite Element Method, Statistical Energy Analysis). The widespread usage of them was blocked through the primitive computer technology. Through the fast development of the computers also numerical methods in general walked through an enormous evolution.

Vehicle NVH became more and more important in the past decades, since the regulations concerning environmental protection (pass-by noise) and also vehicle comfort expectations became more severe. Not only the production quality, the perfect material selection in the passenger compartment, etc., but also the acoustic impression became an important criterion of the quality of a vehicle. Also the low interior noise level, excellent speech intelligibility, excellent sounding of the HiFi system and the specific sound design of a vehicle became more and more important. Also the NVH experts have recognized soon that simulation methods enhance the development process of a vehicle. The goal of the simulations is the calculation, respectively the prediction of the NVH behaviour of a full vehicle or its components. The selection of the proper simulation method is based on the investigated frequency range, one certain method cannot used for the full acoustic frequency

K. Jálics (✉)
University of Miskolc, Miskolc, Hungary
e-mail: machijk@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_9

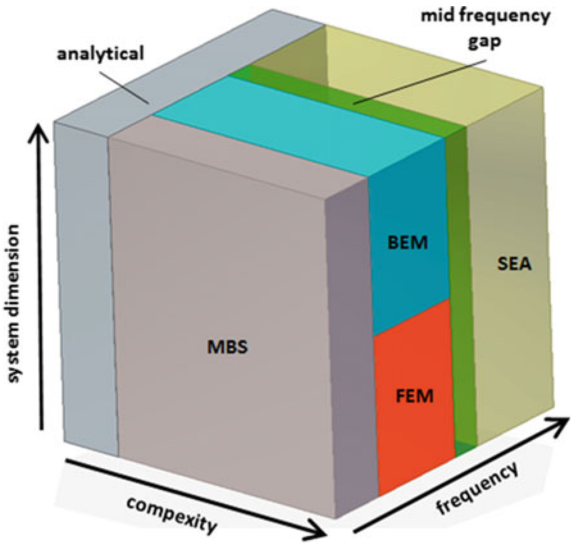


Fig. 1 Methods of the vehicle NVH simulation depending on the complexity, boundary conditions a frequency range [1]

range (0–20 kHz). Therefore 4–5 methods have to be involved successively, or the usage of the co-simulation has to be considered in order to cover the full domain. However, in general the simulation methods are currently not able to predict the perceived noise of the passengers even for a single operating condition.

A general brake down of the simulation methods used for vehicle NVH simulation depending on the frequency range of interest, complexity and system dimensions is shown in Fig. 1.

As an example for serial usage of different the NVH calculation methods of a rail vehicle can be given. In that case the rail—wheel interaction is calculated with MBS. The obtained forces are serving as excitations for the FEM calculation of the bogie. The vibration of the bogie exciting the vehicle body and the internal passenger cavity. The sound pressure level can be finally calculated by the SEA model. The main goal of this chapter is to represent the utilization of this method for NVH applications. Also the advantages and disadvantages of the individual methods will be pointed out.

2 Simulation Methods

2.1 General Aspects of a Simulation Method

In general the main task the simulation is the conversion of a real object, a complex system or a physical problem into a simplified mechanical/mathematical substitution model. Thereby the following terms should be fulfilled, which is basically a balance act between model accuracy and prediction quality:

- idealized assumption
- as accurate as possible
- as simple as possible.

The range of applications of the methods is of course limited through the simplification. As the next step mathematical equations will be set up. By discrete masses, damper and stiffness finite number (e.g. a few hundred) of differential equations or by continuous masses, damping and stiffness theoretically infinite number (e.g. in the practice even some million) of partial differential equations. Depending on the described aspects, the following methods can be used for the NVH simulation.

2.2 Multi-Body Simulation

As a non-continuous method for low frequency calculations the Multi-Body Simulation (MBS) has been established. MBS is generally used for the calculation of the dynamic system performance in the automotive engineering. The field of applications is vehicle dynamic, comfort (NVH), powertrain and engine design. The method uses typically discrete masses (rigid or flexible bodies) which are connected each other or to the surrounding with linear or non-linear coupling elements (spring or damper) or with kinematic constrains. The bodies may undergo large translational and rotational displacements. The bodies have max. 6 independent degrees of freedom which can be limited through boundary conditions or constraints. The displacement conditions are described with generalized coordinates with the number of the degrees of freedom. Mathematically described the system contains regular 2nd order differential equations as seen bellow:

$$M\ddot{q}(t) + D\dot{q}(t) + Kq(t) = F(t) \tag{1}$$

where ***M, D, K*** are time invariant matrices (n x n) for mass, damping and stiffness, $q = (q_1;\ q_2;\ \ldots\ q_n)$ time dependent general coordinates and $F(t)$ includes the excitation forces. By large systems or by higher frequencies the assumption of rigid bodies is no more acceptable. In that case the Finite Element Method (FEM) can be used.

2.3 Finite Element Method (FEM)

FEM is well known since several decades, and widely used in the engineering sciences. The FEM delivers reliable absolute results in the buckling, strength, heat transfer and fatigue analysis. FEM is also a state of the art method of the vehicle NVH simulation. It also provides reliable results concerning natural frequencies, transfer functions for a single part or for a not too complex system as well as in the higher frequency range ($\sim$1000 Hz). However, by a complex vehicle FEM model, which contains a structure (body, chassis, etc.) and also air cavities (passenger compartment, trunk, etc.) a coupled simulation (acoustic-structure) calculation is needed. In that case the frequency range is limited to a few hundred Hz (<300 Hz), since the complexity of the model is already enormous, the model can have several million degrees of freedom, which makes the computational effort expensive. In this chapter only the NVH aspects will be pointed out.

The simulation process starts with the division of the body into many small elements (ashlar, tetrahedron etc.). The discretization is very important at the NVH calculations [2]. The element length determines basically the upper frequency limit of the calculation. Equivalent to the Nyquist sampling theorem, every mode shape (waves) should be sampled with a sufficient number of degrees of freedom (or elements). The maximum element length depending on the requested frequency (f), the bending wave propagation speed (c_B) and N number of elements per wavelength is given as follows for structures:

$$d_{\max} = \frac{\lambda_B}{N} = \frac{c_B}{N f_B} \tag{2}$$

In the practice at least 6 elements per wavelength should be used which delivers a good agreement between accuracy and computation time. An example is shown in Fig. 2 for a steel plate with 2 mm thickness. If e.g. a calculation up to 500 Hz should be performed, with $N = 6$ and 10 elements per wavelength the maximum element length are 33 mm respectively 20 mm. The continuity conditions at the

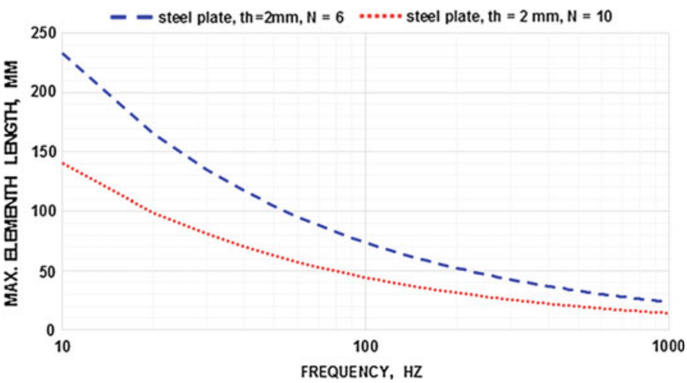


Fig. 2 Maximum element length depending on the frequency (for steel plate with 2 mm thickness, by 6 and 10 elements per wavelength)

bounding surfaces of the surface or volume elements lead to a linear equation system, similar to the Eq. (1) in the previous chapter. The solution of the homogenous equation system delivers the natural frequencies and the mode shapes.

2.4 Statistical Energy Analysis (SEA)

Through the statistical averaging of the modes the problem can be solved simpler. For that case the Statistical Energy analysis (SEA) delivers good results for the higher frequency range. SEA was developed already in the sixties of the last century for the calculation of dynamic behaviour of complex structures at high frequencies. Compared to deterministic methods (e.g. FEM) SEA derives average energetic values which means a major limitation of the method. As the method's name indicates, SEA provides a statistical framework and does not give the exact deterministic solution. In SEA 2 types of averaging will be considered:

- Frequency averaging: averaging of the modes, velocities, powers, etc. in a certain frequency band, general in third octave band
- Spatial averaging: averaging over the points of observation and excitation. All phase relevant information will be lost.

The mathematical principles of SEA can be illustrated as shown in Fig. 3. A system with 2 subsystems is given by the following power equilibrium equations [3, 4]:

$$P_i + P_{ji} = P_{ij} + P_{L,i} \tag{3}$$

$$P_j + P_{ij} = P_{ji} + P_{L,j} \tag{4}$$

where P_i and P_j are the input power to the subsystems, $P_{L,i}$, $P_{L,j}$ power losses (internal) and finally P_{ij} and P_{ji} are the transmitted power from one subsystem to the

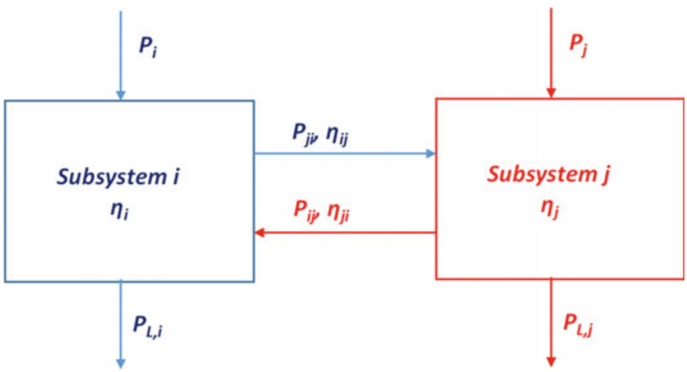


Fig. 3 Typical frequency ranges in the vehicle NVH calculations

other. Equation 2 can be written also in a generalized matrix form with the introduction of the modal density N_i:

$$\begin{bmatrix} P_i \\ \vdots \\ P_n \end{bmatrix} = \omega \begin{bmatrix} \eta_i N_i & \cdots & -\eta_{ni} N_n \\ \vdots & \ddots & \vdots \\ -\eta_{in} N_i & \cdots & \eta_n N_n \end{bmatrix} \begin{bmatrix} E_i/N_i \\ \vdots \\ E_n/N_n \end{bmatrix} \tag{5}$$

where $[P_i]$ is the vector of the input powers; $[E/N]$ is the vector of the modal energies of the subsystems; η_n are the internal loss factor (ILF), η_{nn} are the coupling loss factors (CLF) and N_i is the modal density. For most systems, especially large systems, many subsystems are not connected. This means that the matrix will contain many zero off-diagonal elements and therefore a sparse matrix. The input powers are assumed to be known. If the modal densities of the subsystems, the ILF and the CLF, are known, the total energies of subsystems can be determined. The ILFs are experimentally identified in most cases or extracted from a database. CLF also can be determined experimentally.

This method is mathematically easy and fast compared with FEM. There is a matrix to be solved and the elements will be directly related to the number of the subsystems. In this method, the number of modes per frequency band (or modal density) acts the major role and the accuracy of obtained results largely depends on it. Regarding the minimum required number of resonant frequencies in a certain frequency band, different authors suggested different values. In general several modes per frequency band deliver accurate results, however, with only one mode per band accurate results can be achieved, if the other interacting subsystems have a sufficient number of modes (≥ 3) in the same frequency band.

3 Hybrid Methods

For the low frequency range, the calculation of the modes is for example done using the Finite Element Method. In the high frequency domain, where high modal overlap occurs, SEA is better suited as mentioned in the introduction. A predicting gap exists in the mid frequency range. Furthermore, the structure-borne transmissions are not well predicted. To cover the mid frequency gap, e.g. a hybrid FEM/SEA method can be applied. (Also other combinations of methods, e.g. FEM/MBS, FEM/BEM, etc. existing.) The subsystem with the low modal density is modelled e.g. with FEM. The results will be coupled to the SEA calculation. An example is given in Fig. 4, where a vehicle firewall is shown for a transmission loss analysis. The exact shape of the firewall is built up in FEM, which is fully coupled to neighbouring chambers modelled with SEA. For this test the average SPLs of both chambers can be simply calculated in SEA.

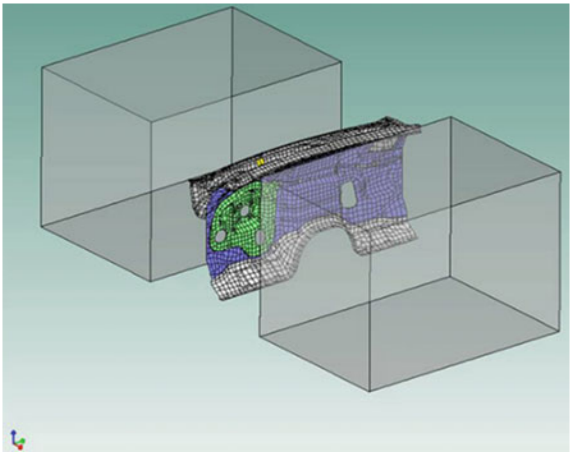


Fig. 4 Hybrid FEM/SEA model of a vehicle firewall for transmission loss calculations [5]

4 Summary

In the past chapters simulation methods were introduced, which are widely used in the vehicle NVH simulation. FEM is one of the most powerful and all-round method also in the vehicle NVH. Although there is a frequency limitation by approx. 300 Hz for the full vehicle coupled acoustic-structure simulation. Also the computational efforts (millions of DOFs) are quite high. SEA is mathematically easy and fast in comparison with FEM. One of the main advantages of SEA is that it can help to identify the major contributor in the overall energy of the system. Also it is easy to monitor the effects of changes in the design. The main disadvantages of the method are the loss of phase information, the results are valid for a subsystem and not for a certain point and low frequency problems because of the low modal density at low frequencies of the structures. MBS is widely used in the low frequency range, but the implementation of elastic parts condensed with the FEM can enhance the frequency range.

References

1. Zeller P (2012) Handbuch Fahrzeugakustik, Springer Verlag
2. Knothe K, Wessels H (2008) Finite Elemente—Einführung für Ingenieure, 4th edn. Springer Verlag, Berlin, Heidelberg
3. Lyon RH (1975) Statistical energy analysis of dynamical systems. MIT Press, Cambridge, Massachusetts
4. Lyon RH, Dejong RG (1995) Theory and application of statistical energy analysis, 2nd edn. Butterworth-Heinemann, Boston
5. https://www.esi-group.com/software-solutions/virtual-performance/va-one/va-one-modules/hybrid-module

Optimal Damping of Random Excited Systems

Ferenc Knopp

Abstract Through the case study of a simple mechanical system (mass, springs, damper) we analyse the effect of the damper coefficients on the output variance. The system described by a continuous differential equation is substituted with a discrete time (sampled) autoregressive moving average ARMA (3, 2) time series model. The displacement input (road profile) also characterized by an ARMA (p, q) stationary stochastic process. A variance transformation formula is derived which uses the (simulated) discrete impulse response function and the autocorrelations of the input signal. This formula is applied to the mechanical system searching the optimal damping which gives the minimal output variance.

1 Signal Transmission of Time Series

A discrete time **A**_uto_-**R**_egressive_ **M**_oving_ **A**_verage_ ARMA(n, m) time series model is defined below. The "a-members" at time k look back to the previous values of the y_k output signal and weigh them, while the "b-members" do the same with the u_k input signal.

$$y_k = (a_1 y_{k-1} + \ldots + a_n y_{k-n}) + (b_0 u_k + b_1 u_{k-1} + \ldots + b_m u_{k-m}) \quad (1)$$

Its stability condition and other features can be found in the literature [2, 5, 8, 9]. In our technical case study we shall see that a continuous ordinary differential equation also gives an ARMA model, if we apply discrete numerical approximations for the derivatives.

The output is the w_k ***impulse response*** *function* (weighing sequence) if $u_0 = 1$ and $u_1 = 0$, $u_2 = 0$, … namely the input is the discrete *unit impulse*. Every input

F. Knopp (✉)
Budapest University of Technology and Economics (BME), Budapest, Hungary
e-mail: knopp@mogi.bme.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_10

value starts an u_k-*stretched* impulse response and the superposition of them at a time k builds up the y_k output signal value.

$$y_k = w_k u_0 + w_{k-1} u_1 + \ldots + w_0 u_k = \sum_{i=0}^{k} w_{k-i} u_i = \sum_{i=0}^{k} w_i u_{k-i} \tag{2}$$

This is the so called ***convolution*** in which the direction of summation is changeable, and extendable to the infinite past. In the special case if the input is the $u_k = 1$ *unit jump* than we get the v_k ***step response*** *function* which is the sum of the impulse response function till the actual time k.

$$y_k = \sum_{i=0}^{\infty} w_i u_{k-i}\,(convolution) \quad v_k = \sum_{i=0}^{k} w_{k-i} \cdot 1\,(step\,response) \tag{3}$$

In simple cases the impulse response function can be expressed by an explicit formula, but generally is numerically created "whirling" the model and storing the values in a vector large enough. For example at the $y_k = a_1 y_{k-1} + b_0 u_k$ ARMA (1, 0) system $w_k = a_1^k b_0\,(k = 0,\ 1,\ 2, \ldots)$.

2 Transformation of Variances

Assuming *stationary* random signals with *zero mean values*: $\mathrm{M}[u_i] = 0$, $\mathrm{M}[y_i] = 0$. Then the l-distance *auto-covariance* and the normalized ***auto-correlation*** for the input series:

$$R_l = R_{uu}(l) = \mathrm{M}[u_i u_{i-l}] \quad R_{uu}(0) = D_u^2 \quad r_{uu}(l) = R_{uu}(l)/D_u^2 \tag{4}$$

Writing the *output* signal—created with infinite convolution—into the first row and first column of a table, multiplying and summing the table, we just get the *square* of it. (*Assuming the existence of the limit*)

y_k^2	$w_0 u_k$	$w_1 u_{k-1}$	$w_2 u_{k-2}$	$w_3 u_{k-3}$	…
$w_0 u_k$	$w_0^2\ u_k u_k$	$w_0 w_1\ u_k u_{k-1}$	$w_0 w_2\ u_k u_{k-2}$	$w_0 w_3\ u_k u_{k-3}$	…
$w_1 u_{k-1}$	$w_0 w_1\ u_{k-1} u_k$	$w_1^2\ u_{k-1} u_{k-1}$	$w_1 w_2\ u_{k-1} u_{k-2}$	$w_1 w_3\ u_{k-1} u_{k-3}$	…
$w_2 u_{k-2}$	$w_0 w_2\ u_{k-2} u_k$	$w_1 w_2\ u_{k-2} u_{k-1}$	$w_2^2\ u_{k-2} u_{k-2}$	$w_2 w_3\ u_{k-2} u_{k-3}$	…
$w_3 u_{k-3}$	$w_0 w_3\ u_{k-3} u_k$	$w_1 w_3\ u_{k-3} u_{k-1}$	$w_2 w_3\ u_{k-3} u_{k-2}$	$w_3^2\ u_{k-3} u_{k-3}$	…
…	…	…	…	…	…

Applying the ***main value*** operator on the cells of the table, the auto-covariances of the shifted signals appear. $\mathrm{M}[u_{k-i}u_{k-j}] = R_{uu}(i-j) = R_{i-j}$.

D_y^2					
	$w_0^2 \cdot R_0$	$w_0 w_1 \cdot R_1$	$w_0 w_2 \cdot R_2$	$w_0 w_3 \cdot R_3$	…
	$w_0 w_1 \cdot R_1$	$w_1^2 \cdot R_0$	$w_1 w_2 \cdot R_1$	$w_1 w_3 \cdot R_2$	…
	$w_0 w_2 \cdot R_2$	$w_1 w_2 \cdot R_1$	$w_2^2 \cdot R_0$	$w_2 w_3 \cdot R_1$	…
	$w_0 w_3 \cdot R_3$	$w_1 w_3 \cdot R_2$	$w_2 w_3 \cdot R_1$	$w_3^2 \cdot R_0$	…
	…	…	…	…	…

Summing this infinite table, we get the ***output variance*** D_y^2. From the main diagonal $R_0 = D_u^2$ can be taken out, and in the first two neighbour sub-diagonal R_1 is common, in the second two R_2, and so on… Taking out D_u^2 the general ***variance-transformation*** formula is ready. In the case of *independent* input (*white noise*) only the first sum remains, summing the main diagonal. (See also [4].)

$$D_y^2 = D_u^2 \cdot \left[\sum_{i=0}^{\infty} w_i^2 + 2 \cdot \sum_{l=1}^{\infty} \left(\sum_{i=0}^{\infty} w_i w_{i+l} \right) \cdot r_{uu}(l) \right] \tag{5}$$

Practically speaking a finite 50 × 50 table—assumed to be large enough—was used, because the impulse response function of a stable system goes to zero fast. A dynamically increasing table size is applicable with checking the relative accuracy.

3 Dynamic Models of a Mechanical System

The system in Fig. 1 can be seen as a (rotated) *vehicle suspension system* which is excited by the $u(t)$ road displacement as input signal. Applying the Newton's II. law of motion [1, 6, 7] for the mass and the weightless node, we get a system of second order differential equations for the $y(t)$ and $z(t)$ output signals:

$$\begin{aligned} m \cdot \ddot{y} &= c_2 \cdot (z-y) + d \cdot (\dot{z} - \dot{y}) \\ 0 \cdot \ddot{z} &= c_2 \cdot (y-z) + d \cdot (\dot{y} - \dot{z}) + c_1 \cdot (u-z) \end{aligned} \tag{6}$$

Ordering these left the $u(t)$ input signal appears on the right hand side.

$$\begin{aligned} m \cdot \ddot{y} + d \cdot \dot{y} + c_2 \cdot y - d \cdot \dot{z} \quad - c_2 \cdot z &= 0 \\ - d \cdot \dot{y} - c_2 \cdot y + d \cdot \dot{z} + (c_1 + c_2) \cdot z &= c_1 \cdot u \end{aligned} \tag{7}$$

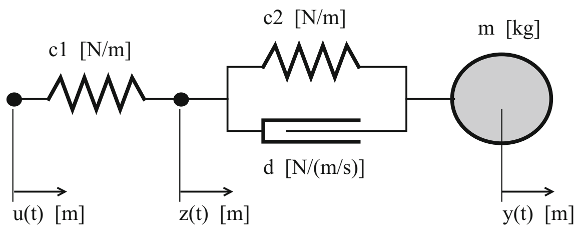


Fig. 1 A mechanical vibrating system

After Laplace transformation [3], there exists a system of algebraic equation for the transformed functions. From here the $y(s)$ global output can be expressed easily with the help of the Cramer's rule.

$$\begin{bmatrix} m\cdot s^2+d\cdot s+c_2 & -(d\cdot s+c_2) \\ -(d\cdot s+c_2) & d\cdot s+(c_1+c_2) \end{bmatrix}\begin{bmatrix} y(s) \\ z(s) \end{bmatrix}=\begin{bmatrix} 0 \\ c_1 \end{bmatrix}\cdot u(s) \tag{8}$$

$$y(s)=\frac{\begin{vmatrix} 0 & -(d\cdot s+c_2) \\ c_1 & (d\cdot s+c_2) \end{vmatrix}}{\begin{vmatrix} m\cdot s^2+d\cdot s+c_2 & -(d\cdot s+c_2) \\ -(d\cdot s+c_2) & d\cdot s+(c_1+c_2) \end{vmatrix}}\cdot u(s)=\frac{N(s)}{D(s)}\cdot u(s) \tag{9}$$

The denominator of the transfer function is the $D(s)$ *characteristic polynomial*. The coefficients of it are the output (left) side coefficients of the continuous differential equation. Similarly $N(s)$ shows the structure of the input (right) side.

After simplifications we find a *proportional-differentiating* (PD) system with *third order delay* (T3) after all our system is a (PD-T3) type.

$$\begin{aligned} &md\cdot\dddot{y}(t)+m(c_1+c_2)\cdot\ddot{y}(t)+c_1 d\cdot\dot{y}(t)+c_1c_2\cdot y(t)= \\ &=c_1c_2\cdot u(t)+c_1 d\cdot\dot{u}(t) \end{aligned} \tag{10}$$

The discrete numerical approximation over a $\Delta=\Delta t$ width four member grid y_{k-3}, y_{k-2}, y_{k-1}, y_k at its (virtual) centre:

$$\dot{y}(\,)\cong\frac{y_{k-1}-y_{k-2}}{\Delta}\qquad \dot{u}(\,)\cong\frac{u_{k-1}-u_{k-2}}{\Delta} \tag{11}$$

$$\ddot{y}(\,)\cong\frac{y_k-y_{k-1}-y_{k-2}+y_{k-3}}{2\cdot\Delta^2} \tag{12}$$

$$\dddot{y}(\,)\cong\frac{y_k-3y_{k-1}+3y_{k-2}-y_{k-3}}{\Delta^3} \tag{13}$$

Substituting these into the (10) continuous differential equation and grouping the coefficients of the grid values:

$$\begin{aligned}
& y_k \cdot \left(\frac{md}{\Delta^3} + \frac{m(c_1+c_2)}{2\cdot\Delta^2}\right) + \\
& + y_{k-1} \cdot \left(-\frac{3md}{\Delta^3} - \frac{m(c_1+c_2)}{2\cdot\Delta^2} + \frac{c_1 d}{\Delta} + \frac{c_1 c_2}{2}\right) \\
& + y_{k-2} \cdot \left(\frac{3md}{\Delta^3} - \frac{m(c_1+c_2)}{2\cdot\Delta^2} - \frac{c_1 d}{\Delta} + \frac{c_1 c_2}{2}\right) \\
& + y_{k-3} \cdot \left(-\frac{md}{\Delta^3} + \frac{m(c_1+c_2)}{2\cdot\Delta^2}\right) = \\
& = u_{k-1} \cdot \left(\frac{c_1 c_2}{2} + \frac{c_1 d}{\Delta}\right) + \\
& + u_{k-2} \cdot \left(\frac{c_1 c_2}{2} - \frac{c_1 d}{\Delta}\right)
\end{aligned} \tag{14}$$

Expressing y_k the derived discrete ARMA (3, 2) time series model:

$$y_k = (a_1 y_{k-1} + a_2 y_{k-2} + a_3 y_{k-3}) + (b_1 u_{k-1} + b_2 u_{k-2}) \tag{15}$$

4 Seeking the Optimal Damping

Figure 2 shows the running result of a QUICK BASIC computer program (under DOS, see Appendix). There are no comments on it, but the subroutine (SUB) names help to associate them to the functions and formulas. In *the Visual Basic environment, these* can be inserted as *modules* (without any graphical statement). They can be realized in other high level programming language too (C/C++ etc.).

Choosing the d [N/(m/s)] damping as ***free parameter*** we get a curve for the output variance with a ***minimum*** (Fig. 3). The smallest value $1.85 > 1$ so at these mechanical and road features the output variance always exceeds the input variance. At zero damping the two springs (c_1, c_2) are in serial switch (softer system), but if the damping goes to infinity (rigid rod) the first spring (c_1) works only (harder system).

Our purpose was to explain the theoretical basis of the problem. Further executed and desired investigations are not described here. The important question is the characterization of the road profile and connections to some “stochastic resonance”.

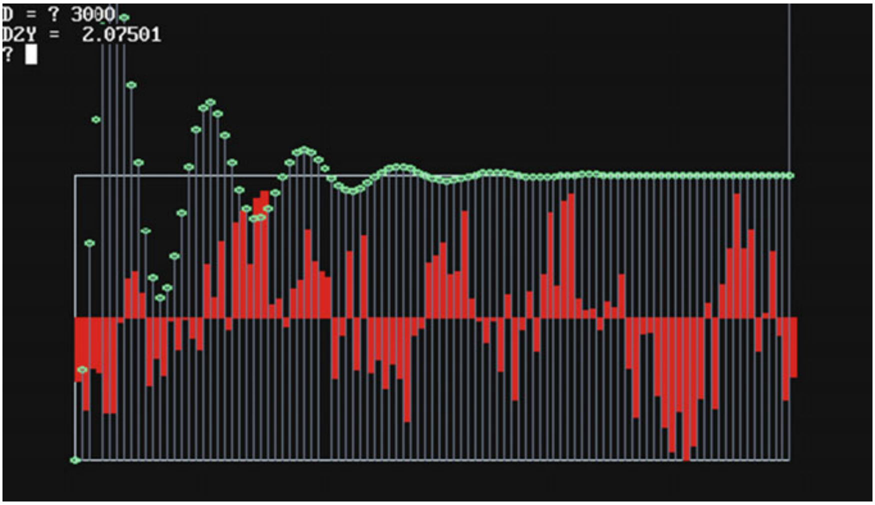


Fig. 2 Simulation of the discretized mechanical vibrating system. Step response (*green*): $m = 1000$, $c_1 = 10000$, $c_2 = 10000$, $d = 3000$, $\Delta = 0.2$, Input signal (*red*): $y_k = 0.5 \cdot y_{k-1} + 0.2 \cdot y_{k-1} + 2.65 \cdot u_k$

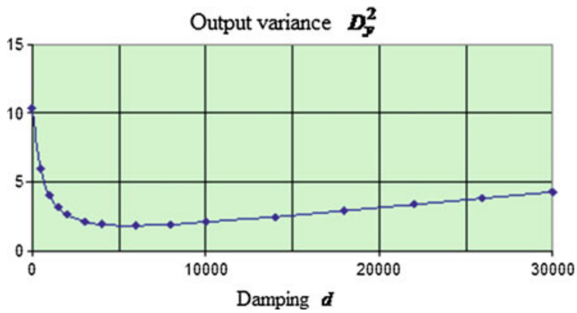


Fig. 3 Output variance in function of the damping

Appendix

```
'SZORAS.BAS
        'DEFDBL A-Z
        N%=3: M%=2: REDIM A(N%),B(M%):
        LONG%=100:  REDIM W(LONG%),V(LONG%),RUU(LONG%)
'       ...............................................
        A1=0.5: A2=0.2
        RUU(1)=A1/(1-A2)
        RUU(2)=(A2*(1-A2)+A1*A1)/(1-A2)
        B0=SQR(12*(1-A1*RUU(1)-A2*RUU(2)))
        FOR L=3 TO LONG%
            RUU(L)=A1*RUU(L-1)+A2*RUU(L-2): NEXT L
'           ---------------------------------------------
        SCREEN 9
        X0=50: Y0=120: X1=550: Y1=320
        DX=5
'           ---------------------------------------------
    110 REM
        CLS
        LINE (X0,Y0)-(X1,Y1),7,B
        FOR I=0 TO LONG%
            X=X0+I*DX
            U=(A1*U1+A2*U2)+B0*(RND-0.5):
            YU=Y1-100-U*50
            LINE (X,YU)-(X+DX,Y1-100),4,BF
            U2=U1: U1=U: NEXT I
        PRINT "D = ";: INPUT D
        M = 1000   '[kg]
        C1=10000   '[N/m]
        C2=10000   '[N/m]
        'D=  500   '[N/(m/s)]
        DT=0.2     '[s]

        'CALL ARMA22 (DT, M,C,D, A(),B())
'             ======          -   -
        CALL ARMA32 (DT, M,C1,C2,D, A(),B())
'            ======                 -   -
        GOTO 120
        FOR I=0 TO 2
            PRINT USING_
            "###   ##.####^^^^   ##.####^^^^";_
            I,A(I),B(I): NEXT I
    120 REM
        CALL WEIGHT (LONG%,N%,M%, A(),B(), W())
'            ======                      -
        CALL STEPR  (LONG%,N%,M%, A(),B(), V())
'            =====                       -
        FOR I=0 TO LONG%
            'PRINT USING_
            '"###   ##.####^^^^   ##.####^^^^";_
            'I,W(I),V(I):
            LINE   (X,YV)-(X,Y1),8
            X =X0+I*DX
            YV=Y1-V(I)*200
            YW=Y1-W(I)*200
            LINE   (X,YV)-(X,Y1),8
            FOR R=2 TO 3: CIRCLE (X,YV),R,10: NEXT R
            'CIRCLE (X,YW),3
            NEXT I

        D2U=1
        CALL OUTVAR (D2U,W(),RUU(), D2Y)
'       ---- ======                 *
        PRINT "D2Y = "; D2Y
        INPUT V$
        GOTO 110
```

```
'          ==================================================

           SUB ARMA22 (DT, M,C,D, A(1),B(1)) STATIC
'          --- ======                -    -
'              ----------------------------------------------
               A0= M/(DT*DT)+D/(2*DT)
               A1=-2*M/(DT*DT)+C
               A2= M/(DT*DT)-D/(2*DT)
               B0= D/(2*DT)
               B1= C
               B2=-D/(2*DT)
'              ----------------------------------------------
               A(1)=-A1/A0
               A(2)=-A2/A0
               B(0)= B0/A0
               B(1)= B1/A0
               B(2)= B2/A0
'              ----------------------------------------------
           END SUB

'          ==================================================

           SUB ARMA32 (DT, M,C1,C2,D, A(1),B(1)) STATIC
'          --- ======                    -    -
'              ----------------------------------------------
               A0= (M*D/DT+M*(C1+C2)/2)/(DT*DT)
               A1=-3*M*D/(DT^3)-M*(C1+C2)/(2*DT^2)_
                  +C1*D/DT+C1*C2/2
               A2= 3*M*D/(DT^3)-M*(C1+C2)/(2*DT^2)_
                  -C1*D/DT+C1*C2/2
               A3=-M*D/(DT^3)+M*(C1+C2)/(2*DT^2)
               B1= C1*C2/2+C1*D/DT
               B2= C1*C2/2-C1*D/DT
'              ----------------------------------------------
               A(1)=-A1/A0
               A(2)=-A2/A0
               A(3)=-A3/A0
               B(0)= 0
               B(1)= B1/A0
               B(2)= B2/A0
'              ----------------------------------------------
           END SUB

'          ==================================================

           SUB OUTVAR (D2U,W(1),RUU(1), D2Y) STATIC
'          --- ======                    *
               I2%=50: L2%=50
'              ----------------------------------------------
               SUM0=0
               FOR I%=0 TO I2%
                    WI=W(I%)
                    SUM0=SUM0+WI*WI: NEXT I%
'              ----------------------------------------------
               SUM2=0
               FOR L%=1 TO L2%
                   SUM1=0
                   FOR I%=0 TO I2%-L%
                       SUM1=SUM1+W(I%)*W(I%+L%): NEXT I%
                   SUM2=SUM2+SUM1*RUU(L%): NEXT L%
'              ----------------------------------------------
               D2Y=D2U*(SUM0+2*SUM2)
'              ----------------------------------------------
           END SUB

'          ==================================================
```

```
'        ==================================================

         SUB WEIGHT (LONG%,N%,M%, AN(1),BM(1), W(1)) STATIC
'        --- ======                            -
'            ----------------------------------------------
             NM%=N%: IF  M%>NM% THEN NM%=M%
             'PRINT "NM "NM%
             REDIM A(NM%),B(NM%)
'                  -      -
             FOR I%=1 TO N%
                 A(I%)=AN(I%): NEXT I%
             FOR I%=0 TO M%
                 B(I%)=BM(I%): NEXT I%
'            ----------------------------------------------
             W(0)=B(0)
             FOR K%=1 TO NM%
                 SUM=B(K%)
                 FOR I%=1 TO K%
                     SUM=SUM+A(I%)*W(K%-I%): NEXT I%
                 W(K%)=SUM: NEXT K%
             FOR K%=NM%+1 TO LONG%
                 SUM=0
                 FOR I%=1 TO NM%
                     SUM=SUM+A(I%)*W(K%-I%): NEXT I%
                 W(K%)=SUM: NEXT K%
'            ----------------------------------------------
         END SUB

'        ==================================================

         SUB STEPR (LONG%,N%,M%, AN(1),BM(1), V(1)) STATIC
'        --- =====                            -
             REDIM W(LONG%)
'            ----------------------------------------------
             CALL WEIGHT (LONG%,N%,M%, AN(),BM(), W())
'                 ======                        -
             V(0)=W(0)
             FOR I%=1 TO LONG%
                 V(I%)=V(I%-1)+W(I%)
             NEXT I%
'            ----------------------------------------------
         END SUB

'        ==================================================
```

References

1. Bokor J, Nándori E, Várlaki P (2000) Studies in vehicle engineering and transportation science. Hungarian Academy of Sciences and BME, Budapest (Hungary)
2. Box EPG, Gwilym MJ Time series analysis forecasting and control, Holden-Day
3. Fodor G (1967) Lineáris rendszerek analizise. Műszaki Könyvkiadó, Analyse of Linear Systems (in Hungarian)
4. Lantos B (2001) Theory and design of control systems I. Akadémiai Kiadó, (in Hungarian)
5. Michelberger P, Szeidl L, Várlaki P (2001) Applied process statistics and time series analysis. Typotex Kiadó, (in Hungarian)
6. Sályi B, Michelberger P, Sályi I (1991) Kinematics and kinetics Tankönyvkiadó, (in Hungarian)
7. Thomson TW (1972) Theory of vibration with applications. Prentice-Hall, Inc.
8. Tusnádi G, Ziermann M (1986) Time series analyses. Műszaki Könyvkiadó, (in Hungarian)
9. Várlaki P (1986) Introduction to the statistical system identification. Műszaki Kiadó, (in Hungarian)

Application of Knowledge-Based Design in Computer Aided Product Development

György Hegedűs

Abstract In the recent years, it can be observed that the products are launched by rapidity and short time. The lifecycles of user equipment has been shortened, the flow of information accelerated, the freely available information more widely accessible. One of the conditions of competitiveness for the arisen problems is the rapid and optimal solutions. Using a software or system, which supports the knowledge based design (KBD) the time and costs of the product development phase can be reduced. In this chapter the results of a product design and development process implemented in a PLM system will be presented on a ball screw drive mechanisms focusing on its returning guide.

Keywords Knowledge-Based design (KBD) · Product development · Computer aided design (CAD)

1 Introduction

In the last decades the computer aided design was the answer for quicker solutions. These systems and software supported the geometric modelling and manufacturing planning, especially (CAD/CAM). Due to the development the integrated systems were launched, which included the different simulation and analysing capabilities of the products. Nowadays the application of product lifecycle management (PLM) systems is an essential part of the product development. One of the advantages of these systems that knowledge-based design supporting modules are available, it is possible to create user-defined macros and algorithms, furthermore decision making and selection criteria arise during the design can be defined previously. Taking advantage of these properties the time of product design and development can be reduced significantly.

G. Hegedűs (✉)
University of Miskolc, Miskolc, Hungary
e-mail: hegedus.gyorgy@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_11

2 Brief Overview of Knowledge-Based Design

Knowledge-based design focuses on the product design, development and its related procedures [1]. In a KBD environment the design process can be accelerated by reusing previously defined techniques, methods and procedures and these can be integrated into the desired tasks or workflows which are involved in the product development processes. The flexibility and complexity of knowledge-based design environment largely depend on the reused geometrical data, previously defined expressions and algorithms (see Fig. 1).

Formerly these environments included non-parametrised geometries (for example: neutral *.stp, *.iges, or converted CAD files). Complex CAD software support the parametrised surface, solid and assembly models, thus the knowledge based design and development can be achieved by the reused parametrised geometry. In product data and product lifecycle management software the mathematical expressions and logical relations or functions extend the capability of the knowledge-based environment. Most CAD, PDM and PLM software support user defined applications are developed in different programming languages such as C/C++, C#, Visual Basic, Java and specific built-in programming language [3]. Users can create their own applications that automatically perform repetitive tasks, accelerate design procedures, and automatically generate complex geometries. In such a case the user defined software obtains the parameters, equations, logical relations and functions from the CAD system and evaluates the numerical computations, algorithms end geometry creations. After that the evaluated parameters will be set into the CAD system. The flexibility, interactivity and time saving of repetitive tasks are the advantages of these KBD applications, however the disadvantages of these KBD environments are that all of the computations, numerical algorithms and communications between the CAD system and user defined software have to be programmed [4–7].

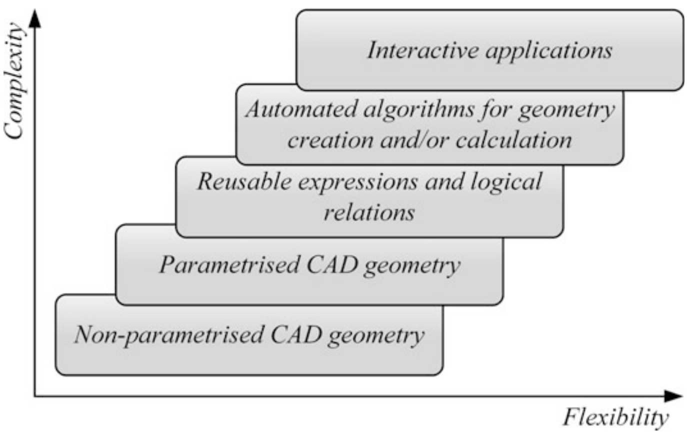


Fig. 1 Flexibility and complexity of a KBD application [2]

Knowledge-based design methods and tools can include rigid or variable geometry data, the integration of calculation and simulation procedures into the design process, or the application of problem-oriented software solutions that can be integrated into the design environment. Although one main advantage of knowledge-based design is that it makes existing, proven solutions available for specific tasks, this also involves some drawbacks that must be acknowledged [2, 8].

3 Application of Ball Screw Drives in Automotive Equipment

Gothic-arc profile ball screw motion transforming mechanisms are widely used to transform the rotational movement into linear movement and vice versa. Due to the automotive development these motion transformation mechanisms have been appeared in the electrically operated steering systems.

The electrically operated steering systems (Electric Power Steering, EPS) have introduced in passenger car steering systems during the last years. The use of these systems was originally limited to smaller vehicles, because the power density of the electronic parts and the energy available from the on-board wiring was not sufficient to serve bigger vehicles and higher steering powers.

New technologies enable the general use of EPS in the superclass now. The advantage of electric power-assisted steering compared to hydraulic power steering is that it is activated only when needed (energy is fed only when the car is steered). Due to this the energy waste and CO_2 emission are less.

The paraxial drive (EPSapa) is characterized by low system friction and high efficiency. The possible customer applications range from dynamic sports cars and upper mid-size cars to high load vehicles such as off-road vehicles and vans. Due to the combination of ball screw mechanism and toothed belt drive with paraxial drive is ideally suited for the customer's differing performance requirements. The wide range of positioning possibilities of the servo unit allows optimum use of the installation space on the vehicle. The linear movement of the steering rack is generated from the rotational movement of the electric motor combined with a toothed belt drive and a ball screw mechanism (see Fig. 2). The ball screw mechanism generates low noise and rigid connection between the steering gear and the vehicle body is possible [10].

Another solution of the application of ball screw drive mechanisms in EPS systems is the so-called rack-concentric steering system (EPSrc). The toothed belt drive is eliminated and the ball nut is connected with a hollow shaft and directly driven by an electric motor (see Fig. 3). The concentric arrangement of the hollow shaft requires a special servo motor, because the rack of the steering passes through the motor. The main advantage of the EPSrc system is its compactness, the disadvantage that due to the missing gear multiplication of the toothed belt drive the electric motor of an EPSrc has to produce a twice as high torque at the same output power level.

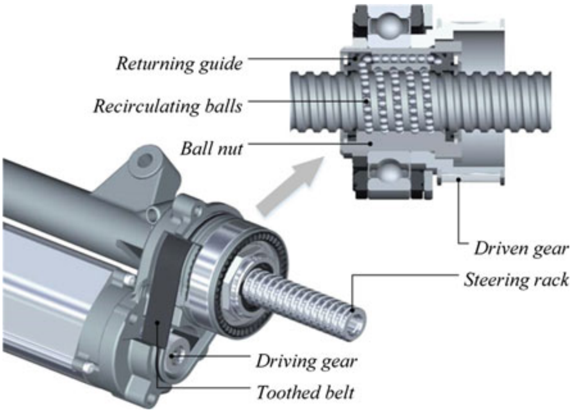


Fig. 2 Electric power steering with paraxial drive (EPSapa) [9]

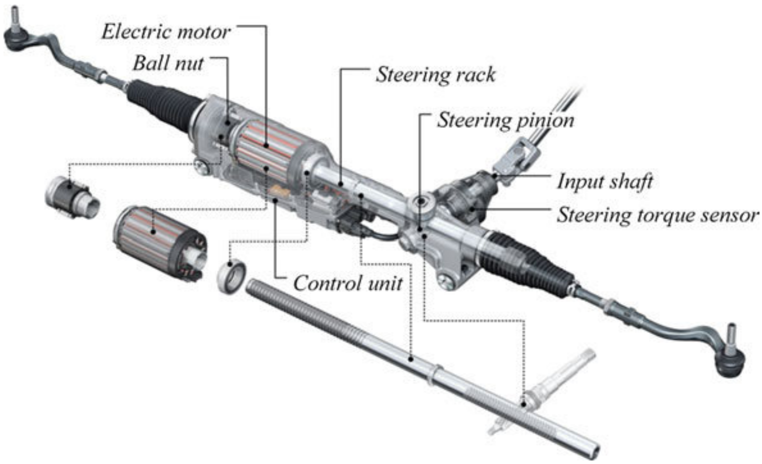


Fig. 3 Electric power steering with rack concentric drive (EPSrc) [11]

4 Application of Ball Screw Drives in Automotive Equipment

The customer's demand requires the development of ball screw drive mechanisms. Gothic-arc profile ball screw motion transforming mechanisms are commonly used in machine tools and the demand for high-lead ball screws is increasing due to high-speed manufacturing. The one of the problem of the high-lead ball screw drive is its manufacturing (determination of the profile of form grinding tool, collision avoidance of the quill and ball nut during manufacturing). The lead of ball screw mechanisms in an EPS system has a typical value of 5–10 mm. On low-lead ball

Table 1 The main design parameters of ball nut

Parameter	Description	Unit
P	Pitch of ball screw drives	mm
D_3	Internal diameter of ball nut body	mm
L_n	Length of thread	mm
D_{pw}	Pitch circle diameter	mm
R_{pr}	Radius of profile arc	mm
c	Centre shift of gothic arcs	mm
D_w	Ball diameter	mm

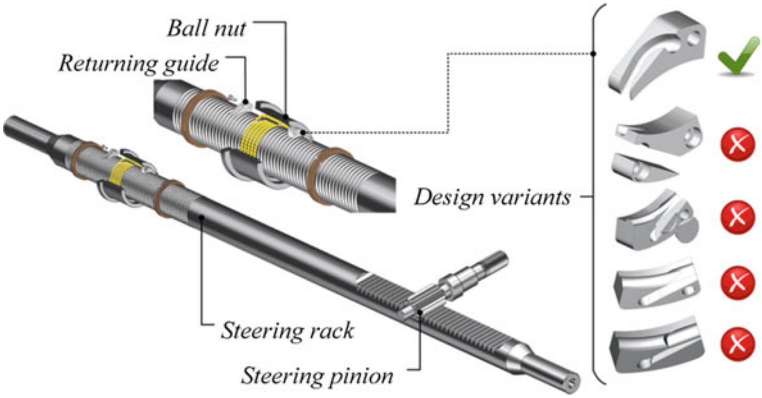


Fig. 4 Redesigned returning guides with KBD

screw drive mechanism the main problem is the arrangements of returning guides in the ball nut. Therefore new returning guides were designed to create a knowledge-based environment.

The KBD environment consists the reused design parameters of ball nut (see Table 1), the logical rules and functions.

The results of the new returning guide design variants are shown in Fig. 4, where the optimal solution was selected by value analysis.

5 Conclusions

This chapter has presented a product development process on the returning guide of a ball screw drive mechanism, which would be applied in automotive steering system. To reduce the time and cost of the development process a knowledge-based design environment was developed in a PLM software. Due to the favourable advantages of KBD system the time of the returning guide design variants generation is reduced and the selection of optimal solution can be carried out easier.

Acknowledgements This research was carried out as part of the TÁMOP-4.2.1.B-10/2/KONV-2010-0001 project with support by the European Union, co-financed by the European Social Fund. This research was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP-4.2.4.A/2-11/1-2012-0001 'National Excellence Program'.

References

1. Hirz M, Harrich A, Rossbacher P (2011) Advanced computer aided design methods for integrated virtual product development processes. Comput-Aided Des Appl 8(6):901–913. doi:10.3722/cadaps.2011.901-913
2. Hirz M, Dietrich W, Gfrerrer A, Lang J (2013) Integrated computer-aided design in automotive development: development processes, geometric fundamentals, methods of CAD, knowledge-based engineering data management. SpringerLink, Bücher. Springer Berlin Heidelberg. ISBN:9783642119408
3. Rocca GL (2012) Knowledge based engineering: between AI and CAD. Review of a language based technology to support engineering design. Adv Eng Inform 26(2):159–179. ISSN:1474-0346. doi:http://dx.doi.org/10.1016/j.aei.2012.02.002
4. Tzivelekis CA, Yiotis LS, Fountas NA, Krimpenis AA (2015) Parametrically automated 3D design and manufacturing for spiral-type free-form models in an interactive CAD/CAM environment. Int J Interact Des Manufac (IJIDeM) 1–10. ISSN:1955-2505. doi:10.1007/s12008-015-0261-8
5. Hegedűs Gy (2016) Newton's method based collision avoidance in a CAD environment on ball nut grinding. Int J Adv Manuf Technol 84(5):1219–1228. ISSN:1433-3015. doi:10.1007/s00170-015-7796-5
6. Gomes S, Bluntzer JB, Mahdjoub M, Sagot JC (2007) Collaborative, functional and knowledge based engineering using a PLM environment. In: DS 42: Proceedings of ICED 2007, the 16th international conference on engineering design, Paris, p 12
7. Catic A, Malmqvist J (2007) Towards integration of KBE and PLM. In: DS 42: Proceedings of ICED 2007, the 16th international conference on engineering design, Paris, p 12
8. Jayakiran Reddy E, Sridhar CNV, Pandu Rangadu V (2015) Knowledge based engineering: notion, approaches and future trends. Am J Intell Syst 5(1):1–17. doi:10.5923/j.ajis.20150501.01
9. Lenksysteme ZF (2013) ZF servolectric—electric power steering system for passenger cars and light commercial vehicles. Brochure
10. Harrer M, Pfeffer P (2017) Steering handbook. Springer International Publishing. ISBN:978-3-319-05449-0. doi:10.1007/978-3-319-05449-0
11. Bochert A, Cäsar T, Eichler J (2010) Steering system. ATZextra worldwide 15(11):140–143. ISSN:2195-1489. doi:10.1365/s40111-010-0252-5

Elementary Calculations for Deflection of Circular Rings

Géza Németh

Abstract The traction drives requires initial tension. Let us imagine such kind of epicyclic drive or harmonic drive. In both cases the annular wheel is made of spring steel and its load is bending. The stiffness of the ring is large enough to produce the required normal force between the annular and the planetary wheels or between the ring and the flexible, cup shaped wheel. The transmission ratio depends on the ratio of planetary and sun wheel in the planetary drive, or the difference between the circumferences of ring wheel and the cup shaped wave wheel on the harmonic drive. Both the deflection of ring wheel and the number of planetary wheels or the number of deflection wave increase the load carrying capacity and decrease the bending moment. The problem of deflection can be managed by differential equation or any kind of energy method. The analysis of equivalence of the outcomes closes the paper.

1 Introduction

The well-known epicyclic gear drive and harmonic gear drive are produced in large series by many manufacturers. Their load carrying capacity, accuracy and the large transmission ratios suggest their application fields. These drives are used mainly in heavy drives, robots and vehicles, and also in kinematic drives. The latter application field usually requires less torque capacity and less accuracy.

The kinematic drives can be epicyclic traction drives and also traction drives with generated deflection waves that similar to the harmonic drives. The gear drives are shape closing ones, and the traction drives are force closing. The gear drives do not require initial tension and can transmit large tangential forces. It is true significantly in the case of harmonic gear drives, where the number of teeth meshing simultaneously is huge. It is worth to notice, that the amplitude of deflection waves

G. Németh (✉)
University of Miskolc, Miskolc, Hungary
e-mail: machng@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_12

necessary for the pairs of gears to leave or enter the meshing is relatively large. In case of traction drive there are no such constraints to increase the wave amplitude.

There are two distinct problems where the friction drive is analysed in this paper. There is an epicyclic traction drive with one outer and one inner connection and there is a harmonic traction drive. Both drives have a ring wheel with relatively large stiffness and usually fixed to the stand. The ring wheel is loaded radially by N number of radial forces, F_r and the same number of tangential forces F_t which are less than the maximum frictional forces, $F_t < \mu_0 F_r$, where μ_0 is the coefficient of static friction between the ring wheel and the mating element $N = 2, 3, 4, \ldots$. The mating elements are the planet wheels or the flexible cup shaped wheel, depending on the type of drive. The number of planet wheels, but the number of waves influences the transmission ratio. A friction drive can transmit sufficient load only with large compressive forces or greater number of compressive forces.

Our purpose is to develop a kinematic traction drive with a relatively large transmission ratio also for a relatively large torque. The compressive force between the mating elements should be assured by the elastic deflection of the ring wheel. Similar problem was analysed in our previous papers [1, 2]. The ring wheel was a helical torsion spring there and it was analysed by a spatial model. When the cross-section $b \cdot h$, the mean diameter d and the material properties E and ν of the ring wheel is known, and also the number of mating planet wheels (in case of epicyclic drive) or the number of waves (in case of harmonic drives), we can estimate the necessary range of deflecting on.

Let us analyse a circular ring with rectangular cross-section, loaded equally with an equally distributed arbitrary number of concentrated radial forces F_r, shown in Fig. 1. Do find the relation between the force and the deflection. We suggest two distinct solutions for this problem. Both are well known in the literature [3]. The first one is the deflection curve for a bar with a circular centre line, the second one is the deflection of the curved bar calculated by the Castigliano's theorem. The rate of thickness and radius of the ring is relatively small, so the ring is considered as a thin one.

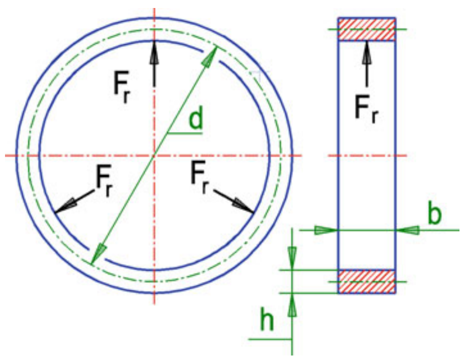


Fig. 1 Circular ring loaded by equally distributed radial forces

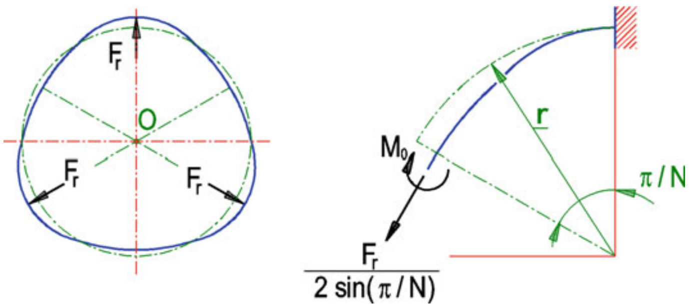


Fig. 2 Models for deflection curve and for Castigliano's theorem

It is also interesting to find relationship between the calculated results. The equation of the deflection curve obtained from differential equation is available in a polar coordinate system, and its origin is the centre of the circular centre line O. The Castigliano's theorem provides its result also in polar coordinate system, but the problem is triple statically indeterminate, and due to symmetry we analyse only a range of π/N arc. One end of the arc is built in, the other is free. In this model the fixed end has no any deflection or slope of the deflection curve. Otherwise the Castigliano's theorem can be applied more conveniently, but its results are not in the appropriate form, that is why the conversion is advisable. The two models are shown by Fig. 2.

2 The Mechanical Model

2.1 The Bending Moment Along the Circular Centre Line

A thin closed bar with circular centre line is loaded by an N number of equally distributed, equal concentrated forces, F acting radially outwards. The maximum value of radial deflection is expected at the locations of forces and the maximum negative radial deflection is expected in the middle of two equal forces. Due to symmetry it is advisable to analyse a π/N arc instead of the full circle. Our assumption is that the strain energy due to shear is negligible to that of the bending.

In the model, one end of the arc of the range of π/N angle is built in, its radius, r is constant, the other end is free and loaded by a tangentially acting force, $F/[2\sin(\pi/N)]$ and a concentrated bending moment, M_0. The arc coordinate, $s = r\varphi$ is started at the free end of the arc. The bending moment function along the arc is

$$M = -M_0 + \frac{Fr}{2}\frac{(1 - cos\,\varphi)}{\sin\left(\frac{\pi}{N}\right)}. \qquad (1)$$

The free end has a deflection but it has no slope. Considering this constraint, the bending moment, M_0 is computable and the bending moment function (1) can be expressed as

$$M = \frac{Fr}{2}\left[\frac{N}{\pi} - \frac{cos\,\varphi}{\sin\left(\frac{\pi}{N}\right)}\right], \tag{2}$$

and its value at $\varphi = \pi/N$ is the maximum one,

$$M_{max} = \frac{Fr}{2}\left[\frac{N}{\pi} - cot\left(\frac{\pi}{N}\right)\right]. \tag{3}$$

2.2 The Deflection Curve

The differential equation of the deflection curve for a bar with a circular centre line, considering the results of [1] and the bending moment function (2), where the radial deflection, u is the wanted variable, is

$$\frac{d^2u}{d\varphi^2} + u = \frac{Fr^3}{2IE}\left[\frac{cos\,\varphi}{\sin\left(\frac{\pi}{N}\right)} - \frac{N}{\pi}\right]. \tag{4}$$

Having solved the Eq. (4) we obtain the function of radial deflection as

$$u(\varphi) = \frac{Fr^3}{2IE}\left[-\frac{N}{\pi} + \frac{\varphi\,sin\,\varphi}{2\sin\left(\frac{\pi}{N}\right)} + \frac{cos\,\varphi}{2\sin\left(\frac{\pi}{N}\right)}\left(1 + \frac{\pi}{N}cot\left(\frac{\pi}{N}\right)\right)\right]. \tag{5}$$

Due to symmetries, it is sufficient to consider the range of arc coordinate as $\varphi\epsilon(0,\ \pi/N)$. The positive values represent the radial increment of the points of bar measured from the centre of the circular bar.

2.3 The Deflection by Strain Energy

The Castigliano's theorem is suitable for the solution of statically indeterminable problems of bars. Due to symmetry, only a part of the full circle should be considered, visualized by Fig. 3. The angle of arc is $2N$th part of the whole 2π.

To obtain the deflections at any point of the beam in the function of φ, beside the bending moment from the real loads (2), denoting by M_{real}, The other part of the imagined loads should also be considered. The sum of these components is

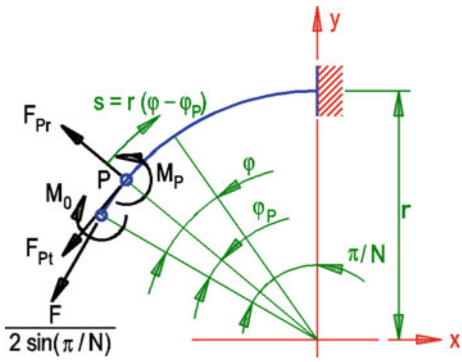


Fig. 3 The cantilever beam model for generating the deflection and slope functions

$$M = M_{real} + F_{Pt} m_t + F_{Pr} m_r + M_P m_\psi, \tag{6}$$

where the so called relative bending moment functions are the partial derivatives of the whole bending moment and their values are

$$\begin{aligned} m_t &= r[1 - \cos(\varphi - \varphi_P)], \\ m_r &= r\,\sin(\varphi - \varphi_P), \\ m_\psi &= 1. \end{aligned} \tag{7}$$

Designating the generalized deflections by δ_i and the relative bending moment functions by m_i, where $i = 1, 2, 3$, the general formula for the deflections at any point of φ_P is

$$\delta_i(\varphi_P) = \frac{r}{IE} \int\limits_{\varphi=\varphi_P}^{\frac{\pi}{N}} M_{real}(\varphi) m_i(\varphi) d\varphi. \tag{8}$$

Having been carried out the integrations when one end of the beam at the angle of π/N is built in, and the independent variable is φ again, instead of φ_P, the radial and tangential deflections are

$$\delta_r(\varphi) = \frac{Fr^3}{2IE}\left\{\frac{1}{2}\sin\left(\frac{\pi}{N} - \varphi\right) - \frac{N}{\pi}\left[1 - \cos\left(\frac{\pi}{N} - \varphi\right)\right] - \left(\frac{\pi}{N} - \varphi\right)\frac{\sin\,\varphi}{2\sin(\pi/N)}\right\}, \tag{9}$$

$$\delta_t(\varphi) = \frac{Fr^3}{2IE}\left\{-\frac{N}{\pi}\varphi + \left[\frac{1}{2}\cot\left(\frac{\pi}{N}\right) - \frac{N}{\pi}\right]\sin\left(\frac{\pi}{N} - \varphi\right) + \frac{\sin\,\varphi}{\sin(\pi/N)} + \frac{\left(\frac{\pi}{N} - \varphi\right)}{2\sin(\pi/N)}\cos\,\varphi\right\}, \tag{10}$$

and the slope at any cross section is

$$\delta_\psi(\varphi) = \frac{Fr^2}{2IE}\left\{\frac{\sin\varphi}{\sin(\pi/N)} - \frac{N}{\pi}\varphi\right\}. \tag{11}$$

2.4 The Equivalence of the Formulae

Comparing the formulae (9) and (5) of the radial deflections, the differences between them is obvious. And this is the expected result, too. The model of the Castigliano's theorem having a built in end and another constraint for the slope on the free end, cannot supply the same values as the solution of the differential equation. The equivalence of the formulae is supposed. It is important to prove this fact. The task is interesting, because the handling of the problem is easier by Castigliano's theorem and on the other hand, the deflections can be more conveniently used, when the fixed point is the centre of the original circular bar.

Let us designate the latter deflections by $u(\varphi)$ and $v(\varphi)$. There are three curves of angle range π/N should be compared and located into the same coordinate system, shown in Fig. 4. The first one is a circle with the centre point of O and constant radius, r. The second one is the deflection curve with the same fixed point and variable radius. The third one is the deflected cantilever of the built in circular bar. Studying these curves the following geometrical formulae can be written for the radial and tangential deflections.

$$\cos\left(\frac{\pi}{N} - \varphi\right) = \frac{u(\varphi) - \delta_r(\varphi)}{\frac{\delta_t(0)}{\sin(\pi/N)}}, \tag{12}$$

$$\sin\left(\frac{\pi}{N} - \varphi\right) = \frac{\delta_t(\varphi) - v(\varphi)}{\frac{\delta_t(0)}{\sin(\pi/N)}}, \tag{13}$$

where $\delta_t(0)$ is the tangential deflection at the free end of the built in model that is shown at Fig. 4. Substituting these transforming formulae in (9) and (10) the

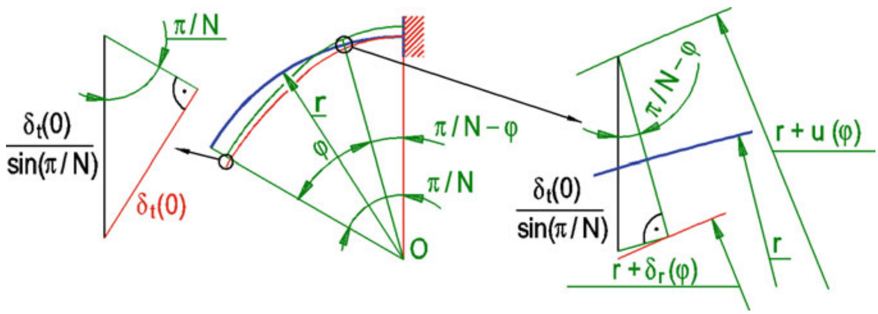


Fig. 4 Deflections at the different coordinate systems

deflections are available in polar coordinate system with starting point of the centre of circular line, O.

$$u(\varphi) = \frac{Fr^3}{2IE}\left\{-\frac{N}{\pi} + \left[1 + \frac{\pi}{N}\cot\left(\frac{\pi}{N}\right)\right]\frac{\cos\,\varphi}{2\sin\left(\frac{\pi}{N}\right)} + \frac{\varphi\,\sin\,\varphi}{2\sin\left(\frac{\pi}{N}\right)}\right\}, \tag{14}$$

$$v(\varphi) = \frac{Fr^3}{2IE}\left\{-\frac{N}{\pi}\varphi + \left[1 + \frac{\pi}{2N}\cot\left(\frac{\pi}{N}\right)\right]\frac{\sin\,\varphi}{\sin\left(\frac{\pi}{N}\right)} - \frac{\varphi\,\cos\,\varphi}{2\sin\left(\frac{\pi}{N}\right)}\right\}, \tag{15}$$

$$\vartheta(\varphi) = \frac{Fr^2}{2IE}\left\{\frac{\sin\,\varphi}{\sin(\pi/N)} - \frac{N}{\pi}\varphi\right\}. \tag{16}$$

The formula (14) is really the same as formula (5) and the (15) also fulfil the constraints and the experiences. The maximum value of the curvature, κ_{max} is expected at the location of the original force application, at $\varphi = \pi/N$. Approximating the deflected curve as $\rho(\varphi) = r + u(\varphi)$, and considering the value of the first derivative at that point to be zero, $\rho'(\pi/N) = 0$, the equation of curvature is fairy simple.

$$\kappa_{max} = \frac{\rho - \rho''}{\rho^2} = \frac{\rho + \left(\frac{\pi}{N} - \frac{N}{\pi}\right)a + a\frac{\pi}{N}\cot^2\left(\frac{\pi}{N}\right)}{\left\{\rho - \frac{N}{\pi}a + \frac{a}{2}\left\{\cot\left(\frac{\pi}{N}\right) + \frac{\pi}{N}\left[1 + \cot^2\left(\frac{\pi}{N}\right)\right]\right\}\right\}^2}, \tag{17}$$

where $a = Fr^3/(2IE)$. Equation (17) gives useful geometrical data, too.

3 Conclusion

The deflection of discrete points or the deflection curve for a bar with the circular centre line, but for arbitrary number of loads are available in the literature. The necessity of these equations is important to design a ring wheel for a simple epicyclic traction drive or a harmonic traction drive. The analyses suggested the equations in the convenient form. The formulae obtained from the model is suitable for a preliminary calculation of some machine elements both in epicyclic and harmonic traction drives.

Acknowledgements The research work presented in this paper/study/etc. based on the results achieved within the TÁMOP-4.2.1.B-10/2/KONV-2010-0001 project and carried out as part of the TÁMOP-4.1.1.C-12/1/KONV-2012-0002 project in the framework of the New Széchenyi Plan. The realization of this project is supported by the European Union, and co-financed by the European Social Fund. The research work was supported by the Hungarian Scientific Research Found grants OTKA 29326 and Fund for the Development of Higher Education FKFP 8/2000 project.

References

1. Németh G, Németh N, Péter J (2015) Strength calculation of elastic elements in epicyclic traction drive (in Hungarian). In: Sándor B (ed) Műszaki Tudomány az Észak-Kelet Magyarországi Régióban 2015 Konferencia előadásai, Debreceni Akadémiai Bizottság Műszaki Szakbizottsága, pp 213–219. ISBN:978-963-7064-32-6
2. Németh G, Péter J, Németh N (2013) A new type of epicyclic traction drive. Adv Mech Eng 1 (1):137–142. ISBN:978-963-473-623-3
3. Timoshenko S (1968) Strength of materials, elementary theory and problems. Princeton, New Jersey, pp 442

Position- and Speed-Dependent, Power-Absorbing Hydraulic Cylinder with Mathematically Predictable Characteristics

Lajos Albrecht, Ferenc Mészáros, Szilárd Szabó and Balázs Barna

Abstract Within the framework of the research a position- and speed-dependent, power-absorbing hydraulic cylinder were constructed based on laboratory tests and finite element numerical simulation studies. The elaborated cylinder construction has a mathematically predictable characteristics and based on this feature can be used for different applications in the field of kinetic energy absorbing.

Keywords Hydraulic power absorbing · Mathematically predictable characteristics · Articulation system

1 Introduction

Utilizing the funds won in tender no. KMOP-1.1.1-08/1-2008-0050, between 2009 and 2013, the associates of SOLID4D Kft. carried out the development of a hydraulic cylinder that is suitable for application as a power-absorbing unit of special characteristics depending on the displacement, with a very simple structure [1]. Its special features and simple structure makes is possible to use this product in a wide range of application areas; for example, in the machine and automotive

L. Albrecht (✉)
Mechanical and Transport Engineering, University of Miskolc, Miskolc, Hungary
e-mail: lajos.albrecht@gmail.com

F. Mészáros
Industrial Engineering, University of Miskolc, Miskolc, Hungary
e-mail: ferenc.meszaros@innospectrum.hu

S. Szabó
Department of Fluid and Heat Engineering, University of Miskolc, Miskolc, Hungary
e-mail: aram2xsz@uni-miskolc.hu

B. Barna
Mechanical Engineering, Department of Machine Tools, University of Miskolc, Miskolc, Hungary
e-mail: barna.balazs@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_13

industry, too. In the last 3 years (2013–2016), progress was made in the field of utilization concerning development and production, furthermore, a new type of articulation systems for trailers and sliding articulated buses was tested.

2 Theoretical Bases

The clarification of theoretical questions and the mathematical formulation of the solution took place as a solution for a problem occurred in connection with a research and development topic. The basic idea was to utilize the dynamic pressure forming during the movement of the piston in the hydraulic cylinders, and to make the dynamic run of the pressure, occurring in the cylinder, predictable by regulating the outflow by a variable gap in the cylinder. The preliminary mathematical analyses revealed the special behaviour according to which, by applying a properly designed, changing throttle gap led by piston displacement, a power-absorbing hydraulic cylinder (bumper, shock absorber) can be created which automatically adapts to the load (depending on the speed and position of the piston). The run of the braking force—as a function of the current speed of the piston—changes automatically so that during its stroke length, the cylinder absorbs the kinetic energy of the colliding object under all circumstances. The forces occurring during the collision process are the smallest possible, i.e. it behaves ideally as a power-absorbing bumper (shock absorber or damper). According to the proposed solution, longitudinal grooves of varying cross section are prepared for the inner wall of the cylinder. Upon loading (displacement of) the piston, the oil flows through these as throttles between the cylinder wall and the piston. By properly creating the cross section of gaps and their axial size distribution, the dynamic load caused by the force affecting the piston bar is of predictable rate, and the displacement and its time course (function) can be determined in advance according to the needs and adjusted to the given technical task.

3 Hydraulic Basic Research

As the first step, the numeric and laboratory test of the cross flow factor of gaps in various shapes took place, in order to be able to determine the cross flow factor as a function for hydraulic throttles of different cross section and length; under variable

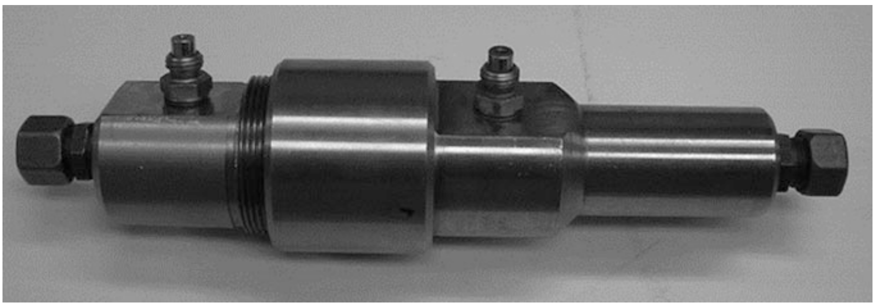


Fig. 1 The apparatus serving for receiving resistance bodies with different throttles to be tested, in an assembled status

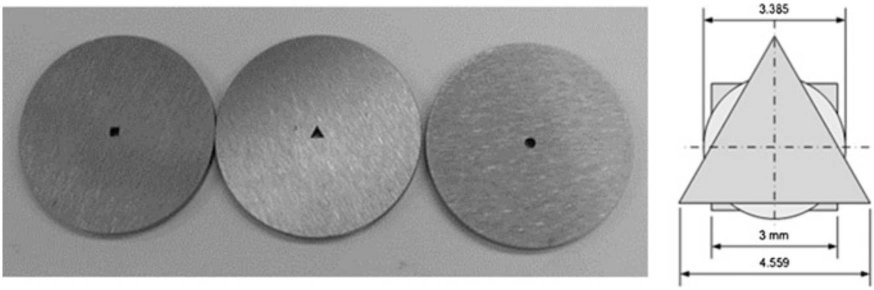


Fig. 2 Resistance bodies with boreholes of different shape (cushion) and the cross sections of throttles

parameters. For these tests, a measurement tool was prepared with removable discs (Fig. 1), to which we machined gaps of a length of 5, 10 and 15 mm with three different cross sections (triangle, rectangle and circle) (Fig. 2).

The final purpose of the joint performance of the numeric tests and measurements was to control all of the initial parameters for the further numeric simulations. Thus, at the end of the validation process, it is possible to perform the preliminary calculation of new variations by applying the calculations, avoiding the costly manufacture and measurement of a high number of prototypes. The three-dimensional hydrodynamic simulation calculations took place by applying the Ansys Fluent program system [2].

The simulation calculations and the measurements were prepared with the throttles of three different shapes, of lengths of 5, 10 and 15 mm but of the same cross section (9 mm^2). In Fig. 3, we present the measurement section established for the reception of throttles, which contains the pressure tappings, as well—these are located 50–50 mm before and after the throttles. In order to avoid cavitation, we

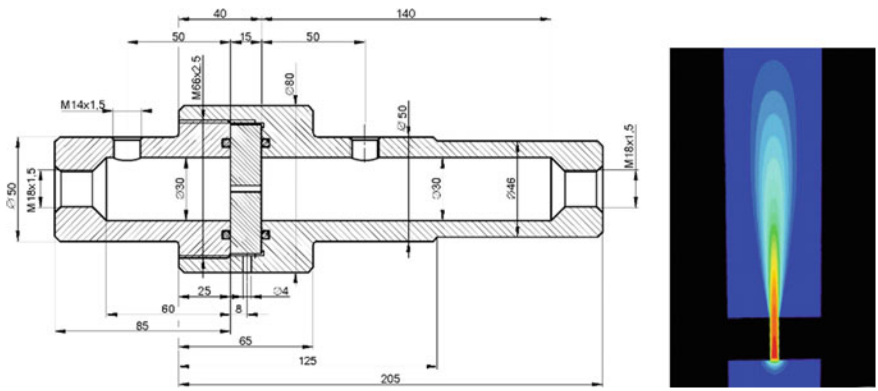


Fig. 3 The drawing of the measurement section containing the throttling element and the flow picture obtained during the simulation calculations

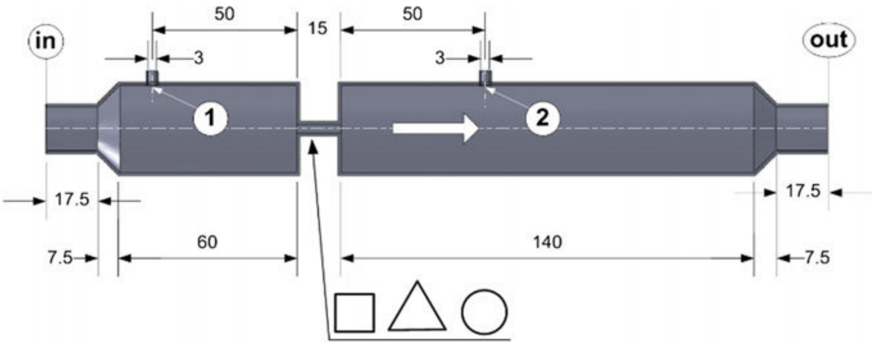


Fig. 4 Geometry of the calculation model

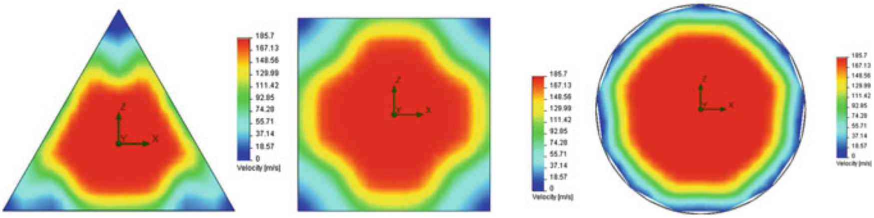


Fig. 5 The speed distributions emerging in the middle plane of boreholes in case of a pressure difference of 60 bar

kept the exiting pressure at a level of $p_k \approx 30$ bar. Based on the features of the university's testing equipment, pressure differences between $p \approx 10 \div 80$ bar could be established on the throttles. (Later on, we had the opportunity to perform the control measurements also up to a pressure of 40 bar in the laboratory of a company that deals with the manufacture of hydraulic parts).

The calculation performed with the Ansys-Fluent flow simulation program [2] took place on the numeric model with the geometry drafted in Fig. 4, which fully corresponds to the surfaces of the measurement section—drafted in Fig. 3—in contact with oil.

During the simulation calculations, we obtained very informative results, for example, in the relatively small gaps, the speed distribution was calculated, as well; showing the dead zones of the flow, as shown in Fig. 5.

The analyses were supported also by the determination of flow conditions formed in the spaces before and after the gaps. The pictures of Fig. 6 show examples of this.

We summarized the results obtained with the calculations and during the measurements in Figs. 7 and 8, with comparative tests.

Figure 7 shows the pressure difference-volume flow diagram prepared based on the measurement and calculation results. Based on the figure, it can be stated that the measured $Q(\Delta p)$ curves run very close to each other. A slightly higher volume flow is detected only in case of the circle-shaped borehole. In Fig. 8 we can see the

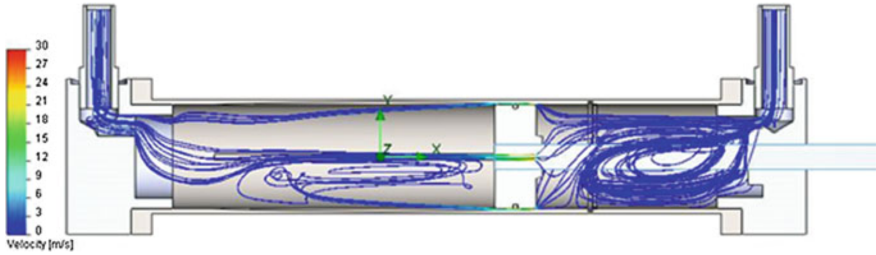


Fig. 6 Manifestation of speed distributions by trails

Fig. 7 The measured pressure difference-volume flow curves

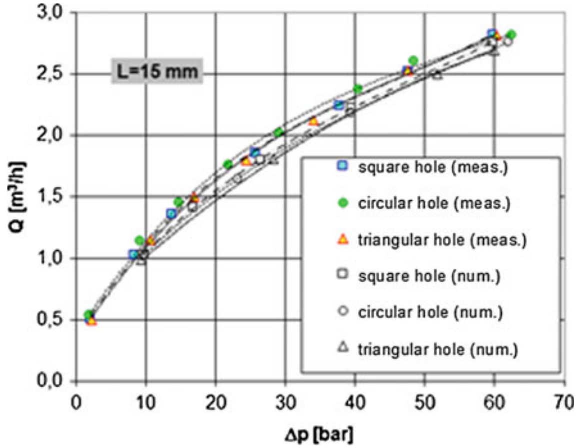
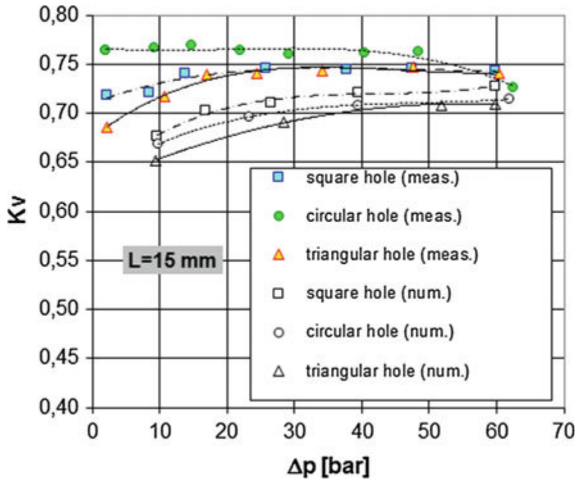


Fig. 8 The measured pressure difference-cross flow factor curves



dependence of the K_v cross flow factor from the pressure difference. In accordance with Figs. 7 and 8 shows that in case of the rectangular and triangular borehole, the measured cross flow factors differ slightly from each other only in case of a small pressure difference. In case of the triangle, here its value is smaller, which can be justified by the corner design of the triangle—unfavourable from the aspect of the flow.

In case of the circular cross section, the cross flow factor moves decisively above the tested range compared to the other cases, and the difference decreases with the increase of the Δp pressure difference. The calculated cross flow factors are $4 \div 14\%$ smaller that the measured values. The bigger difference occurs at low volume flows, and the difference gradually decreases with the increase of the volume flow. The differences are not significant, probably during the manufacturing procedures the dimensions are slightly different from the exact values applied at the calculations. A further research task may be the revision and clarification of the geometries applied at the measurement and the calculations and of other determining factors.

In Figs. 9 and 10, in case of a hydraulic cylinder equipped with a gap of changing cross section, it shows the cross-flown quantities and the cross flow factor

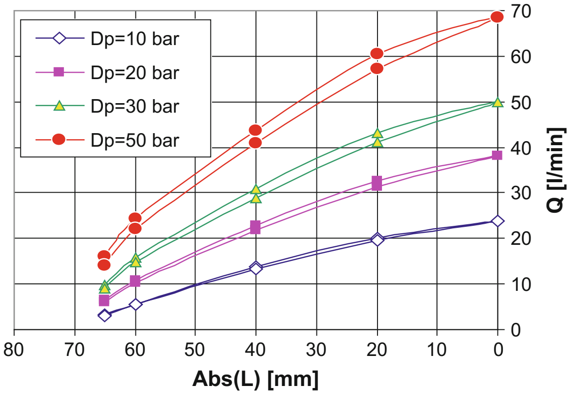


Fig. 9 Correlation of the L stroke and the Q flow rate in case of various pressure differences

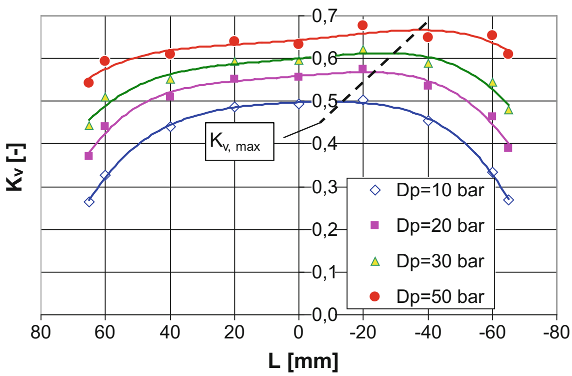


Fig. 10 The change of the cross-flow factor as a function of the L stroke

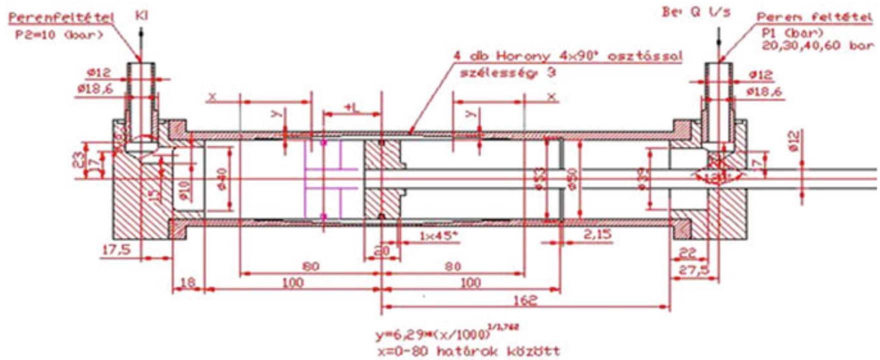


Fig. 11 Draft of a measurement arrangement enabling variable piston position

as a function of the stroke (piston position) in case of various pressure differences. These measurement results provided the basis for the establishment of the final construction.

In order to be able to draw the diagrams of Figs. 9 and 10, we collected data by using a measurement aid shown in the draft of Fig. 11 and the photo of Fig. 12, furthermore, by measurement arrangement. It was possible to set the position of the piston by the threaded spindle attached to the shaft. After carrying out the cross flow test of the throttling gap of a cross section variable according to the given function, established in the wall of the cylinder and of a length of 160 mm; it became possible to draw the diagrams of Figs. 9 and 10 and to get a picture on the evolution of the cross flow factor depending on the variable geometry of the throttling gap.

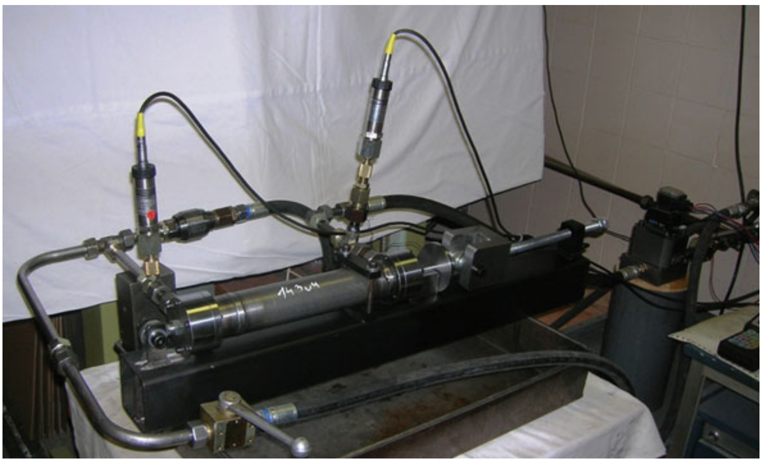


Fig. 12 The measurement tool and measurement arrangement of cross flow tests

The above described tests—at which the hydraulic simulation was performed by the Department of Flow and Heat Engineering Machines of the University of Miskolc, and the validation was conducted by the Department of Machine Tools on a hydraulic load bench—were completed successfully. The aim of the study was to get information on the dampening effect of the throttling gaps created in the wall of the cylinder, the relationship between the gap size (shape) and the dampening effect, i.e. the behaviour of the so called "cross-flow factor"—usually indicated as a constant—as a function and on its values depending on the changing test parameters. The purpose of the joint performance of measurements and calculations was to find out in what way and with what level of precision is it possible to estimate the throttling effect with numerical simulation.

4 Testing of Prototypes and Steps of Utilization

By using the results of the research performed, several kinds of prototypes were prepared and tested. In addition to the power absorbing hydraulic cylinders of smaller diameter (28 and 50 mm) a complete articulation system was prepared for buses, as well, which is able to replace or exchange those applied currently, with a much simpler and more reliable design. Application for the industrial property right protection has been submitted for this structure, as well. The successful testing of the new-type articulation system took place in July this year (2016), built into a 435-type IKARUS articulated bus. Figure 13 illustrates its sketch. The tests performed are detailed in a related article.

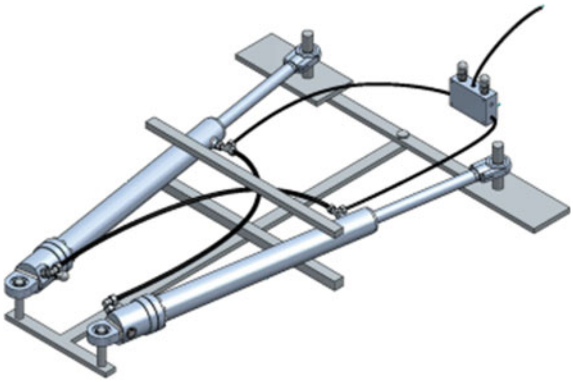


Fig. 13 Assembly draft of the articulation system

References

1. Bencs P, Barna B, Makó I, Szabó Sz (2011) Effect of gap geometry on flow losses. In: Proceedings MicroCAD '11 international computer science conference, Section D, Miskolc, Hungary, pp 13–18
2. FLUENT: Fluent V6.2 User's guide. Fluent Inc., Lebanon, New Hempshire, USA, 2005

Part II
Technology

Utilization of the GD OES Depth Profiling Technique in Automotive Parts Analysis

Tamás I. Török and Gábor Lassú

Abstract Vehicles often operate in rather harsh and even extreme environmental conditions, so their many parts all should resist well against corrosion, wear and other outer impacts, which surface phenomena need continuous developments and frequent testing. One rather fast and quite effective surface analytical technique is the Glow Discharge Optical Emission Spectrometry (GD OES), the applicability of which is demonstrated here by its several laboratory uses to detect the elementary composition in depth of some contaminated then plasma cleaned, pre-treated then metal and/or organic coated (painted or varnished) or just properly surface modified samples originating from different segments of automotive manufacturing or dis-assembled cars put under maintenance and/or repair.

1 Introduction

There are quite many new and novel developments in surface engineering by means of which one can enhance the surface properties (i.e. resistance to wear and corrosion, decorative appearance, lubrication, solderability etc.) of several important functional and structural parts of a vehicle. The two key methods of surface engineering are surface coatings and surface modification as illustrated in Fig. 1.

In both surface treatment techniques the coated or modified outermost zones of any component must be thoroughly controlled during their manufacturing via non-destructive testing techniques, however, time to time their modified surface properties must also be checked in depth through applying some additional and

T.I. Török (✉) · G. Lassú
Laboratory of Surface Techniques/Technologies, Institute of Metallurgy,
Faculty of Materials Science and Engineering, University of Miskolc, Miskolc, Hungary
e-mail: fektt@uni-miskolc.hu

G. Lassú
e-mail: feklassu@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_14

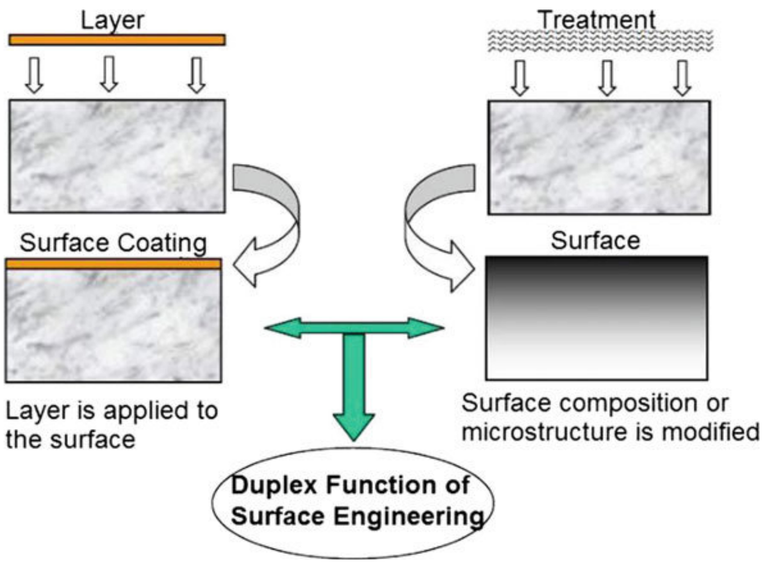


Fig. 1 The two major types of surface treatments, i.e. so-called surface engineering techniques widely used to finish many automotive parts [1]

highly sophisticated destructive type techniques. One such and very important property is the elementary composition of the surface zones which can be analysed in depth with high accuracy and in a rather fast manner by a special surface analytical spectroscopic technique called Glow Discharge Optical Emission Spectroscopy (GD OES).

2 Glow Discharge Optical Emission Spectrometry (GD-OES)

The phenomenon of cathode sputtering is known for a long time, but the use of glow discharge (GD) in spectrometry started when Grimm [2] built his first glow discharge spectrometer with a hollow anode source. By now the glow discharge spectrometry has become a widely used tool for surface and interface analysis with the Grimm type glow discharge source. The Grimm glow discharge source is a flat type source, which consists of an anode tube and a flat sample playing the role of the cathode. There is a spacer (ceramic cathode block and O-ring) between the sample surface and the anode tube to maintain a fixed distance (d = 0.1–0.3 mm) and assure the vacuum tightness [3]. The volume in front of the flat sample is

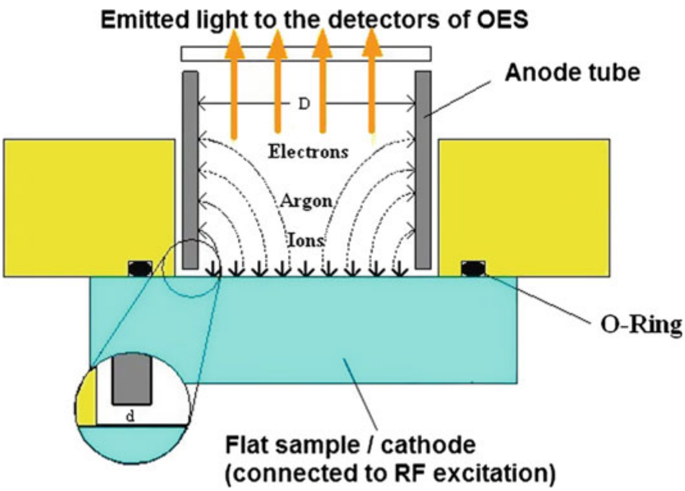


Fig. 2 Schematic of the Grimm-type glow discharge source attached to the sample sealed with an O-ring. A circular spot with a diameter of 4 mm is sputtered with high purity argon forming a crater the bottom of which is so analysed continuously in depth

pumped down to the vacuum ~0.1–1 Pa then filled up by high purity Ar gas up to 300–1300 Pa. (See Fig. 2.)

The flat sample properly attached to the Grimm-type GD source is then sputtered and atomized by the high purity argon plasma and the constituting atoms of the sample so ablated into the plasma will emit light which is detected by the optical emission spectrometer (OES). In this way and using RF (radio frequency) plasma excitation one can sputter and erode electrically conductive and non-conductive materials alike, which is a great advantage of the given system. For demonstrating this special attribute of our RF powered GD OES apparatus, the first example is a spectrum recorded while analysing a vitreous enamelled steel sheet (Figs. 3 and 4).

In the qualitative GD OES spectrum (Fig. 3) all the major constituting elements (Na, Si, O; Ca, Ba, K, Al, B etc.) of this ceramic (silicate) type coating could be detected and analysed down to the base steel sheet (metal substrate) where the signal of iron (Fe) is started increasing at a crated depth of around 140 µm, and it took only about 25 min sputtering, i.e. testing time.

The quantitative GD in depth spectrum (Fig. 4) reveals even more information about the distribution of the elements of the coating, for example, the increasing concentrations of iron closer to the substrate (between about 100–140 µm) is the consequence of the evolution of a bonding zone between the steel and the enamel-coating.

Otherwise, it is worth mentioning here that using such ceramic coatings on some automotive parts, like on cast exhaust manifolds is becoming a novel and

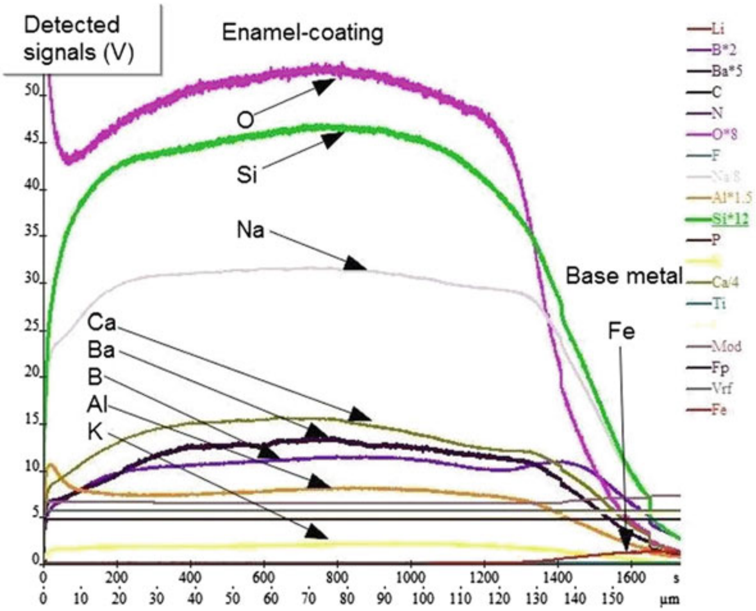


Fig. 3 GD OES depth profile analysis of a glassy (vitreous) enamel coated steel sample. Light intensities of the elementary constituents (detected by the photomultipliers and measured in volts) are shown in the function of both the sputtering time and crater depth as the sample was ablated/eroded by the argon plasma

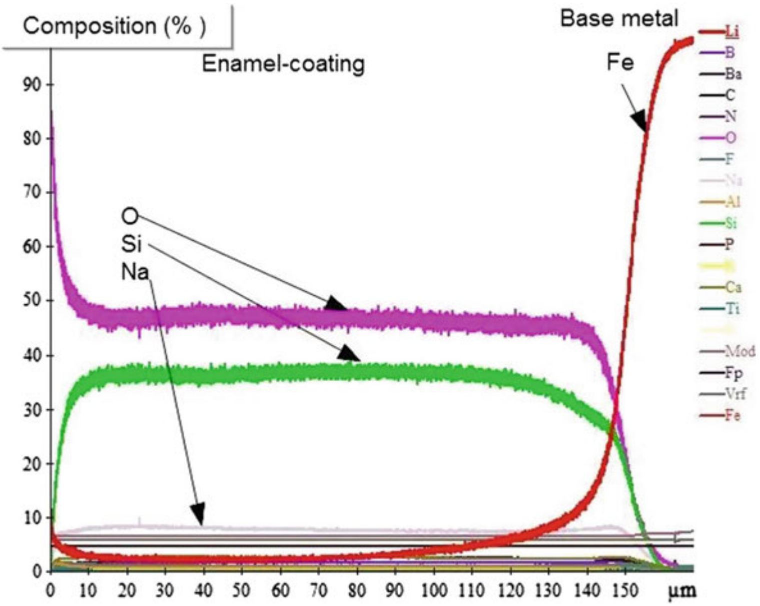


Fig. 4 Quantitative in depth elementary composition (in weight percentages) of the same vitreous enamel coated sample as in Fig. 3

acceptable idea of today [4]. And so, the traditional vitreous (glassy) enamels are finding new applications in the modern car making industry as a special and highly heat and corrosion resistant composite coating material.

3 Testing Multilayered Coatings (Car Body Panels) with GD-OES Depth Profiling

Car body panels of today can be viewed as a well-known example of the very successful application of one of the most advanced surface coating technologies applied both for protection and aesthetical reasons. Such multilayered coating systems consist of several thin sublayers as it is shown in Fig. 5.

As the overall thickness of such car body coatings can be much above 0.1 mm, which is close to the maximum crater depth accessible by our GD OES apparatus, so some of the upper part of the coating must be previously removed prior to starting the GD sputtering if one wishes to analyse the multilayers down and/or into the substrate as well of such samples [6]. The present GD OES record (See in Fig. 6) was obtained by analysing a painted and lacquered (clearcoat) car body panel where the in depth qualitative spectrum shows many important features of the multilayered coating. In Fig. 6 it is clearly seen that below the organic (C, O) clearcoat (Lacquer) there is an effect basecoat (Al), below which the detected elements (Ba, S, Ti, O etc.) are related to the applied colouring pigments and fillers, while the Zn and Ni peeks indicate that the steel panel must have been coated

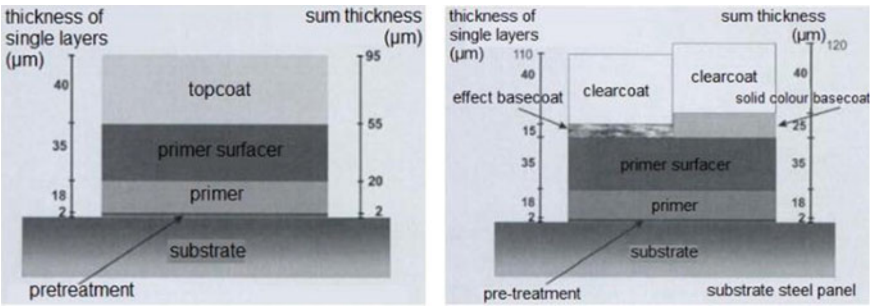


Fig. 5 Structure of three-layer automotive coating/left/and four-layer coating with effect basecoat and solid colour basecoat/right/[5]

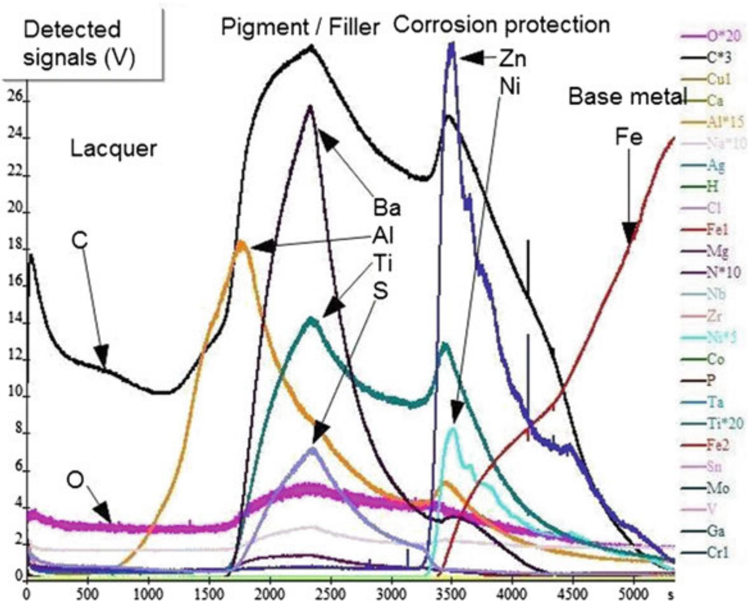


Fig. 6 Qualitative in depth GD OES elementary analysis of a multilayer painted car body panel with marking the major constituting elements down to the steel sheet (base metal)

(electroplated) with a thin zinc-nickel alloy layer as well to enhance the overall corrosion resistance of the finished steel car body panel. Otherwise, such a surface engineering approach is one the most sophisticated ones of today.

4 Plasma Cleaning and GD OES Testing of the Cleanliness

As it is well-known, there are many proven and widely used washing-cleaning methods applied in the car manufacturing industry to clear away even thick contaminant layers and corrosion products from automotive parts. However, more recently there is a growing demand for the removal of even the last traces of residual contaminants, that is, there is growing demand for ultrafine cleaning. And, to answer that demand some newly emerging surface cleaning technologies like plasma treatments working with partially ionized (plasma) gases are also coming to be introduced in the car industry. Then, in connection to that, the surface cleanliness must also be monitored with high sensitivity, for which task the GD OES surface analytical technique can also be a suitable tool to check and evaluate the pre- and post-treatment surface cleanliness and surface activation state [7].

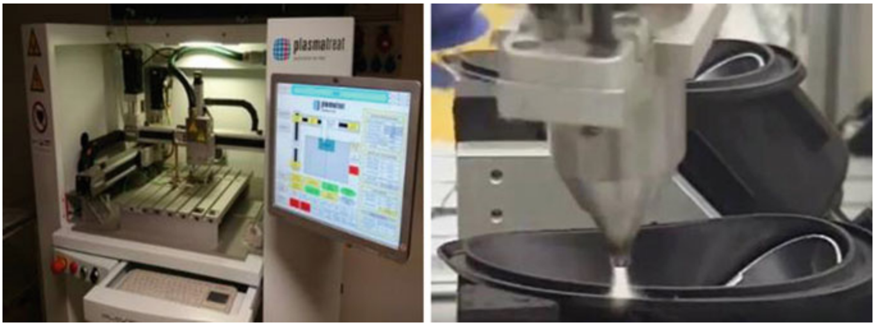


Fig. 7 The open air atmospheric plasma jet laboratory apparatus (Plasmatreater AS 400) installed in the Surface Techniques Laboratory (Institute of Metallurgy)/left/. An industrial application example of using plasma jet for surface fine cleaning and surface modification/activation before gluing and sealing a car headlight housing/right/[8]

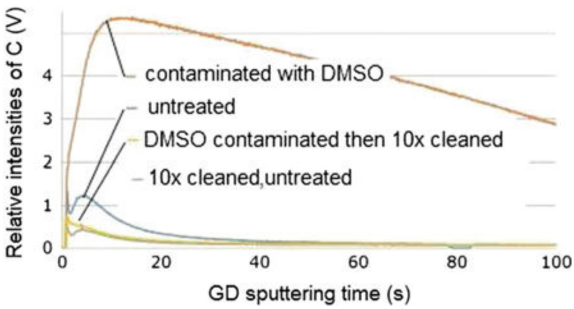


Fig. 8 Results of open air plasma jet fine cleaning monitored/tested by GD OES analysis through checking the surface carbon (C) indicative of the surface residual contamination (or cleanliness) level after scanning the surface with the cleaning plasma jet [9]

The effectiveness of our laboratory plasma system (shown on the left in Fig. 7) for fine surface cleaning was demonstrated through performing repeated plasma cleaning experiments of artificially contaminated surfaces (stained with liquid DMSO) in several laboratory tests [8], and a kind of a visual summary of the results is shown in Fig. 8.

Though the results presented in Fig. 8 are only qualitative, but the effectiveness of the plasma fine cleaning is quite convincing if one compares the low levels of the C intensities detected after scanning the original (already quite clean) surface and that of the artificially DMSO (dimethyl sulfoxide) contaminated then plasma cleaned one.

5 Summary

Almost all automotive parts are surface treated by one way or another, and this paper highlights a few examples of the utilization of a highly sophisticated RF GD OES surface analysis technique which can be used for in depth analysis of inorganic or organic surface coatings alike as well as to check/detect the surface contamination after fine cleaning or follow the elementary composition of the close-to-surface zones down to about 200 μm. As the given GD OES equipment has the capacity of simultaneously detecting more than 40 chemical elements in a very wide range (often between 100% and 1 ppm), it can be efficiently used to analyse rather thick nonmetallic vitreous glassy enamel layers, multilayers like that of the modern surface coatings of car body panels as well as very thin "nanolayers" of residual surface contaminations after even the most thorough traditional surface cleaning procedures. In connection to the latter example, one novel application for fine surface cleaning by our open air laboratory atmospheric plasma jet was also demonstrated.

Acknowledgements This research was (partially) carried out in the framework of the Center of Applied Materials Science and Nano-Technology at the University of Miskolc. The described work was carried out as part of the TÁMOP-4.2.2.A-11/1/KONV2012-0019 project in the framework of the New Széchenyi Plan. The realization of this project is supported by the European Union, co-financed by the European Social Fund.The described article/presentation/study was carried out as part of the EFOP-3.6.1-16-00011 "Younger and Renewing University – Innovative Knowledge City – institutional development of the University of Miskolc aiming at intelligent specialisation" project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund.

References

1. Kennedy DM (2008) Exploitation of surface engineering technology research. In: Conference papers, Dublin Institute of Technology
2. Grimm (1967) Glimmentladungsröhre. German patent DE1589389
3. Nelis T, Pallosi J (2006) Glow discharge as a tool for surface and interface analysis. Appl Spectroscopy Rev 41:227–258
4. Method of making a ceramic coated exhaust manifold and method, US patent 5718046 (1998)
5. Poth U (2008) Automotive coating formulation. Vincentz, Hannover (Germany), p 43
6. Török TI, Lévai G, Szabó M, Pallósi J (2010) Characterizing coatings of car body sheets by glow discharge optical emission spectrometry (GD-OES), Chap 11, pp 335–349. In: Makhlouf N (ed) High performance coatings for automotive and aerospace industries. N.Y.
7. Török T, Urbán P, Lassú G (2015) Surface cleaning and corrosion protection using plasma technology. Int J Corros Scale Inhib 4(2):114–122
8. Bonding of headlight housing with Openair Plasma. https://www.youtube.com/watch?v=V3wagCyiatg/made public on 16 Oct. 2013/
9. Urbán P, Kun É, Ferenczi T, Sós D, Török T (2014) Fine cleaning of metal surfaces by air plasma jet treatment, BKL Kohászat. J Metall 147/2:25–28 (in Hungarian)

Analysis of Surface Topography of Diamond Burnished Aluminium Alloy Components

Gyula Varga and Viktória Ferencsik

Abstract Nowadays diamond burnishing, which belongs to the cold plastic manufacturing procedures, is used more frequently for final finishing operations of parts. By its application, the surface roughness and the micro-hardness in the sub-layers of the components can be increased. The procedure of diamond burnishing can be performed for final finishing manufacturing of outer and inner cylindrical surfaces, and shaped surfaces (e.g. conical, spherical and even statue like) too. The parameters which effect to the surface features during manufacturing are burnishing speed, feed rate, burnishing force, the number of passes, material and geometrical data of the working part of the burnishing tool, furthermore the lubricant applied to burnishing. During our experiments we have chosen from the above mentioned parameters the burnishing speed, the feed rate and the burnishing force and we examined what is the effect of these parameters to the surface topography when manufacturing outer surface of cylindrical components by burnishing tool having given geometrical dimensions. The experiments were executed by the Factorial Experiment Design method. On the base of the evaluated experiment data the improvement ratio of surface roughness was determined by empirical formulas. The technological parameter and burnishing force values were shown out, which provided the highest improvement ratio of surface roughness.

1 Introduction

Cold plastic manufacturing is a dynamically developing forming technology. The main endeavour of modern plastic forming is to create the shape and dimension of the planned part accurate, next to the assuring of appropriate values of strength and strain characteristic. This process does not require high volume of coolants and

G. Varga (✉) · V. Ferencsik
University of Miskolc, Miskolc, Hungary
e-mail: gyula.varga@uni-miskolc.hu

V. Ferencsik
e-mail: ferencsik.viktoria@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_15

lubricants, so that type of manufacturing operation can be realised which environmental load is very small. The surface strain hardening belongs to these procedures, which is characterised by cold hardening and residual stresses [1].

The technological solutions of mechanical surface toughening are differing from each other in the feature of relative displacement of the working surfaces of the workpiece and the tool. From this point of view, it can be differed:

- burnishing, applying sliding relative displacement,
- roller burnishing, based on the rolling relative displacement, and
- surface hardening technological procedures applying impact body combined with dynamic effects.

The main types of the above mentioned procedures are shown in Fig. 1.

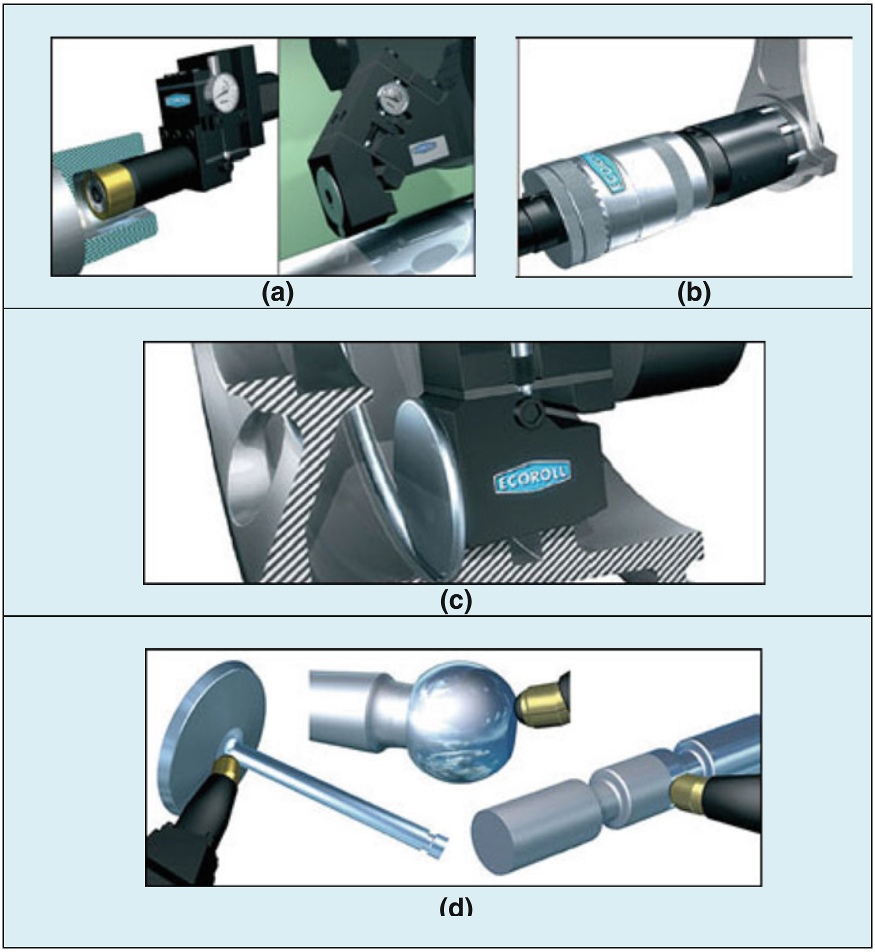


Fig. 1 Tools for cold working operations: **a** single-roller mechanical tools, **b** multi-roller mechanical tool, **c** deep rolling tool, **d** burnishing tools [2]

Life time of workpieces can be reached economically by different precision finish machining procedures e.g. precision grinding [3]. Among the manufacturing procedures we intend to examine the diamond burnishing. If diamond burnishing is used in manufacturing companies, it requires the environmental oriented monitoring [4]. By the application of wear-resistant complexes for modifying its surface properties the improvement of working efficiency of cutting tools can be determined [5]. In the innovative manufacturing processes, the evaluation of machine tool quality is a very important activity [6] and different machining procedures such as hard turning [7], next to it the optimization of milling manufacturing operation [8, 9].

Several studies deal with the theoretical and experimental examination of burnishing process, more exactly, how the different burnishing parameters have effects on surface roughness of burnished part.

During their examinations Luo et al. dealt with diamond burnishing of components made from aluminium alloy type LY12 and bronze type H62 [10, 11]. They created empirical formulas and experienced that from the point of view micro-hardness of the burnished surface the depth of burnishing is the most effective parameter which is followed by the burnishing force, feed rate and the speed of burnishing. Among the burnishing force, feed rate and burnishing speed Yu and Wang took into account the effect of nose radius of the tool, furthermore the rigidity of the tool spring to the surface roughness when executed experiments on aluminium alloy [12]. According to the obtained results, it was stated that the use of greater tool noise radius and higher tool spring rigidity is more advantageous. Stalin John et al. applied the low plasticity burnishing (LPB) [13] to their experiments those workpiece materials which widely applied in the industry as EN24, AISI 1020, Al606, and bronze.

According to the examinations related to surface micro-hardness and circularity generally it can be stated that the increase of the feed has an unfavourable effect on both of them. Tadic et al. experimentally examined the rigidity of burnishing tool when the material of the burnished workpice was aluminium alloy, type EN-AW-6082 [14]. Their experiments verified their teoretic statement: among the surface roughness and burnishing parameters, as burnishing force, feed rate and number of passes, the extents of the correlations are high. Majzoobi, Jouneghani and Khademi examined the lifetime [15] against fatigue by numeric simulation of material type Al7075. They concluded that by increasing of the number of passes and decreasing the feed rate at the same time the depth of zone of residual stress and its homogeneity is increasing. Randjelovic et al. approached the examination of burnishing procedure theoretically by applying numeric simulation [16]. For creating of Finite Element Model they took the kinematics of peripheral milling as a basis and their examined material was aluminium alloy type EC1350. They stated that the best surface roughness can be achieved when the burnishing interpenetration depth almost approaches the maximum peak height (Rp).

Taking into account the literatures mentioned above, the Taguchi type full experiment design method [17] was used in our investigations when burnishing aluminium alloy workpiece. The aim of the examinations is to reveal connection among burnishing parameters and surface roughness in case of workpieces made from aluminium alloy.

2 Experimental Investigations

2.1 Workpiece Characteristics

The material and the hardness of the workpiece to be burnished can be differed for a very wide range. For the experiments we have chosen lightly alloyed aluminium material. The purpose of examination of aluminium alloy material was that non-ferrous materials such as aluminium alloys are more often applied in a variety of industrial sectors such as automobile, aerospace and astronautics. The reason of this, is that most of the non-ferrous materials have low density and have good mechanical properties [10]. The examination of the chemical composition of the lightly alloyed aluminium was executed on a scanning electron microscope type Apollo X. The results of the measurements on three points are shown in Table 1, where: Wt (%)—weight ratio, At (%)—atomic ratio.

The shape and dimension of the workpiece are shown in Fig. 2.

2.2 Burnishing of Outer Cylindrical Surfaces

In diamond burnishing the reduction of surface roughness and improving of the shape correctness of the surface is caused by the interaction between the tool, - which material is much harder than the workpiece -, and the workpiece while performing rotational and linear movement with sliding friction between them [18] (Fig. 3). Generally, the material of the working part of the burnishing tool can be hardened steel, carbide, mineral ceramics, and natural or artificial diamond. The working part of the burnishing tool made from artificial polycrystalline

Table 1 Measured values of chemical composition of the aluminium alloy

Measurement 1			Measurement 2			Measurement 3		
Elements	Wt (%)	At (%)	Elements	Wt (%)	At (%)	Elements	Wt (%)	At (%)
Al	92.97	97.05	Al	92.36	96.82	Al	91.00	96.23
Si	0.25	0.25	Si	0.19	0.19	Si	0.15	0.15
Fe	0.73	0.47	Fe	0.68	0.34	Fe	1.12	0.57
Cu	5.01	2.13	Cu	5.59	2.49	Cu	6.36	2.86
Bi	0.41	0.02	Bi	0.48	0.06	Bi	0.49	0.07
Pb	0.63	0.08	Pb	0.71	0.10	Pb	0.89	0.12

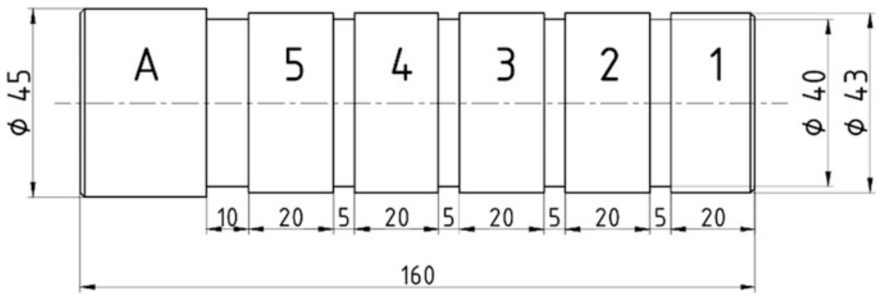


Fig. 2 Geometrical dimensions of the workpiece

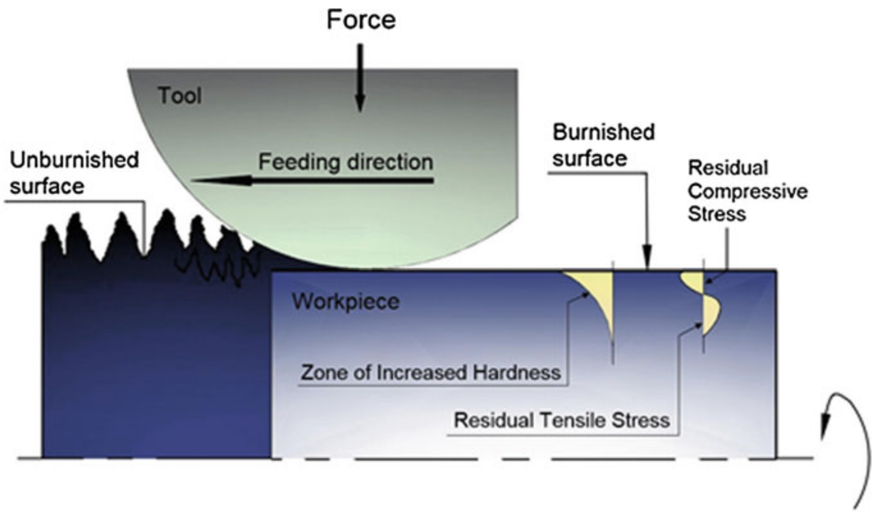


Fig. 3 Schematic illustration of burnishing process [5]

diamond in our experiments with the radius R = 3.5 mm. Big advantage of using this type of tool is that it provides higher stability for the manufacturing system [11].

The pressure necessary for realizing of cold shaping originates from the overlapping being in between the working part of the tool and the surface of the workpiece to be formed.

Plastic deformation is realised in 0.01–0.2 mm thick layer in the subsurface of the workpiece because of the static contact shaping element (burnishing tool) and the outer surface of the workpiece [19, 20].

Burnishing of outer cylindrical surfaces can be performed on conventional universal lathes or up-to-date CNC lathes [21, 22]. During our burnishing experiments a CNC lathe with flatbed produced by the firm OPTIMUM type OPTIturn

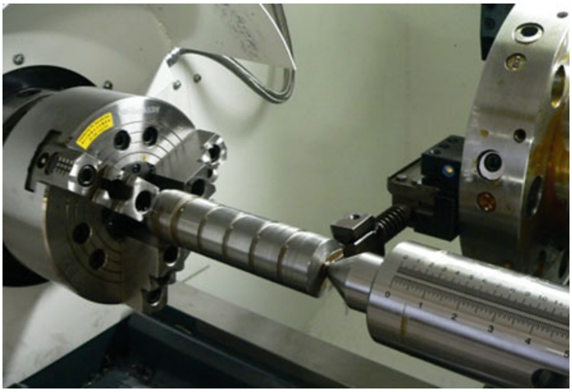


Fig. 4 Execution of diamond burnishing process

Table 2 Applied burnishing parameters

Sign of experiments	Burnishing parameters			Transformed parameters		
	F_b (N)	f (mm/rev)	v_b (m/min)	x_1	x_2	x_3
1	10	0.001	15	−1	−1	−1
2	20	0.001	15	+1	−1	−1
3	10	0.005	15	−1	+1	−1
4	20	0.005	15	+1	+1	−1
5	10	0.001	30	−1	−1	+1
6	20	0.001	30	+1	−1	+1
7	10	0.005	30	−1	+1	+1
8	20	0.005	30	+1	+1	+1

S600 (Fig. 4) was used. During our burnishing experiments the applied oil kinematic viscosity was $\nu = 70$ mm^2/s.

2.3 *The Applied Burnishing Parameters*

Examined parameters: burnishing force, feed rate, and burnishing speed, which parameter ranges are as follows:

Burnishing force	$F_{b1} = 10$ N and $F_{b2} = 20$ N
Feed rate	$f_1 = 0.001$ mm/rev and $f_2 = 0.005$ mm/rev
Burnishing speed	$v_{b1} = 15$ m/min and $v_{b2} = 30$ m/min

The matrix of the Taguchi type Factorial Experiment Design can be seen in Table 2, which contains the burnishing parameters in natural dimensions and their transformed values as well.

During the examination the improvement of the surface roughness parameter (ρ_{Ra}) of the Arithmetical mean deviation of the profile (Ra) was created which can be seen in the formula (1):

$$\rho_{\mathrm{Ra}} = \frac{\mathrm{Ra}_b}{\mathrm{Ra}_a} \cdot 100\,\%, \tag{1}$$

where:

- ρ_{Ra} the improvement ratio of surface roughness parameter Ra. This is a dimensionless parameter which characterises the improvement of parameter change because of manufacturing,
- Ra_b Ra parameter before burnishing,
- Ra_a Ra parameter after burnishing.

If the value of ρ_{Ra} is greater than 100, then the value of the Arithmetical mean deviation of the profile (Ra) is improving in consequence of burnishing.

2.4 Measurement of Surface Roughness

Before and after burnishing the measurement of surface roughness of the specimen were executed on a 3D surface roughness measuring machine type AltiSurf 520 (Fig. 5).

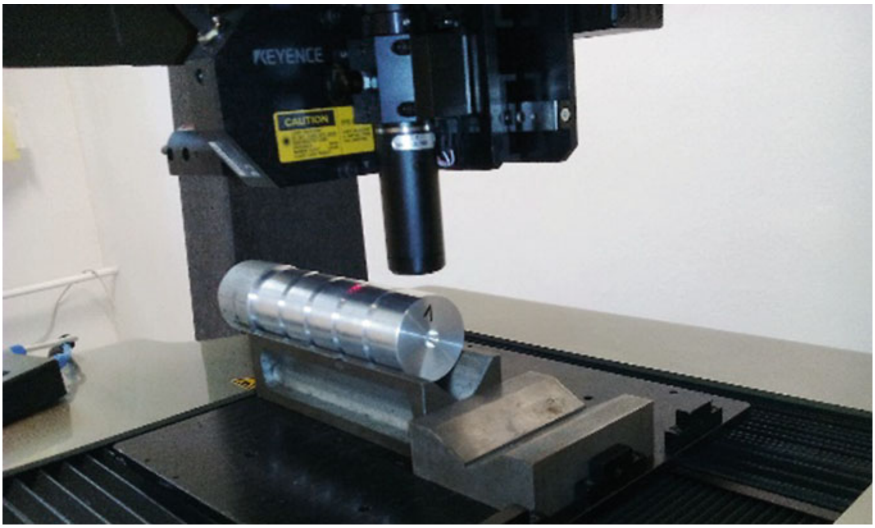


Fig. 5 Measuring of surface roughness on the AltiSurf 520 3D measuring machine

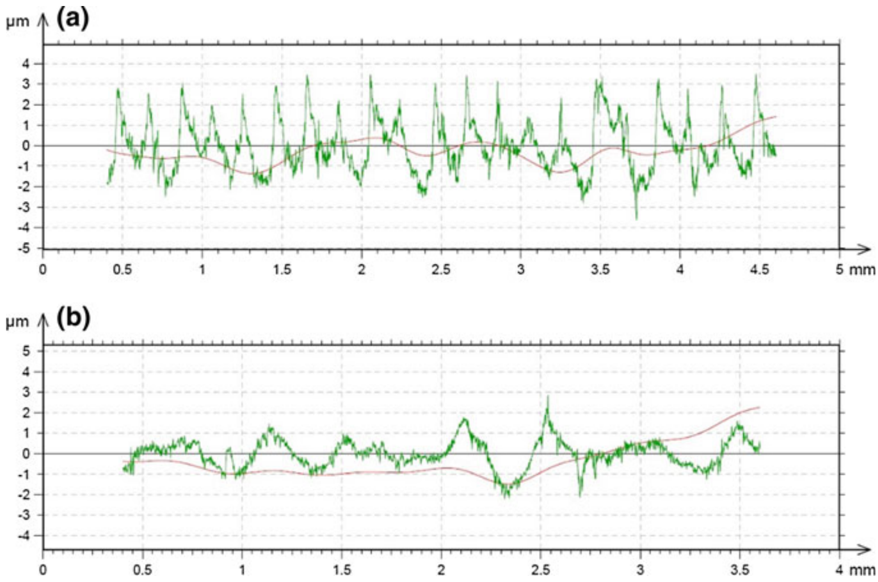


Fig. 6 2D Surface roughness profiles of the Arithmetical mean deviation (Ra). **a** Experiment number 1 (Table 2) before burnishing and **b** after burnishing

During measuring the optical sensor was used and the evaluation of the measuring data was done by the own software of the measuring machine (PheNix). The measurements were done in three different positions rotated by 120°. Figure 6 shows the changing of the Arithmetical mean deviation of the profile (Ra) for the 1st measuring set up (Table 2) (a) and (b). Figure 6 shows how the Arithmetical mean deviation of the profile reduced after burnishing.

3 Results

The values of the measurements can be seen in Table 3. We emphasise again that improvement in the burnishing process was realised when the value of ρ_{Ra} is greater than 100.

From the values of Table 3 those burnishing parameters could be determined where in most cases the value of the surface roughness of the burnished specimen improved related to its state before burnishing. By the use of Factorial Experiment Design empirical formulas can be determined [21]. Calculations and demonstrations of the functions were done by the use of MathCad software.

Table 3 Measured values of the surface roughness R_a

Sign of the workpiece	R_a (μm)				ρ_{Ra}
	After turning		After burnishing		
1/1	1.0151	1.0117	0.5339	0.4231	239.11
	0.9324		0.4337		
	1.0877		0.3016		
1/2	0.8596	0.9299	0.2081	0.2631	353.44
	0.9703		0.3432		
	0.9599		0.2380		
1/3	0.9117	0.9374	0.2688	0.3040	308.36
	1.0416		0.3299		
	0.8589		0.3134		
1/4	0.8977	0.8834	0.2610	0.3017	292.81
	0.8745		0.2966		
	0.8780		0.3476		
1/5	0.9472	0.9524	0.3677	0.4891	194.73
	0.9423		0.3533		
	0.9676		0.7462		
2/1	1.0228	1.1319	0.3010	0.4141	273.34
	1.1294		0.4605		
	1.2434		0.4807		
2/2	1.1285	1.0559	0.7023	0.6421	164.44
	1.0405		0.5868		
	0.9986		0.6371		
2/3	0.9532	0.9814	1.3605	1.2703	77.26
	1.0050		1.1822		
	0.9860		1.2681		

$$\rho_{Ra} = k_0^{Ra} + k_1^{Ra} F_b + k_2^{Ra} f + k_3^{Ra} v_b + k_{12}^{Ra} F_b \cdot f + k_{23}^{Ra} F_b \cdot v_b + k_{23}^{Ra} f \cdot v_b + k_{123}^{Ra} F_b \cdot f \cdot v_b \tag{2}$$

where:

$k_0^{Ra} = 67.75$	$k_1^{Ra} = 17.354$	$k_2^{Ra} = 6.569 \cdot 10^4$
$k_3^{Ra} = 0.483$	$k_{12}^{Ra} = -2.349 \cdot 10^3$	$k_{13}^{Ra} = -0.178$
$k_{23}^{Ra} = -1.061 \cdot 10^3$	$k_{123}^{Ra} = -59.85$	

Figure 7 shows the extent of the improvement of the application of sliding burnishing. It can be stated that the application of the smaller burnishing speed (v_b = 15 m/min) served the higher improvements. When the burnishing force was F_{b2} = 20 N and the burnishing speed v_{b1} = 15 m/min than the reduction of feed

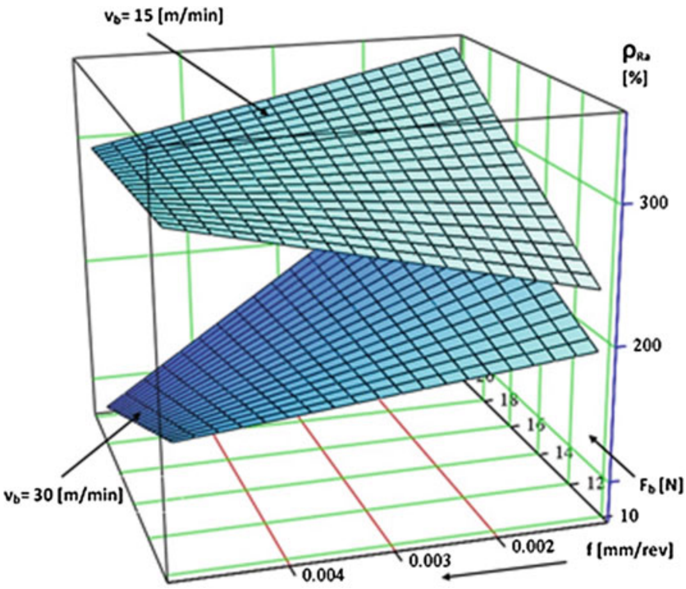


Fig. 7 Evaluated experimental result

rate from the value of f_2 = 0.005 mm/rev to f_1 = 0.001 mm/rev caused a 60.63% increase in the value of improvement ratio surface roughness parameter (ρ_{Ra}). When the value of feed rate was f_1 = 0.001 mm/rev, and the burnishing speed v_{b1} = 15 m/min than the increase of the value of improvement ratio surface roughness parameter (ρ_{Ra}) was 114.33% while the burnishing force was increased from the value of F_{b1} = 10 N to F_{b2} = 20 N.

4 Summary

The paper deals with the experimental analysis of sliding burnishing when the material of the workpiece was lightly alloyed aluminium material. Experimental parameters were the burnishing force, feed rate and the burnishing speed.

The aim of the experiments was how these parameters have effect on the Arithmetical mean deviation of the profile (Ra). The experiments were executed and evaluated on the base of Taguchi type full Factorial Experiment Design. The evaluation was more visible by the use of the dimensionless improvement ratio of surface roughness parameter (ρ_{Ra}). After determining the empirical formula its 3D demonstration made the evaluation more spectacular.

On the base of the present research work, it can be stated:

- Among the examined parameters, the effect of the burnishing speed is the most dominant, as while the burnishing force was F_{b2} = 20 N and the burnishing feed rate f_2 = 0. 005 mm/rev than the reduction of burnishing speed from the value of v_{b1} = 30 m/min to v_{b1} = 15 m/min caused a 215.58% increase in the value of improvement ratio surface roughness parameter (ρ_{Ra}).
- The most appropriate improvement ratio surface roughness parameter (ρ_{Ra}) resulted when the burnishing parameters were as follows:
  - F_{b2} = 20 N,
  - f_1 = 0.001 mm/rev,
  - v_{b1} = 15 m/min.

Acknowledgements The described study was carried out as part of the EFOP-3.6.1-16-00011 "Younger and Renewing University—Innovative Knowledge City—institutional development of the University of Miskolc aiming at intelligent specialisation" project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund.

References

1. Tisza M, Balogh A, Schäffer J (2007) Mechanical technologies. Miskolc (Hungary), p 251
2. http://www.transet-tool.com/index.php?option=com_content&task=view&id=43&Itemid=38. Accessed 21 June 2016
3. Kundrák J, Markopoulos A (2016) FEM/AI models for the simulation of precision grinding. Manuf Technol 16(2):384–390
4. Rusko M, Kralikova R (2013) Implementation of environmental oriented monitoring in the manufacturing company. Adv Mater Res 816–817:1225–1230
5. Vereschaka AA (2013) Improvement of working efficiency of cutting tools by modifying its surface properties by application of wear-resistant complexes. Adv Mater Res 712–715:347–351
6. Kuric I (2011) Evaluation of machine tool quality. Int J Quality Res 5(4):269–275
7. Zębala W, Kowalczyk R (2014) Cutting data influence on cutting forces and surface finish during sintered carbide turning. Key Eng Mater 581:148–153. ISSN 1013-9826. doi:10.4028/www.scientific.net/KEM.581.148
8. Lancea C, Ivan NV, Chicoş LA, Oancea Gh (2008) Optimization of CNC milling files since CAD phases. In: Annals of DAAAM for 2008 & Proceedings 19th international DAAAM symposium intelligent manufacturing and automation: focus on next generation of intelligent systems and solutions, Oct. 22–25, Trnava, Slovakia, ISSN:1726-9679, pp 741–742
9. Kundrák J, Felhő Cs (2016) 3D roughness parameters of surfaces face milled by special tools. Manuf Technol 16(3):532–538
10. Luo H, Liu J, Wang L, Zhong Q (2006) The effect of burnishing parameters on burnishing force and surface micro-hardness. Int J Adv Manuf Technol 28(7):707–713
11. Luo H, Liu J, Wang L, Zhong Q (2005) Investigation of the burnishing process with PCD tool on non-ferrous metals. Int J Adv Manuf Technol 25:454–459
12. Yu L, Wang L (1999) Effect of various parameters on the surface roughness of an aluminium alloy burnished with a spherical surfaced polycrystalline diamond tool. Int J Mach Tools Manuf 39:459–469
13. Stalin John MR, Suresh P, Raguraman D, Vinayagam BK (2014) Surface characteristics of low plasticity burnishing for different materials using lathe. Arab J Sci Eng 39:3209–3216

14. Tadic B, Todorovic MP, Luzanin O, Miljanovic D, Jeremic MB, Bogdanovic B, Vukelic D (2013) Using specially designed high-stiffness burnishing tool to achieve high-quality surface finish. Int J Adv Manuf Technol 67:601–611
15. Majzoobi GH, Zare Jouneghani F, Khademi E (2016) Experimental and numerical studies on the effect of deep rolling on bending fretting fatigue resistance of Al7075. Int J Adv Manuf Technol 8(9):2137–2148. doi:10.1007/s00170-015-7542-z
16. Randjelovic S, Tadic B, Todorovic MP, Vukelic D, Milarodovic D, Radenkovic M, Tsiafis C (2015) Modelling of the ball burnishing process with a high-stiffness tool. Int J Adv Manuf Technol 81(9):1509–1518. doi:10.1007/s00170-015-7319-4
17. Fridrik L (1989) Chosen chapters from the topic of planning of experiments of production engineering. Budapest (Hungary), pp 109
18. Akkurt A (2011) Comparison of roller burnishing and other methods of finishing treatment of the surface of openings in parts from tool steel D3 for cold forming, Metal Science and Heat Treatment 53(3–4), (Russian Orig. Nos. 3–4, March–April, 2011), pp 145–150
19. Luca L, Neagu-Ventzel S, Marinescu I (2005) Effects of working parameters on surface finish in ball-burnishing of hardened steels. Precision Eng (Elsevier) 29:253–256. doi:10.1016/j.precisioneng.2004.02.002
20. El-Taweel TA, El-Axir MH (2009) Analysis and optimization of the ball burnishing process through the Taguchi technique. Int J Adv Manuf Technol 41:301–310. doi:10.1007/s00170-008-1485-6
21. Varga G (2016) Possibility to increase the life time of surfaces on parts by the use of diamond burnishing process. Key Eng Mater 686:100–107. ISSN:1662-9795. doi: 10.4028/www.scientific.net/KEM.686.100
22. Varga G, Sovilj B, Pásztor I (2013) Experimental analysis of sliding burnishing. Acad J Manuf Eng Editura Politechnica 11(3):6–11. ISSN:1583-7904

Investigation of Tyre Recycling Possibilities with Cracking Process

Viktória Mikáczó, Andor Zsemberi, Zoltán Siménfalvi and Árpád Bence Palotás

Abstract The field of vehicle tyres is a key pillar to the Vehicle Engineering BSc launched in September 2016 at the Faculty of Mechanical Engineering and Informatics of the University of Miskolc and the Tyre manufacturing postgraduate course in the technological specialisation on which work is in progress. Generating a yearly amount of several 100 millions of tyres as waste of the automotive industry is, almost 80% of them as passenger car tyres and 20% as truck tyres, whose management creates a huge load to bear on society. These days a relevant task of this field is to find a solution that is reducing environmental loads and sustainable to the solution. Vehicle tyres contain many organic and inorganic compounds: natural and artificial caoutchoucs (NR, SBR, BR, IIR, EPDM), silica, zinc oxide, sulphur, steel and artificial fibres, anti-ageing agents, carbon black etc. whose production requires a significant use of fossil energy carriers. There are several ways of recycling tyres lost their original function: incineration, recycling in its material (rubber-based pavements, roads, sporting grounds) or chemical conversion (energy carrier, chemical raw material), respectively. These days cracking in combined material flow embodies one of the main research directions of chemical conversion. The bottom line is that several raw materials are decomposed in parallel during catalyst-assisted thermal cracking: blends of different ratios of biomass, plastics, rubber tyre. This publication presents options of chemical conversion and its

V. Mikáczó (✉) · A. Zsemberi · Z. Siménfalvi
Institute of Energy Engineering and Chemical Machinery,
University of Miskolc, Miskolc, Hungary
e-mail: mikaczo.viktoria@uni-miskolc.hu

A. Zsemberi
e-mail: zsemberi.a@gmail.com

Z. Siménfalvi
e-mail: simenfalvi@uni-miskolc.hu

Á.B. Palotás
Department of Combustion Technology, Institute of Energy and Quality
Affairs, University of Miskolc, Miskolc, Hungary
e-mail: arpad.palotas@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_16

optimisable parameters. We investigated thermo-catalytic thermal cracking (cracking) of rubber and polystyrene waste producing thereby valuable petrol- and gas oil-type hydrocarbon products.

1 Introduction

Entailing many enterprises and organisations, automotive industry covers the fields of design, development, manufacturing, marketing and sales in relation to a broad scale of vehicles used in a number of walks of life. It is one of the most lucrative economic sectors, which has undergone a rapid evolution after the world wars and, although set back by the oil crises of the 1970s, is still in constant progress in our days. Hungary also has had a vivid presence in this process: various automotive associations have been established in the last years (such as Bakony-Balaton Mechatronics and Automotive Cluster) and large car makers consider Hungary as a potential labour market target that played a huge role in the improvement of the domestic export already in 2011. In response to the need for expertise arising in relation to these, the Faculty of Mechanical Engineering and Informatics of the University of Miskolc elaborated and launched the curriculum of the Vehicle Engineering BSc starting in September 2016, moreover, the Tyre manufacturing postgraduate course in the technological specialisation on which work is in progress, offers professionals coming from other technical areas advantageous orientation opportunities. The automotive industry as a leading industry branch and along with this, the field of vehicle tyres is a relevant pillar of the two previous courses.

The use of various artificial polymer derivatives has risen significantly in the last 100 years, the issue is, however, that several types of hydrocarbon derivatives are increasingly prevalent in the resulted waste. It is estimated that 1.6–1.8 billion rubber tyres and 350–370 million tonnes of plastic waste were generated globally in 2015. Their raw material being crude oil, due to also this it is important to consider polymers that have become waste to be a secondary source of raw materials as thereby the environmental load of harmful landfills on nature can be reduced and global CO_2 emission diminished as well.

The main research direction of recycling rubber tyres is thermo-catalytic conversion processes for energy purposes. However, it is a relevant aspect that tyres have many materials of different quality and quantity such as styrene-butadiene artificial caoutchouc, butyl caoutchouc, natural caoutchouc, butadiene caoutchouc, ethylene-propylene-diene caoutchouc, zinc oxide, metal oxides, sulphur, silica, barium sulphate, carbon black, stearic acid, IPPD, phenolic resin, CSB, hexamine, textile (polyester, polyamide, aramid), steel.

Due to the many heterogeneous components, utilisation of rubber tyres by means of thermo-catalytic thermal cracking is a highly complicated task posing a serious challenge to both researchers and engineers of the current era.

2 Energetical Significance of Recycling

Mankind obtains the majority of its energy demand mainly from fossil energy carriers. The percentage distribution of the world's energy demand broken down into sources is shown in Fig. 1.

Figure 1 reveals that only 13.8% of the produced energy arises from renewable sources that imply an ever increasing threat to nature. Another relevant aspect is that motorisation in developing countries (e.g. China, India) is on a hectic rise because of which hydrocarbon-based engine fuels are demanding more and more, covered today mainly still from crude oil.

Mankind has accumulated a rubber waste and plastic waste base of several thousand billion tons in the last 100 years that can be converted in a so-called cracking in combined material flow having adequately optimised operational parameters into a high-quality liquid energy carrier.

2.1 Products of Cracking of Rubber Tyre

By means of thermo-catalytic process of rubber tyre, energy carriers and many raw materials for the chemical industry can be produced: gas products (methane, ethane, propane, hydrogen), petrol-, gas oil- as well as heavy oil-type hydrocarbon fractions. Based on professional literature data, mostly the opportunity of utilising in energetics emerges regarding further utilisation [1].

The economic competitiveness of these technologies can be considerably improved by increasing the yield of so-called volatile products (gas, petrol-, gas, oil-type fractions) therefore our publication focuses on this area as well. It can be said on the basis of the above thoughts that the hidden exploitation of polymer derivatives by the formed valuable products can promote the contribution to ensure

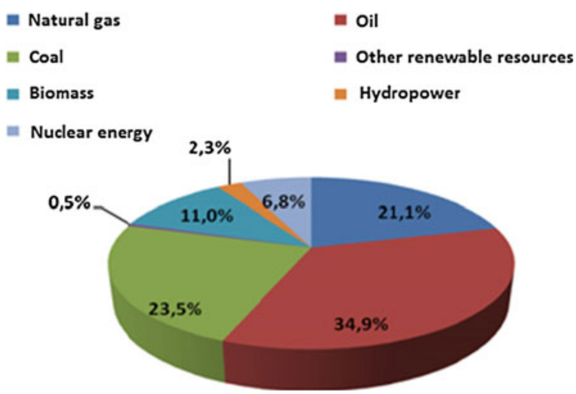


Fig. 1 Percentage distribution of the world's energy demand, according to sources (2015); US Department of Energy

the increasing energy demand and thereby reducing the number of landfills extremely harmful to nature.

2.1.1 Challenges of the Cracking Process

The liquid product produced purely from rubber tyre has a significant content in olefin, which is mainly due to the fact that the macromolecules of caoutchouc are unsaturated carbon chains. The most relevant disadvantages of liquid fraction produced via the process:

- High content in olefin: causes thermal instability inhibiting the storage possibilities of bio oil.
- High content in ash and solids.
- The produced medium is corrosive.
- High content in heteroatoms.

The latest worldwide research activities aimed also at being able to improve disadvantageous properties already during the cracking process so that the economic competitiveness of the technology can be raised too. Such an alternative solution is the so-called thermo-catalytic thermal cracking in combined material flow researched by us too during which the plastic admixtures at an adequate ratio to rubber waste, improves the quality of generated liquid fraction due to its significant hydrogen content.

3 Recycling Options of Rubber Tyres and Polymer Derivatives

Environmental requirements and the related measures at many points of the world do not work properly because of difficult and often non-transparent economic situation. Ever more stringent regulations with regard to waste management come into force in the European Union. Member states are expected to carry out technological developments that are able to manage the increasingly more voluminous waste to obtain the best efficiency.

3.1 Mechanical Recycling of Waste

The mechanical method can be used for recycling a material flow composed of a homogeneous polymer. These types of processes require high-purity collection or precise selection therefore their application is not overly popular. After proper pre-treatment and selection, both process and real polymer waste can be recycled by mechanical recycling. Raw materials undergo first shredding as the form of powder or pellet is more advantageous in further processing.

3.2 Waste Incineration

Incineration is an exothermic process during which the organic content of the waste is turned into gases and water by burning into water leaving the combustion chamber as flue gas. The non-combustible materials remain as ash or fly ash. Such procedures may contribute to the achievement of volume reduction of 85–95% of starting waste and mass reduction of 60–70% on average. The time-frame of the method is relatively low, but its disadvantage is that its investment and operational costs come are high.

3.3 Chemical Utilization of Polymers

During chemical utilisation polymers are converted into shorter molecules thus, in general liquid and gas products develop. The product fractions can be used both as a petrochemical raw material or a fuel. With the appropriate knowledge of the kinetics of these processes has been often incomplete, thus, the production of the product of the desired quality poses a high challenge as we already mentioned before.

Caoutchoucs of various types and polyethylene, polypropylene and polystyrene have several advantageous properties arising from their properties for chemical recycling. However, it presents difficulties if the raw material to be processed has contaminants as these can appear in the products as well. Such can be physical contaminations (dust, oil, etc.) occurring on the surface of rubber and plastic waste or hetero atoms in the structure of polymer (O, S, Cl etc.).

3.3.1 Theory of Thermal and Thermo-Catalytic Thermal Cracking

Cracking and pyrolysis can be a purely thermal or thermo-catalytic process, in which the long carbon chains (polymers) break up into a shorter carbon chain saturated and unsaturated hydrocarbons by cracking of C, C- and C, H-bondings. It is important that thermal cracking performed below 650 °C is called cracking, whereas processes above 650 °C are called pyrolysis [2].

Catalytic cracking and pyrolysis technologies employ principally aluminium-silicate (zeolite) type catalysts, thus, it is accompanied by isomerisation, ring closure and hydrogen transfer as well. The products obtained so contain mainly more valuable branched alkanes (isoparaffins) as well as hydrocarbons and only some small quantities of alkenes occur [3]. In the thermal cracking processes performed below 650 °C mostly a complicated mixture of liquid hydrocarbons develops, therefore they can be more relevant for mainly the preparation of engine fuels [2]. In summary, the below processes must be reckoned with during the thermal and thermo-catalytic thermal cracking [3]:

- Initiation, during which radicals, or ions come into being.
- Formation of further unstable compounds, targeted or random chain cleavage: depolymerisation, formation of monomers, hydrogen-transfer reactions, intermolecular hydrogen migration (formation of paraffin and diene), vinyl group-assisted isomerisation.
- Ring closure by the recombination, the disproportion of formed radicals and ions.

During the catalyst-assisted thermal degradation of hydrocarbons, the processes resulting in the deterioration of polymer structure occurring due to the thermal and catalytic effect competing with each other [3].

3.3.2 Composition of Rubber Tyres

The application area of rubber tyres and various polymer derivatives range across a broad spectrum. Therefore, to produce an oil of adequate quality, the composition of potential raw materials must be researched too.

Table 1 presents the material composition of rubber tyres of an average passenger car and a truck [4].

Both passenger car and truck tyres are the outcome of highly complex engineering activities designed to meet dynamic loads from aggressive environmental and application on the long term and safely. Table 1 reveals as well that carbon-based material to be used as raw material for cracking makes up 67–74% of the whole mass. That means that maximum 50–55 kg liquid hydrocarbon can be produced from 100 kg of rubber tyre that—also arising from the character of the raw material can have many non-desired contaminants.

Table 1 Material composition of rubber tyres of an average passenger car and a truck [4]

Components	Passenger car tyre (m/m%)	Truck tyre (m/m%)
Rubber/elastomer	47	45
Carbon black	21.5	22
Metal	16.5	25
Textile (synthetic fibre)	5.5	–
Zinc oxide	1	2
Sulphur	1	1
Fillers	7.5	5
Carbon-based substances, in total	74	67

4 Most Relevant Parameters of Thermal Cracking Processes

Thermal cracking (cracking, pyrolysis) is carried out in a suitably designed reactor in oxygen-poor or oxygen-free atmosphere, during which the material decomposes chemically. The simplified process of thermal cracking is illustrated in Fig. 2.

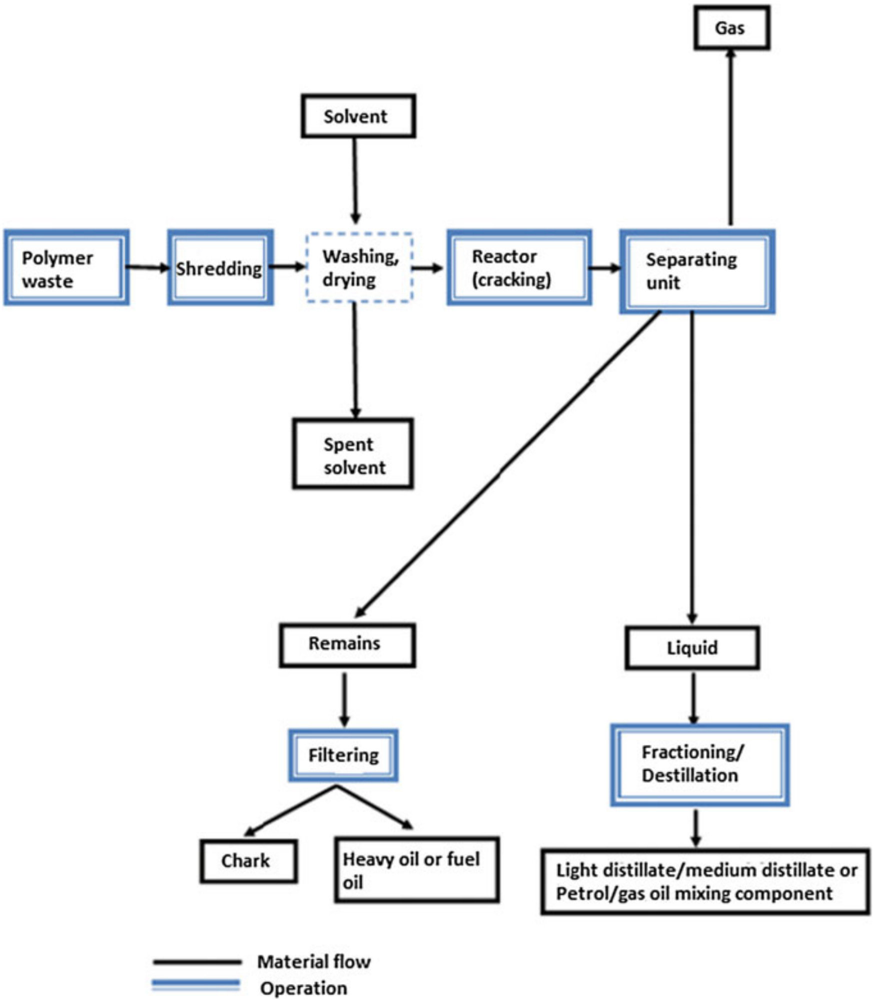


Fig. 2 The simplified workflow of thermal cracking process

During the processes executing thermal cracking the following operational variables play the most important role [5, 6]:

- Temperature
- Residence time
- Operating pressure
- Selection of catalyst
- Mixing conditions
- Granular size
- Heating rate

Their appropriate selection is indispensably important to ensure proper product distribution. The pyrolysis is an endothermic process, therefore it is essential to ensure proper temperature, which is one of the most important operational variables. It can be generally said that more gas state product develops in high temperature processes (over 600 °C). In low temperature processes (below 400 °C) less gas develops, but the formed fluid product is much more viscous and more secondary product develops. The majority of the processes is made occur at 400–500 °C [7].

4.1 Effect of Temperature on Thermal Cracking Processes

Plastics, artificial caoutchoucs and other organic polymers break down into smaller carbon chain entities as temperature rises. Simply put, this constitutes the basis for the development of valuable volatile fractions. Without applying a catalyst, the degradation of polymers during thermal cracking occurs via a radical mechanism, therefore, the share of unbranded hydrocarbons shall be larger in the volatile products compared to thermo-catalytic cases, thus the latter is more beneficial economically as well. 'Today and on the medium term (about until 2040) mainly engine fuels and lubricants with high isoparaffin content shall constitute economic, environment-friendly and safe fuels [8].

The rise in temperature has a significant effect on the decomposition of polymers as the reaction-rate constant rises with temperature significantly. It is to be observed that a lot more volatile product develops at a higher temperature in a much shorter time [7].

4.2 Effect of the Catalyst in the Thermal Cracking Process

During catalytic processes decomposition processes happen via ionic mechanisms. In the presence of a catalyst, the share of gas yield grows considerably, further, the average molecular weight of fluid phase developed so is lower than in the thermal case. This phenomenon is due to the catalyst to increase secondary cracking. The products developed so lead to more valuable hydrocarbon fractions. It is also beneficial that less energy must be input to the system in the catalytic process as this decreases activation energy [7].

5 Presentation of the Experimental Process

The goal of our research was to develop an experimental equipment that is able to analyse thermo-catalytic thermal cracking processes with an optimum parameter combination.

By designing a so-called multi-stage catalyst attachment an opportunity opens up to make qualitative and quantitative analysis of developed products available in the widest possible range. The formed share of quantities of various products allows for the conclusion of kinetic rate parameters too. Our goal in the establishment of the adequate mathematical model is to model these processes—in a way coherent with measurement points—using computer modelling in the future.

5.1 *Decisive Processes and Components of the Experiments*

Figure 3 shows the workflow of the measurement equipment. The unit marked with number 1 is a nitrogen cylinder by means of which the system is inertialised. By infeeding nitrogen, a slight overpressure, oxygen exclusion during the operation is ensured as well as the residence time of product steams in the tube reactor system is controlled.

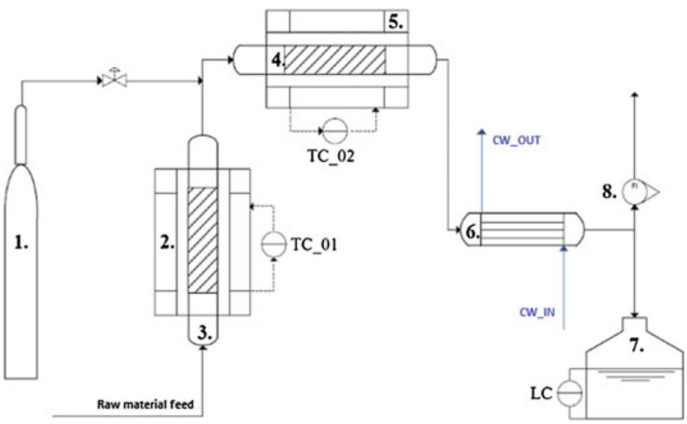


Fig. 3 Workflow of the experimental equipment: 1. nitrogen bottle; 2. furnace_01; 3. vertical reactor; 4. horizontal reactor (catalyst attachment); 5. furnace_02; 6. condenser; 7. liquid collector; 8. rotameter [9]

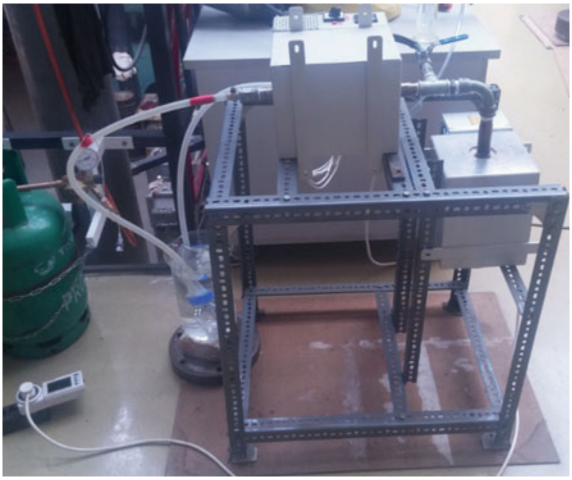


Fig. 4 The insulated reactor system in top view

The raw material was infed in unit 3, the vertically positioned tube reactor. The reactor tube was heated with *Hőker Cső 250/900* (2) electrically heated furnace whose power is 650 W. Unit 4 is a so-called horizontally positioned tube reactor (thermo-catalytic tar decomposer) heated by operational unit 5 (Hőker Cső 250/900), which is just as an electrically heated furnace with a power of 400 W.

The selection of proper structural material is extremely important for appliances operating at high temperatures. Based on the analytical aspects, Wnr. 1.4845 (H9) grade austenitic heat resisting steel was selected as most appropriate for the structural materials of the externally heated reactor (3) and thermo-catalytic tar decomposing tube reactor and the solid particle separation unit. Figure 4 shows the uninsulated tube reactor system along with the installed furnaces.

The right hand side in Fig. 4 the nitrogen infeed valve can be seen, which is a very important unit of the experimental process.

The device marked with number 6 in Fig. 3 is the condenser, which cooled down the high-temperature (350–450 °C) hydrocarbon steams and this way a liquid phase is produced that is gathered in the liquid collector tank 7. Unit 8 is the rotameter, which measured the quantity of gas fraction.

5.2 *Presentation of Raw Materials Applied*

3 parallel series of experiments were performed in the measurements each as per the following:

- homogeneous polystyrene waste, 40 g
- homogeneous rubber waste, 40 g
- PS and rubber waste blend, 20 g each

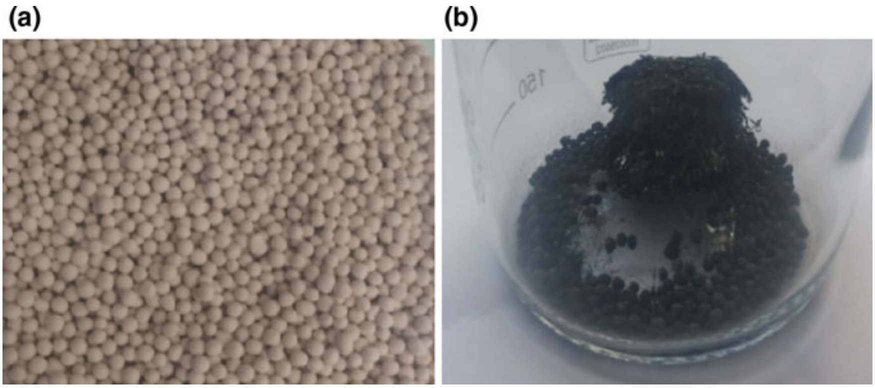


Fig. 5 The zeolite catalyst before the measurements (**a**), the nickel oxide metal and zeolite catalyst remained after the measurements (**b**)

In the measurements 5 g of a zeolite and nickel oxide-type catalyst were used by blending it in the raw material in the vertical tube reactor. Figure 5a shows the zeolite catalyst before the measurements. Figure 5b illustrates the nickel oxide metal and zeolite catalyst remained after the series of measurements.

Figure 5b reveals that the catalyst during the measurements undergoes a serious carbonisation that brings about deactivation. Because of this, catalysts must be regenerated at 800 °C in oxygen atmosphere. The exact tendency to drop in useful efficiency is still unknown to us. Another of our future goals for research is to reveal the rate of deactivation quantitatively as aromatisation and isomerisation reactions are determined by the catalyst primarily and this way it has direct influence on the composition of the more valuable liquid fraction as well.

5.3 The Steps of Measurements

40 g raw material and 5 g catalyst were infed in the vertical tube reactor in each case, operated at 450 °C by us. The experiments ran for 50 min, out of which heating to operating temperature made up 22 min. The horizontal operational unit was running at 300 °C. In the first step the deoxygenation of the reactor system was performed by means of nitrogen. The inflow of nitrogen was constantly ensured throughout the operation. Oil formation started between 300 and 305 °C in minutes 17–18 and lasted until minutes 35–37 in the case of polystyrene.

5.4 Presentation of Measurements

This publication covers the quantitative evolution of hydrocarbon fractions generated during cracking as many qualitative analysis data are not yet available for all fractions.

5.4.1 Results of Cracking of Homogeneous Polystyrene Waste

Figure 6 shows the evolution of the quantities of fractions produced from purely polystyrene waste.

Figure 6 shows that the quantity of gas, liquid and solid fractions developed on average one-by-one during measurements are the following: 15, 23.6, 1.4 g, which are in average mass%: 37.6, 58.9, 3.5 m/m%. The quantity of valuable volatile fractions arisen as 96.5% in mass% on average that can be regarded as a particularly good outcome. The liquid product is illustrated in Fig. 7.

The average result of gas chromatography analyses is listed in Table 2 broken down in volume% distribution.

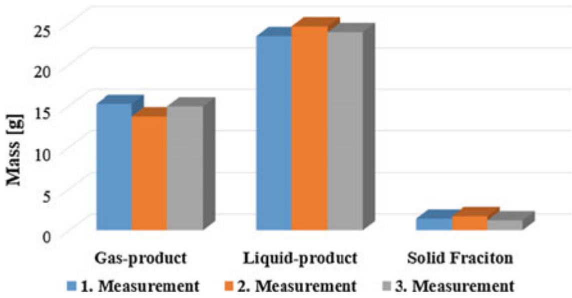


Fig. 6 Quantity of hydrocarbon products produced from purely polystyrene waste

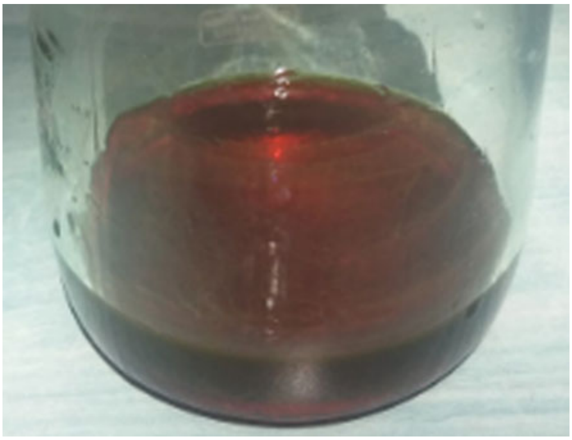


Fig. 7 The liquid hydrocarbon fraction produced from purely polystyrene waste

Table 2 Average gas composition of gas samples produced during cracking of polystyrene (V/V%)

Component	Quantity (V/V%)
CO_2	2.4
C_2H_4	9.1
C_2H_6	1.9
H_2	45.15
O_2	0.2
N_2	1.2
CH_4	30.85
CO	9.2

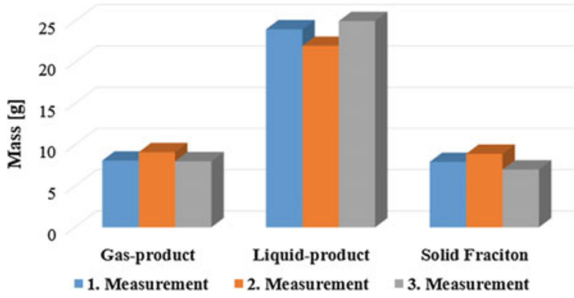


Fig. 8 Quantity of hydrocarbon products produced from purely rubber waste

Table 2 shows that the gas composition is a high-quality combustible hydrocarbon fraction whose more than two third is made up by methane and hydrogen. We have reached the conclusion due to the high hydrogen content of the gas fraction that polystyrene blended with rubber waste can be able to 'in situ' hydrogenate the olefin-rich fraction in the steam phase in the catalytic horizontal tube reactor.

5.4.2 Results of Cracking of Homogeneous Rubber Waste

Figure 8 summarises the quantities of cracking-products produced from purely rubber waste.

As can be seen in Fig. 8, the average quantities of gas, liquid as well as solid hydrocarbon fractions produced from rubber waste were the following: 8.4, 23.6, 8 g, which are in average mass%: 21, 59, 20 m/m%. In contrast to the cracking products of polystyrene waste, the quantity of gas fraction has dropped significantly, by 16.6 mass% in total. The liquid product evolved nearly identically in both cases. The weight of the solid remains (chark, tar) increased considerably in the case of rubber waste cracking, which is a disadvantage for it. The explanation for the phenomenon is that rubber contains a significant amount of carbon black, UV and thermal stabilisers as well as anti-ageing agents due to which long carbon chains can be more difficult to crack off at the same temperature. Furthermore, it can

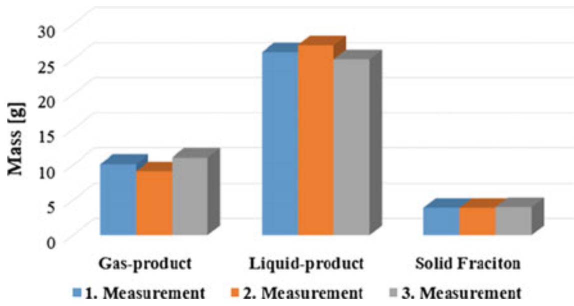


Fig. 9 Quantity of hydrocarbon products produced from polystyrene and rubber waste blend

be said that the system tends to shift more towards the formation of tar due to the unsaturated carbon chains.

We tried to optimise the quantity and quality of more valuable gas and liquid fractions to be produced from rubber waste by admixing polystyrene.

5.4.3 Results of Polystyrene and Rubber Waste Blend

Figure 9 summarises the results of cracking in combined material flow.

As can be seen in Fig. 9, the average quantities of gas, liquid as well as solid hydrocarbon fractions produced from polystyrene and rubber waste blend were the following: 10.1, 26, 3.9 g, which are in average mass%: 25.17, 65, 9.83 m/m%.

As can be seen from measurement results, the weight of solid decomposition remains (chark, tar) can be reduced by the combined cracking of rubber and polystyrene waste—in comparison to homogeneous rubber fraction while the quantity of valuable gas and liquid products grows. The reason for the phenomenon can be that hydrogen-richer polystyrene raw material hydrogenates unsaturated carbon chains at high temperature in the presence of a catalyst and this way the quantity of tar and chark drops in the reactor body.

The key goal of our future research is furthermore to put the formed petrol- and gas oil-type hydrocarbon fractions to a qualitative analysis in each case and on a broad scale. We have not shed light on how polystyrene admixed to rubber waste affects quality parameters precisely and according to what mechanism.

6 Summary

Our work composed of researching recycling options of various polymer derivatives (rubber and polystyrene waste) by means of thermo-catalytic degradation.

We determined that the gas fraction forming during thermo-catalytic thermal cracking of polystyrene contains a significant amount of hydrogen that can be suitable for 'in situ' hydrogenation of unsaturated carbon chains in the case of rubber waste.

We shed light on that the quantity of petrol- and gas oil-type hydrocarbon fractions produced from rubber waste could be increased if rubber waste was converted with polystyrene in combined material flow. The quantity of gas and liquid fraction on average rose 4.17 and 6 m/m%, respectively, in comparison to homogeneous cases. The formation of tar and chark dropped 10.17 m/m% on average, which can be considered a good outcome, but still too low compared to the optimum industrial operation.

References

1. Miskolczi N, Nagy R (2012) Hydrocarbons obtained by waste plastic pyrolysis: comparative analysis of decomposition describe by different kinetic models. In: Fuel processing technology
2. Markó L Organic Chemistry I, http://www.tankonyvtar.hu/hu/tartalom/tkt/szerves-kemia-szerves/index.html, (Szerves Kémia I.)
3. Miskolczi N, Bartha L (2005) Investigation of thermal recycling of plastics and further utilization of products. PhD dissertation
4. Evans A, Evans R (2006) The composition of a tyre: typical components. The Waste & Resources Action Programme
5. Akhtar J, Amin NAS (2012) A review on operating parameters for optimum liquid oil yield in biomass pyrolysis. Renew Sust Energy Rev 16(7):5101–5109
6. Velghe I, Carleer R, Yperman J, Schreurs S (2011) Study of the pyrolysis of municipal solid waste for the production of valuable products. J Anal Appl Pyrol 92:366–375
7. Kaminsky W, Scheirs J (2006) Feedstock recycling and pyrolysis of waste plastics: converting waste plastic into diesel and other fuels. Wiley series in polymer science, John Wiley & Sons, LTD
8. Hancsók J (2008) Modern engine fuels. Public habilitation presentation, Veszprém
9. Zsemberi A, Siménfalvi KZ, Palotás AB (2016) Thermal and thermo-catalytic co-cracking. GÉP Technical periodical of the Scientific Society of Mechanical Engineering 2016/3, vol LXVII, pp 49–52

Utilisation of Various Hydro-Carbon-Based Wastes by Thermo-catalytic Conversion

Andor Zsemberi, Zoltán Siménfalvi and Árpád Bence Palotás

Abstract The global need for energy and raw materials is constantly on the rise as mankind's technology progresses. Due to more and more environmental load and fossil energy carriers exhausted, processes designed for thermo-catalytic conversion of various hydrocarbon-based wastes (plastics- and rubber waste, biomasses) and fuels with a low calorific value (lignite, brown coal) have come into focus in the last decades. The essence of these processes is that solid raw materials forming long carbon chains can be converted at medium-high temperatures (410–450 °C) by means of a special reactor system into more valuable hydrocarbon fractions of liquid and gas state such as petrol-, gas oil-, fuel gas-type products. We examined in our work, how low-quality rubber waste and/or brown coal, plastic waste raw materials can be converted into better quality products—of primarily liquid state. The problem raising a number of open points is a complicated optimisation issue as various heterogeneous components and their content in aggressive contaminants (sulphur, chlorine, nitrogen, oxygen, oxides, carbonates etc.) can largely affect decomposition kinetics thus the quality and quantity of hydrocarbon products formed so as well. This publication covers the system modelling techniques in detail that can be used as a foundation for the basis of mathematical modelling of high-complexity technical systems.

A. Zsemberi (✉) · Z. Siménfalvi
Institute of Energy Engineering and Chemical Machinery,
University of Miskolc, Miskolc, Hungary
e-mail: zsemberi.a@gmail.com

Z. Siménfalvi
e-mail: simenfalvi@uni-miskolc.hu

Á.B. Palotás
Department of Combustion Technology, Institute of Energy and Quality Affairs,
University of Miskolc, Miskolc, Hungary
e-mail: arpad.palotas@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_17

1 Introduction

Fluid, solid and gas state products produced thermal-catalytically from solid organic-based raw materials such as rubber, polymer and biomass are drawing more and more attention globally as valuable energy carriers can be generated from them. Their highly significant disadvantage is however, that they may contain a number of contaminants: solid particles, toxic compounds, tar, heavy metal damps etc. Because of this, the technical elaboration of the process poses a considerable challenge. Many a publication back that heterogeneous components affect each other's thermal decomposition during operations [1].

Based on the data gathered by the Energy Agency of the USA and British Petroleum, it can be presumed that the demand for coal, natural gas and liquid fuels is expected to increase significantly further. The tendency is supported by Fig. 1.

Figure 1 reveals that crude oil, coal and natural gas shall may remain the main energy sources of the world for a few decades to come, therefore it is relevant that remaining reserves must be handled by mankind in an optimal way.

Computer and/or mathematical simulation is an indispensable element in today's engineering practice. These tools provide an opportunity to understand the processes taking place better and as a result of this make manufacturing safer, more efficient and economical. The amount of kinetic rate and mathematical publications with regard to the topic is, however, fairly poor and only small quantities or 'factory-grade' raw materials are used in the investigations in most cases. Under industrial circumstances, parameters determined this way can be employed to a limited extent only.

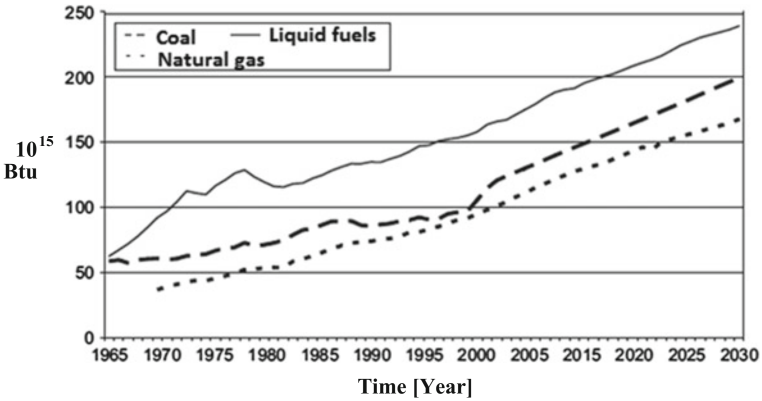


Fig. 1 Growing tendency of demand for coal, natural gas and liquid fuels globally [US Department of Energy, British Petroleum]

2 Significance of Cracking Technologies

Figure 2 confirms and details the undoubted advantages of processes in combined flow compared to direct incineration.

Besides the storability and multi-faceted applicability of energy carriers, no dioxins or furans form in these processes [2].

As it can be seen in Fig. 2, active carbon and bio oil fractions obtained during pyrolysis can be safely and economically stored, thus, these energy carriers can be put to use at a later point. It merits mention that particular products can be utilised in direct energy production in addition to their application as secondary raw materials for the chemical industry.

2.1 *Characters of the Products of Technology*

The oil fraction obtained during the thermo-catalytic technology is basically a brownish-black viscose liquid containing a number of acidic corrosive substances as well. Due to the continuous and rapid depletion of fossil reserves, the energy carriers produced by the combined material flow of solid rubber, plastic waste and/or biomass fractions—mainly the volatile products—can imply a long-term, predictable alternative energy source in the future:

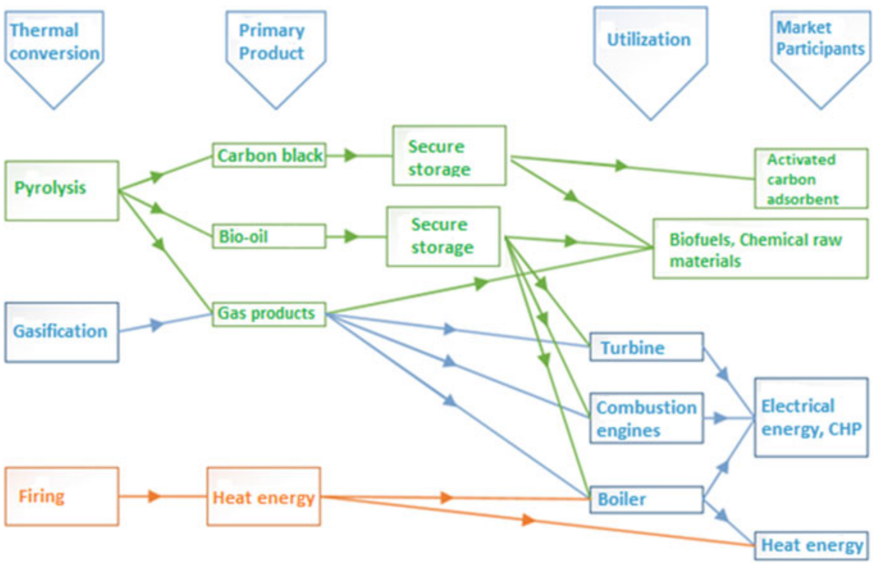


Fig. 2 The thermal processes [1]

- Oil fraction: a mixture of petrol- and gas oil-type hydrocarbons.
- Gas products: methane, ethane, propane, ethene, propene, hydrogen etc.
- Solid products: chark, carbon black, tar.

The produced products have, however, many disadvantages, because of which their utilisation is limited:

- High content in oxygen: causes thermal instability inhibiting the storage possibilities of bio oil.
- Low heat of combustion.
- High content in ash and solids.
- High water content (strongly acidic medium forms, mineral acids form from heteroatoms: HCl, H_2SO_4 etc.).
- Corrosive: in part due to this, special structural materials needs to be applied for processes designed for pyrolysis.
- Poor combustion properties.

The improvement of physico-chemical properties of producible products means a challenge to researchers as this concerns a material of highly complex composition. Many a process is yet to be clarified in order to be able to produce various hydrocarbon fractions economically and utilise safely [3]. Acidic medium is expressively harmful as they exert a catalytic effect on the polymerisation of unsaturated unstable compounds, thus, viscosity as well as quality oil further deteriorates [4].

2.1.1 Most Relevant Contaminants of Products

The qualitative and quantitative detection of the most important contaminants of the produced gas, liquid and solid fractions is relevant not only due to their usability. By the comprehensive analysis of data it will be possible:

- To select the appropriate cleaning process (optimal, energy-efficient solution).
- To be able to select the structural material of the equipment to be applied too.

70–80 m/m% of the infed raw material in combined material flow—in the presence of proper reactor parameters and catalyst—can be converted to oil fraction [5], on top of which ~10–10 m/m% gas and solid-state hydrocarbon fractions develop as well. However, because of the issues covered in the previous chapter, these cannot be directly used in internal combustion engines. Product containing-components developing in the thermo-catalytic processes are considerable due to the heteroatom content of the raw material and depending on this they contain the following to various degrees [6]:

- solid contaminations
- tar
- H_2O

- NH_3
- H_2S; COS; SO_2
- HCl
- alkali and alkaline-earth metals, heavy metals

A number of sources confirm that the produced hydrocarbon fraction contains the following to a certain extent (of course, depending also on the raw material): compounds containing K, S, P, Cl, Ca, Zn, Fe, Cr, Br and Sb as well.

It is advised to remove the above mentioned components—before use—as they damage significantly the contacting structural materials as well.

3 The Role of Applied Mathematical Models

The role of mathematical modelling is indisputable in science and modern engineering applications. The computer evolution of the last 50 years allowed faster manageability and better understanding of difficult and complicated issues for many questions arising in practice. Chemical systems are complicated systems, changing one element can have a significant impact on others, thus, even in the case of the simplest existing chemical reactors one needs to know by all means the heat, mass, and pulse-transfer processes of which almost all include chemical reactions and the related generated/removed reaction heat too [7].

3.1 System Models

System models are a common language of different engineering fields. They are characterised by a wide range of application areas, mutual convertibility and broad scale. A model (models) follow(s) up technologies throughout their lifecycle. A stationery model can be employed only during design. The control of the system, however, means a timely activity, therefore, the relevance of system dynamics is a key issue. The model used in final engineering practice is a formal model, which is a mathematical one (variables, correlations). Selecting the type most suitable for the solution to the task is not simple, it may be the result of weighing up several aspects [7, 8].

3.1.1 A Priori Models

The most important group of models is the class of a priori model (theoretical model, white box model etc.) whose bottom line is that they reflect the (internal) structure of the modelled object in part or in full, all of their variables can be assigned to a physical sense, their correlations are based on physical, chemical laws, furthermore, they are able to transmit information both in space and time. They are suitable to model both existing and imaginative objects. In design, a priori model may be used exclusively. They can be employed in technology, design, the design of technology control system, the design of instrumentation, process modelling, process optimisation and the foundation of technology operations [7, 8].

3.1.2 A Posteriori Models

The other large group of system models is the class of a posteriori model (empirical model, black box model etc.). The feature of these is that they do not reflect the structure of the modelled object not even in part, the so-called input and output variables have a physical content only, their correlations are formal that connect the input and output variables, the expected adequacy can be ensured by the use of measurement data. They can be used neither for spatial nor timely information transmission. They are suitable for modelling a working object, therefore can be used primarily in control. If, for one, the intervention signal changes, how the controlled characteristic has changed, this shall explain how dynamically the system changes [7].

3.2 Reaction Kinetics Approach of Thermal Cracking Processes

In the case of various combined material flows it is worth examining processes taking place between specific components and reactions taking place (at both micro and macro level). This way the value of process products can be increased significantly, meaning the reduction of ROI time of the scaled up processes too. To optimise reaction parameters and understand processes taking place better many reaction models have been created, which, however, describe mainly processes taking place in pure homogeneous material flows. Very little information is available on the modelling of blended and real-life waste.

The reaction kinetic description of cracking of rubber-, polymer waste and/or coal is a fairly difficult task even separately as very complex processes happen and significant differences emerge both in thermal and thermo-catalytic cases. The type of reactor is a relevant aspect, but the quantity of raw material to be processed plays an important role because of the blending ratios. Based on literature data, most of cracking experiments are performed in batch reactors and fairly few are involved in dealing with continuous or semi-continuous technologies. The reason for this is that continuous yield needs raw material of higher volume and the construction of the reactor is fairly complicated to make proper blending and heat transfer relations happen. Furthermore, it can be said that even the removal of contaminants formed during the process with more tendency to deposition and adhesion [3, 5].

3.2.1 Reaction Kinetic Examination of Rubber, Plastic and Biomass Waste

Limin Zhou et al. analysed the combined cracking of various plastics (artificial caoutchouc, HDPE, LDPE and PP) and low volatile substance-containing carbon (LVC) by means of TGA. Experiments were carried out in a nitrogen atmosphere at a heating rate of 20 °C/min. between temperatures of 20 and 750 °C. The degradation of plastics occurs between 438 and 521 °C, whereas that of carbon between 174 and 710 °C (the reason for this is the similar molecular structure of polymer [9]). They preassumed first-order reactions also in these cases, similarly to publications mentioned above. Just like most of researchers—they determined reaction kinetic parameters by means of the Arrhenius equation [9]. Equation (1) means the description of experimental results via conversion, meaning actually the mass reduction curve.

$$x = \frac{W_0 - W}{W_0 - W_\infty}, \tag{1}$$

where W_0 is the initial weight of the specimen, W_∞ is the final weight of the specimen having passed cracking, W is the weight of the specimen at the specific time. To determine reaction rate, the following Eq. (2) was employed [9]:

$$\frac{dx}{dt} = k(T) \cdot f(x) = A \cdot e^{\left(-\frac{E}{R \cdot T}\right)} \cdot (1 - x)^n. \tag{2}$$

The left side of the equation expresses the conversion belonging to the specific time. This correlation was made linear such that natural logarithms of both sides were taken, illustrated by Eq. (3) [9]:

$$\ln\left(\frac{dx}{dt}\right) = \ln A + \mathrm{n} \cdot \ln(1 - x) - \frac{E}{R} \cdot \frac{1}{T}. \tag{3}$$

The left side of the equation was depicted versus 1/T whose slope was used to determine the activation energy (E) and frequency factor (A) for each conversion. First-order mechanism was presumed for the reaction, obviously [9]. Results obtained via the mentioned methods are illustrated in Fig. 3.

The numerical value of concrete reaction kinetic parameters can be generated by means of Fig. 3. In the light of this, the mass-, component- and heat balance of the system can be described as well.

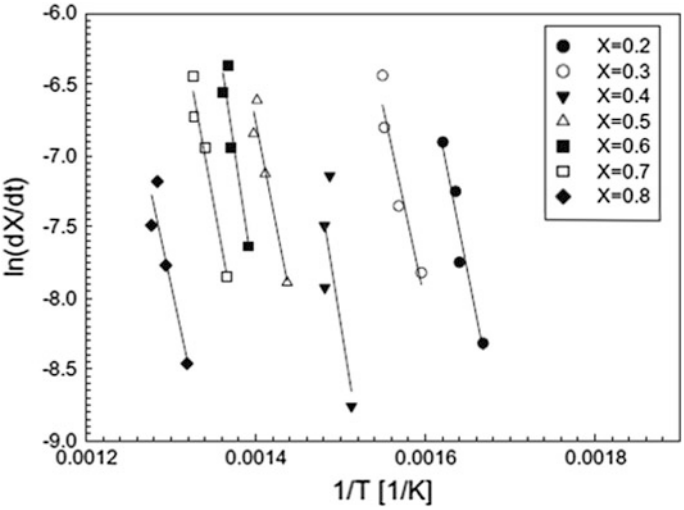


Fig. 3 Steps of determination of capitalisation energy and frequency factor [9]

4 Measured and Calculated Results

The kinetic model employed by us presumes an approach based on first-order reaction kinetics to describe the degradation of polymers, illustrated by Eq. (4).

$$-\frac{dW}{dt} = k \cdot w^n. \tag{4}$$

In this relation the mass reduction curves of the polymer must be determined and from these the general reaction rate constants can be calculated. In addition to the latent activation energies and reaction rate constants, also the pre-exponential factors can be determined by the quantity of the formed volatile products (gas and liquid).

By using the adequate kinetic model, the yield and key properties of the products can be predicted with a high degree of accuracy. The main issue is posed by the determination of the reaction rate constant as cracking is significantly affected by many physical, geometrical and the steric factor. The activation energy and pre-exponential factor are to be determined by means of Eq. (5):

$$k = A \cdot e^{-\frac{E}{R \cdot T}}, \tag{5}$$

where '*E*' is activation energy, '*A*' is the frequency factor (pre-exponential factor), 'R' is the gas constant and '*T*' is the absolute temperature. The dimension of frequency factor and 'k' reaction rate constant is the same as they are based on the concentration unit and therefore depend on the reaction order.

Correlations may become simpler by using thermogravimetry and various coupled methods, but in this case a new parameter needs to be introduced. This new parameter is the heating rate.

The main problem in the cases of thermo-catalytic cracking is that the activity of catalysts deteriorates during the process. Therefore, first-order kinetics may be used via simplifications in such cases. According to the other option, there exists a model for fluid catalytic cracking, which considers the deactivation of the catalyst by Eq. (6).

$$\eta = \exp[-\alpha C(c)], \tag{6}$$

where η is the reduction of activity of the catalyst, $C(c)$ is the carbonisation of catalyst and α is the catalyst type-dependent constant. Equation (6) can be applied to the thermo-catalytic cracking of polyethylene and polypropylene. In this case the value of η must be calculated in each case of catalyst [10].

4.1 The Experimental Equipment and Steps of Measurement

The goal of our research was to develop a mathematical model for our measurement results that can be used for the determination of reaction kinetic parameters. In the light of these data the model can become suitable to allow for the calculation of mass-, component- as well as energy flow of the developing gas, liquid and solid products, when processing high-volume material flow in an experimental equipment.

By means of the well-operating model one can perform measurements on the experimental equipment quicker and safer.

4.1.1 Steps of Experiments

We employed a tube reactor for the measurements. We started with 40 g of raw material in each case, blending in 5 g zeolite and a nickel oxide-coated metal grid. The experiments ran for 50 min. The raw material was infed in a batch operation in the vertically positioned tube reactor. In the first step the raw material was dried at 105 °C for 15 min, while the entire system was being ‘flushed’ with nitrogen (the nitrogen flow was ensured during the totality of the operation). In the second step the temperature of the reactor was heated to 450 °C in 22 min.

During operation the hydrocarbon steams left the vertically-positioned reactor body to the horizontal tube reactor that was kept at 300 °C. Product steams left from this into the condenser where they were cooled down to 20 °C. Afterwards, the liquid phase was gathered in the liquid collector, and the gas phase was flared off.

Table 1 Results of thermo-catalytic cracking

Products	Measurement 1 quantity (g)	Measurement 2 quantity (g)	Measurement 3 quantity (g)
Gas fraction	10.1	9.1	11
Liquid fraction	26	27	25
Solid fraction	3.9	3.9	4
	Measurement 1 composition (m/m%)	Measurement 2 composition (m/m%)	Measurement 3 composition (m/m%)
Gas fraction	25.25	22.75	27.5
Liquid fraction	65	67.5	62.5
Solid fraction	9.75	9.75	10

4.1.2 Experimental Results

The values of remaining gas, liquid and solid fractions obtained during measurements are listed in Table 1 expressed in g.

As can be seen in the first table, the average quantities of gas, liquid as well as solid hydrocarbon fractions produced from polystyrene and rubber waste blend were the following: 10.1, 26, 3.9 g, which are in average mass%: 25.17, 65, 9.83 m/m%.

4.2 The Mathematical Model Fitted to the Experiments

The reaction kinetic features of the gross cracking reaction were determined by means of the following Eqs. (4, 7–8). The activation energies of thermal cracking processes can be calculated by means of the Arrhenius Eq. (5).

$$w = w_0 \cdot e^{-k \cdot t}, \tag{7}$$

$$ln \frac{w}{w_0} = -k \cdot t, \tag{8}$$

where 'w' is the weight of the specimen, 'w_0' is the weight of initial plastic waste, 'k' is the reaction rate constant, 'n' is the order of the reaction, and 't' is the time for cracking.

Using Eq. (8) the resulting reaction rate constants can be determined in the light of the quantitative data of gross processes. Next, we performed further calculations by resolving a multi-variable differential equation system according to the scheme modelling the cracking in Fig. 4 to determine the reaction kinetic parameters of specific partial reactions.

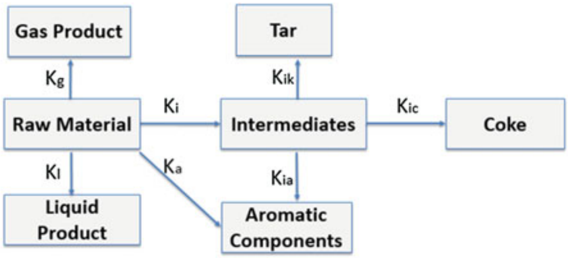


Fig. 4 The scheme is modelling the cracking

Legend for the designations in Fig. 4: k_g: reaction rate constant of gas formation, k_l: reaction rate constant of liquid formation, k_i: reaction rate constant of intermediate product formation, k_a: reaction rate constant of aromatic product formation, k_{ik}: reaction rate constant of tar formation, k_{ia}: reaction rate constant of aromatic from intermediate product, k_{ic}: reaction rate constant of chark formation. The reaction taking place can one-by-one be described by the following differential Eqs. (9–15).

$$\frac{dw_M}{dt} = -k_g \cdot w_M - k_i \cdot w_M - k_a \cdot w_M - k_l \cdot w_M, \tag{9}$$

$$\frac{dw_g}{dt} = k_g \cdot w_M, \tag{10}$$

$$\frac{dw_l}{dt} = k_l \cdot w_M, \tag{11}$$

$$\frac{dw_i}{dt} = k_i \cdot w_M - k_{ik} \cdot w_i - k_{ia} \cdot w_i - k_{ic} \cdot w_i, \tag{12}$$

$$\frac{dw_a}{dt} = k_{ia} \cdot w_i - k_a \cdot w_M, \tag{13}$$

$$\frac{dw_t}{dt} = k_{lk} \cdot w_i, \tag{14}$$

$$\frac{dw_{ic}}{dt} = k_{lc} \cdot w_i, \tag{15}$$

where ‘w_M, w_l, w_g, w_t, w_i, w_{ar}, w_c’ are weights of raw material, liquid phase, gas phase, tar, intermediate products, aromatic components and chark. The ‘k’ index represents the reaction rate constants of the mentioned components as were detailed previously. The analytic solution of differential equation systems (9)–(15) was performed by MATLAB software (with the Runge–Kutta solution method).

In the light of the reaction rate constants of the processes, the activation energies were calculated by means of the Arrhenius Eq. (5).

Table 2 Main components of rubber tyre waste

Main components of rubber tyre waste	Composition (m/m%)
Natural caoutchouc SMR5CV	29.59
Styrene butadiene caoutchouc	29.59
Carbon black ISAF N220	29.59
Stearic acid	0.59
IPPD (antiozonant)	0.89
Zinc oxide	2.96
Phenolic resin	2.37
Sulphur	0.89
CSB (accelerator)	0.89
Hexamin H-7	0.18
PVI (phthalimide)	0.12
Aromatic oil	2.34

Table 3 Values of rate constants calculated via the model at 450 °C

Rate constants	Value (1/s)
k_i	0.072
k_g	0.022
k_l	0.010
k_a	$4.71 \cdot 10^{-9}$
k_{ia}	0.075
k_{ik}	0.023
k_{ic}	0.002

Table 2 presents the components of rubber tyre waste taken into consideration when describing the mathematical model.

4.2.1 Results Obtained via the Mathematical Model

Reaction-rate constants at 450 °C obtained via the model are listed in Table 3.

Results of the reaction-rate constants indicated in Table 3 at 450 °C calculated via the model are presented in Fig. 5.

It can be seen in Fig. 5 that the conversion of the raw material marked with light blue curve evolves according to fidelity as it has almost fully transformed into a product in 50 min. The quantity of gas-, liquid (aromatic fraction as well)—as well as solid products, provided one-by-one the following: 9.2, 25.4, 5.4 g, which values correlate with experimental results at an accuracy of 95%. Therefore, it can be stated that the applied reaction kinetic model functions well, it is suitable for conducting examinations at a broader temperature interval.

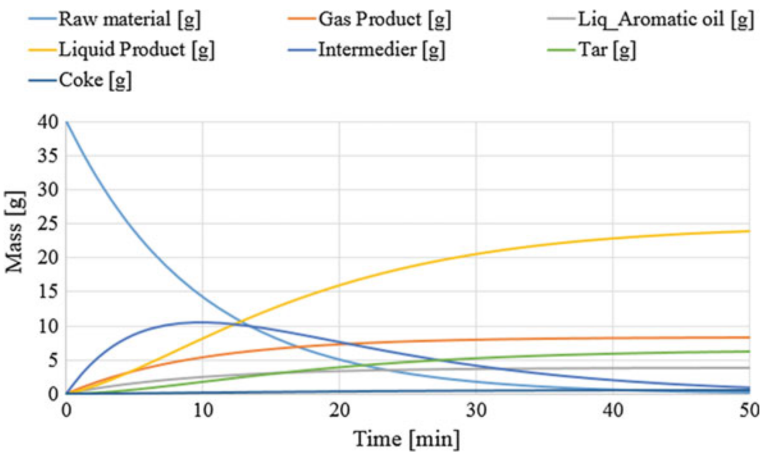


Fig. 5 The results calculated via the mathematical model at 450 °C

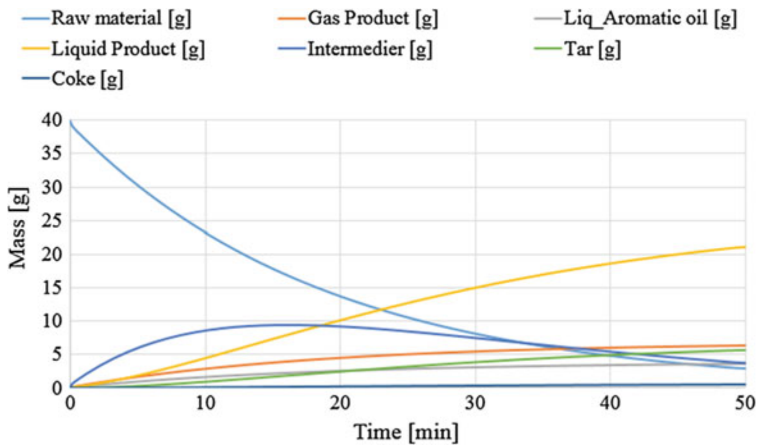


Fig. 6 The results calculated via the mathematical model at 400 °C

A relevant input parameter of the model is the temperature that has a large impact on product distribution, thus we performed an investigation at 400 °C. We were interested in what quantities of valuable components (mainly liquid- and aromatic phase) develop. The results are illustrated in Fig. 6.

The results disclose well that not even the conversion of the raw material is sufficient at 400 °C, thus, the infed solid phase converted at 93%, which appears as a loss in the process. In addition, the combined quantity of liquid- and aromatic phase lags behind the results measured at 450 °C as it resulted in 24.6 g. It can be seen that it is not worth conducting experiments for parameter combinations and raw material tested by us below 450 °C.

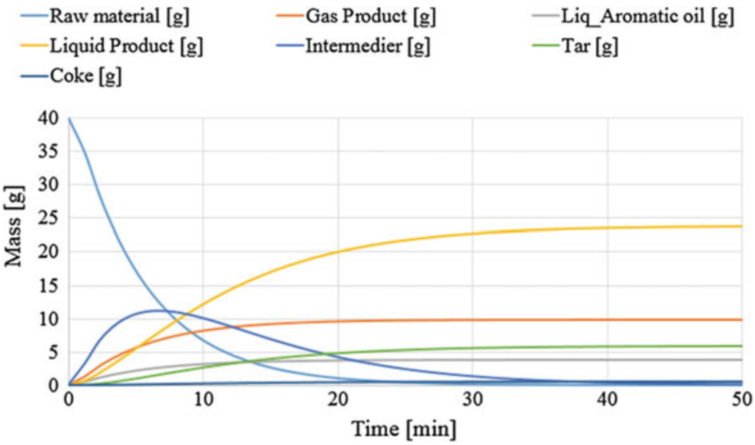


Fig. 7 The results calculated via the mathematical model at 500 °C

In the next step we performed investigations at 500 °C, whose results are illustrated in Fig. 7.

Figure 7 reveals that the conversion of raw material is completed already in 40 min, besides, the quantity of the formed liquid- and aromatic phase added up to 27.4 g in total, which is better than the results measured and calculated at 450 °C. Therefore, it is worth conducting further measurements also at 500 °C in order to operate the reactor system at an optimum level as the goal is to gain the highest possible quantity of liquid- and aromatic phases in the shortest possible time.

5 Conclusion and Determinate of Research Goals

Our work has focused on thermo-catalytic cracking technology, which is suitable for the conversion of rubber, plastic waste and other solid hydrocarbon-based raw material into valuable mainly liquid- and gas-state fraction. Conducting the measurements and revealing the reaction kinetic parameters are complicated, costly and not least a time-consuming task. Consequently, taking our results as a basis, we created a general mathematical model by means of which high-safety examinations can be performed.

The results disclose that the elaboration of the mathematical model can largely facilitate scientific research as analyses can be conducted even without concrete physical measurements to make optimum operational conditions more precisely. We determined that it is not worth performing experiments at temperatures lower than 450 °C in future. Conversely, we reached better results at 500 °C by means of the created mathematical model. We wish to confirm the calculated results from specific measurements in the future.

On top of that, another goal of our future research is to implement the current model into a so-called 'Flowsheet' simulator, with which calculation results and sensitivity analyses can be further refined.

References

1. Ruiz JA, Juárez MC, Morales MP, Muñoz P, Mendívil MA (2013) Biomass gasification for electricity generation: review of current technology barriers. Renew Sustain Energy Rev 18:174–183. doi:http://dx.doi.org/10.1016/j.rser.2012.10.021
2. Stiegel GJ, Clayton SJ (2001) Gasification technologies. In: A program to deliver clean secure and affordable energy, US Department of Energy, Office of Fossil Energy, National Energy Technology Laboratory, 2001
3. Qi D, Zhonyang L, Kixiang Z, Jun W, Wen C, Yi Y (2012) Experimental study on bio-oil upgrading over Pt/SO4-2/ZrO2/SBA-15 catalyst in supercritical ethanol. Fuel 103:683–692
4. Huiyan Z, Rui X, He H, Gang X (2009) Comparison of non-catalytic and catalytic fast pyrolysis of corn cobina fluidized bedreactor. Bioresour Technol 100(3):1428–1434
5. Ying X, Tiejun W, Longlong Ma, Qi Z, Wang L (2009) Upgrading of liquid fuel from the vacuum pyrolysis of biomass over the Mo–Ni/γ-Al2O3 catalysts. Biomass Bioenergy 33 (8):1030–1036
6. Xu C, Donald J, Byambajav E, Ohtsuka Y (2010) Recent advances in catalysts for hot- gas removal of tar and NH3 from biomass gasification. Fuel 89(8):1784–1795. doi:http://dx.doi.org/10.1016/j.fuel.2010.02.014
7. Rasmuson A, Andersson B, Olsson L, Andersson R (2014) Mathematical modeling in chemical engineering. Cambridge University Press; 2014
8. Szeifert F, Chován T, Nagy L (1998) System models-System analysis (Rendszermodellek-Rendszeranalízis). Egyetemi jegyzet, Veszprém (Hungary)
9. Nikrityuk PA, Meyer B (2014) Gasification processes modeling and simulation. Wiley-VCH Verlag GmbH & Co
10. Miskolczi N, Bartha L (2005) Investigation of thermal recycling of plastics and further utilization of products. PhD dissertation, Veszprém (Hungary)

Development of Nitrided Selective Wave Soldering Tool with Enhanced Lifetime for the Automotive Industry

Zsolt Sályi, Zsolt Veres, Péter Baumli and Márton Benke

Abstract The aim of our research is to develop an economic material combination exhibiting good wetting with solder alloy melts and enhanced lifetime against erosion in lead-free soldering applications. Iron-nitrides have strong bonds compared to metals, therefore they are expected to have increased resistance against erosion in solder melts compared to iron soldering tools. The first step of our research is to find the proper substrate material for nitride coating to obtain the desired wetting behaviour. The present paper focuses on the substrate material selection through wetting examinations of nitrided W302, 42CrMo4 and C45 steels with SAC 305 solder alloy melt. It was found that the substrate composition strongly affects the wetting contact angle. The best wetting was achieved with the C45 type steel substrate. The correlation between substrate composition, compounds formed during nitriding and wetting behaviour is discussed.

1 Introduction

In selective soldering applications the solder alloy melt is driven to the soldering location with a soldering tool. In case of selective wave soldering, the tool is a nozzle, while for hand soldering it is the tip of the soldering iron. For all cases, the soldering tools must have good wettability with the solder alloy melt to ensure a stable contact with the component to be soldered. In most cases, the tool is

Z. Sályi (✉) · Z. Veres · P. Baumli · M. Benke
Institute of Physical Metallurgy, Metalforming and Nanotechnology,
University of Miskolc, Miskolc, Hungary
e-mail: zsolt.salyi92@gmail.com

Z. Veres
e-mail: vezso3838@gmail.com

P. Baumli
e-mail: peter.baumli@gmail.com

M. Benke
e-mail: fembenke@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_18

produced of iron with low alloying element/impurity concentration or high purity Armco iron. For those tools, the metallic type bonding in iron ensures the desired wettability with the molten solder alloy. At the Fe-solder alloy melt interface, $FeSn_2$ intermetallic layer is formed [1]. The wettability is further enhanced through the $FeSn_2$ intermetallic layer at the tip/solder melt interface. The $FeSn_2$ intermetallic layer is either produced intentionally prior to soldering (hand soldering) or during the first soldering cycle (selective wave soldering). During application, the thickness of the intermetallic layer gradually increases with soldering cycles. Fe atoms are provided by atom transport from the tool as the $FeSn_2$ intermetallic layer thickens. This atom extraction appears macroscopically as the erosion of the soldering tool. The erosion produces a rough surface on which ditches also form in the later stages. The erosion is further enhanced in the ditches where the solder melt is in contact with unharmed Fe surface. After a certain degree of erosion, the good wetting between tool and solder melt is destroyed and the tool must be replaced. The erosion is accelerated by the application of lead-free soldering alloys such as SAC and INNOLOT due to the increased Sn content compared to previously used Sn-Pb alloys and the elevated soldering temperature [2–12]. The erosion process is more pronounced in selective wave soldering applications where the tool is exposed to the continuous flow of solder alloy melt [9]. In such applications, the erosion effect of the melt flow contributes to the diffusion driven Fe extraction. Finally, the degradation of the soldering tool leads to the need to replace the soldering tool which, in turn, requires the interruption of the soldering process, calibration of the new tool and cause an overall loss of production time for soldering companies.

The erosion reaction of metals in Sn based solder alloys can be reduced by composition modification of the solder alloy melt. Nishikawa et al. showed that Co addition to Sn-Ag solder alloy effectively reduces the erosion reaction of iron plated copper in Sn-Ag solder melt [11]. However, if the solder composition cannot be modified, the erosion reaction must be suppressed by designing erosion resistant solder tools. Watanabe et al. examined the erosion resistance of Fe-MWCNT (Multi Wall Carbon Nanotube) composite in lead free solder melt. They found that the wetting conditions decreased, but the resistance against erosion was increased [12]. Although a promising material might have been found, the fabrication of the developed composite material is quite costly and its production can only be carried out in laboratory conditions.

The long-term objective of our research is to develop an economic and commercially obtainable substrate-coating material combination that exhibit proper wetting with the most commonly used lead-free SAC solder alloys and increased erosion resistance compared to the currently applied iron tools. The first step of our research is the selection of the proper substrate material and the wetting examinations of nitride coatings. For substrates, three steel types were chosen with varying alloying contents, being W302, 42CrMo4 and C45. For coatings, Fe-nitrides were produced at the surface of the substrates with two different nitriding processes. In this manuscript, the results of wetting contact angle measurements are presented and the correlation between contact angle and substrate composition is discussed. The potential substrate-nitride combination of selective soldering tool applications is selected based on the results of the wetting examinations. The examination of the erosion behaviour of the potential substrate-nitride combination will be the next step of the research and is not yet performed.

2 Materials and Methods

Three steel types were chosen for substrates with decreasing alloying element contents being W302, 42CrMo4 and C45, respectively. The standard compositions of the examined steels are given in Table 1.

Plates with dimensions of 8 mm × 10 mm × 1 mm were cut from the bulk alloys. The plates were austenized in an air furnace at 860 °C for 30 min, then quenched in room temperature water. The plates were subsequently annealed at 600 °C for 20 min in air furnace. According to Böhler's recommendation, at least 2 h of annealing is required in the case of W302 steel. The annealing time was chosen to be shorter, since the subsequent nitriding ensured the completion of the annealing process. After removing the decarburized surfaces the plates were divided into two groups. The two groups were subjected to different nitriding processes. For the first group, before the nitrocarburising a preoxidation at 350 °C for 30 min in an air medium was carried out. Then the plates were heated up to the nitriding temperature of 550 °C in nitrogen. This heating type is the standard heating process in conventional gas nitriding that removes contaminations and produces an oxide layer on the surfaces. After 6 h of the nitriding, the furnace was cooled down and the samples were removed. Samples of such treatment are called here as "oxidized" For the second group, the plates were heated up to 550 °C in nitrogen atmosphere. Thus, no oxide layer was produced. After 6 h of nitriding the furnace cooling down and the samples were removed. Samples of this group are called as "oxide-free".

X-ray diffraction (XRD) phase analysis was performed with a 40 kV 40 mA Bruker D8 Advance diffractometer using Co tube.

Before wetting angle measurements, the surfaces of the samples were cleaned with a commercial flux, Lux-Tools DIN EN 29454. The samples were placed in an air furnace with a small piece of SAC 305 solder alloy on the top. The composition of the SAC 305 alloy is given in (Table 2) After holding the samples at 320 °C for

Table 1 The standard compositions of the examined steels, wt%

	C	Si	Mn	Cr	Mo	Ni	V	W	Co	P	S
W302	0.39	1.10	0.40	5.20	1.30	–	0.95	–	–	–	–
42CrMo4	0.38–0.45	<0.40	0.60–0.90	0.90–1.20	0.15–0.30	–	–	–	–	<0.035	<0.025
C45	0.42–0.50	<0.40	0.50–0.80	<0.40	<0.10	<0.40	–	–	–	<0.045	<0.045

20 min, they were removed and cooled down to measure the equilibrium contact angles. If the contact angles measured on the two sides of the same sample happened to deviate with 10°, the results were neglected. Such deviations originated from surface roughness caused by improper sample machining. The presented results are the average of contact angles measured on the two sides of the same sample.

Glow Discharge Optical Emission Spectrometry (GD-OES) examinations were carried out with a GD Profiler 2 (Power: 25 W, Module: 6 V, Phase: 5 V, Pressure: 500 Pa, flushing time: 5 s, pre-integration time: 100 s).

3 Results

Figure 1 shows the X-ray diffraction patterns of the oxide-free nitride samples. Bragg reflections of $Fe_{2\text{-}3}N$ (ε) and Fe_4N (γ′) nitrides and the ferrite substrate can be seen. No reflections of any other phases are present. Note that Co tube was applied, which excites Cr atoms. Due to the fluorescence radiation originating from the excitement of the Cr atoms, the sign/background ratio of W302 alloy is lower

Table 2 Composition of the SAC 305 alloy used for wetting angle measurements, wt%

Sn	Ag	Cu
96.34	2.95	0.59

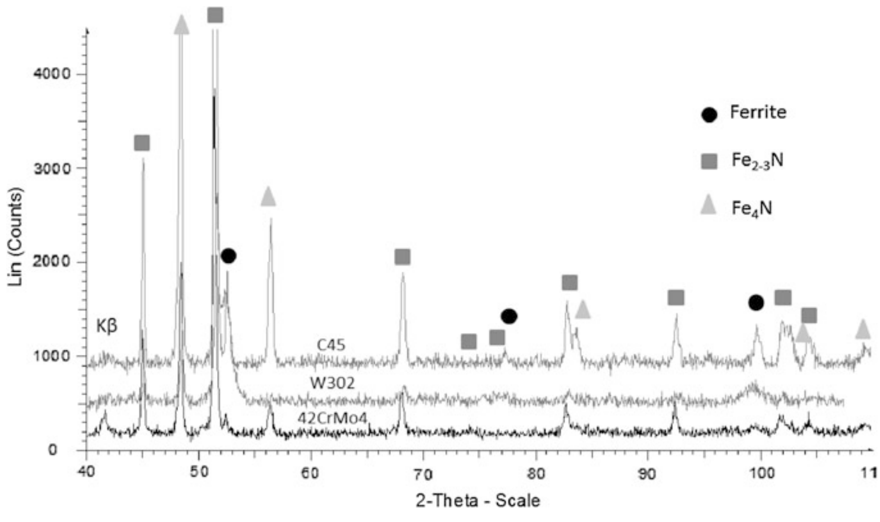


Fig. 1 XRD patterns of the oxide-free nitride coated W302, 42CrMo4 and C45 samples

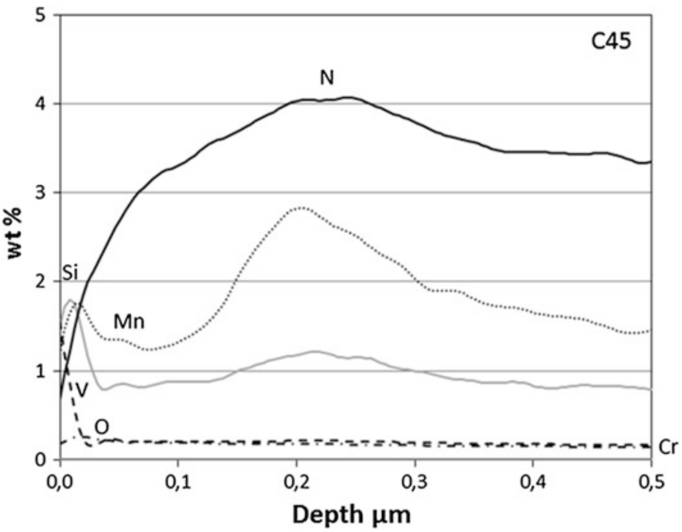


Fig. 2 GD-OES spectra of the C45 alloy with oxide-free nitride coating

compared to 42CrMo4 and C45, therefore lower intensity peaks are not visible on the spectrum of the W302 sample.

Figure 2 shows the near-surface region of GD-OES spectra of alloying elements of the C45 sample with oxide-free coating. Since the equipment was not calibrated for N and O elements, their absolute value could not be determined. However, the distributions of these elements are well represented. For other elements, both their distributions and absolute values could be determined. It can be seen that the increase of the N content is monotonic with increasing depth. There is a small peak of O in the sub-surface region, meaning that some O atoms were present during the nitriding process. The Si and Mn show definite peaks near the surface. The relative positions of Si, Mn and O peaks indicate that compounds/complexes of Si, Mn, and O formed on the surface with an amount less than 2 wt%. There is a small quantity of V in the alloys, which diffused from the bulk to the surface.

Figure 3 shows the near-surface variation of elements of the 42CrMo4 sample with oxide-free coating. Again, the N content increases monotonically with depth in the near surface region. Only a small peak appears in the Mn content with less than 1 wt%.

Figure 4 shows the GD-OES spectra of the W302 sample with oxide-free coating. It can be seen that Cr, Si, Mn, O have peaks at the same positions. Furthermore, the distribution of N is not monotonic, but having a local peak exactly where Cr, Si, Mn and O. This means that compounds/complexes of Cr, Si, Mn, O and N formed at the surface. The net amount of these compounds exceeds 5 wt%. The V diffused from the bulk to the surface.

Figure 5 shows the measured equilibrium contact angles of the oxidized samples and molten SAC 305 solder alloy systems. The W302 sample has the highest

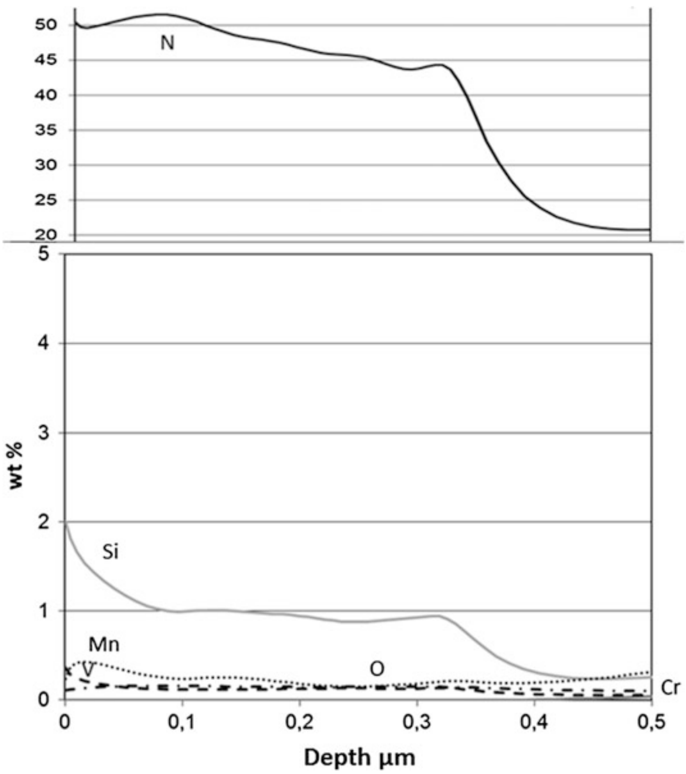


Fig. 3 GD-OES spectra of the 42CrMo4 alloy with oxide-free nitride coating

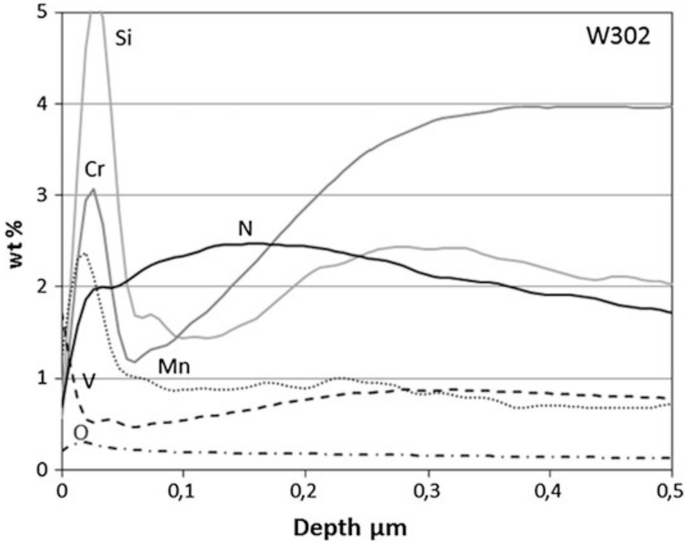


Fig. 4 GD-OES spectra of the W302 alloy with oxide-free nitride coating

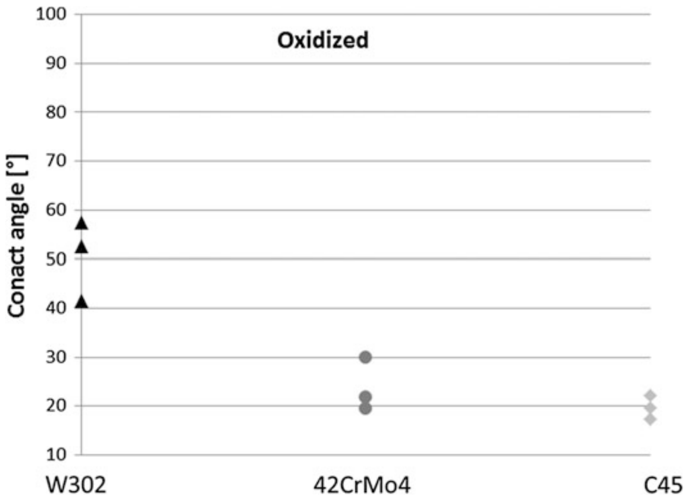


Fig. 5 Contact angles of the SAC305 alloy on the oxidized nitride coated W302, 42CrMo4 and C45 substrate

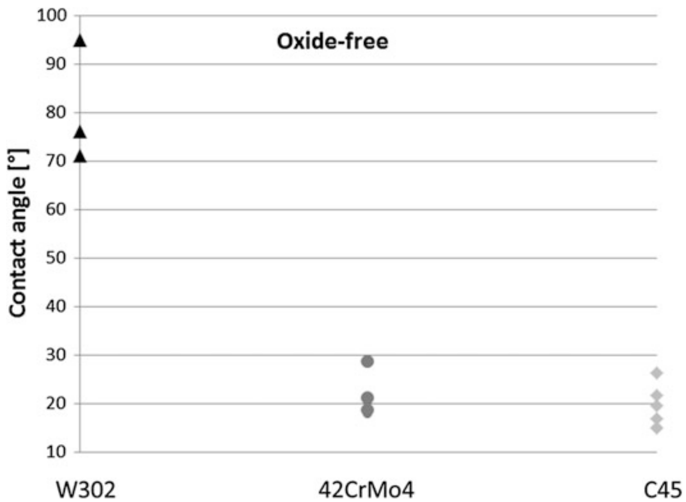


Fig. 6 Contact angles of the SAC305 alloy on the oxide-free nitride coated W302, 42CrMo4 and C45 substrate

contact angles of ~40°–50°, thus, it has the poorest wetting with the SAC 305 alloy. The contact angle for the 42CrMo4 samples are ~20°–30°, while those of the C45 samples are ~20°.

Figure 6 shows the contact angles of the oxide-free samples and molten SAC 305 solder alloy systems. The contact angles for the W302 samples are ~70°–100°.

As for the 42CrMo4 and C45 samples, the contact angles are the same, being $\sim 20°$. It is worth to note that the oxide-free nitriding notably increased the wetting angles of the W302 samples compared to the oxidized nitriding. For the 42CrMo4, there is only a slight increase in contact angles, while none is for the C45 samples.

4 Discussion

XRD examinations confirmed that $Fe_{2\text{-}3}N$ (ε) and Fe_4N (γ′) nitrides were produced at the surface of the oxide-free nitrided samples. For the C45 and 42CrMo4 alloys, GD-OES examinations revealed that no notable quantity of oxides/nitrides of alloying elements or impurities formed during the oxide-free nitriding. The formation of such compounds was inhibited because of the low alloying element concentration of the C45 and 42CrMo4 steels. It was concluded that the good wetting of SAC 305 soldering alloy melt on the oxide-free nitrided C45 and 42CrMo4 samples were due to the lack of additional oxides/nitrides, i.e. the desired wetting originated from the favourable adhesion between the $Fe_{2\text{-}3}N$ (ε) and Fe_4N (γ′) nitrides and the SAC 305 solder alloy melt. As for the W302 alloy, high wetting angles were measured after the oxidized nitriding, furthermore, wetting angles were even higher after the oxide-free nitriding. According to GD-OES examinations, nitrides of alloying elements such as Cr, Si and Mn formed during the oxide-free nitriding at the surface of the W302 alloy. It is well known that CrN is non-wettable for most metal melts, and, because of this character, it is used as non-wetting coating in soldering and aluminum casting processes [13–15]. It was deduced that the poor wetting of molten SAC 305 alloy on the nitrided W302 alloy was due to the presence of non-wetting compounds on the surface during nitriding. Finally, the obtained results showed that low alloyed steel substrates such as C45 or 42CrMo4 with nitride coatings can be candidates for wetting coatings within lead-free soldering applications.

5 Conclusions

The suitability of steel substrate-nitride coating combinations as wettable materials with enhanced erosion resistance for lead-free soldering applications are being investigated. This paper reports about the correlation between substrate composition and wetting behaviour. Samples of W302, 42CrMo4 and C45 steels were covered with $Fe_{2\text{-}3}N$ (ε) and Fe_4N (γ′) nitride layers using gas-nitriding. For the first set of alloys, conventional oxidizing was performed prior to nitriding, while for the second sets, the oxidizing step was skipped. The wettability was examined in the pre-oxidized and oxide-free samples/SAC 305 solder alloy systems. It was found that the composition of substrates strongly affects the wettability through formation of nitrides and oxides of alloying elements at the surface. The best wetting with contact angles of $\sim 20°$ was achieved with the C45 type steel substrate. It was deduced that low alloyed steel substrates with nitride coatings can be candidates for wettable materials for lead-free soldering tools.

Acknowledgements The authors are grateful for Tibor Kulcsar for the GD-OES examinations. The described article was carried out as part of the EFOP-3.6.1-16-00011 "Younger and Renewing University—Innovative Knowledge City—institutional development of the University of Miskolc aiming at intelligent specialisation" project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund. Marton Benke was further supported by the Postdoctral Researcher Fellowship of the Hungarian Academy of Sciences and Peter Baumli by the Janos Bolyai Research Fellowship of the Hungarian Academy of Sciences.

References

1. Nishikawa H, Takemoto T, Kifune K, Uetani T, Sekimori N (2004) Effect of iron plating conditions on reaction in molten lead-free solder. Mater Trans 45(3):741–746. doi:10.2320/matertrans.45.741
2. Information on http://www.elexp.com/Images/Weller_Coping_with_Lead_Free.pdf
3. Shangguan D (2005) Lead-free solder interconnect reliability. ASM International. ISBN:978-0-87170-816-8
4. Henshall G, Bath J, Handwerker CA (2011) Lead-free solder process development. Wiley. ISBN:978-0-470-41074-5
5. Ipser H (2007) European cost action 531: basic research on lead-free soldering. J Mining Metallurgy 43B(2):109–112. doi:10.2298/JMMB0702109I
6. Kroupa A, Dinsdale AT, Watson A, Vrestal J, Zemanova A (2007) COST531 Project—study of the advanced materials for lead-free soldering. J Mining Metallurgy 43B(2):113–123. doi:10.2298/JMMB0702113K
7. Kroupa A, Dinsdale A, Watson A, Vřešťál JJ, Zemanova A, Broz P (2012) The thermodynamic database cost MP0602 for materials for high-temperature lead-free soldering. J Mining Metallurgy 48(3)B:339–346. doi:10.2298/JMMB120711043K
8. Takemoto T, Uetani T, Yamazaki M (2004) Dissolution rates of iron plating on soldering iron tips in molten lead-free solders. Soldering Surface Mount Technol 16(3):9–15. doi:10.1108/09540910410562473
9. Nishikawa H, Kang S, Takemoto T (2009) Erosion behavior and interfacial reaction of stainless steels in molten lead-free solder. Trans JWRI 38(2):53–56
10. Information on http://www.circuitspecialists.com/blog/soldering-tip-care-tips
11. Nishikawa H, Komatsu A, Takemoto T (2005) Interfacial reaction between Sn–Ag–Co solder and metals. Mater Trans 46:2394–2399. doi:10.2320/matertrans.46.2394
12. Watanabe J, Sekimori N, Hatsuzawa K, Uetani T, Shohji I (2010) Study on erosion resistance characteristics of Fe-MWCNT composite plating with respect to lead-free solder. J Phys: Conf Ser 379(012025):1–10. doi:10.1088/1742-6596/379/1/012025
13. Wang Y (1997) A study of PVD coatings and die materials for extended die-casting die life. Surf Coat Technol 94–95:60–63. doi:10.1016/S0257-8972(97)00476-3
14. Gulizia S, Jahedi MZ, Doyle ED (2001) Performance evaluation of PVD coatings for high pressure die casting. Surf Coat Technol 140:200–205. doi:10.1016/S0257-8972(01)01037-4
15. Panjan P, Cekada M, Kirn R, Sokovic M (2004) Improvement of die-casting tools with duplex treatment. Surf Coat Technol 180–181:561–565. doi:10.1016/j.surfcoat.2003.10.119

The Effect of Tensile Strength on the Formability Parameters of Dual Phase Steels

Gábor Béres and Miklós Tisza

Abstract The importance of dual phase (DP) steels in the automotive industry was continuously spreading in the last decade. With their special microstructure—containing ferrite and martensite in particular ratio—high strength and increased formability is available. That's the reason why the application of DP steels is providing to exceed the 50% in a modern car body structure, according to the European program of Ultra-Light Steel Auto Body Advanced Vehicle Technology. This paper presents the experimental results of hemispherical dome tests and uniaxial tensile tests of three types of DP steels: DP 600, DP 800 and DP 1000. The effect of the tensile strength on the formability was investigated. It was described by the total and the ultimate tensile elongation, the average anisotropy and the limiting dome height (LDH). Based on our results it can be concluded that both the total and the uniform elongations are nearly linearly decreasing as the tensile strength is increasing. The slope of total elongation is more sensitive to the strength growth. However, it is no longer true for the plastic anisotropy. The reduction rate of average anisotropy stops over 800 MPa, and does not change until 1100 MPa. According to the dome tests results, the formability is also influenced by the sample geometry—through the deformation path—besides the tensile strength. The LDH values in biaxial stretch strain conditions are less dependent on the tensile strength. They are within a 3 mm interval for all three strength classes. In plane or stretch-press strain conditions, higher reduction can be observed. The characteristics of dome height curves are similar for all samples, regardless of their strength.

G. Béres (✉)
Department of Materials Technology, Faculty of GAMF Technical and IT, University of Pallasz Athéné, Kecskemét, Hungary
e-mail: beres.gabor@gamf.kefo.hu

M. Tisza
Department of Mechanical Technologies, Faculty of Mechanical Engineering and IT, University of Miskolc, Miskolc, Hungary
e-mail: tisza.miklos@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*, Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_19

1 Introduction

Weight-loss developments in the vehicle industry resulted in relevant research results in the science of structural materials. Spreading of non-ferrous materials—like aluminium alloys and composites—can be observed, but the steels still remain determinative materials for car body manufacturing. To reducing weight, newly developed, advanced high strength steels (AHSS) and ultrahigh strength steels (UHSS) have been used as car body elements. Earlier [1] and present-day results [2]—validated with FE modelling simulations—showed that using these steels allows thinner components for lightweight constructions with the same, or even with improved stiffness and crashworthiness. Furthermore, a significant reduction of fuel consumption and greenhouse gas emission can be available in this way. 57 kg weight reduction is equivalent to 0.09–0.21 l/km fuel saving [2].

Application of AHSS materials significantly increased formability beside higher strength compared to conventional steel types. AHSS includes dual phase (DP), complex phase (CP) and transformation induced plasticity (TRIP) steels [3].

DP steels are typically multi-phase steels, widely used in the automotive industry due to their favourable high strength and good formability. The application of these multi-phase steels for car body panels production exceeds 50% according to the European ULSAB-AVC (Ultra Light Steel Auto Body—Advanced Vehicle Concept) consortium in 2011 [4].

DP steels are commonly manufactured by hot rolling, which followed by fast cooling from the inter-critical temperature interval below to Ms temperature. In this way, their microstructure consists of hard martensite islands in soft ferrite matrix. The mechanical properties of DP steels are mainly depending on the amount, the morphology, the dispersion and the cohesion of these two phases [3–6]. With key factors of the manufacturing, the characteristics of the ferrite and martensite particles can be significantly controlled. These key factors are the holding temperature or the holding time after the hot rolling, and the cooling rate respectively. Many studies [4–7] present the effect of the grain structure for the formability of DP steels. Some of them [4–6] use microstructure-based modelling, some other work [7] deals with experimental investigations. Microstructure-based models use representative volume elements (RVEs) to describe the multiphase material features. The definition of such model is time-consuming, however can lead to a more precise solution.

This paper is intended to present the experimental results of the effect of the tensile strength on the formability parameters of DP steels. Measuring the tensile strength is easier than that of the microstructure properties and provides sufficiently accurate benchmark for industrial users.

Table 1 Chemical composition of the investigated DP steels (%)

	C	Si	Mn	P	S	Cr
DP 600	0.085	0.171	0.872	0.013	0.005	0.014
DP 800	0.161	0.187	1.520	0.012	0.003	0.027
DP 1000	0.148	0.181	1.500	0.012	0.004	0.042

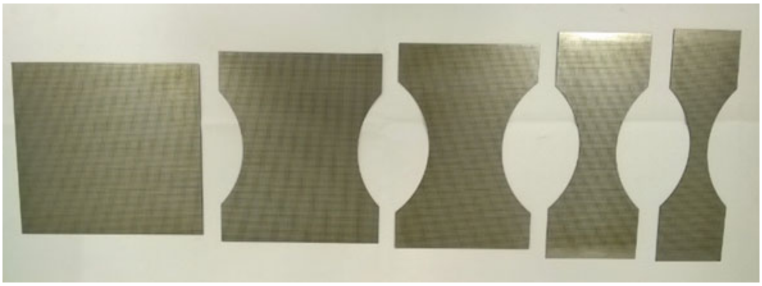


Fig. 1 Modified Nakajima specimens

2 Samples and Materials

Three types of commercially available dual phase steels—DP 600, DP 800 and DP 1000—were investigated by mechanical testing. Initial table sizes were 2000 × 1250 × 1 mm. The chemical compound of the applied materials is summarized in Table 1.

The given values were acquired by optical spectroscopy measurements. The results are in good agreement with the supplier's certification [8]. It is worthwhile to note that DP 800 has slightly higher carbon content than DP 1000. Nevertheless, both types fall into the manufacturing tolerances.

Tensile tests and limiting dome height (LDH) tests were applied as mechanical investigations. Tensile tests were carried out in accordance with MSZ EN ISO 6892 standard. The crosshead speed was 25 mm/min. Both the tensile and doming samples were cut by laser beam from 1 mm thick blanks, with tensile sample orientation 0°, 45° and 90° angles with respect to the rolling direction. Their dimensions also aligned to the standard mentioned above. Doming specimens called modified Nakajima specimens (Fig. 1) were manufactured based on literature [9]. The longitudinal axes of the samples were perpendicular to the rolling direction, since it is known that this direction shows the worst results in terms of stretch formability [9]. The applied blank holder force was 120 kN, which totally prevent the movement of the sheets under the blank holder. The diameter of hemispherical punch was 100 mm and it passed with 50 mm/min stroke speed. Surface conditions were prepared without lubrication.

3 Results

3.1 Tensile Tests

Primary values of tensile test results can be seen in Table 2. The engineering stress-strain curves are shown in Fig. 2. It can be seen, that the ultimate tensile strength values are close to the values of denomination signs of each steel grade. The highest tensile strength is measured at the DP 1000 and then it continuously decreases toward the DP 600. Conversely, the total elongation increases with the strength reduction.

Plastic anisotropy ($\bar{r}$) cannot be deduced from the figure, but reports the thinning tendency of the materials. The calculation method of it can be described by Eq. (1):

$$\bar{r} = \frac{r_0 + r_{90} + 2 \cdot r_{45}}{4}. \tag{1}$$

In this context r_0, r_{90} and r_{45} are the measured values in the direction 0°, 90° and 45° compared to the rolling direction. The relationship between each r values in different directions can be followed on Fig. 3. It is observed that in case of the DP 600 the anisotropy increases linearly and the highest r value appears at 90°. The

Table 2 Basic mechanical parameters

	E (N/mm^2)	Re (N/mm^2)	Rm (N/mm^2)	A_{80} (%)	$\bar{r}$
DP 600	2.1*10^5	445	656	20.6	0.92
DP 800	2.1*10^5	571	879	16.0	0.75
DP 1000	2.1*10^5	767	1099	10.6	0.76

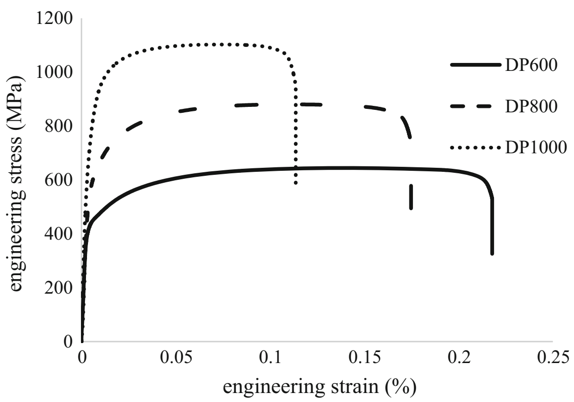


Fig. 2 Engineering stress-strain curves

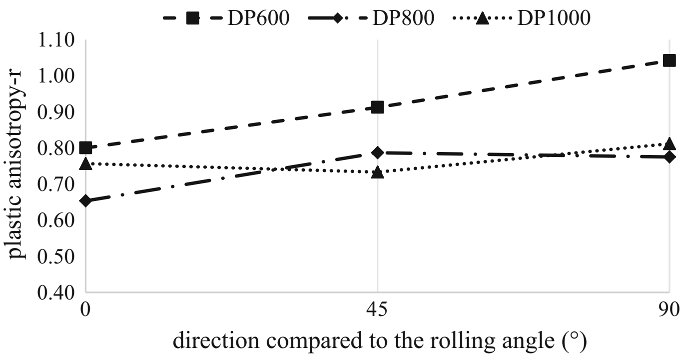
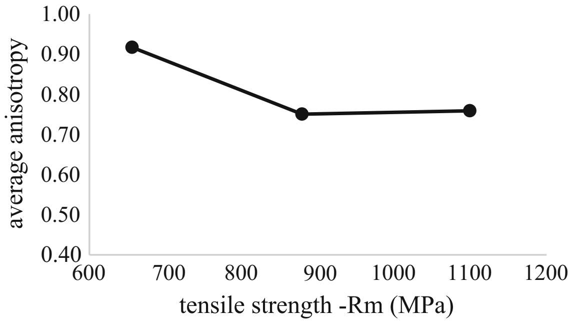


Fig. 3 Plastic anisotropy values in each specimen direction

Fig. 4 The average anisotropy dependence on tensile strength



functions' forms are not linear and the changing tendency shows opposite behaviour for the DP 800 and DP 1000.

The dependence of the average anisotropy on the tensile strength is added by Fig. 4. Based on the tensile strength—which is a function of the microstructure properties, like martensite volume fraction, martensite carbon content or grain size—the average anisotropy does not significantly change over 880 MPa. It increases below that, up to 660 MPa surely. This fact means, that such DP steels could behave similarly during deep-drawing operation, which has higher strength than 880 MPa. It is also confirmed by our previous research work [10].

The changing attitude of the total and uniform elongations deviate from the anisotropy in the function of tensile strength. Both parameters are decreasing linearly, as the tensile strength increases (Fig. 5). From these results, it can be concluded that the stretch formability of such steels also decrease linearly with the increase of strength.

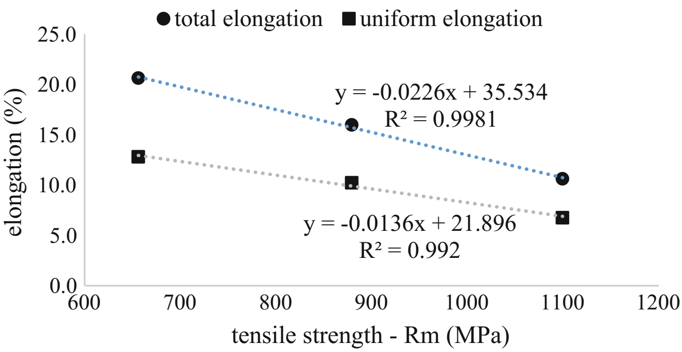


Fig. 5 Changing of the elongation in the function of strength

Table 3 LDH values at different specimen bridge-widths

Bridge-width (mm)	20	40	80	125	200
DP 600	21.3	27.3	24.9	26.9	37.2
DP 800	19.3	23.9	23.9	24.4	37.4
DP 1000	15.1	20.0	20.2	22.1	35.1

3.2 *Limiting Dome Height Tests*

Limiting dome height (LDH) tests were performed to measure the maximal dome height for different sample geometries. The applied geometries (shown in Fig. 1) make possible the investigation of the effect of deformation path on stretch formability. The measured limiting dome height values in mm, and the applied sample bridge-width dimensions are shown in Table 3.

The LDH curves in the function of bridge-width are given by Fig. 6. Note that the LDH values are decreasing with the increase of the strength. The characteristics are similar, the lowest is at 20 mm width, and the highest is at 200 mm width for all materials. From these results it can be concluded that these steels resist better against stress in biaxial stretch strain condition than in-plane or press-stretch strain conditions.

If the LDH values are taken into account as the function of the tensile strength, the biaxial stretch strain condition seems to provide the best formability likewise (Fig. 7). Samples marked by “LDH_200” refers to the LDH of the specimens with 200 mm bridge-width, but different strength. It can be stated that the changing characteristic is nearly linear, and the dome height results are within 3 mm interval, while the tensile strength is almost doubled. Higher changing—roughly 5 mm—appeared in the already lower LDHs in plane strain condition. It is displayed by “LDH_80” nomination in the figure. The biggest deviation is discovered in

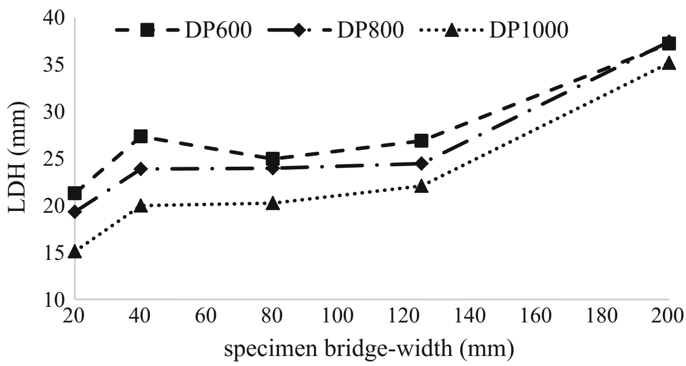


Fig. 6 LDHs at different sample geometry

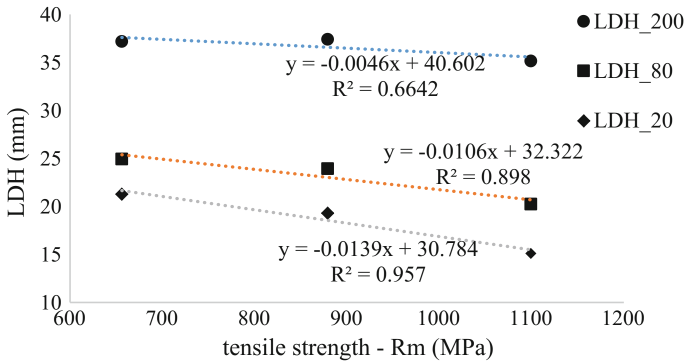


Fig. 7 The effect of the tensile strength and deformation path for the stretch formability

compress-stretch deforming path (LDH_20). Thus the LDHs less dependent on the tensile strength in biaxial stretch processes, but its influencing effect intensifies with the approaching of the compressing deformation states.

4 Summary

Mechanical investigations of three types of dual phase steels were performed by tensile and limiting dome height tests. The exact material types were commercially available DP 600, DP 800 and DP 1000. Since the strength properties are in direct connection with their microstructure, formability parameters were investigated according to their tensile strength as the basis of analysis. Total and ultimate elongation, anisotropy and limiting dome height behaviour were also examined. From these investigations the following main conclusions may be drawn:

- The changing characteristic of plastic anisotropy in different deforming directions is similar for DP 800 and DP 1000. Average anisotropy doesn't decrease further over 800 MPa.
- Total and uniform elongation reduces linearly with the increase of the strength. The slope of the total elongation reacts stronger for the strength modification.
- LDH values are significantly influenced by both the tensile strength and the deformation paths. Favourable deformation capability can be observed in biaxial stretch strain conditions, almost independently the materials' strength. In plane and compress conditions, lower LDH values and stronger strength dependence can be observed.

Acknowledgements The presented research work was supported by Ministry of Human Resources, in the program of National Talent Program NFTÖ-16—scholarships to provide individual development.

References

1. Li Y, Lin Z, Jiang A, Chen G (2003) Use of high strength steel sheet for lightweight and crashworthy car body. Mater Des 24:177–182
2. Cui X, Zhang H, Wang S, Zhang L, Ko J (2011) Design of lightweight multi-material automotive bodies using new material performance indices of thin-walled beams for the material selection with crashworthiness consideration. Mater Des 32:815–821
3. Kuziak R, Kawalla R, Waengler S (2008) Advanced high strength steels for automotive industry. Arch Civil Mech Eng VIII:103–117
4. Uthaisangsuk V, Prahl U, Bleck W (2011) Modelling of damage and failure in multiphase high strength DP and TRIP steels. Eng Fract Mech 78:469–486
5. Surajit Kumar Paul (2013) Real microstructure based micromechanical model to simulate microstructural level deformation behavior and failure initiation in DP 590 steel. Mater Des 44:397–406
6. Sun X, Choi KS, Soulami A, Liu WN, Khaleel MA (2009) On key factors influencing ductile fractures of dual phase (DP) steels. Mater Sci Eng, A 526:140–419
7. Movahed P, Kolahgar S, Marashi SPH, Pouranvari M, Parvin N (2009) The effect of intercritical heat treatment temperature on the tensile properties and work hardening behavior of ferrite–martensite dual phase steel sheets. Mater Sci Eng, A 518:1–6
8. Official certificate of material testing laboratory of SSAB EMEA AB, Borlange, Sweden
9. Tisza M, Kovács PZ, Lukács Zs (2015) Formability of high strength sheet metals with special regard to the effect of the influential factors on the forming limit diagrams. Mater Sci Forum 812:271–275. doi:10.4028/www.scientific.net/MSF.812.271
10. Danyi J, Végvári F, Béres G (2016) Járműipari célú acéllemezek mélyíthetőségi és mélyhúzhatósági problémái. Miskolci Egyetem Közleményei: Anyagmérnöki Tudományok 39(1):19–28

Comparison of Two Laser Interferometric Methods for the Study of Vibrations

Miklós Béres and Béla Paripás

Abstract We performed a comparison between two laser interferometric methods. One of the methods is used by our modernized Michelson type laser interferometric motion analyser (LIMA), which measures primarily the displacement, while the other method is the one used by an LDV (Laser Doppler Vibrometer) device, which primarily measures the velocity on the basis of the Doppler effect. The motion that is measured is the vibration of the corner cube reflector (retroreflector) of the measuring arm of the LIMA. During the measurements the LDV measured the back plate of the house of the retroreflector. These simultaneous analyses of the vibrations made possible the comparison of these two methods for the precise velocity studies. In general, it was found that the velocity values, obtained by the two methods correspond to each other within 2% accuracy.

1 Introduction

Laser measurement techniques, especially the laser interferometrical methods, are widely used in automotive developments [1–3]. The Institute of Physics at the University of Miskolc and the Wigner Research Centre for Physics developed a precision laser interferometric motion analyser (LIMA) [4, 5] more than two decades ago. At the present, it serves demonstration purposes, but the goal is to make it appropriate for the completion of practical tasks (e.g. certifying industrial equipment). Since the construction was relatively old, the technical part had to be converted in accordance with the requirements of the present day. Some operating units of the equipment remained—such as the laser, optics, detectors, etc., but the control of the analyser and the analysis of the measured data had to be upgraded. Due to the operating principle of the device—which we describe later—it measures the displacement

M. Béres (✉) · B. Paripás
University of Miskolc, Miskolc, Hungary
e-mail: mechbere@uni-miskolc.hu

B. Paripás
e-mail: fizpari@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_20

accurately even without calibration, and the velocity of the point can be calculated accurately from the sub-micron resolution data (with the appropriate software background). The tests and developments were carried out in the Institute of Physics [6] with the support of the Wigner Research Centre for Physics.

Our other motion analyser that uses the interferometric principle is a Polytech PDV-100 type LDV (Laser Doppler Vibrometer) device [7]. This device primarily measures the velocity, therefore, based on its operating principle it fundamentally differs from the displacement meter we developed. Measuring simultaneously with the two precision devices that use different principles, comparing the measurement results, and analysing the differences is certainly an exciting engineering task. This measurement comparison can be considered as the calibration of the LDV device, which is thought to be less precise.

2 Interferometric Measuring Systems

2.1 The Laser Interferometric Motion Analyser (LIMA)

The device is essentially a Michelson interferometer in which, by exploiting the polarisation, we doubled the accuracy and made the direction of motion also measurable. The increased accuracy is $\lambda/8$. Numerous such solutions can be found in the literature [8–10].

The opto-mechanical drawing of the interferometer is shown in Fig. 1. The heart of the interferometer is a temperature-stabilised single-mode He–Ne laser (UNIPHASE-1007). The frequency of its mode is $4.7375 \cdot 10^{14}$ Hz, which corresponds to a wavelength of 632.82 nm in vacuum and 632.64 nm in normal air. The laser beam formed by the beam expander is split into two perpendicular, approximately the same intensity beams by the polarizing beam splitter prism in such a way that the polarisation of the two beams are also perpendicular to each other. One beam is reflected from the reference corner cube prism (at the top of the diagram), while the other is reflected from the movable measuring corner cube prism (on the right side of the diagram). The beams meet again in the same polarizing beam splitter prism where they previously split up. The phase relationship of the two beams is determined by the path length difference between the two arms [4]. Naturally, this path length difference has to be within the range of the coherence length of the laser (which can be a few meters).

The combined beam is led to the detector unit by mirrors. There the laser beam is split into three parts by beam splitter prisms. After creating the appropriate phase shifts (0, $\lambda/2$, $\lambda/4$), these are detected by three independent photodiodes (BPx61) behind polarizers with the appropriate orientation. Out of the signals of these photodiodes, A and B are in opposite phase, while between A and C there is a 90° phase difference. Therefore, in total the detector system detects a maximum (or minimum) at every 90° (quarter period).

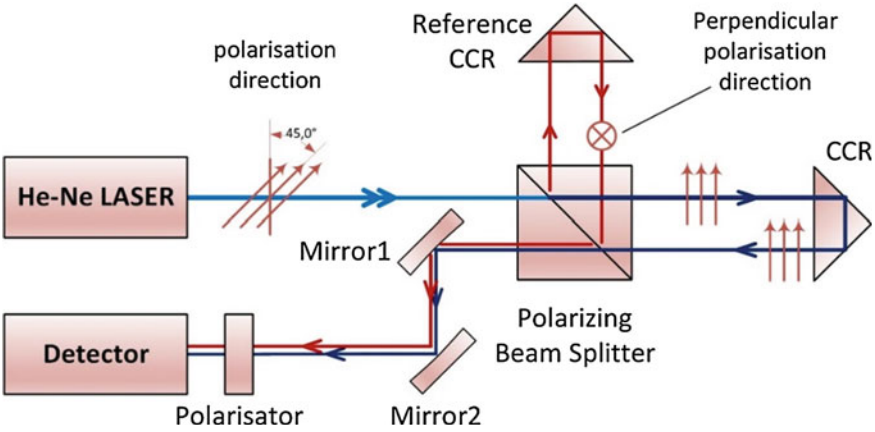


Fig. 1 Opto-mechanical outline of the interferometer (LIMA) at the institute

In this device the measurement is based on the detection of these, therefore the achievable accuracy in terms of change in the path length of light is λ/4. This in turn corresponds to a displacement of λ/8 for the corner prism in the measuring arm (since the light travels there and back). The signal order of the three detectors also determines the direction of motion of the measuring arm.

A measuring card and a LabView program written for the measurement are connected to the measuring system. The signals are transferred from the detector unit to the central unit. This requires an appropriate measuring card that can handle the signals of the three detectors (which in principle can follow a motion of 15 cm/s in the measuring arm, corresponding to 3 million pulses) and can also determine the signal order of the detectors. The National Instruments NI USB-6341 OEM type measuring card is connected to the computer (laptop) through a USB port. On the card the operation of the analogue, digital and counter subsystems are coordinated by the NI-STC3 timing and synchronisation technology, providing independent timers for the analogue and digital I/O subsystems that are on the same card. For us the four improved 32 bit counters are extremely important, which we can use for frequency measurement, pulse width measurement, and encoder operations. This latter is necessary for determining the signal order of the detectors, i.e. the direction of motion. In principle the counters can count at a speed of 100 MHz, which can be read by the computer at a frequency of 1 MHz, therefore the data handling and data transfer speed of the card is more than enough for our photo detectors.

The numerical and graphical visualisation of the metering data is also performed by the measuring program (Fig. 2). This system can be used to study various movements that we can generate for example by a differential screw micrometer, or by gently pushing and hitting the vibration-free steel table. The study of these is aided by the built-in Fourier analyser (FFT) that shows the frequency spectrum of the created vibrations.

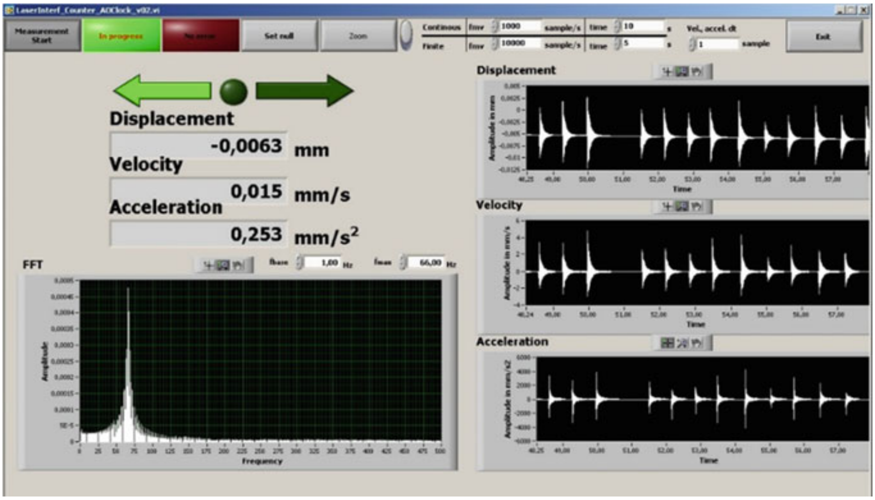


Fig. 2 The continuous measurement mode of the measuring program

2.2 *The LDV (Laser Doppler Vibrometer)*

The LDV (Laser Doppler Vibrometer) is a device that is suitable for contactless measurement, which we can use to examine the vibration of various objects. The laser beam exiting the LDV equipment has to be focused onto a surface of the object to be examined, and the velocity function of this surface can be inferred using the frequency of the reflected laser light based on the Doppler-effect [10]. If we are detecting the wave that is reflected from an object moving at velocity v (in case of a stationary wave source), then the detected frequency f differs from the original f_0 frequency by an f_D Doppler shift:

$$f = f_0 \frac{c + v}{c - v} \cong f_0 \left(1 + 2\frac{v}{c}\right) = f_0 + f_D \tag{1}$$

Therefore

$$f_D = 2f_0 \frac{v}{c} = 2\frac{v}{\lambda} \tag{2}$$

In case of a receding reflecting object the signs are opposite; c is the phase velocity of the wave in the given medium. Naturally, this is only true if there is no cosine error, i.e. if the laser beam and the velocity vector are parallel.

Figure 3 shows the schematic setup of the vibrometer. The laser beam arriving from the helium–neon (He–Ne) laser (with frequency f_0) is split by the beam splitter into a reference and a measuring beam. The measuring beam arrives onto the target

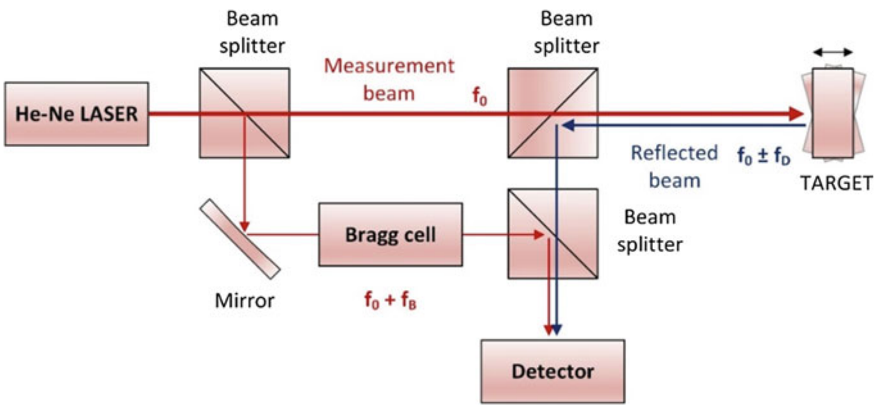


Fig. 3 Schematic setup of the vibrometer

object, from which it reflects and changes its frequency by an f_D Doppler shift because of the motion of the object in accordance with Eq. 2.

The light is reflected from the target object in every direction, but a portion of the light is collected by the LDV optics, and it is reflected onto the photo detector through the beam splitter. The frequency of this light is $f_0 + f_D$. The reference beam passes through the Bragg cell, which adds a frequency shift of f_B (in our case 20 MHz). The scattered light interferes with the reference beam on the photo detector. The frequency difference of the two interfering waves also appears ($f_B - f_D$), and this value falls into the 10 MHz range. This is such an intensity variation that the photo detector can already detect (it cannot follow the original 10^{14} Hz variation). The output of the photo detector is a standard frequency modulated (FM) signal, with the Bragg cell as carrier frequency and the Doppler shift as modulation frequency. From this signal the time dependence of the velocity of the target object can be determined by demodulation. The LDV equipment we use is a Polytec PDV-100 (Portable Digital Vibrometer) type vibrometer.

3 Measurements Results

In the previous chapters we showed that the LIMA is able to analyse the motion (its component that is in the direction of the laser beam) of the corner prism in the measuring arm. The LDV on the other hand does not require a corner cube prism (that is fixed onto the point to be measured), although a reflective film dot that is stuck onto the point greatly helps the measurement. Therefore, both devices are able to analyse the vibrations of the gently knocked corner prism in the measuring arm of the LIMA. The simplest way for this is when the two lasers of the two devices shine onto the two opposite sides of the corner prism. Thus the devices measure the projection of the motion along a common line, but from opposite directions.

Consequently, if the corner cube prism is receding from the laser of the LIMA, then it is approaching the laser of the LDV. It should be noted that the velocities for the two instruments still have the same sign, because the LIMA considers moving away to be a positive displacement, while for the LDV approach has positive velocity.

We performed our earlier comparison measurements on the old "vibration-free" table of the laser interferometry laboratory of the Institute of Physics, on the same table that the laser interferometric motion analyser was originally assembled. The quotation marks are warranted because the table mainly protected the analyser from the vibrations of the floor, however, it mutually transmitted the vibrations of the analyser units (laser, beam splitter prism, measuring prism) mounted to different points on the table. The situation was also worsened by the fact that we could only place the LDV equipment on a stand (tripod) that was independent from the table. All these resulted complex, hardly analysable vibration images that were slightly different on the two devices, which made the calibration unreliable as well.

During the last few months we managed to acquire a vibration-free Nexus Breadboard table top with a size of 900 × 1800 × 110 mm, which we placed onto our custom built frame and air springs. The vibration-free table configured this way, in addition to easily decreasing the amplitude of the vibrations of the environment below the measurement threshold of our laser interferometric motion analyser (0.1 μm), it effectively separates the units that are mounted at different points on the table. Therefore on this table the vibrations of e.g. the gently knocked corner prism in the measuring arm does not spread onto e.g. the beam splitter prism. With the newly acquired units we could also solve the mounting of the LDV onto the table. The measurements presented in this paper were already carried out on this table (Fig. 4).

As we have indicated in the preceding chapters, the LIMA primarily measures displacement, and the LDV measures velocity. Therefore the obtained graphs cannot be compared directly, e.g. from the displacement data we have to derive the

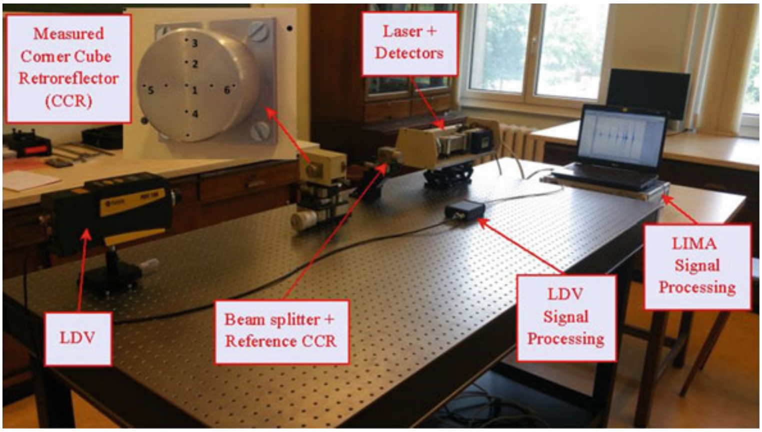


Fig. 4 Positioning of the devices

velocity. This is done by the LabView program with the frequency that can be set in its graphical interface (it divides the displacement measured in the previous time interval with the length of the interval). For these measurements we typically set a frequency of 1200 Hz (T = 833.33 μs), because the 1200 sample/s (and its multiples) can also be chosen on the LDV. If we want to compare the two velocity graphs, we also have to match the timescale of the two graphs. Since both measuring programs are run by the same machine, the factor of the two timescales are matching in principle, therefore the time differences should also match precisely. However, this is still not the case, in a 10 s measurement interval, the LDV velocity graph shifts by approximately 1 ms (one or two channels) relative to the one of the LIMA. We could not figure out the reason for this, initially we considered it as an error to be eliminated, but later it even helped our work. Namely, it helped with aligning the starting values of the scales. This was needed because we could not arrange for the μs precision simultaneous start of the two types of measurements. The rough aligning (that is about 0.5 ms precision) of the starting values of the two graphs can be done manually, because the knocking of the corner prism in the measuring arm starts the vibrations in a well determined way. The best match (possibly the best two matches) can be found visually, then our inaccuracy is therefore T/2 (≈0.5 ms) at most. The further increase of precision is helped by the 10-4 relative deviation of the timescales, because due to this there will surely be such a knock, where the alignment is more precise by an order of magnitude. This knock will be the subject of our comparative studies. Finding the best alignment was also helped by the time graph of the deviation of the velocity values measured by the two methods.

Figure 5 shows the result of a 12 s long knocking sequence, that is the damped vibrations following the knocks, on the aligned timescale. In the upper part of the figure the LDV (black) and LIMA (red) velocity data are represented together. (The LDV data are corrected, see later.) The two data sequences completely overlap, the differences cannot really be seen at this magnification.

Therefore, we calculated the differences as well, and are showing those at the bottom part of the figure. They show it quite well that the largest discrepancies are always at the first oscillation following the knock. Moreover, it can also be seen that after this the differences are smaller, but they have approximately constant amplitude all the way until the vibrations completely die down.

This is especially true after the knocks at around the 4th and the 7th second, where the alignment is obviously the best. In these ranges the discrepancies (therefore independently of the amplitude of the vibration, and excluding the first oscillation) steadily fall into the region between −0.1 and +0.1 mm/s. The reason for this definite maximum can be easily understood. Namely, the measurement precision of the LIMA (as we indicated earlier) is $\lambda/8 \approx 0.08$ μm, and the length of a time window is 1/1200 s. The ratio of these two gives 96 μm/s ≈ 0.1 mm/s, therefore this is the velocity measurement precision of the LIMA (the random error in the applied measuring program and sampling frequency). This velocity measurement precision could be improved by decreasing the sampling frequency, but then the equipment would be less sensitive to the faster vibration components.

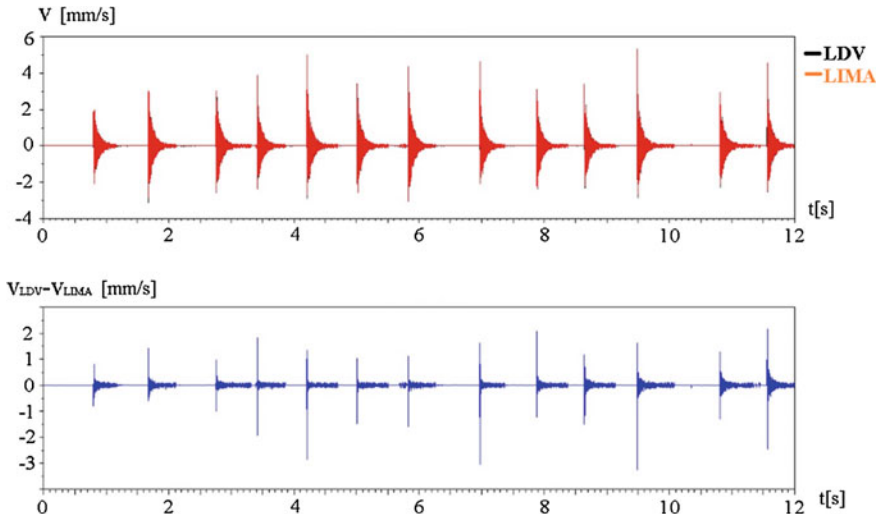


Fig. 5 The velocity data of the damped vibrations following a knocking sequence: *upper* LDV (*black*) and LIMA (*red*) velocity data together, *lower* their difference

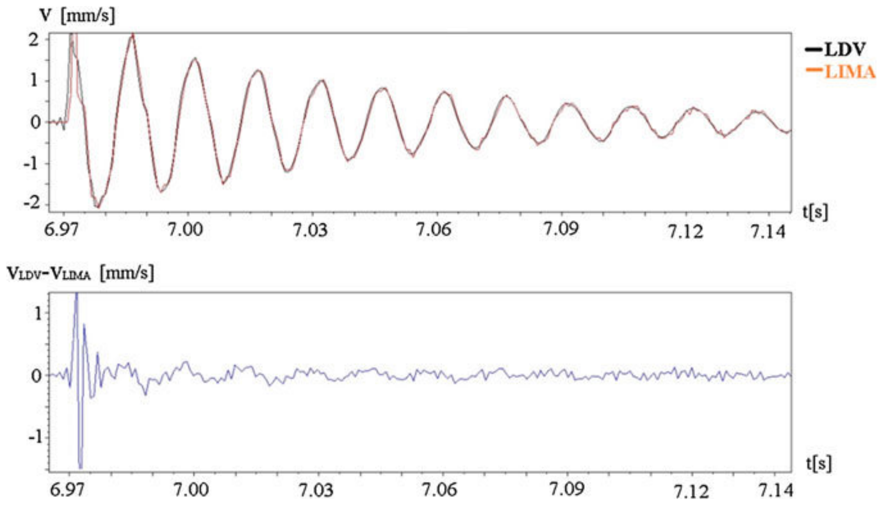


Fig. 6 The first few oscillations after the 8th knock zoomed in (The data are as in the previous diagram.)

In Fig. 6 we can also look at the few oscillations after the knock that happened at 6.97 s (otherwise the eighth) zoomed in. Starting from the second period, both devices measure a damped vibration of 67.8 Hz, with an almost perfect alignment. The significant difference is in the first half period. There the LDV measures an approximately 260 Hz quickly decaying component more definitely than the LIMA.

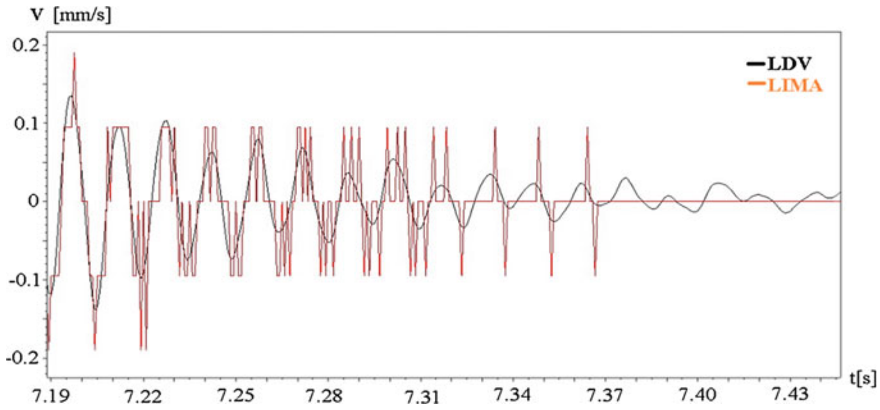


Fig. 7 The cessation of the oscillations after the 8th knock (The data are as in the *upper part* of Fig. 6.)

It's definitely because the slow vibration is the vibration of the whole corner prism, while the fast vibration is the vibration of the elements of the prism relative to each other. The light of the LDV laser is reflected at the back and in one point, while that of the LIMA is reflected at the front and in three points (on the three sides of the prism one after another). Therefore the LIMA somewhat averages the fast vibrations.

The monitoring of the dying down of the vibrations is also instructive. Figure 7 is essentially the continuation of Fig. 6 (with the omission of a few periods, and on a finer velocity scale). It can be seen that the LDV nicely follows the oscillations the whole time, which are not purely sinusoidal towards the end; they consist of the sum of multiple vibrations. However, in case of the LIMA the velocity becomes very spiky, essentially only three velocity values occur: zero and ±96 μm/s, depending on the fact if there was a signal arriving in the previous time window, and which sign displacement it belonged to (see above). With the decrease of the amplitude, the values that are different from zero become more and more scarce, and eventually they completely disappear. According to what has been written earlier, the LIMA cannot yet see the vibrations with amplitudes less than 50 nm. In the case of 67.8 Hz this measurement threshold corresponds to a ($v = A\omega$) 20 μm/s velocity measurement threshold, and this is also confirmed by the diagram.

Based on the foregoing, we can clearly state that with the current measuring program the LIMA is less suitable for the study of movements with a velocity of 0.1 mm/s than the LDV. This is so because although the LIMA measures the displacement very accurately (in λ/4 units), it derives the instantaneous velocity (and acceleration) inaccurately from this, because the time windows are not synchronised to the arrival of the signals. Therefore, this derivation method (which we have already described at the beginning of the chapter) should be developed by all means. We emphasize that the "inaccuracy" of the LIMA applies only to the instantaneous velocity values; it does not apply to the average velocities

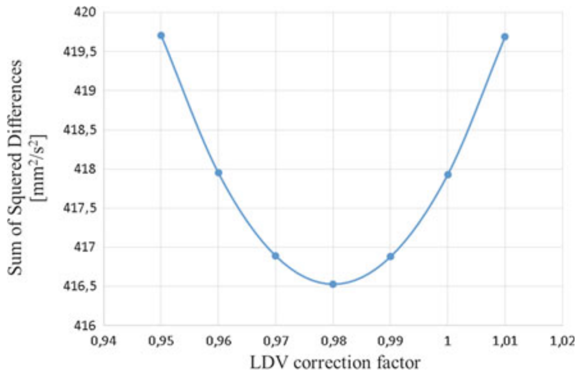


Fig. 8 Determination of the optimal LDV correction factor

Table 1 The LDV correction factors corresponding to the measurement points on the corner cube prism

Point number	Coordinates (mm)	Correction factor	LDV excess (%)
1	(0;0)	0.98	2.0
2	(0;7)	0.91	9.9
3	(0;14)	0.86	16.3
4	(0;−7)	1.05	−4.8
5	(−14;0)	1.00	0
6	(14;0)	1.00	0

corresponding to longer time periods. Those are very precise, because we calculate them from high precision data.

The key objective of the comparison measurements was to determine if the two measurement methods provide the same measurement results, and if there is a possibility for making the calibration of the LDV, which is thought to have a less precise scale, more precise. Figure 6 suggests that the alignment, and thereby the calibration are impeccable. This is roughly true, but Fig. 6 nevertheless contains a correction factor. The LDV velocity data had to be multiplied by 0.98, so that the alignment becomes the best (i.e. the sum of the squares of the discrepancies becomes minimal). This is shown in Fig. 8.

Does this mean that the LDV method is wrong by 2% (relative to the LIMA), and the “LDV Scaling factor” should be modified from the factory set 5 mm/s/V to 4.9 mm/s/V? The answer is clearly no. Namely, let’s see if the result of the correction depends on at which point the laser beam of the LDV hits the backplane of the corner prism. We have already indicated the measurement points at the backplane of the enlarged prism in the upper right part of Fig. 4; we measured the graphs showed in Fig. 5, 6 and 7 with the laser beam reflected from measurement point 1 (the centre of the backplane). We summarized the multiplication factors that we obtained in a similar way for the other measurement points in Table 1.

According to the data in Table 1, moving vertically upwards along the centreline of the prism the correction factor changes by about 1% for every mm. (The top of the corner prism swings out more, therefore at the top the LDV data have to be reduced more in order to obtain alignment with the LIMA.) Moving horizontally (also along the centreline) the required corrections are much smaller; the two edges of the prism vibrate with an amplitude that is about 2% greater than the middle. This indicates that the applied knock creates a vibration that has a small torsion component as well.

The laser beam of the LIMA enters and leaves the prism on its other side, at a distance of 15 mm from each other in its horizontal centre plane. Therefore, the motion of the reflecting points of the prism could be around the average of the centre (point 1) and the two edges (points 5–6), for which the corresponding correction factor could be 0.99. This means that the velocity data measured by the two devices differ by about 1% if we consider the points of the centreline of the prism with coordinates (7;0) and (−7;0). At the same time we think that the setting precision of the spot of the laser beam is approximately 1 mm on the backplane of the prism, which—as described above—leads to an approximately 1% error for the correction factor. That is in the end the discrepancy of the data is not larger than our estimated measurement error. Therefore the scaling factor of the LDV is accurate within a 1–2% margin of error. This (quite small) measurement error may be caused mainly by the fact that the LIMA measures the averaged vibrations of the whole corner prism, while the LDV only measures one point.

4 Conclusions

Based on the measurements carried out we can state that both the LIMA and the LDV work excellently given the appropriate circumstances and settings. The modernized data collection system fits the old opto-mechanical unit well.

The new vibration-free table with the air springs significantly improved the accuracy of the measurements. It was repeatedly demonstrated that laser interferometric measurements should only be carried out on vibration-free tables of adequate quality otherwise the vibrations of the precision elements caused by the environment or another element can falsify the measurements.

We found that the LDV is more suited to determine the instantaneous velocities than the LIMA, which primarily measures displacement. This is especially true for the vibrations where the velocity remains under 0.1 mm/s. The development of the measuring program of the LIMA, working out a better method for calculating instantaneous velocity and acceleration is certainly warranted.

The measurements demonstrated that the different points of the gently knocked corner prism in the measuring arm vibrate differently. These differences are greater than the differences of the data provided by the LIMA and the LDV (in the central setting). Based on our measurements we can state that within the margin of error the LIMA and the LDV measure the velocities to be the same (with the constraints mentioned above). We estimate this margin of error to be 1–2%. The reason for even this (small) error is not within the devices, but it is the fact that the devices employing completely different principles do not “see” the same exact vibration.

Acknowledgements This work was carried out within the framework of the Centre for Excellence in Mechatronics and Logistics operating in the strategic research field of the University of Miskolc.

References

1. Beeck MA, Hentschel W (2000) Laser metrology—a diagnostic tool in automotive development processes. Opt Lasers Eng 34:101–120. doi:10.1016/S0143-8166(00)00077-4
2. Chen F, McKillip D (2007) Measurement and analysis of vibration and deformation using laser metrology for automotive application. Proc Institution Mech Eng Part D—J Automobile Eng 221:725–738. doi:10.1243/0954470JAUTO67
3. Castellini P, Montanini R (2002) Automotive components vibration measurements by tracking laser doppler vibrometry: advances in signal processing. Measur Sci Technol 13:1266. doi:10.1088/0957-0233/13/8/314
4. Citrovszky A (1998) LI-03 típusú szubmikronos lézerinterferométer kezelési és karbantartási könyv, Budapest
5. Jani P, Czitrovszky A, Szótér L, Tajnafoi J, Barna B, Molnár L (1999) A laser interferometric motion analyser system and some of its applications. Publications of the University of Miskolc, Series D-Natural Sciences Physics 1:17
6. Béres M, Paripás B, Majár J, Bodolai T, Rónai L, Illavszky V (2016) Comparative study of two laser interferometric methods for the study of vibrations. In: Proceedings 30th MicroCAD international multidisciplinary science conference, Section C3: natural sciences, ISBN:978-963-358-113-1, Miskolc (Hungary)
7. (2012) Polytec GmbH: PDV-100 Portable Digital Vibrometer Datasheet
8. Hussain B, Muhammad T, Rehan M, Aman H, Aslam M, Ikram M, Raja MYA (2013) Fast processing of optical fringe movement in displacement sensors without using an ADC. Photonic Sensors 3(3):241–245. doi:10.1007/s13320-013-0088-5
9. Ripper GP, Garcia GA, Dias RS, Zindeluk M (2005) Primary calibration of single-point laser doppler vibrometers (LDV). In: Inter-Noise Congress, Rio de Janeiro
10. Mark Cs (2004) Fundamentals of light sources and lasers. Wiley

Deburring of Polimer and Metal Matrix Composites

János Líska

Abstract Nowadays have developed more and more such materials, which mechanical and physical properties are extremely better compared to commonly used materials. These composite materials, which were used before a few years ago just in the construction industry for stiffening or as a decorative element, are used also often is the automotive industry. This fact opened a new way in machining, where we must comply to increasingly stringent precision and visual requirements. The burr as attendant phenomenon is known as conventional metal cutting, is also present at machining of special materials (e.g. composites) as well. For high time for piece deburring increase the costs of parts production, therefore industry places much emphasis on reduction of it. The article investigates various options of deburring after machining of metal and polymer-based composite materials.

1 Overview of Deburring

What constitutes a "burr-free" part varies among companies and quality control departments. For some, it means having no loose materials at an edge. For others, it means having nothing visible to the naked eye or an edge condition that will not cause any functional problem in the next assembly process. Missing material or a hump of rounded metal at an edge may or may be called a burr [1].

Edge quality is of concern for the performance, safety, cost, and appearance of a part. The following is a reasonably complete list of the problems caused by improperly finished edges:

J. Líska (✉)
Pallasz Athéné University, Kecskemét, Hungary
e-mail: liska.janos@gamf.kefo.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_21

- Cut hands in assembly or disassembly.
- Interference fits in assemblies.
- Jammed mechanisms.
- Scratched mating surfaces that allow seals to leak.
- Increased or changed friction.
- Increased wear on moving or stressed parts.
- Electrical short circuits.
- Cut wires from sharp edges and sharp burrs.
- Unacceptable high-voltage breakdown of dielectric.
- Irregular electrical and magnetic fields.
- Clogged filters and ports from loose burr accumulation.
- Cut rubber seals and O-rings.
- Turbulence and nonlaminar flow.
- Reduces formability.

Figure 1 illustrates the most commonly used deburring processes. Unfortunately, no single machine or process produces all the required edge quality on every edge for every burr without side effects [1].

Our aim is to decrease the material cost of floor structure by changing the profile, dimensions and number of cross members as well as the thickness of the deck plate.

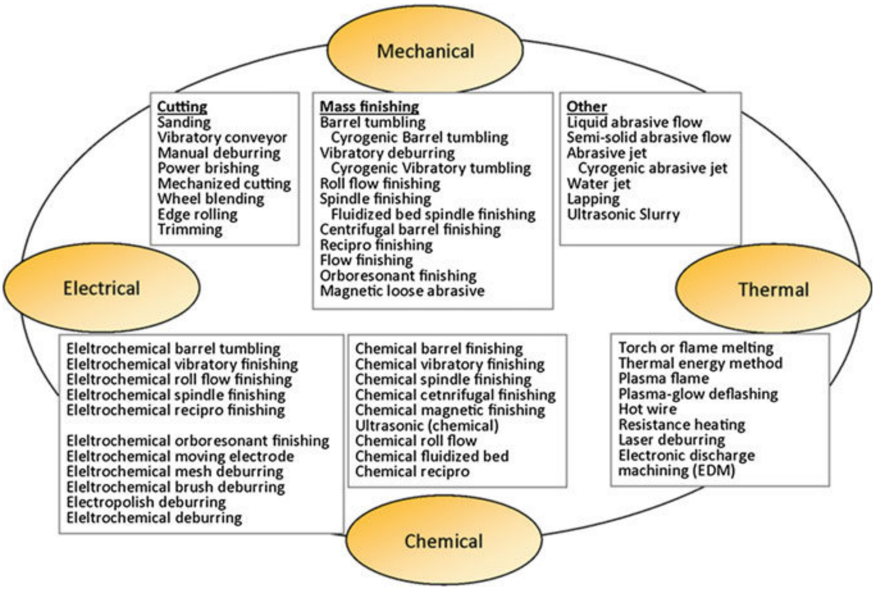


Fig. 1 Principal deburring processes and their removal mechanisms [1]

2 Deburring Economics

Few examples of comparative deburring economics have been published. The following equations provide a reason been estimate of costs using various processes. The equations use following terminology [1]:

C	Deburring cost per part
CD	Depreciation cost per hour = Machine cost per operating hours
CM	Maintenance cost per hour of operation
CL	Labour cost per hour to run the machine
CP	Cost of power used
CA	Cost of cleaning per hour after deburring
CE	Cost of media per hour = media cost × percentage of hourly attrition
CC	Cost of compound per hour
CW	Cost of water per hour
DO	Overhead as percentage of labour rate
N	Number of parts run per hour = n/t
n	Number of parts run per cycle
t	Time (hours) per cycle
Cg	Cost of gas per cycle
W	Power used, in kilowatts
Ct	Total tool cost
Np	Total number of parts
Cs	Total cost of solution
K1	Percentage of cycle time that an operator actually spends controlling deburring operation
K2	Percentage of cycle time that operator spends cleaning parts
CB	Cost of cleaning materials per hour

For vibratory (loose abrasive) processes, use

$$C = [C_D + CM + WCP + CB + CE + CC + CW]/N + [CL(1 + DO)(K1 + K2)]/N \tag{1}$$

For thermal energy methods, use

$$C = [CD + CM + CL(1 + DO)WCP + CA]/N + Cg/n + Ct/Np \tag{2}$$

For brush deburring, use

$$C = [CD + CM + CL(1 + DO)WCP + CA]/N + Cb/Np1 \tag{3}$$

For flame deburring, use

$$C = [CD + CM + CL(1 + DO)WCP + CA]/N + Cg/n \quad (4)$$

For manual deburring, use

$$C = CL(1 + DO) + CA]/N + Ct/Np \quad (5)$$

For mechanical deburring, use

$$C = [CD + CM + WCP + CA + CL(1 + DO)]/N + Ct/Np \quad (6)$$

For chemical deburring, use

$$C = [CD + CM + WCP + CA + CL(1 + DO)]/N + Cs/Np \quad (7)$$

For electropolish deburring, use

$$C = [CD + CM + WCP + CA + CL(1 + DO)]/N + Ct/Np + Cs/Np1 \quad (8)$$

These equations have some important limitations. First, they assume that the conventional form of the deburring process is used. As mentioned earlier, it is frequently possible to alter the process slightly to obtain faster or better results. Such alterations may insert another cost term into the equation, however. Unless the conventional approach is used, the equation provides only initial estimates of cost.

A second limitation of these equations is that they assume that the one knows the value of each component and the time required to remove the burr. Although it is possible to use “rule of thumb” costs for media, compounds, and the like, only a few publications provide any information on the time required to remove a burr of a specific size. As additional research is reported, this will become less of limitation. In the interim, information can be extrapolated from the results produced by other parts subjected to the same process.

At third limitation of these equations is that they ignore the costs of floor space, area heating, lighting, maintenance, insurance, and supervision [1].

3 Burr Formation Mechanism

Burrs, flash, and related protrusions are formed by the six physical processes listed in Table 1. Burrs formed by the first three processes involve plastic deformation of the workpiece material. Solidification of material on the working edges, the fourth processes of formation, forms a burr-like projection. The fifth process, incomplete cutoff, occurs when the workpiece is allowed to fall from the part before the cut is completed. Flash forms whenever the pressure on molten material is sufficient to force the material between the two halves of a die or mould. The examples of burr types we can see on Figs. 2 and 3 shown in the theoretical definition of burr characteristics [1].

In end-milling operations, Poisson burrs are formed on edge 1. Depending upon cutter geometry, they also form on edges 2, 4 and 10 (Figs. 4 and 5). An entrance burr occurs on edge 6. Rollover burrs are produced on edges 3, 7, and 9. On half of

Table 1 The physical processes involved in the formation of burrs, flash, and related protrusions [1]

Process	Name of protrusion
Lateral flow of material	Poisson burr
Bending of material (such as chip rollerover)	Rollover burr
Tearing of chip from the workpiece	Tear burr
Redeposition of material	Recast bead
Incomplete cutoff	Cutoff projection
Flow of material into cracks	Flash

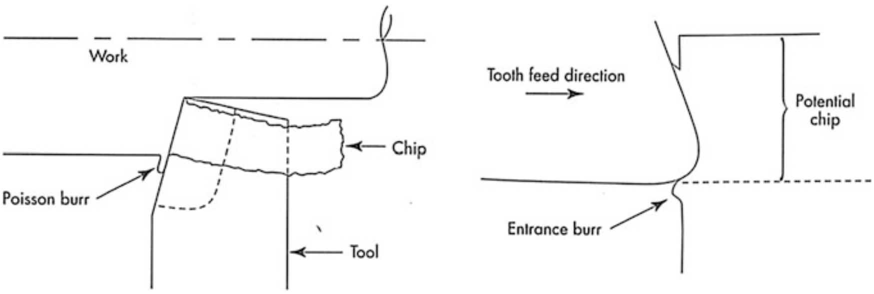
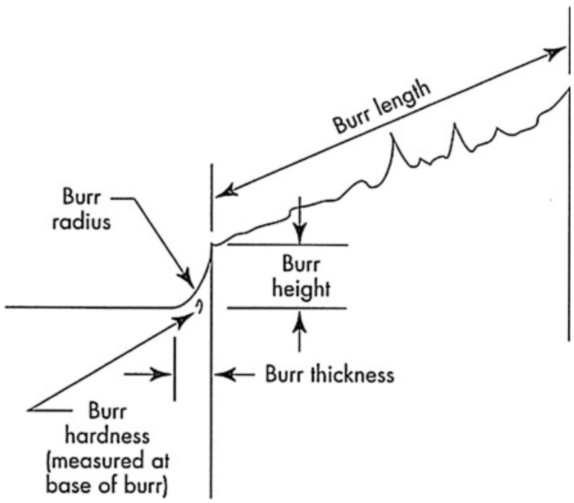


Fig. 2 Examples of burr types; poisson and entrance burr

Fig. 3 Definition of burr characteristics [1]



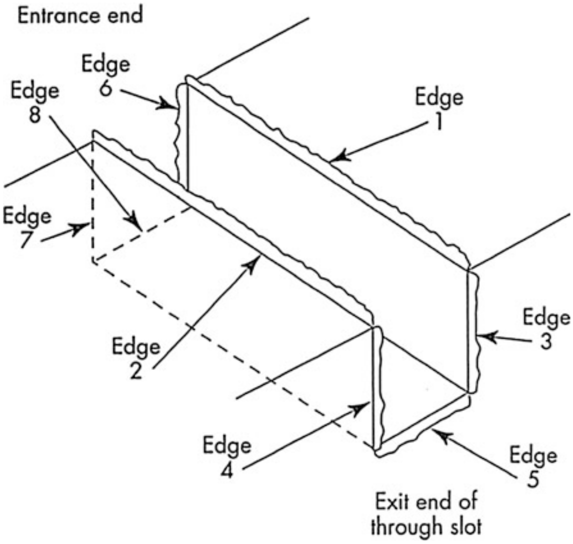


Fig. 4 Burr locations on slots produced by side or end mills

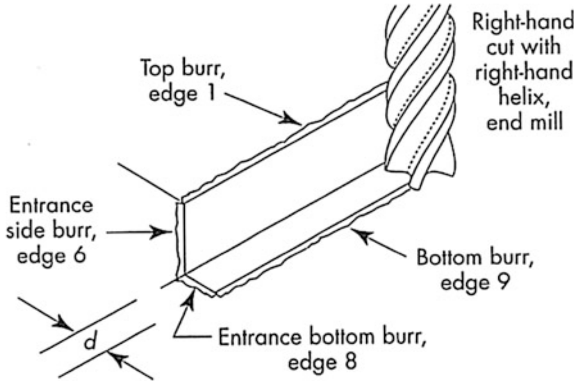


Fig. 5 Identification of burr locations in end-milling edges of workpieces [1]

edges 5 and 8 an entrance burr is produced, whereas a rollover burr equals the radial depth of cut until the cut exceeds 0.6 times the tool diameter. For deeper cuts in the height tends to remain constant at 0.6 times the tool diameter, although it may vary with workpiece properties and tool forces [1].

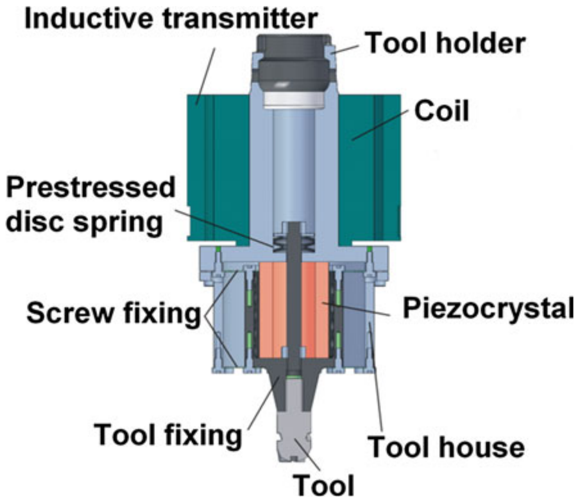


Fig. 6 DMG—Sauer Ultrasonic technology [2]

4 Examples of Composite Machining

The central element of ultrasonic-based machine is specially, for this purpose developed spindle with tool system, known as an appropriated tool system for HSK standard. It takes across the high-frequency of ultrasound technology by induction from the spindle to the Ultrasonic HSK clamping, which puts the tool into vibrational mode. The contact between tool and work-piece creates and terminates fifty thousand times by oscillation. By using of the technology, the chips become detached from the surface of work-piece only in small pieces. Figure 6 shows the conceptual drawing of technology.

The advantage of ultrasonic machining showed at most on work-piece: while material smeared (negative burr formation) at conventional machining machines, the surface lost its cell structure. At ultrasonic machining surface preserved its original appearance.

At designing of experiments is never clean, which results are expected. This case is a completely good example of that mechanism of stock removal of polymer based composite materials are very different compared to stock removal of metallic materials. It is interesting from several aspects, because momentarily do not exists model, which would accurately describe this process. During the experiments and at interpreting of results we can conclude, that in that case is no ductul cutting. It is clearly visible from both images (Figs. 7 and 8). The chip breaks in shear plane and there is no plastic surface deformation. It can be also observed that two machined materials behave differently. While at material GFRP shapes of the chip are elongated, shapes of chips at materials CFRP are fragmented. This can be explained by the fact, that the modulus of elasticity of material GFRP is larger.

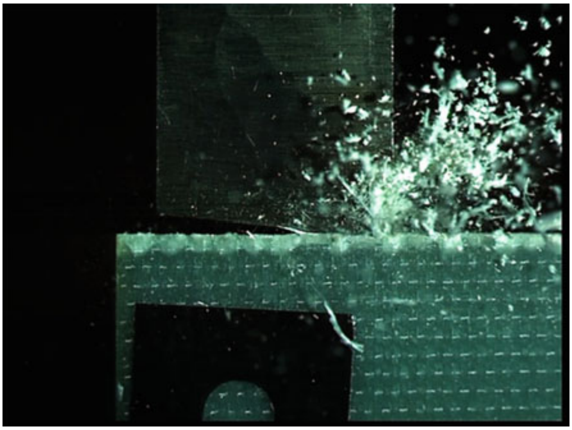


Fig. 7 GFRP high-speed camera image

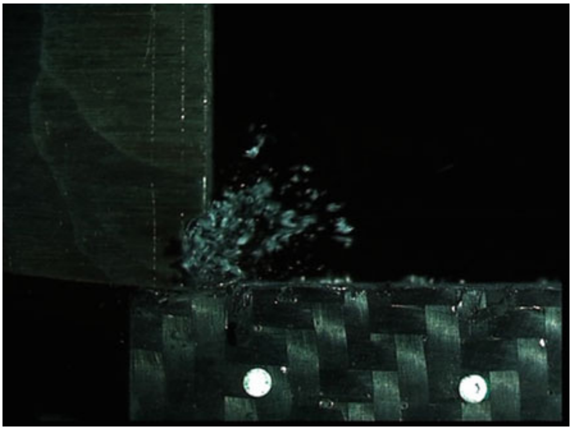


Fig. 8 CFRP high-speed camera image

References

1. Gillespie LK (1999) Deburring and edge finishing handbook. Society of Manufacturing Engineers. ISBN 0-87263-501-5
2. Uni-Hannover Teil projekt. http://www.sfb653.uni-hannover.de/en-us/Pages/Teilprojekt-T4.aspx. Accessed 26 Feb 2016

The Past, Present and Future of the Training of Internal Combustion Engines at the Department of Energy Engineering of BME

Ákos Bereczky

Abstract Vehicle industry plays an important role in the GDP production and employment of present-day Hungary. In the June of 2016 the share of vehicle manufacturing subsection within processing industry was 31.4% [1] and the sector employed 135 thousand persons in 2014 [2]. Thus higher education takes a significant role in the provision of adequately trained students and the co-operation with the industry has also a great role. The author reviews the training of internal combustion engines, which is the most frequently applied resource today—related to his own narrower field within vehicle industry and the joint research at the Department of Energy Engineering of BME. The author presents the Department in the field, the present areas of research, the industrial co-operations and the available infrastructure. Finally the further plans are being outlined.

1 Introduction

The target of training at University level is to issue appropriate engineers for industry professionals and to educate junior lecturers, the next generation of professors. The training is based on lectures, problem solving and laboratories. The output levels are different for a manufacturing or a research and development company and different for a Ph.D. study. But the relevant theoretical background is needed to support practices and laboratory measurements. The practices can be a good complement with modelling practices, while labs are necessary for proper infrastructure and good industrial support with demonstration, joint projects and research. The Author presents the present infrastructure and research activities of the past and present.

Á. Bereczky (✉)
Department of Energy Engineering, Budapest University of Technology and Economics, Bertalan Lajos u. 4–6, Budapest 1111, Hungary
e-mail: bereczky@energia.bme.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_22

2 Infrastructure

The infrastructure can be separated into different areas in, the Laboratory of the Department of Energy Engineering. The first important part is the complex engine test systems. There are three complex systems installed, they are based on 257–300 kW Eddy-current brakes, maximum speed is 8000 RPM. All of these systems have gravimetric fuel consumption and blow-by measuring systems and closed cooling systems. One of them is the AUDI Container Engine Brake System. Two of the systems are for industrial projects, one of them a spark ignition type engine is installed with an open source ECU for mainly educational purposes and for student projects. The laboratory has a gas engine based generator set for the demonstration and test of the energy systems, this is mainly to determine the heat balance at different conditions and fuels.

We have different small test systems; an Octane Rating Engine, and a CFR Cetane rating engine (Fig. 1). The Octane Rating Engine is made by the BASF, on this system performable to the octane number measurement of gasolines according to a standard necessary for the examinations. We compare the knock intensity of the fuel to be examined at a constant compression ratio with the similar value of two reference fuels, beside different air-ratios. The Cetane rating engine used for measurements is a single cylinder, four-stroke cycle pre-chamber type diesel engine. The displacement volume is 610 cm^3, the speed is 900 RPM which is the value determined by the ASTM standard, the adherence of which is ensured by an asynchronous electric engine belt connected to the engine. These systems are very

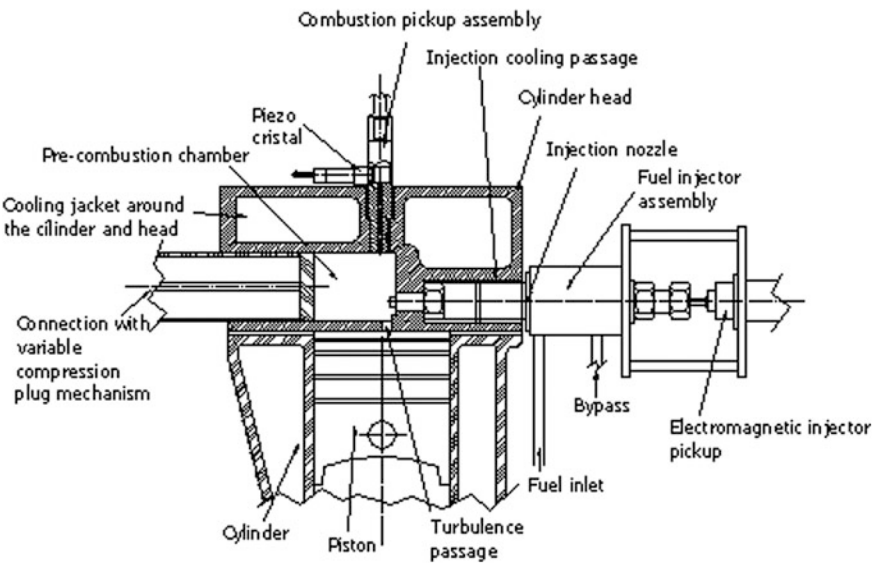


Fig. 1 Combustion chamber of the cetane rating engine [3]

Fig. 2 The portable emission measuring laboratory

useful for educational purposes, for example students can investigate parameters like excess air ratio influence on the combustion process. The Department has indication systems to measure the pressure in the combustion chamber too.

The third element of the laboratory is the emission measuring system, we have one portable emission measuring laboratory (Fig. 2). With this system we can measure VOC (T.HC.), $NO/NO_2/NO_x$, CO, SO_2, O_2, CO_2 and CH_4 concentration in exhaust channels in the field tests. The second system is the HORIBA exhaust gas analyser, for VOC (T.HC.), $NO/NO_2/NOx$, CO, O_2 and CO_2 emission.

3 Research Activities

The main targets of the research activities are focusing on renewable energy resource utilization in compression ignition (Diesel) engines. This work can be separated to biodiesel fuels investigation and utilization of alcohol fuels.

Alcohol fuels cannot be used directly in Diesel engines, because of their very low cetane number and due to other problems, therefore they are usually incorporated in Diesel oil. The alcohol fuel‘s higher amount of oxygen in a combustion process reduces emission rates of incomplete combustion products (PM, HC, CO) and helps to improve the combustion process. The investigated alcohol fuels were methanol, ethanol and n-butanol. Methanol is a highly toxic alcohol, there are several methods known for its production:

- Carbon monoxide and hydrogen react over a catalyst
- From carbon dioxide with catalytic hydrogenation of CO_2 with H_2 (Oláh methanol economy)
- Direct catalytic conversion of methane.

Methanol has a low Cetane number, and it constitutes a problem that it creates an unstable blend, therefore either emulsifiers are used, or a triple mixture is created e.g. with biodiesel [4]. There are several methods to produce ethanol:

- Fossil base: hydration of ethylene
- Microbial fermentation of sugars, distillation, and dehydration.

Its utilisation in Diesel engines is problematic similarly to that of methanol, a stable mixture can be produced only with additives [5]. The n-butanol production of can be carried out in the following ways:

- commercially produced from fossil fuels,
- Acetone–butanol–ethanol (ABE) fermentation (Clostridium Beijerinckii bacteria).

The advantage of butanol is that it can be incorporated in diesel oil [6]. However, its production from renewable raw materials requires further developments.

Methanol tested with diesel fuel, rape seed methyl ester and methanol blend up to 30 V/V% FAME with 10 V/V% methanol ratio [7]. We found high engine load values of CO concentration reduced by 2–13%, while HC concentrations were changing in the range of −8 to +18%, through all experimental load points, the NOx emission increased by 1.5–10% depending on the load and the soot concentration values were reduced by 13–45% compared to reference Diesel fuel [7]. The n-Butanol tested with diesel blend up 5, 10 and 20 V/V%. We found that 1500 rpm at 75% load for B05, B10 and B20 mixtures the CO emissions reduced by 35.7, 57.1 and 71.4%, NOx emission increased in this study by 10.3, 32.3, and 54.4%, the HC emission concentration increased by 21.4, 71.4, and 214%, and soot reduction relative to DF at 55.5, 77.8 and 85.1% respectively.

Wet-ethanol investigated in a Dual fuelled engine [8, 9] (Fig. 3). The investigated parameters were premixed ratio (rp), pilot timing and engine load [10]. At 100% load dual-fuel operation with wet ethanol (except for rp = 20%) is beneficial from the point of view of emission. The hardly after treatable components like NOx and especially PM emissions radically decrease. The elevated HC emission can be easily after treated by the means of conventional oxidative catalytic converter. At higher premixed ratios CO emissions are in the same range, the decrement of NOx emission is 61% compared to diesel operation. Particulate matter (PM) emission radically decreases by 95% compared to diesel operation. HC emission increases significantly and monotonously with increasing premixed ratios.

On the field of biodiesels the laboratory investigated many different methyl esters and pure oils like: rape seed oil, sunflower oil, waste vegetable oil and very

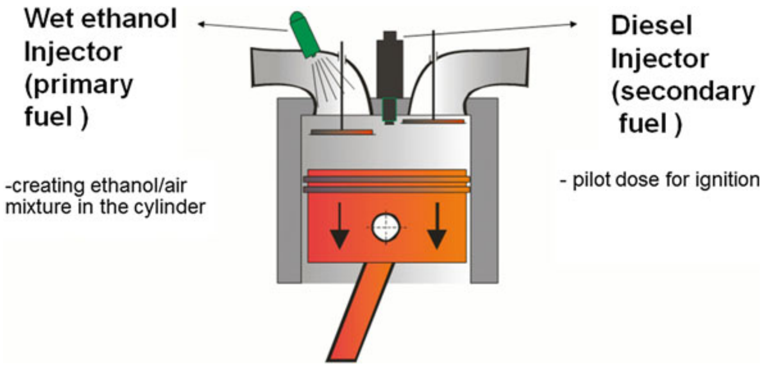


Fig. 3 Work schema of the dual fuelled engine (based on [10])

special non-edible oils, like croton megalocarpus oil methyl ester and jatropha oil methyl ester of Eastern Africa origin. Croton megalocarpus oil is a less common biodiesel feedstock [11], it is a tree borne non-edible oilseed of African origin. The plants are indigenous to East Africa, and are widely found in the mountains of Tanzania, Kenya and Uganda. Croton megalocarpus seeds contain approximately 32% by weight of oil [12]. Jatropha oil, will become a significant source for biodiesel production now or in the future. Jatropha is a fast growing plant, which requires little water or fertilizer; it can survive in infertile soils [13]. Jatropha oil is mostly found in developing countries, especially in Africa and Asia. Jatropha plants have a high seed yield which can be continuously produced for 30–40 years. The oil content in the jatropha seeds is approximately 30–40% by weight [14].

The tests (Fig. 4) with the utilization of Croton oil and Jatropha oil methyl ester presented that the emissions of CO and NOx were slightly higher for biodiesel samples than for mineral diesel. THC was found to be reduced at intermediate load and full load it was similar, the PM emission was lower for biodiesels at all loads than diesel fuel [11, 12]. In the field of biodiesel, we started the investigation of the TBK or MOME fuel in a cooperation with the Engine and Vehicle Emission Test Laboratory, Centre for Environment Protection and Sustainability Research, Institute for Transport Science Nonprofit Ltd. [15]. TBK is based on rapeseed oil, trans esterificated with methyl acetate. It is very important to highlight that the TBK biodiesel is a pure material; it doesn’t have any additive contrary to the two other standardized fuels, which contain many additives in order to improve their physical and chemical properties and make them suitable for use in internal combustion engines. It is also important to mention that the production procedure of TBK is an invention of three Hungarian engineers (János Thész, Béla Boros, Zoltán Király).

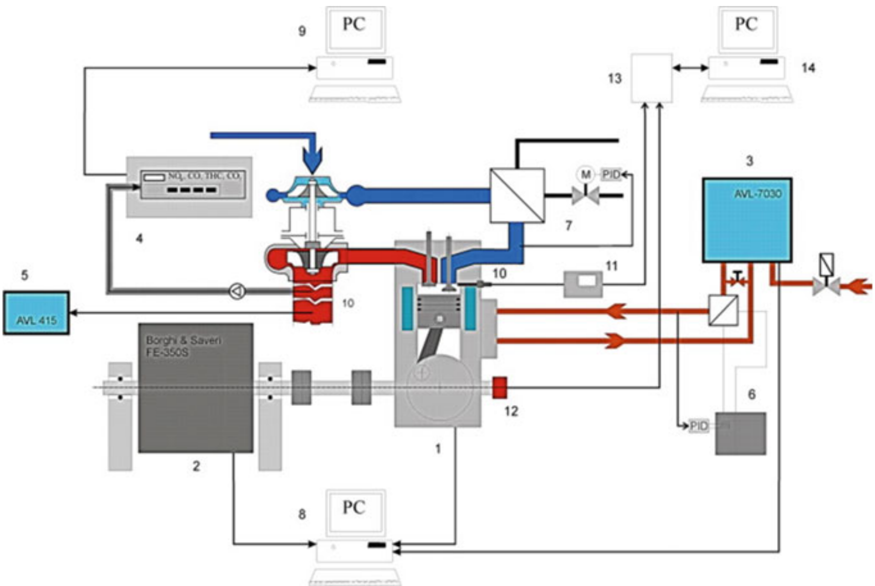


Fig. 4 Test system for research activities [7]; (*1* VW 1.9 TDI (1Z) Engine, *2* Eddy current dynamometer, *3* Fuel consumption measuring system, *4* Exhaust gas analyser system, *5* Smoke meter, *6* Fuel temperature controller, *7* Intake air temperature controller, *8–9* Computers of the controlling and measuring systems, *10* Piezo transducer, *11* Charge amplifier, *12* Crank angle speed encoder, *13* High speed data acquisition system, *14* Computers of the indication system)

4 Modelling

Modelling of the working process is very important for educational and development purposes. At problems solving we use a one dimensional system, like the AVL Boost system and the AVL FIRE CFD software package for CFD calculations.

4.1 *The AVL Boost Software System*

The Boost software is a program using a one dimensional model, with the help of which the processes of the given engine can be easily simulated. To run the Boost software the built up of a 1D model is needed (Fig. 5), and then the adequate selection of the parameters of the model, the careful setting of the initial and the boundary conditions is required [16].

The adequate timing of the valve openings and the start-up of the combustion has a great impact on the results.

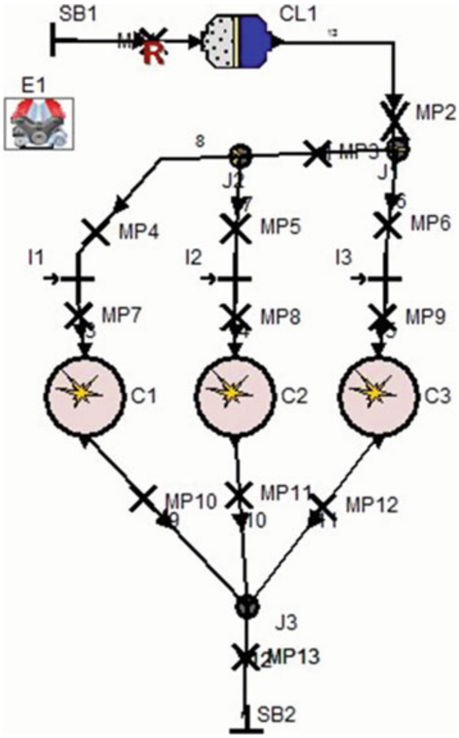


Fig. 5 The model built up in the Boost software for dual fuelled engines [16]

In the case of dual fuel engine we use triple "Vibe" (i = 1, 2, 3) the resultant combustion law of the fuel in the certain burnt distances can be obtained with the summary of the previous functions in the proportion of the transmitted heat

$$X_\varphi = a^* X_{1,\varphi} + b^* X_{2,\varphi} + (1 - (a + b))^* X_{3,\varphi}, \quad (1)$$

where a, b—are constants to be determined, X_1, X_2, X_3—for the chosen burnt distances, are "Vibe" approximate functions [16]. In the case of a 24 kW load more peaks can be found in the operation period of greater alcohol proportion (Fig. 6) thus the two parameter model (dual "Vibe") is not able to approximate the original combustion law any more. With the help of the triple "Vibe" the original combustion law is easy to approximate, the model generated with the help of the so obtained parameters is expectedly to approximate the real processes (Fig. 6).

4.2 CFD Simulations (Based on [18])

For CFD simulation of internal combustion engines we use the AVL FIRE software, this is a multi-purpose thermo-fluid CFD software package. AVL's Engine

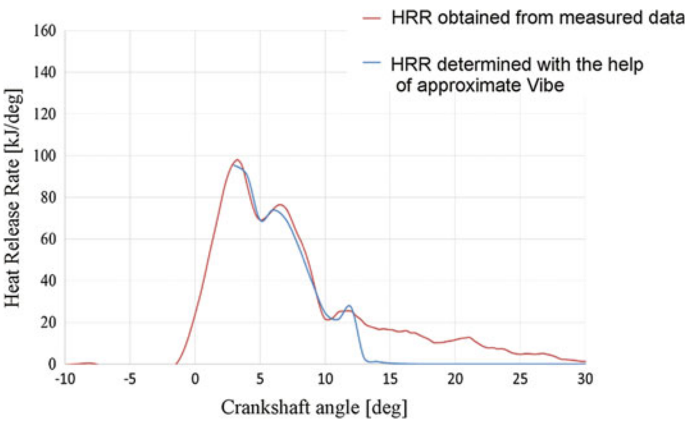


Fig. 6 Measured and approximated "Vibe" heat release by a load of 24 kW in dual fuel operation [17]

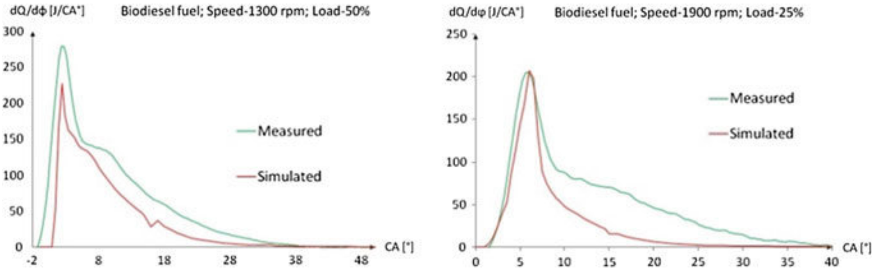


Fig. 7 The calculated and measured heat release rates in case of biodiesel fuel at different speeds and loads [18]

Simulation Environment (ESE-Diesel) in the FIRE module is used to define the geometry and create the meshing for the combustion chamber. To reduce the computational demand of the simulations, the 3D model is simplified and a less detailed model is created. The simulations were accomplished with the Extended Coherent Flame Model (ECFM-3Z) model, which is developed for diesel combustion simulation. This is a combustion model based on a flame surface density transport equation and a mixing model that can describe inhomogeneous turbulent premixed and diffusion combustion. The model relies on the ECFM combustion model, with an improved burnt gas chemistry description compared to ECFM [19]. With this CFD software package the combustion process of different fuels investigated (Fig. 7).

5 Conclusions

The Author presented a triple base of training of internal combustion engines at the Department of Energy Engineering of BME. On lectures, we should give a strong background that is based on high level courses of thermodynamics and fluid dynamics on B.Sc. and M.Sc. level. For problem solving at B.Sc. level students can work with one dimensional system, like the AVL Boost system and at M.Sc. level with AVL FIRE CFD calculation. The laboratories are separated, for B.Sc. level the heat balance test on the gas engine and the investigation of the combustion process in the Octane rating engine are requirements. On M.Sc. level the complex engine tests, ECU application are the main targets.

On Ph.D. level a complex renewable fuel test goes on like the presented different biodiesel and dual fuel engine tests, and may give a basis for the thesis work.

The present infrastructures are not new, but with the help of the University and the Industrial partners the infrastructures are sustainable and we can further develop the systems.

Acknowledgements This study was conducted at Budapest University of Technology and Economics (BME), Hungary and was supported by the Hungarian/South African Intergovernmental S&T Cooperation Program (Grant: ZA-17/09). The authors are grateful to the AVL GmbH Company for the possibility to use their simulation software (AVL Boost and FIRE) under the university partnership program and for the help of the industrial partners, like AUDI Hungaria Motor Ltd., AVL-Autokut Ltd., Hungarian Suzuki Ltd. and many other companies.

References

1. First release of the Hungarian Central Statistical Office, Industry, June, 2016 (second estimation). Publication: 12 August 2016
2. Association of the Hungarian Automotive Industry (in Hungarian: Magyar Gépjárműipari Egyesület, MAGE) Skilled labour is the obstacle of vehicle industry, 10 April 2015
3. Laza T, Bereczky A (2009) Influence of higher alcohols on the combustion pressure of diesel engine operated with rape seed oil. Acta Mechanica Slovaca 13(3):54–61
4. Lukács K, Bereczky Á (2015) Investigation of utilization potential of different alcohol fuels in compression-ignition engines. In: Proceedings of the 12th international conference on heat engines and environmental protection, pp 57–62
5. Matuszewska A, Odziemkowska M (2011) Study on the impact of co-solvent on selected properties of mixtures of diesel oil with bio-ethanol. CHEMIK 65(6):543–548
6. Siwalea L, Lukács K, Bereczky A, Mbarawa M, Kolesnikov A (2014) Performance, combustion and emission characteristics of n-butanol additive in methanol–gasoline blend fired in a naturally-aspirated spark ignition engine. Fuel Process Technol 118:318–326
7. Žaglinskis J, Lukács K, Bereczky Á (2016) Comparison of properties of a compression ignition engine operating on diesel–biodiesel blend with methanol additive. Fuel 170:245–253
8. Emőd I, Tölgyessi Z, Zöldy M (2006) Alternative vehicle drives: alternative motor fuels—hybrid cars—fuel-cell drive. Maróti könyvkerekedés és könyvkiadó Kft, Budapest, 232 pp. ISBN:963 9005 738 (in Hungarian)

9. István E (2003) The use of fuels as the fuels produced from agricultural products and waste from internal combustion engines, technical development, environmental studies (KMFP-00031/2002) (in Hungarian)
10. Lukács K, Bereczky Á (2011) Experimental investigation of dual-fuel diesel engine with wet ethanol. In: The role of renewable in energy generation: 10th international conference on heat engines and environmental protection, pp 185–190
11. Kivevele T, Lukács K, Bereczky Á, Mbarawa M (2011) Engine performance, exhaust emissions and combustion characteristics of a CI engine fuelled with croton megalocarpus methyl ester with antioxidant. Fuel 90(8):2782–2789
12. Kivevele TT, Mbarawa MM (2010) Comprehensive analysis of fuel properties of biodiesel from croton megalocarpus oil. Energy Fuels 24:6151–6155
13. Sarin R, Sharma M, Sinharay S, Malhotra RK (2007) Jatropha-palm biodiesel blends: an optimum mix for Asia. Fuel 86(10):1365–1371
14. Kivevele T, Huan Z, Lukács K, Bereczky Á, Mbarawa M (2016) Impact of antioxidant additives on the engine performance and exhaust emissions using biodiesel made from Jatropha Oil of Eastern Africa origin. Res Dev J 32:1–8
15. Szabados Gy, Bereczky Á (2015) Comparison tests of diesel, biodiesel and TBK-biodiesel. Periodica Polytech-Mech Eng 59(3):120–125
16. Lukács K, Szurdoki Z, Bereczky Á (2013) Dual-fuel engine modeling with AVL Boost, ENELKO2013 XIV. Nemzetközi Energetika-Elektrotechnika konferencia. Nagyszeben, Románia, pp 95–100 (in Hungarian)
17. Bereczky Á, Lukács K, Sipos Á (2015) Determine the dual-fuel combustion engine’s law to the model AVL Boost, OGÉT 2015: XXIII. Nemzetközi Gépészeti Találkozó, pp 32–36 (in Hungarian)
18. Szabados Gy, Lovas M (2015) Combustion simulation of diesel fuel and biofuel by the help of AVL’s multi-purpose thermo-fluid CFD software. In: Proceedings of the 12th International Conference on Heat Engines and Environmental Protection, pp 51–56
19. AVL: AVL FIRE® VERSION (2011) Combustion and emission module. Edition 10/2011

Concept of a New Method for Helical Surface Machining on Lathe

Dániel Kiss and Tibor Csáki

Abstract In this paper the concept of a new turning method for helical surfaces will be demonstrated. In mechanical engineering several types of surfaces can be mentioned, which are difficult to machine for example the thread of a ball-nut or small workpieces with high pitch threads or helical surfaces. The main difficulty with these surfaces is that they can't be machined with a tool that generates the desired surfaces, because we can't assure the rigidity for these machining conditions. These surfaces are similar to threads in a certain respect. But we cannot use the exact threading cycle, because in the threading cycle the geometry of the tool defines the geometry of the surface of the thread. This new method is based on the threading NC cycle, but in this case the tool geometry won't be copied to the surface. The surface geometry will be defined by the start points (and/or start angles) of the thread, which can be computed by an appropriate computer program.

1 Introduction

In machinery there are several helical surfaces or helical grooves for example ball-screws and ball-nuts, worm gears, etc. Some of these are difficult to be machined. These difficulties can be originated to the shape, size and the position of these surfaces. Taking a ball-nut for example the inner ball groove is complicated to cut when it has high-lead profile. These helical surfaces are similar to threads but have a non-standard profile. There are several solutions to this problem, but these are not universal solutions in the meaning of tooling. It means that each type or size needs a tool which can only be used for the machining of that workpiece.

D. Kiss (✉) · T. Csáki
University of Miskolc, Miskolc, Hungary
e-mail: kiss.daniel@uni-miskolc.hu

T. Csáki
e-mail: csaki.tibor.dr@gmail.com

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_23

Since the Department of Machine Tools at the University of Miskolc has dealing with machine tool components and design, we started our investigation around high-lead ball-nuts [1, 2].

2 Thread Cutting

In modern CNC machines there are special subroutines or cycles for a specific application. One of these applications is the thread cutting which has several cycles to fulfil the expectations. From these the L97 or CYCLE 97 is the Sinumerik version of the thread cutting cycle, which has a wide parameter range to make it possible the programming of different type of threads.

During thread turning we have to use the appropriate tool for a specific thread. It follows that during thread machining the profile of the tool defines the profile of the workpiece. There are 3 main types of infeed methods, which are the following (Fig. 1).

Based on these we face problems when we want to machine large threads or threads with special profiles, especially at inner thread machining. In this case the thread milling can be a solution. But in this case we have to use special tools for each workpiece, because we have to use modified quill angle [4].

3 Test with Profiled Tool

At the beginning, we wanted to simulate the machining of ball-nuts. We changed the gothic profile to a simple arc to make the tool design easier (Fig. 2). In this case the tool holder had 32 mm diameter. All the tests were conducted on a DMG CTX alpha 500 CNC lathe in the workshop of the Department of Machine Tools. The CNC lathe is equipped with Siemens Sinumerik 840D control.

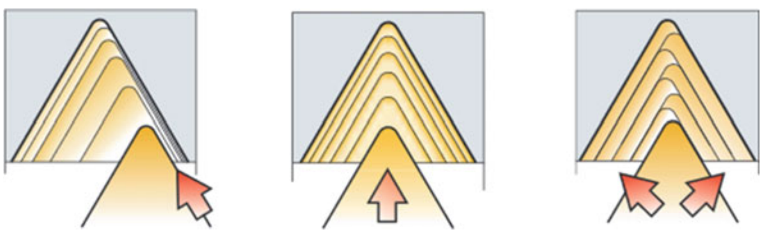


Fig. 1 Infeed methods from *left* to *right*: modified flank infeed, radial infeed, incremental infeed [3]

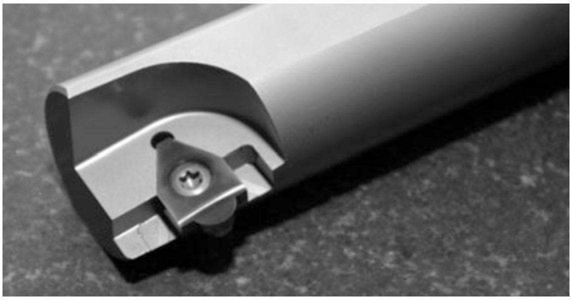


Fig. 2 The tool used for the experiment [5]

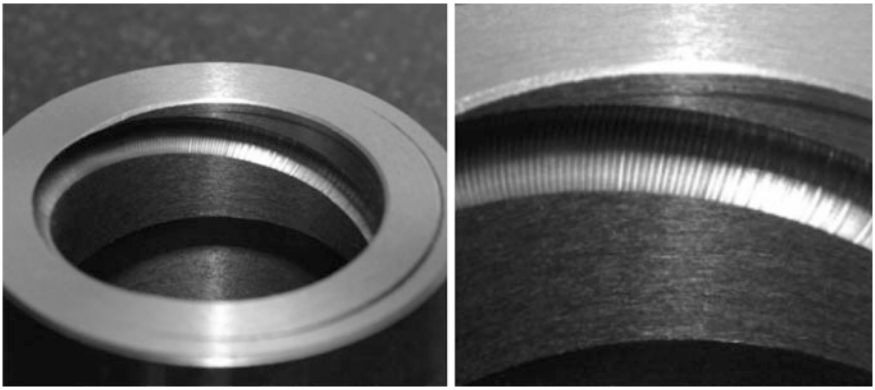


Fig. 3 Marks of vibrations on the surface machined with the profiled tool

During the test we used the Cycle 97 subroutine, which used for the programming of thread cutting. Radial infeed and decreasing ap was programmed pass-by-pass. The workpiece had 53 mm inner diameter and the pitch was 25 mm.

During machining excessive vibrations turned up (Fig. 3), which lead to inadequate surface finish and serrated chips (Fig. 4). It is clear that this kind of manufacturing cannot be realizable at special kind of thread shapes [5].

4 The Concept

Drawing the conclusions based on the previous test, it is clear that there are circumstances when the conventional threading cycle cannot be used. The length of the tool edge was too long and the half circle form is far from the linear. Vibrations appear because of either the rigidity of the machine, workpiece and the tool. A workable option is to use a tool, which doesn't have the thread profile on it but

Fig. 4 The chips in the different stages of the machining [5]

has an elementary geometry. In this case the thread cutting cycle can't be used in the conventional way since we don't use profiled tool. To overcome the problem that the tool doesn't machine the full profile we have to examine the threading cycle parameters. The parameters in the CYCLE97 are the following [6]:

CYCLE97 (PIT, MPIT, SPL, FPL, DM1, DM2, APP, ROP, TDEP, FAL, IANG, NSP, NRC, NID, VARI, NUMT, _VRT):

- PIT (Thread Pitch)
- MPIT (Thread Pitch as Thread Size)
- SPL (Thread Starting Point Longitudinal)
- FPL (Thread End Point Longitudinal)
- DM1 (Thread Start Diameter)
- DM2 (Thread End Diameter)
- APP (Run-In Path)
- ROP (Run-Out Path)
- TDEP (Thread Depth)
- FAL (Finishing Allowance)
- IANG (Infeed Angle)
- NSP (Starting Point Offset)
- NRC (Number of Roughing Cuts)
- NID (Number of Idle Cuts)
- VARI (Machining Type)
- NUMT (Number of Thread Turns)
- _VRT (Retraction Distance)

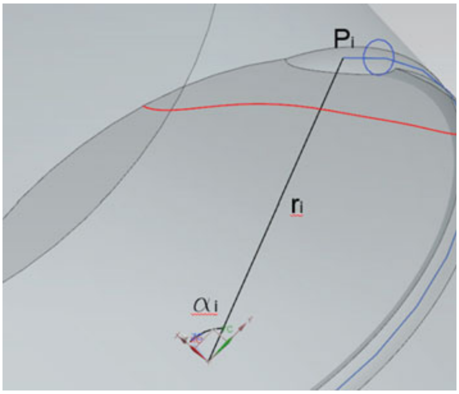


Fig. 5 Start point parameters

Among these parameters by changing of the TDEP and NSP parameters we can approximate the desired thread profile.

To calculate these parameters at first the thread profile has to be approximated with the profile of the tool in the threads normal section. The obtained tool center points, then have to be transformed to the plane, which is normal to the workpiece and positioned at the thread start (Fig. 5).

Then the two parameters can be calculated (parameters from Cycle 97):

$$TDEP_i = DM1 - r_i, \quad (1)$$

$$NSP_i = \alpha_i. \quad (2)$$

5 Testing the Concept

To verify the theory we performed a test at the Department of Machine Tools at the University of Miskolc. In the CNC program we called the threading cycle multiple times to achieve a simple half circle profile on a shaft. The diameter of the workpiece was 45 mm, the pitch was 25 mm and the radius of the profile was 4 mm. For the test we used a V shaped tool. The machined workpiece is shown in the next picture (Fig. 6).

The test was successful, because the machined profile approximated well the theoretical profile, but there's much room for improvement in the aspect of surface finish.

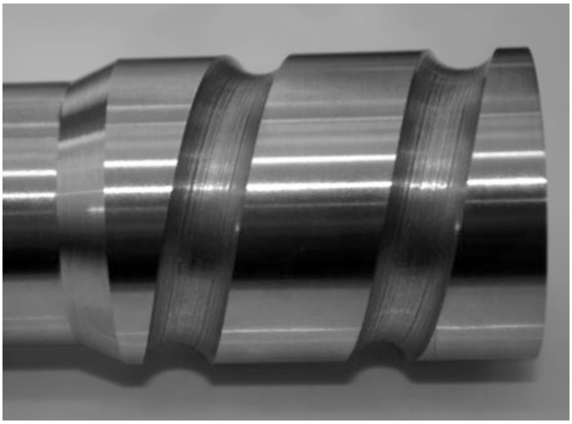


Fig. 6 Test workpiece

6 Conclusion

In this research, we searched for a possible way to machine threaded surfaces with non-standard and deep profiles. The problem and a possible solution were shown in this paper.

The new solution can be used to machine both outer and inner surfaces with the appropriate tool. This solution applies to the CYCLE97 threading cycle, but can be interpreted to other cycles where starting point offset can be programmed.

For the further research it is necessary to create a mathematical algorithm which calculates the required parameters for the cycle and generates the CNC program by taking into account both the tool and thread profiles. With this program both roughing and finishing can be executed.

When this algorithm is done a proper tool or tool holder have to be made where the correct pitch angle can be set.

This new method can be used for the machining of several threaded workpieces like ball-nuts and -screws, worm in worm drives and automotive steering parts.

Acknowledgements This research was carried out as part of the TÁMOP-4.2.1. B-10/2/KONV-2010-0001 project with support by the European Union, co-financed by the European Social Fund, in the framework of the Centre of Excellence of Mechatronics and Logistics at the University of Miskolc.

References

1. Hegedűs G, Takács G (2013) Tool profile generation by Boolean operations on ball nuts. Key Eng Mater 581:462–465. doi:10.4028/www.scientific.net/KEM.581.462
2. Hegedűs G (2016) Newton's method based collision avoidance in a CAD environment on ball nut grinding. Int J Adv Manuf Technol 84:1219–1228. doi:10.1007/s00170-015-7796-5
3. Sandvik Threading Application Guide (C2920-031) (2015)

4. Harada H, Kagiwada T (2004) Grinding of high lead and gothic-arc profile ball-nuts with free quill-inclination. Precis Eng 28(2):143–151
5. Mihályi G, Csáki T, Kiss D (2013) Design of variations of high-pitch ball nut tool suitable for machining, XXI. In: OGÉT conference, Arad (Romania), pp 282–284 (in Hungarian)
6. Siemens Sinumerik 840D Cycles Programming Manual (6FC5398-3BP20-1BAO, 01/2008)

Part III
Electrotechnics, Informatics

Intelligent Transportation Systems to Support Production Logistics

Péter Veres, Tamás Bányai and Béla Illés

Abstract Intelligent transportation systems (ITS) include both the traffic stream control and the intelligent vehicles. Cell phone networks and global positioning systems (GPS) enable the use of geographical information (GI) so that individual vehicles can locate themselves and global transportation systems can be enhanced taking advantages of new information technology solutions and algorithms. One of the major parts of ITS research is the assignment, routing and scheduling of vehicles in global transportation processes. This paper proposes an integrated engineering optimization algorithm to support the solution of assignment and scheduling problems of vehicles in intelligent transportation systems. This novel approach combines the available hardware and software components of an ITS with an algorithm to optimize the transportation processes of a global supply chain. To gain insight into the complexity of the logistic problem, the new model of supply chain including ITS is also described.

1 Introduction

Successful industrial and service processes provide a significant competitive advantage against other members of the market. Logistics related operations such as transportation, warehousing, customer service and inventory carrying account for 10–95% of the total cost, depending on the corporate sector, therefore managers traditionally think; logistics is a mandatory cost bucket. For that very reason it is very important to take the logistics related operations and processes under control and turn logistics into the source of competitive edges. Logistics processes can be

P. Veres (✉) · T. Bányai · B. Illés
University of Miskolc, Miskolc, Hungary
e-mail: altveres@uni-miskolc.hu

T. Bányai
e-mail: alttamas@uni-miskolc.hu

B. Illés
e-mail: altilles@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_24

divided into four important sequences: purchasing, production or service, distribution or sales and recycling. Transportation processes are especially remarkable in the purchasing and distribution processes of supply chain.

The explosion in technologies of ITSs made it possible to build geographical information systems. The improvement of wireless communication, computational technologies, floating cellular data, global positioning systems, sensing technologies and onboard systems leaded to the possibility to locate individual vehicles and communicate with them real time. This real time communication involves the collection and analysis of big data and makes it possible to optimize their operation from the point of view logistics, emission and costs.

This paper is organized as follows. Section 2 presents a literature review, which systematically summarizes the research background of supply chain optimization, ITSs and heuristic optimization. Section 3 describes the problem and hypothesis. Section 4 presents the general model of a supply chain including ITS based vehicles for transportation operations. Section 5 explains the solution algorithm. Conclusions and future research directions are discussed in Sect. 6.

2 Literature Review

This section reviews relevant literature related to supply chain optimization, ITS and heuristic optimization. Due to the large amount of researches on these fields the most relevant scientific results have to be summarized before to elaborate the model, algorithm and solution.

2.1 Research on Supply Chain Optimization

Within the frame of this chapter the author focuses on the supply chain optimisation of service sector. There are two main types of supply chains of service processes. In the case of service only supply chains (SOSC) the products of services are only pure services and no real physical products take part in the supply chain operations. In the case of product service supply chain (PSSC) both services and related physical products are integrated parts of the processes [1]. The research field of service related supply chains includes a wide area of services: energy supply chain [2], supply chain of chemical corporate sector [3], garment industry [4], transportation [5] and local food supply chain [6]. The optimization of supply chain includes the following main topics: scheduling of logistics services in service supply chain [7], routing and service level consideration [8], cost optimization [9], location, inventory and pricing [10]. The researches in the field of routing of supply chain cover both homogeneous and heterogeneous vehicle routing [11]. The widening of heterogeneous supply chain caused a big change of design, control and scheduling of vehicles. This change influences the required transportation and IT infrastructure. To the author's knowledge there has been no research work on real-time routing and fleet size optimization of supply chain, which takes into consideration the geographical information from ITS.

2.2 Research on Intelligent Transportation Systems

Intelligent transportation systems represent a group of tools and technologies, by the aid of which it is possible to improve the management, control and controlling processes of different transportation systems. ITS technology includes hardware and software components. Operation research has a key role to reach the full potential of ITS and there are publications in the field of transportation engineering focusing on the assessment and contribution of operation research [12–14]. GIS technology makes it possible to support the optimization of supply chain and define strategies to minimize the overall costs through the reduction of fossil fuel consumption [15]. ITS include Advanced Traveller Information System (ATIS) and Advanced Driver Assistance System (ADAS) [16] and these systems can support the design and control of transportation supply. Wireless technologies, state of the art electronics and automation makes it possible to build a strong hardware background for methods and algorithms, by the aid of which it is possible to enhance the possibilities of ITS. ITS can be represented not only in the field of freight transport but also in the processes of passenger transport.

2.3 Research on Heuristic Optimization

In computer sciences, mathematical optimization and engineering heuristic is a technique for problem solving and decision making in the field of NP-hard problems.

The heuristic optimization of transportation supply covers not only the functional parts of logistics (purchasing, production, distribution), but also focuses on the development of new models and methods to decrease the economical footprint of logistics services [17]. Heuristic optimization methods can support the solution of problems in the field of internal logistic related problems and external logistic networks. The complexity of supply chains required up to date methods to find near-optimal network design solutions for design, operation and management. The optimization methods of global supply chain are mostly heuristic methods. Much literature has been published in the field of design of large scale supply chain. There are several applications for heuristic algorithms from the sizing optimization of truss structures [18] to the transport energy modelling [19] or nuclear energy production [20]. One of the main streams of literature in the field of hybrid metaheuristic addressed the development of complex algorithms for large systems [21–23].

3 Problem Description

There are many companies, who launch a vehicle that is partially or completely empty, such as an empty return. It has always been a huge problem for companies, especially those who have small vehicle fleet and there is no proper transportation background behind them. In some cases, carrier companies have methods, that can chance the route of a transport vehicle after launch (for example, a traffic problem occurs), but this usually requires human intervention and planning. There are solutions in the form of network-based applications that collects data from joined small and medium companies and about their vehicles. However, most of these applications only collect the basic parameters of the vehicles and when they are not in use so the application is trying to provide work for that vehicle for the specified period. If we could integrate algorithms and extensions into these systems, it would track and calculate automatically with vehicles on the road and give them performable tasks, with which we can achieve further savings. This can be implemented by an expanded set of data, which contains the vehicle's position at every new transporting task, which requires an integrated GPS and identification system. It also needs to provide the vehicle's free capacity, destination, and the delivery deadline to avoid delays. With such a system it makes easy for a company to determinate, to give the new task to a running vehicle (Fig. 1) or to launch a new one to fulfil the demands (Fig. 2).

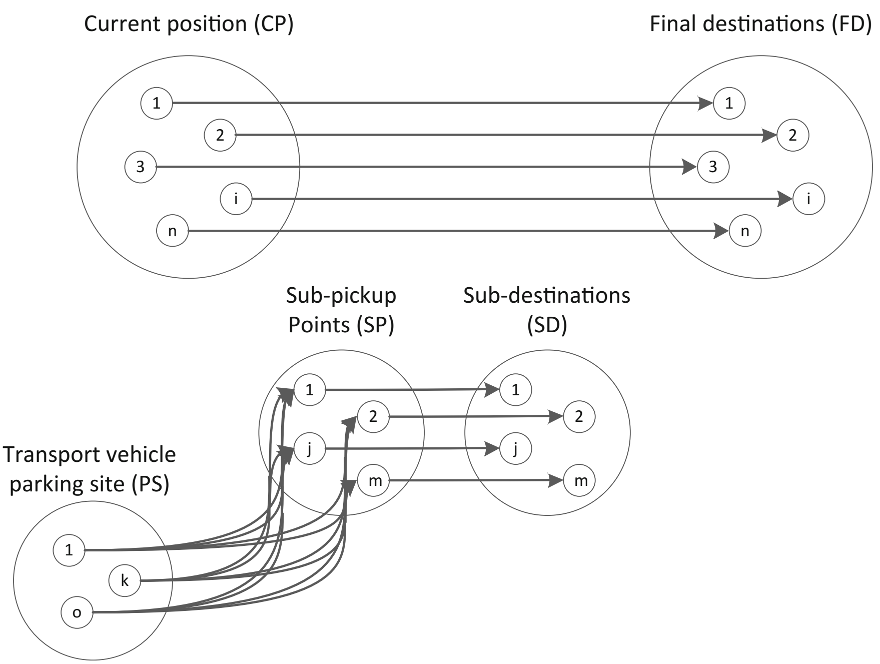


Fig. 1 Model of the supply chain related assignment problem without runtime optimization

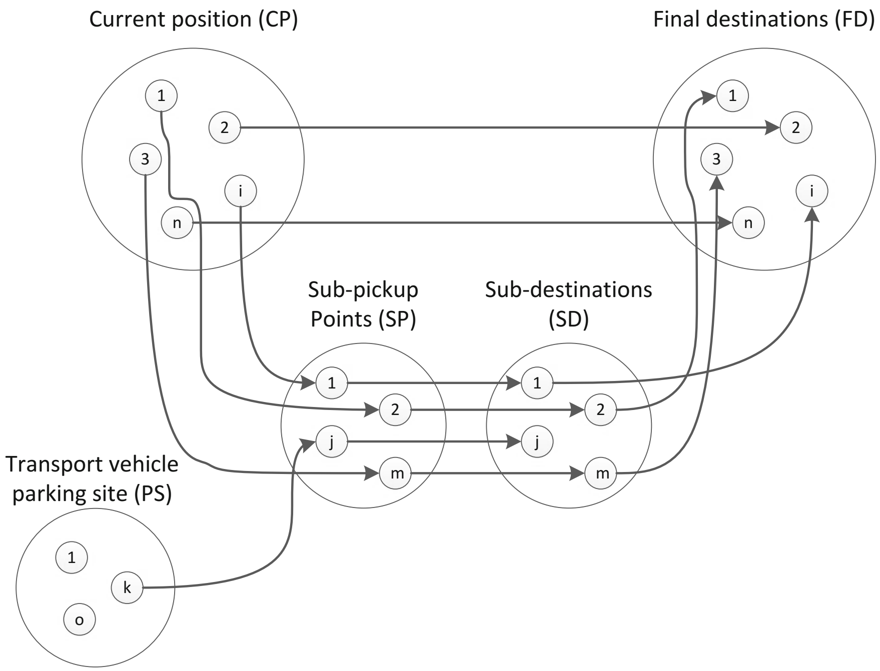


Fig. 2 Model of the supply chain related assignment problem with runtime optimization of free transportation resources

4 Model of Its Based Supply Chain

The model of the geoinformation based supply chain coordination can be divided into four main parts. The first part is the cooperative level. This level includes all networking partners for production and services processes. The suppliers are responsible for the production of components of final products produced by the main company. Trading companies mean business today to business partners. In the case of 3rd party logistic partners, the production company outsources logistic processes, like transportation, loading, warehousing (included bounded warehousing) and forwarding. Other services, like banking, insurance and customs belong to this level of the model too. The corporation level includes the design and operation of own resources. The product design and prototyping is the first important part of this level because this is the initial part of the lifetime of the product, where its logistic related requirements for production can be influenced. Design processes (CAD) are connected to process planning (CAPP) and production planning through material requirement planning (MRP I) and manufacturing resource planning (MRP II). The planning of production processes influences the manufacturing resources, so CAPP is connected to computer integrated manufacturing (CAM). The production planning and scheduling and the design and operation of internal

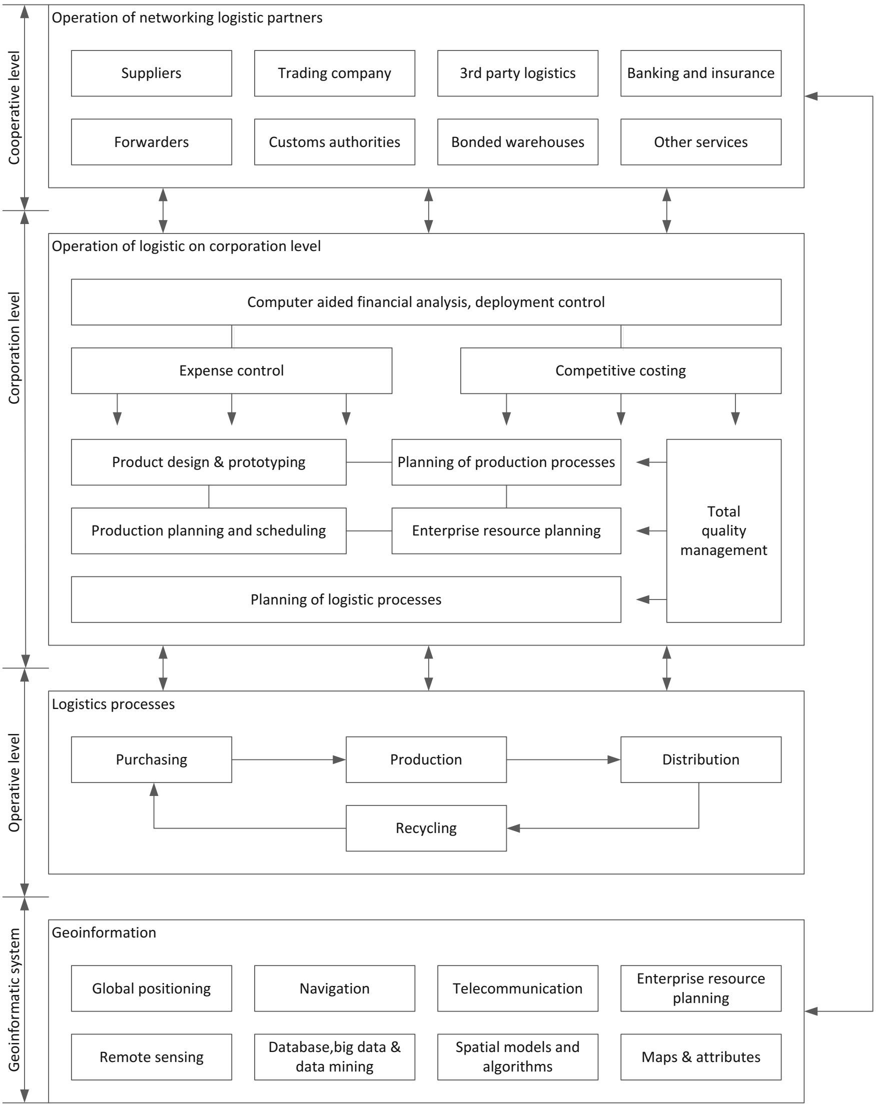


Fig. 3 General model of geoinformation based supply chain coordination

production and logistic resources are connected too. The financial analysis including expense control and competitive costing and the total quality management (TQM) effects on the whole corporation process. On the operation level, the logistics related processes can be divided into four important parts: purchasing, production, distribution and inverse logistics (Fig. 3).

The geoinformatics integrates the technologies and know-how of global positioning, navigation, telecommunication and IT solutions, remote sensing of vehicles and loads, database and related solutions for big data problems and data mining, spatial models, algorithms and decision aspects. The geoinformation based supply chain control optimizes the supply chain processes on the networking level.

The mathematical model takes into account the part of the problem that deals with vehicles on the road but has no effect on other activities, such as vehicle under loading. The model is designed to maximize the savings that can be achieved via inspecting several vehicles on the road by making a detour transport for a new demand beside its original task. If that's not possible or it's too expensive, a new vehicle should be launched to fulfil the new request. The basis of the model given by the distances between each point and it's proportional costs.

As formula (1) shows, savings can be achieved if a vehicle from it's current position (CP) goes to a sub-pickup point (SP), where it picks up goods and then dumps it at a sub-destination point (SD) and continues its original route to the final destination (FD). If the total shipping costs from the current position to the final destination and the costs from the transport vehicle parking site (PS) to SD through SP are significantly lower than the original routing costs than the additional transportation requirement will be performed by vehicles from PS. Along with these, we have to add the new vehicles launching and order processing cost. Furthermore, if a request is fulfilled, we can expect a delivery profit from another company. Formula (3) and (4) determines, how many vehicles do the task needs total. Formula (5) and (6) stipulates that nor can't the volume or the weight of the new demands be more than the free capacity of the vehicle. Furthermore formula (7) describes that the vehicle with the assigned new request has to arrive at its original destination no later than the deadline. The notations used in the paper are described in Table 1.

$$\begin{aligned} CS = \sum_{l=1}^{p} \left[CT_1 * L_{FDi}^{CPi} + CT_2 * \left(L_{SPj}^{PSk} + L_{SDj}^{SPj} \right) \right] \\ - CT_1 * \left(L_{SPj}^{CPi} + L_{SDj}^{SPj} + L_{FDi}^{SDj} \right) + \sum_{s=1}^{q} CO_s + \sum_{t=1}^{m-q} CP_t \rightarrow max. \end{aligned} \tag{1}$$

$$i = 1\ldots n; \quad j = 1\ldots m; \quad k = 1\ldots o \tag{2}$$

$$0 \le q \le m; \quad n \le p \le n + m \tag{3}$$

$$p = n + q \tag{4}$$

$$CAP_{Ti}^{V} - V_{TGi} \ge \sum_{r \in \theta_i} V_{TDrj} \tag{5}$$

$$CAP_{Ti}^{W} - W_{TGi} \ge \sum_{r \in \theta_i} W_{TDrj} \tag{6}$$

$$TC_i + T_{SPj}^{CPi} + T_{SDj}^{SPj} + T_{FDi}^{SDj} \le TR_i \tag{7}$$

Table 1 Notations used in the paper

CT_1	Original transportation cost, which is proportional with the distance
CT_2	Transportation cost of a new transport vehicle
CO_s	Cost of the *s*th order processing and assigning a new vehicle
CP_t	Profit from fulfilling the *t*th demand
L_{FDi}^{CPi}	Distance between the current position of the *i*th vehicle and the *i*th final destination
L_{SPj}^{PSk}	Distance between *k*th vehicle parking site and the *j*th sub-pickup point
L_{SDj}^{SPj}	Distance between the *j*th sub-pickup point and *j*th sub-destination
L_{SPj}^{CPi}	Distance between the current position of the *i*th vehicle and the *j*th sub-pickup point
L_{FDi}^{SDj}	Distance between the *j*th sub-destination and the *i*th final destination
n	Number of original transports and vehicles that are on road
m	Number of incoming transport demands
o	Number of transport vehicles parking sites
q	Number of new vehicles needed to fulfill the transport demands
p	Number of vehicles will be transporting including the new demands
θ_i	Set of elements of j, that the *i*th vehicle can transport due to capacity restriction
CAP_{Ti}^V/CAP_{Ti}^W	Volume/weight capacity of the *i*th vehicle
V_{TGi}/W_{TGi}	Volume/weight of the transported goods in the *i*th vehicle
V_{TDj}/W_{TDj}	Volume/weight of the *j*th transport demand
T_B^A	Time needed to go from A to B
TC_i	Current time of the *i*th vehicle at the CP point
TR_i	Requested arrival time of the *i*th vehicle to it's original destination

5 Heuristic Algorithm for Real-Time Optimization of Its Based Supply Chain

The above mentioned model and a genetic algorithm based heuristic make it possible to analyze different transportation problems represented by networking production companies. Table 2 describes the data set of the transportation demands. The geographical information of the additional transportation demands is described by the GPS coordinates of the sub-pickup points (SP) and the sub-destinations (SD). These GPS coordinates were generated randomly in the eastern part of Hungary. The required transportation demand is given in loading units (LU) and the estimated income is on this value.

The geographical information of the available transportation resources is described by the GPS coordinates of the free vehicles as current positions (CP). The data set includes the available loading capacities for each vehicle in loading units (Table 3).

The solution method of the above mentioned multidimensional optimisation problem is based on Rechenberg's genetic algorithm and includes the following

Table 2 Data set for the transportation demands

Nr.	GPS coordinates of the SPs		GPS coordinates of the SDs		Required capacity in LU	Estimated income in €
	Latitude	Longitude	Latitude	Longitude		
1	48.103	20.142	47.722	20.238	5	1000
2	48.943	20.005	48.459	20.315	4	2000
3	48.356	20.629	48.707	21.049	7	1800
4	48.676	21.403	48.829	21.758	8	1500
5	47.723	21.998	47.908	22.022	2	1200

Table 3 Data set for the available transportation resources

Nr.	GPS coordinates of the CPs		Available loading capacity (loading unit)
	Latitude	Longitude	
1	45.706	20.165	3
2	46.128	20.009	8
3	45.264	21.548	3
4	45.701	21.684	2
5	46.444	20.153	9
6	48.255	20.645	8
7	46.133	22.369	7
8	46.562	20.098	8
9	48.885	20.061	6

steps: (1) generate assignments between free transportation vehicles and additional transportation demands as permutation vectors; (2) initiate the first population of the genetic algorithm; (3) perform the genetic algorithm using reproduction, crossover and mutation operators until termination condition is true.

The operators were created to balance between exploration and exploitation. The reproduction operator gives preference to the solutions with the best fitness, allowing them to pass to the next generation.

The crossover operator produces child solutions from more than one parent solutions. The mutation is responsible for genetic diversity and it supports to avoid the convergence to a local minimum or maximum. The crossover operator is a one-point crossover operator, because the use of multi-point crossover operators is very complicated in the case of permutation vectors and special constraints. Figure 4 shows the convergence of the algorithm in the case of the mentioned data set, where the ratio of the reproduction, crossover and mutation operators is 50:46:4.

The results of the optimisation leaded to increased capacity utilisation of the transportation resources, because the additional transportation demands mustn't be performed by an expensive third party logistics provider, who is represented in the model by the transport vehicle parking site (PS).

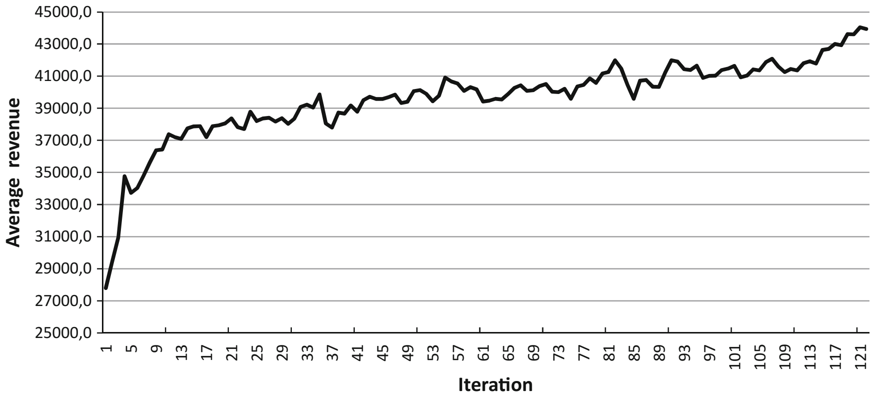


Fig. 4 Convergence of the algorithm

Table 4 Utilization of the free transportation resources

Nr.	Loading capacity	Available free loading capacity	Additional loading (loading unit)	Utilisation (percent)
1	20	3	–	85
2	25	8	–	68
3	20	3	–	85
4	30	2	–	93
5	25	9	–	64
6	15	8	7	100
7	20	7	–	65
8	20	8	–	60
9	25	6	4	40

As Table 4 and Fig. 5 shows not all of the free transportation resources are used to perform additional transportation processes because of the limited available loading capacity, the long distance among free transportation resources and sub-pickup point or among sub-destinations and final destinations, the value of the expected revenue and the short distance among transport vehicle parking site and sub-pickup points.

As Fig. 5 shows, in the case of the given dataset only two additional transportation demands are fulfilled by the available free transportation resources, the remaining seven are performed by the third party logistics provider.

This genetic algorithm based heuristic real-time optimization makes it possible to increase the efficiency of the whole supply chain including networking production companies with more sites and own transportation capacities.

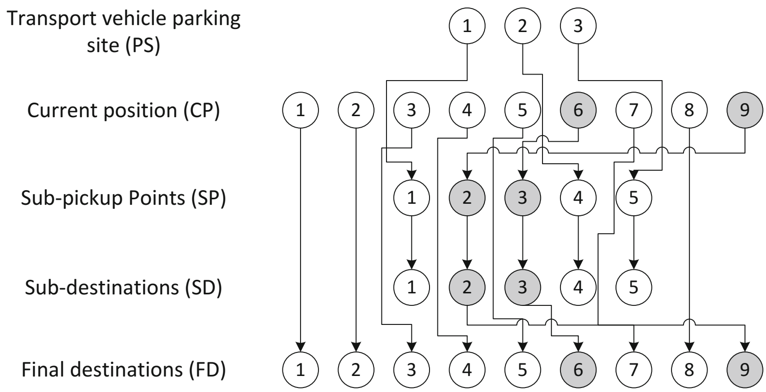


Fig. 5 One possible solution of the problems represented by the above data set

6 Conclusion and Further Research Directions

This study developed a methodological approach for coordination of supply chain of production companies. It proposed a model for companies performing their purchasing related transportations by own resources. Two models were proposed: the first general model shows the levels of supply chain coordination and the second model is a mathematical model which focuses on the optimization of the assignment problems. The algorithm is tested with different data sets. The results of this study can be used to improve the utilization of logistics resources and increase business efficiency. Within the frame of this paper the multiple transportation demands are ignored. This representation is near to the reality, but the analysis of the multiple transportation demands would be useful and this is the direction of our future researches.

Acknowledgements The described study was carried out as part of the EFOP-3.6.1-16-00011 "Younger and Renewing University—Innovative Knowledge City—institutional development of the University of Miskolc aiming at intelligent specialisation" project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund.

References

1. Wang Y, Wallace SW, Shen B, Choi T-M (2015) Service supply chain management: a review of operational models. Eur J Oper Res 247(3):685–698
2. Papapostolou C, Kondili EM, Kaldellis JK (2014) Energy supply chain optimisation: special considerations for the solution of the energy planning problem. Comput Aided Chem Eng 33:1525–1530
3. Sha M, Srinivasan R (2016) Fleet sizing in chemical supply chains using agent-based simulation. Comput Chem Eng 84(4):180–198

4. Chowdhury MMH, Quaddus MA (2015) A multiple objective optimization based QFD approach for efficient resilient strategies to mitigate supply chain vulnerabilities: the case of garment industry of Bangladesh. Omega 57(A):5–21
5. Véronneau S, Roy J, Beaulieu M (2015) Cruise ship suppliers: a field study of the supplier relationship characteristics in a service supply chain. Tour Manag Perspect 16:76–84
6. Martikainen A, Niemi P, Pekkanen P (2014) Developing a service offering for a logistical service provider—case of local food supply chain. Int J Prod Econ 157:318–326
7. Liu W, Wang Q, Mao Q, Wang S, Zhu D (2015) A scheduling model of logistics service supply chain based on the mass customization service and uncertainty of FLSP's operation time. Transp Res Part E: Logistics Transp Rev 83:189–215
8. Alaei S, Setak M (2015) Multi objective coordination of a supply chain with routing and service level consideration. Int J Prod Econ 167:271–281
9. Mvundura M, Lorenson K, Chweya A, Kigadye R, Bartholomew K, Makame M, Lennon TP, Mwangi S, Kirika L, Kamau P, Otieno A, Murunga P, Omurwa T, Dafrossa L, Kristensen D (2015) Estimating the costs of the vaccine supply chain and service delivery for selected districts in Kenya and Tanzania. Vaccine 33(23):2697–2703
10. Kaya O, Urek B (2016) A mixed integer nonlinear programming model and heuristic solutions for location, inventory and pricing decisions in a closed loop supply chain. Comput Oper Res 65:93–103
11. Koç C, Bektaş T, Jabali O, Laporte G (2016) Thirty years of heterogeneous vehicle routing. Eur J Oper Res 249(1):1–21
12. Crainic TG, Gendreau M, Potvin J-Y (2009) Intelligent freight-transportation systems: assessment and the contribution of operations research. Transp Res Part C: Emerg Technol 17(6):541–557
13. Ko YD, Jang YJ, Lee MS (2015) The optimal economic design of the wireless powered intelligent transportation system using genetic algorithm considering nonlinear cost function. Comput Ind Eng 86:67–79
14. Shah N, Kumar S, Bastani F, Yen I-L (2012) Optimization models for assessing the peak capacity utilization of intelligent transportation systems. Eur J Oper Res 216(1):239–251
15. Zhang F, Johnson D, Johnson M, Watkins D, Froese R, Wang J (2016) Decision support system integrating GIS with simulation and optimisation for a biofuel supply chain. Renew Energy 85:740–748
16. Cantarella GE (2013) Day-to-day dynamic models for Intelligent Transportation Systems design and appraisal. Transp Res Part C: Emerg Technol 29:117–130
17. Hoff A, Andersson H, Christiansen M, Hasle G, Lokketangen A (2010) Industrial aspects and literature survey: fleet composition and routing. Comput Oper Res 37:2041–2061
18. Degertekin SO (2012) Improved harmony search algorithms for sizing optimization of truss structures. Comput Struct 92–93:229–241
19. Ceylan H, Ceylan H, Haldenbilen S, Baskan O (2008) Transport energy modeling with meta-heuristic harmony search algorithm, an application to Turkey. Energy Policy 36(7): 2527–2535
20. Aghaie M, Nazari T, Zolfaghari A, Minuchehr A, Shirani A (2013) Investigation of PWR core optimization using harmony search algorithms. Ann Nucl Energy 57:1–15
21. Wu B, Qian C, Ni W, Fan S (2012) Hybrid harmony search and artificial bee colony algorithm for global optimization problems. Comput Math Appl 64(8):2621–2634
22. Fesanghary M, Mahdavi M, Minary-Jolandan M, Alizadeh Y (2008) Hybridizing harmony search algorithm with sequential quadratic programming for engineering optimization. Comput Methods Appl Mech Eng 197(33–40):3080–3091
23. Amini F, Ghaderi P (2013) Hybridization of harmony search and ant colony optimization for optimal locating of structural dampers. Appl Soft Comput 13(5):2272–2280

A New Scheduling Software for Supporting Automotive Component Manufacturing

Mónika Kulcsárné Forrai and Gyula Kulcsár

Abstract This paper shows the modeling and solving of a special production fine scheduling problem of the automotive industry. A new scheduling software has been developed to create execution plans to the production activities that satisfy the customers' demands. The main characteristic of the scheduling problem is that different types of shared resources (e.g. production lines, attachment points, mold carriers, tools) have to be allocated simultaneously, and multi-processing tasks have to be scheduled to meet production orders within strict time limits. To solve the problem we consider not only the primary technological processes but also the tool-preparation processes. This inbuilt sub-problem is converted to a special resource environment by using a problem space transformation procedure. To minimize the tardiness, we schedule the jobs in the given resource environment. We elaborated a new solving algorithm that can create the optimal solution for the sub-problem in polynomial running time. This solution for the sub-problem is applied to meet the tool-preparation constraints of the full production fine scheduling problem. An advanced multi-objective searching algorithm solves the full problem. The paper presents the approach of the developed solving method, the defined objective functions and the applied neighbouring operators. In the solving process, all the decision making sub-tasks (assigning, sequencing and timing) are managed simultaneously. The concrete values of the decision variables are set by a multi-operator and multi-objective local searching algorithm. The fine scheduling software can also support inventory control by using special objective functions that can help to optimize the manufacturing from the point of view of the product type dependent stock levels. The scheduling problem comes from the plant of Fehrer Hungaria Járműipari Kft., specialized in vehicle seat products (Mór, Hungary).

M.K. Forrai (✉) · G. Kulcsár
Department of Information Engineering,
University of Miskolc, Miskolc, Hungary
e-mail: aitkfm@uni-miskolc.hu

G. Kulcsár
e-mail: iitkgy@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_25

1 Introduction

Today's globalized market environment compels manufacturers to continuously improve their competitiveness. The rapid changes in the market conditions and the increasing customer demands make it necessary to improve the production efficiency and the readiness for delivery. However, in order to achieve cost-effective production manufacturers strive to maintain low inventories, achieve high utilization of resources and reduce the associated costs.

Nowadays, modern enterprise management activities cover all operations of the production systems. Production engineers and managers utilize ever more computer integrated application systems to support their decision making. Software systems for discrete production processes can be classified into four hierarchical groups according to the different supported fields and time horizons. These application categories are as follows:

1. Enterprise Resources Planning (ERP)
2. Computer Aided Production Engineering (CAPE)
3. Manufacturing Execution Systems (MES)
4. Manufacturing Automation (MA).

Production planning and control modules of the integrated systems deal with the allocation of limited resources to production activities in order to satisfy given customer demands over a well-defined time horizon. Both planning and control tasks lead to optimization problems, in which the main goal is to create plans that meet the actual demands and maximize the production performance.

The planning and control of discrete production processes usually take place at several hierarchical levels. This approach means that the solving processes run in a hierarchical way by ordering the corresponding decisions according to their relative importance. Hierarchical production planning uses a suitable optimization model at each level of the hierarchy. Each model extends the constraints of the problem at the lower level.

In practice, models for solving these problems cover a very wide range due to the varying characteristics of different production systems and their business environments. However, the results of production planning and scheduling in general cannot be applied directly to the operational production management because the created plans are based on aggregated resources and they consider only the primary processes. In these cases, production fine scheduling can solve the problem. An advanced fine scheduling software is needed to create short-term execution plans (fine schedules) that cover every important detail and process.

In this paper, we present an advanced scheduling software with the applied model and algorithms for supporting the automotive seat component manufacturing.

2 Problem Description and Requirements

The scheduling problem, which will be outlined in this section, is inspired by a real case study concerning the production system of Fehrer Hungaria Járműipari Kft., a firm specialized in vehicle seat components (Mór, Hungary). This company produces different types of seat elements with variable series simultaneously. The customers (business partners) have very strict delivery due dates. The number of the product types is increasing because of market trends. Due to these requirements, it is important to develop better production plans, make more flexible, fine schedules, and use advanced software at the shop floor level.

The plant produces seat components for different vehicle-assembly enterprises. These customer enterprises regularly send their product demands as production orders that include the product types, the numbers of items and the due dates. These orders have to be fulfilled within strict time limits by using the available resource-constrained manufacturing system.

The seat elements are made on circle-shaped production lines (Fig. 1). The system manages all the production orders as a set. In general, the product types can be manufactured on more than one production line (PL), and a PL can carry out a specified number of laps (rounds) in one shift. Each PL has a dedicated number of attachment points that are called positions. Shape carriers are connected to these positions. The construction of the carriers can be one or two-sided and they transport the shapes on the path. The shape means a special tool, which is named mould in practice. The tools can be placed on the left and/or right side of the carrier. The permitted configurations of the devices are specified by master rules and technological plans. These regulations determine the assignments and usage of production lines, positions, carriers, tools, sides and product groups.

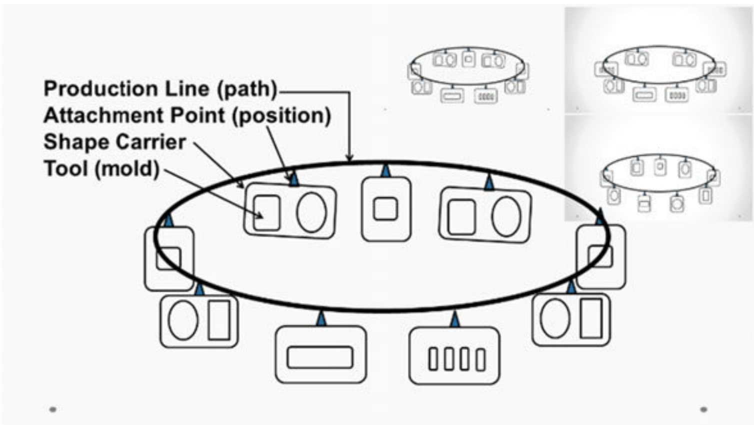


Fig. 1 A simplified scheme of the production lines

The production lines are characterized by different constructions (number and type of positions, manufacturing capabilities, etc.), and the intensities (number of rounds in one shift) can also differ. Each production line has an independent timetable that specifies the enabled shifts as a work calendar on the actual time horizon. The shift is the unit-long time interval of the calendar. In the shifts the settings of the production lines can be slightly modified by exchanging one or more configurations attached to the positions. This exchanging activity is the basic control operation of the manufacturing execution system. A configuration consists of one carrier and its mounted tools. For manufacturing products many suitable carriers and tools are accessible as shared resources. When we want to activate a given configuration in a given position of a given production line, we have to disconnect the actual configuration from the position and connect another prepared configuration to the same position.

To activate a prescribed configuration, the workshop has to prepare it in advance. The assembly and disassembly tasks of the configurations are carried out by skilled workers. Such an additional process takes considerable time, so the number of the performable preparatory tasks is constrained in the full system during one shift. The constraints can vary in time.

In connection with the described production scheduling problem the most important requirements of shop floor management can be summarized as follows:

- The production orders should be filled in time with no exceptions.
- The manufacturing processes should be realized with the least possible number of configuration preparations and exchanges.
- The accumulation of excessive stock should be avoided.
- The product-type dependent minimum stock levels should be maintained.
- The utilization of the production lines should be maximized.

The main purpose of our project was to develop fine scheduling software in order to initiate a detailed execution program for the production lines in order to fill the production orders and to synchronize the configuration preparation and exchange activities on the shop floor in the manufacturing system. The defined problem does not fit into any of the known scheduling problem categories, so we have identified a new combined problem of shop floor scheduling and inventory management.

3 Scope of the Research

In the literature, different scheduling models can be found (e.g. [1–3]). Many survey papers are also focused on modelling and solving production scheduling problems (e.g. [4–6]). One of the main groups of these models is the parallel resource scheme. Detailed reviews for this topic are given, for example, in [7–9]. The parallel resource environment consists of machines that represent the fundamental building block of the system.

Considering production performance, both the allocation of machines and the sequence of jobs are of great importance. Many scheduling models are known in the literature, but most of them use only one objective function. Usually one of the performance indicators related to due date plays the main role in scheduling models for supporting Make to Order (MTO) manufacturing. Only a few of the models (e.g. [10–12]) deal with multi-objective cases, which are very important in practice.

Since the existing models in the operation research topic often disregard the machine processing abilities, limited availability time frames of the machines, limited logistical capacities, and shared tools, the improvement and extension of the models are justified. In order to consider the aforementioned important features of real scheduling problems we have focused on the simulation based scheduling approach. This paper presents our approach, in which a practice-oriented simulation based multi-objective searching method is used to solve the addressed fine scheduling problem of vehicle seat manufacturing by preserving important details.

4 Practice-Oriented Fine Scheduling Based on Simulation

We have elaborated an integrated approach to solve the examined combined problem of production scheduling and inventory control as a whole without decomposition. In our approach (Fig. 2), all the decisions are made simultaneously. Applying the solution in practice, a special searching algorithm has been developed based on simulation, overloaded relational operators and multiple neighbouring

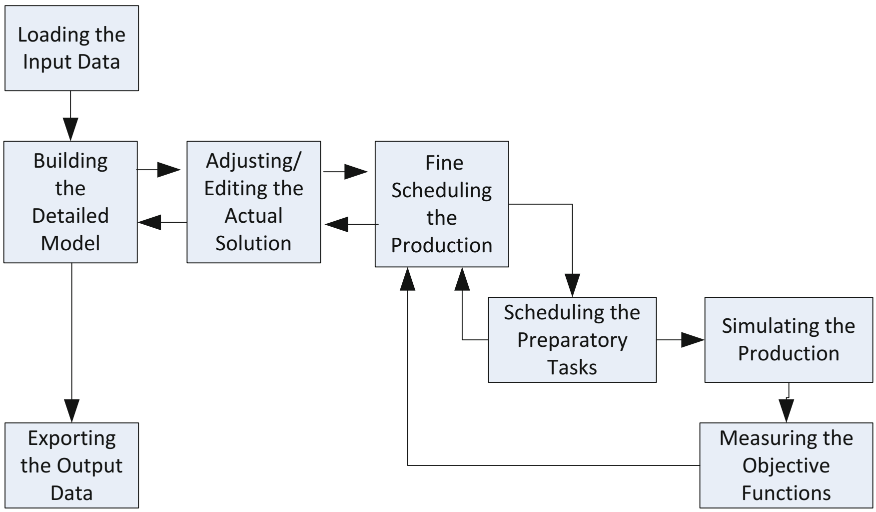


Fig. 2 Fine scheduling approach

operators. The core of the solver explores iteratively the feasible solution space and creates neighbour candidate solutions by modifying the decision variables. The created candidate fine schedule declares the configuration of carriers and tools to be used and the product types to be produced in each position of each production line in each shift.

When the searching algorithm creates a new candidate fine schedule, it has to decide whether the candidate solution is executable. This means that the required preparatory activities have to be scheduled in time by considering the constrained capacities of the skilled workers. If there is no lateness when the needed preparatory tasks are finished, then the production fine schedule is executable; otherwise it is not feasible.

To analyse the candidate schedules, the actual values of the objective functions are evaluated by a simulation that virtually performs the planned operations and processes in the modelled replica of the real resource environment with capacity and technological constraints. In this execution-driven simulation algorithm, the workpieces are passive objects of the model and they are processed, moved, and stored by the active model objects such as production lines, material handling devices, human workers and buffers. The numerical tracking of the product units provides the time data of the manufacturing processes. Consequently the simulation can also calculate exactly the stock level values of the product types in the planned time horizon. This is the part of the approach that encapsulates the dependency of real-world scheduling and inventory control problems. The successful adaptation of the approach into practice is highly influenced by the efficiency of the simulation algorithm.

5 Scheduling the Configuration Preparatory Tasks

The configuration-preparation sub-problem of the production fine scheduling can be modelled independently as a special parallel machine scheduling problem with time-varying resource availability constraints. To solve this problem, we have developed a rule-based algorithm that gives an optimal solution from the point of view of the maximum lateness.

This scheduling sub-problem is converted to a special resource environment by using a problem space transformation procedure. This transformed problem has a time-varying number of virtual parallel resources, independent jobs with unit processing times, integer release times and due dates. To minimize the tardiness, we schedule the jobs in the given resource environment.

The basic idea of the method is that the actual number of the available worker teams is calculated and stored in each shift, than we schedule the releasable task with the earliest due date. The algorithm starts from the earliest shift and goes on with the enabled shifts in chronological order. If any free worker team can be found in an intermediate shift, and there are releasable unscheduled tasks, then the task with the earliest due date will be scheduled by assigning it to any free worker team and to the actual shift. This algorithm can create the optimal solution of the sub-problem in polynomial running time.

The created schedule of the configuration preparations specifies exactly what preparatory task is carried out in which shift. If there is any late preparatory task in the created optimal solution, then the inbuilt scheduling problem is unsolvable, consequently the examined full fine schedule is not feasible. If each preparatory task can be finished in time without lateness, then the solution is suitable to meet the preparation requirements of the planned production processes.

6 Multi-objective Fine Scheduling

6.1 Scheduling Engine

The core of the scheduling software uses an advanced searching algorithm variant that iteratively moves from an actual schedule to a candidate's schedule in the neighbourhood until the stop criterion is satisfied. To reach and examine the unexplored regions of the search space the method creates new neighbours of the base solution. The software provides default searching parameters, but the user can calibrate the actual values on the main form of the solver (Fig. 3).

To avoid a local optimum the method puts the taboo schedules that have been examined in the recent past into a taboo list. The taboo schedules are excluded from the further neighbourhoods. The taboo list works as short-term memory limited by a

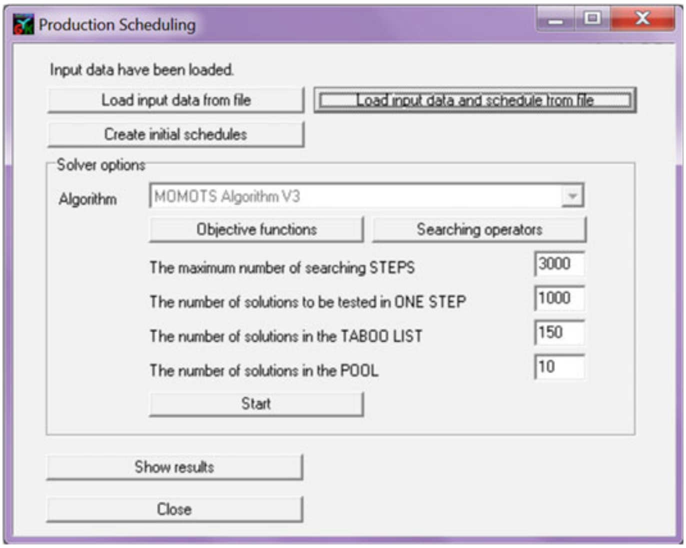


Fig. 3 The main form of the solver

parameter. A pre-defined value specifies the number of the stored elements. A new set of neighbours is generated at random successively by using priority controlled neighbouring operators. An input parameter specifies the number of the neighbours in the neighbourhood.

6.2 Neighbouring Operators

To generate new solutions the operators can carry out permitted modifications on the base solution. The applied neighbouring operators are as follows:

- N_1 operator removes the currently used configuration from a randomly chosen position of a randomly chosen production line in a randomly chosen shift.
- N_2 operator chooses a production order randomly and schedules the manufacturing of its product on the default production line of the current product. If it is not possible, then the operator chooses and uses an alternative production line.
- N_3 operator chooses a late production order randomly and schedules the manufacturing of its product similarly to operator N_2.
- N_4 operator chooses a late production order randomly and schedules the manufacturing of its product on a randomly chosen suitable production line of the current product.
- N_5 operator chooses the production order that has the highest priority from the set of late orders and schedules the manufacturing of its product on a randomly chosen suitable production line of the current product.
- N_6 operator includes three different modification activities. When this operator is called for running, one of three variants is chosen randomly.
  The first variant chooses the production order that can be filled with the least amount of moulds from the set of late orders. The second variant chooses the production order with the earliest due date from the set of late orders. The third variant chooses the production order that can be filled on the least amount of production lines from the set of late orders.
  In all three cases, the operator schedules the manufacturing of the product which is required by the chosen production order. The job execution is assigned to the suitable production line which is loaded with the least number of planned setups.
- N_7 operator chooses randomly one of the two following modifications:
  In the first case, the operator chooses the production order that has the highest priority from the set of late orders. If all moulds of its product are reserved, then the operator frees a randomly chosen mould and uses it to schedule the manufacturing of the product on the suitable production line which is loaded with the least number of planned setups.
  In the second case, the operator chooses randomly a production order from the set of late orders. If all moulds of its product are reserved, then the operator frees

a randomly chosen mould and uses it to schedule the manufacturing of the product on a randomly chosen suitable production line.

- N_8 operator removes a randomly chosen configuration exchange from the schedule.
- N_9 operator chooses the production order that has the maximal lateness and schedules the manufacturing of its product on a randomly chosen suitable production line.
- N_{10} operator chooses the production order that has the maximal lateness and schedules the manufacturing of its product on the default production line of the current product. If this is not possible, then the operator chooses and uses an alternative production line.
- N_{11} operator chooses the product that will have the maximal shortage compared to the required minimal stock level at the end of the time horizon. The operator schedules the manufacturing on the default production line of the chosen product. If this is not possible, then the operator chooses and uses an alternative production line.
- N_{12} operator chooses a product randomly from the set of products which will have shortages compared to the required minimal stock level at the end of the time horizon. The operator schedules the manufacturing of the chosen product similarly to operator N_{11}.

Scheduling the manufacturing of a given product in a given position of the production line means that the algorithm firstly chooses a shift in which a new exchange can be carried out without exceeding the prescribed maximal number of exchanges. The second step looks for the final shift by considering the calendar and the shift in which the next exchange will be carried out. After determining the actual time frame, the neighbouring operators select the suitable free moulds and examine whether there is any prepared configuration that includes suitable moulds. If the answer is yes, then one of the deployable configurations is chosen randomly for scheduling, otherwise the algorithm chooses the most appropriate free mould based on the priorities associated with the moulds previously. If the actual carrier can mount more than one mould, then the corresponding set of moulds is selected for scheduling. In this decision situation, the neighbouring operators take also into account what kind of products can be manufactured in the same group. The methods of the neighbouring operators cover the handler of the shared available carriers and moulds.

6.3 Objective Functions

The performance indicators of the generated solutions are calculated based on an in-built simulation module. This component serves the actual values of the objective functions. To express the shop floor management's goals as the criteria of a multi-objective optimization problem, we use eighteen objective functions to be minimized. These are as follows:

- f_1: The maximum product shortage [pc]
- f_2: The sum of product shortages [pc]
- f_3: The number of tardy production orders
- f_4: The number of set-up activities (changes)
- f_5: The maximum number of set-up activities in one shift
- f_6: The number of product types with surplus
- f_7: The sum of product surpluses [pc]
- f_8: The maximum product shortage at the end of the time horizon [pc]
- f_9: The sum of product shortages at the end of the time horizon [pc]
- f_{10}: The sum of the priorities of tardy production orders
- f_{11}: The maximum priority of tardy production orders
- f_{12}: The number of product types with tardiness
- f_{13}: The maximum product shortage (compared to zero) [pc]
- f_{14}: The sum of product shortages (compared to zero) [pc]
- f_{15}: The maximum tardiness of production orders [shift]
- f_{16}: The sum of the tardiness of production orders [shift]
- f_{17}: The number of configuration preparations
- f_{18}: The sum of unused capacities of the production lines [Standard Production Unit].

In the model we apply new key performance indicators (KPIs) concerning the stock levels as objective functions. The essence of the measure of these indicators is that we merge the manufactured product quantities and the delivered product quantities in a stock-time function, so at any time the actual value of the stock level in the case of each product type can be calculated (Fig. 4). Using these values, we can simply calculate the shortage or the surplus compared to the lower reference

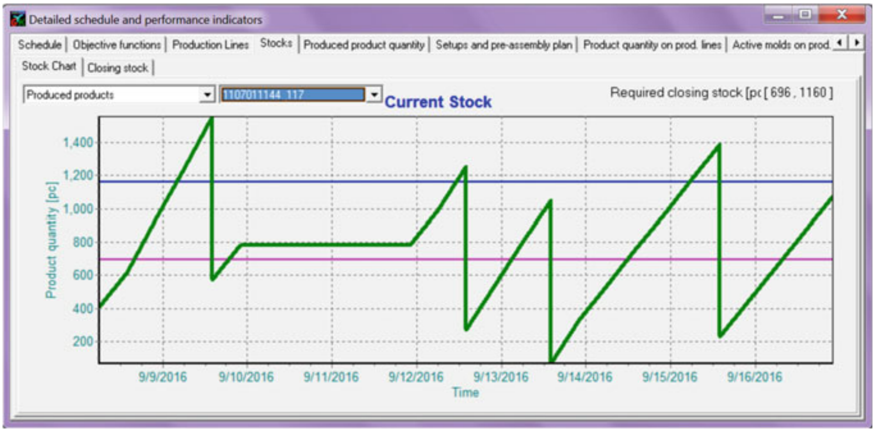


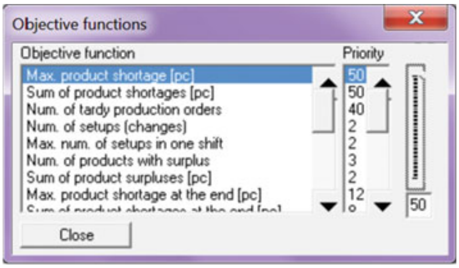
Fig. 4 Stock-time chart

limit or to the upper reference limit. The tardiness of the production orders can also be easily calculated based on the stock-time chart (Fig. 5). We measure the difference between the actual stock level and the required quantity at the time of the accepted delivery due date. This signed value of the product quantity shows the shortage (negative) or the surplus (positive). The lateness of a given production order means the time interval that has elapsed between the prescribed deadline and the real delivery time. If the lateness is negative or zero, then the given production order is satisfied in time (the tardiness is zero), otherwise the production order is tardy and its tardiness is greater than zero. Using this approach based on the individual value of each production order, the maximum and the sum values of the total set can also be interpreted similarly to the usage of job data and optimality criteria in classical scheduling theory.

The importance of the declared goals can vary over time, so the software supports the user in expressing the actual importance of the objective functions by adjusting the priorities (Fig. 6).

Fig. 5 The tardiness of the production order

Fig. 6 User interface for adjusting the priorities



6.4 Comparing and Qualifying the Candidate Solutions

For managing effectively the full system of the objective functions, in the development of the software we used our own previously elaborated method, which is described in detail in [13, 14]. This model is based on the calculation of the relative quality of a given solution by comparing it to another solution, considering multiple objective functions simultaneously.

The formal description of the relative qualification model is as follows:

$$f_k : S \to R^+ \cup \{0\}, \quad \forall k \in \{1, 2, \ldots, K\}, f_k \to \min \tag{1}$$

$$D : R^2 \to R, D(a,b) = \begin{cases} 0, & \text{if } \max(a,b) = 0 \\ \frac{b-a}{\max(a,b)}, & \text{otherwise} \end{cases} \tag{2}$$

$$F : S^2 \to R, F(s_x, s_y) = \sum_{k=1}^{K} (w_k \cdot D(f_k(s_x), f_k(s_y))) \tag{3}$$

where

S	the set of the feasible solutions;
f_k	the kth objective function to be minimized;
K	the number of objective functions;
s_x,s_y	two given solutions;
w_k	the priority of the kth objective function;
$F(s_x,s_y)$	the relative quality of the solution s_y compared to the solution s_x

.

Using the signed value of the function $F(s_x, s_y)$ we extend the interpretation of the relational operators to the solutions s_x and s_y in the S. The definition of this operator overloading is the following:

$$(s_y ? s_x) := (F(s_x, s_y) ? 0) \tag{4}$$

where the question mark ? indicates any of the relational operators (<, $\leq$, >, , =, $\neq$). For example, s_y is a better solution than s_x ($s_y < s_x$ is true) if $F(s_x, s_y)$ is less than zero.

This relative qualification model can effectively solve the comparison of the candidate solutions in the searching process, so the software is able to realize multi-objective optimization by taking into account the actual requirements of the users.

6.5 Advanced Functionalities of the Scheduling Software

The implemented scheduling software based on this approach can automatically create a detailed solution that specifies what the manufacturing system should perform in the planned time horizon.

To increase the flexibility and effectiveness of the production planning, scheduling and control process, we developed an advanced software module for supporting user interactions. Graphical interfaces provide useful charts, diagrams, tables and reports to visualize the results.

The full scheduling process can be scrutinized and checked on the screen. At runtime the user can modify the control priorities and the parameters of the searching algorithm (Fig. 7).

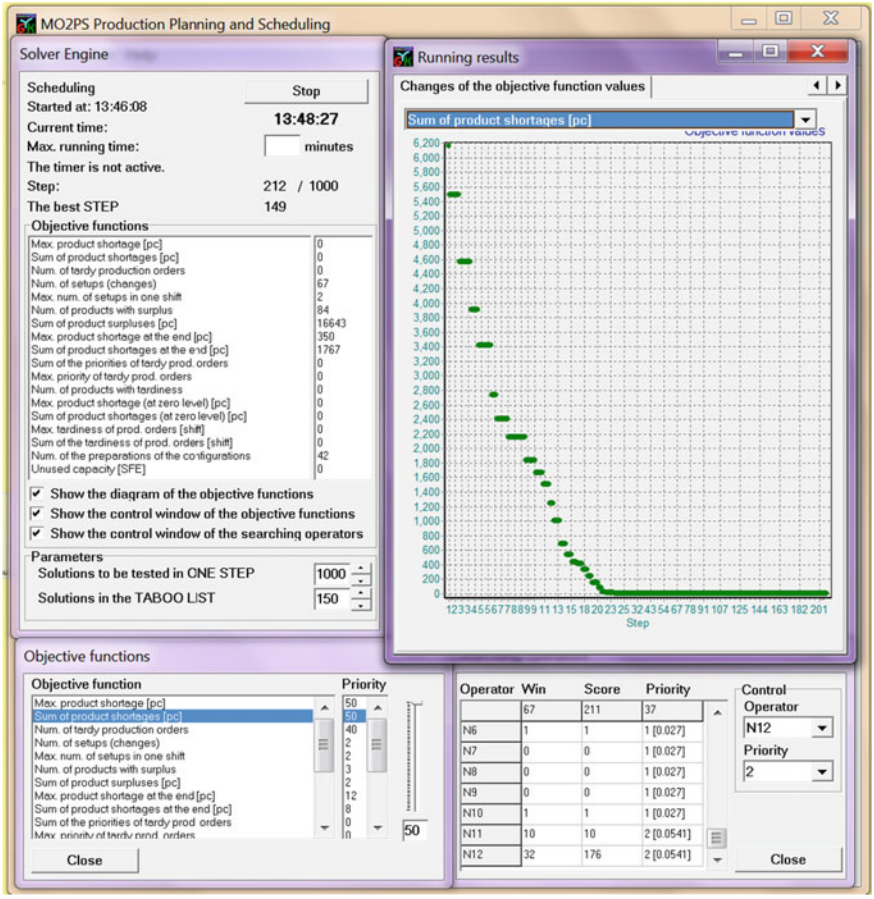


Fig. 7 Graphical user interfaces of the solver engine

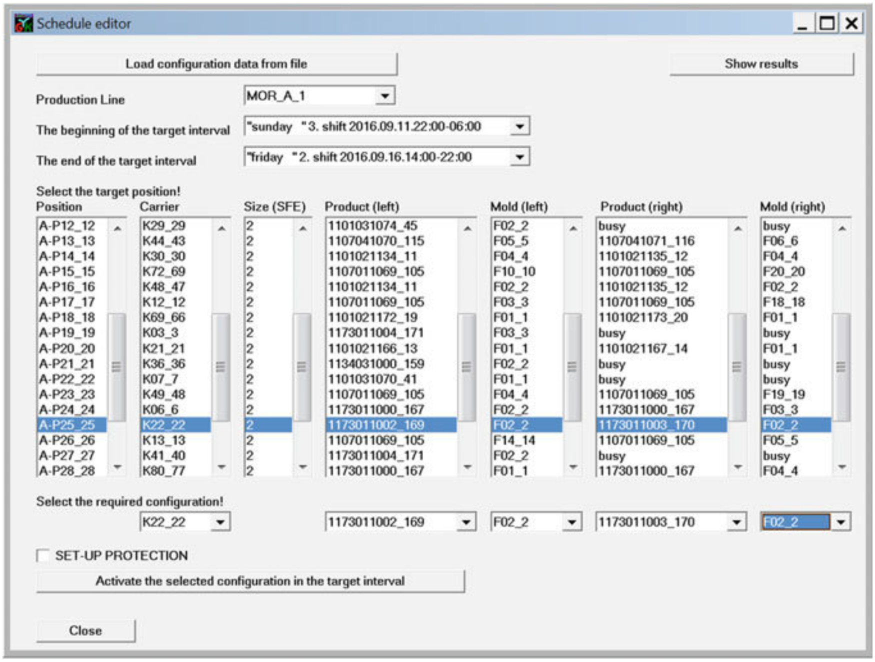


Fig. 8 An editor window for supporting direct user interventions

The user continuously receives full information about the actual status when the scheduling process is running. For example, the forms show the elapsed time, the current and the best searching steps, the current values of the objective functions and their tendencies. The software effectively supports the usage of the human intelligence. It allows the expert user to optionally control the searching process by modifying the priorities of the objective functions and the neighbouring operators.

In addition, the application also makes it possible for the user to edit the actual schedule by using available operating tools manually (Fig. 8). Similarly to the neighbouring operators, the built-in protection ensures that the usage of the manual editing tools can lead only to feasible solutions.

The editor module and its outlined feature can also play a key role when the process engineers want to declare mandatory configurations to be used for testing. In such cases, they can use the same editing tools to express their exact requirements, but the checkbox with the title SET-UP PROTECTION has to be ticked before clicking on the button to activate the modification (Fig. 8). The automatic scheduler has to take into account these new constraints. For this purpose, we equipped the solver engine with blocking techniques that make it possible to recognize and distinguish the modifiable and the protected configuration exchanges.

The software offers many formats to show the aggregated and detailed information of the fine schedule. For example, the fine schedule of the production

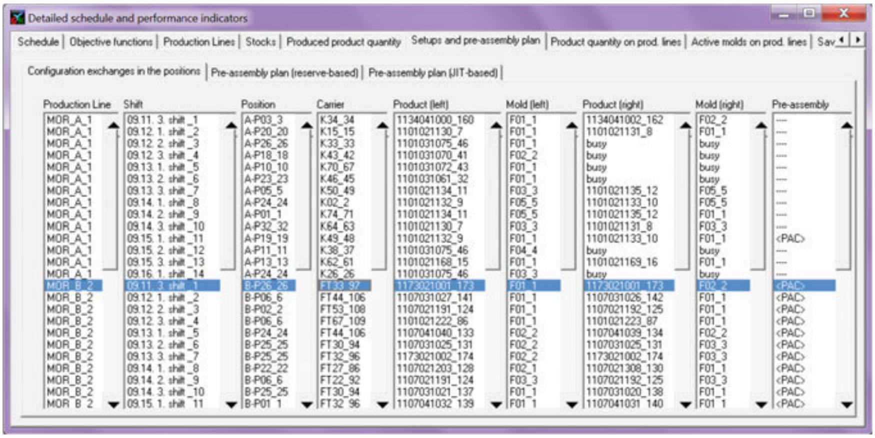


Fig. 9 An appeared format of the created fine schedule of the manufacturing processes

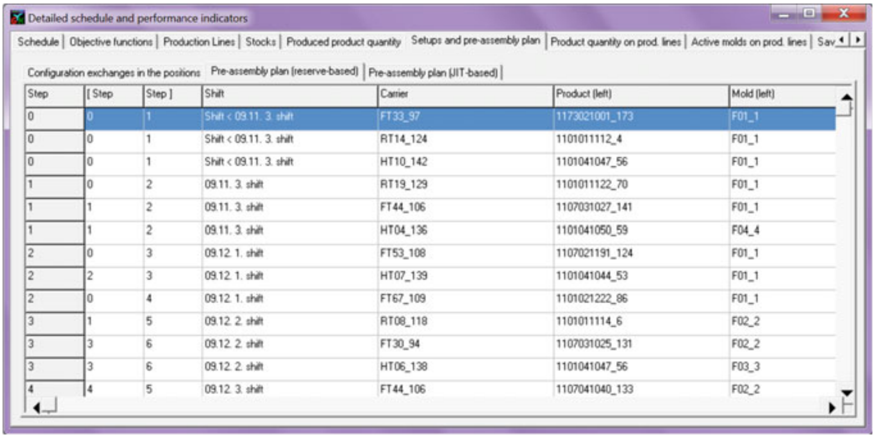


Fig. 10 The slack-oriented schedule of the configuration preparatory tasks

processes can be displayed as a list of the configuration exchanges with the corresponding data that declare which configurations (carriers and tools) have to be attached to which position in which shift and what kind of product types have to be produced (Fig. 9). Each attached configuration is used in the position until it is replaced by another configuration.

The schedule of the configuration preparatory tasks can be shown in a simple table that specifies exactly what pre-assembly activity (PAC) has to be carried out in which shift (Fig. 10). Each row means one PAC task. The numbers in the second and third columns of the table show the constrained time interval bounded by the identifiers of the earliest shifts and the latest shifts in which the task can be done.

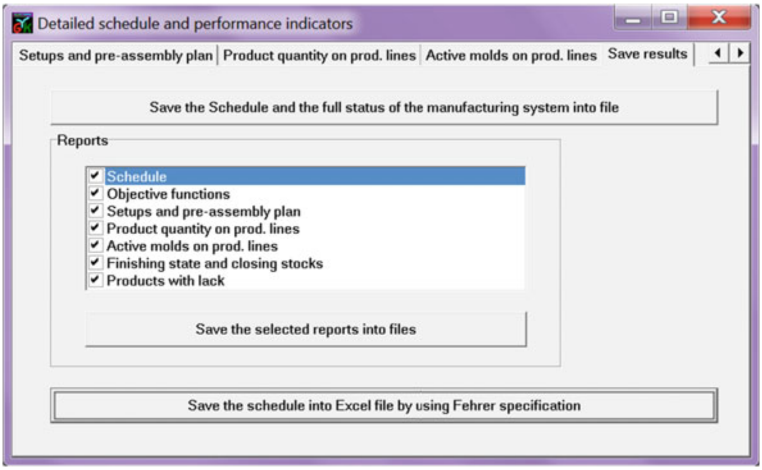


Fig. 11 User interface for exporting the selected reports

The results are in the first column, which gives the identifiers of the chosen shifts, and the same data can be seen in the fourth column in the calendar item (date) format.

The results of the fine scheduling process can also be exported as reports into xls (excel) or csv (comma separated values) file format in order to increase the portability (Fig. 11). The saved data can be loaded into our software or other applications.

In summarizing the research and development work, we can say that our production fine scheduling model integrates many new functional components and sub-systems, of which some essential items are as follows:

- Multi-objective fine scheduling of the production processes
- Managing the time-varying availability-constrained manufacturing resources
- Managing the shared accessible resources
- Managing the predefined configurations which have to be used in given positions and given time intervals for the test manufacturing required individually by process engineers
- Scheduling the preparatory tasks of the manufacturing
- Optimizing the stock levels within the product-typed dependent limits
- Significant expansion of the scope of the objective functions in order to meet industrial demands.

These advanced functionalities of the presented software effectively help to satisfy the requirements of shop floor control systems in practice.

7 Conclusions

This paper is primarily concerned with automotive industrial scheduling problems, which require advanced scheduling software to assign limited available resources to the operations and to schedule the preparatory and the manufacturing activities over time. The problem comes from automotive seat component manufacturing, in which typically variable series of items are produced in parallel. The developed models, algorithms and the implemented software are used in practice.

This paper describes the application of a practice-oriented approach for solving multi-objective scheduling problems. It is based on built-in production simulation and usage of relational operators for comparing the qualities of solutions in advanced searching algorithms. The introduced software has been successfully tested on real problems considering the multiple objectives and constraints originating from the automotive industry. Our new application can schedule predictively what the actual manufacturing system should perform in the planned time horizon to fill the current production orders. Each solution created for the shop floor is feasible because the model considers all of the hard constraints. The software can also support the management in managing the production scheduling and the inventory control tasks in an integrated manner. The obtained results encourage us to apply our approach to solving other planning, scheduling and control problems of the industry.

Acknowledgements This research was partially carried out in the framework of the Center of Excellence of Mechatronics and Logistics at the University of Miskolc. The financial background of the software development was provided by the Fehrer Hungaria Járműipari Kft.

References

1. Brucker P (2007) Scheduling algorithms, 5th edn. Springer, Berlin
2. Pinedo ML (2009) Planning and scheduling in manufacturing and service, 2nd edn. Springer, New York
3. Pinedo ML (2008) Scheduling theory, algorithms, and systems, 3rd edn. Springer, New York
4. Allahverdi A, Ng CT, Cheng TCE, Kovalyov MY (2008) A survey of scheduling problems with setup times or costs. Eur J Oper Res 187:985–1032
5. Aytug H, Lawley MA, Mckay K, Mohan S, Uzsoy R (2005) Executing production schedules in the face of uncertainties: a review and some future directions. Eur J Oper Res 161:86–110
6. Lei D (2009) Multi-objective production scheduling: a survey. Int J Adv Manuf Technol 43 (9–10):926–938
7. Cheng TCE, Sin CCS (1990) A state-of-the-art review of parallel-machine scheduling research. Eur J Oper Res 47:271–292
8. Koulamas C (1994) The total tardiness problem: review and extensions. Oper Res 42:1025–1041
9. Mokotoff E (2001) Parallel machine scheduling problems: a survey. Asia-Pac J Oper Res 18:193–242

10. Baykasoğlu A, Özbakir L, Dereli T (2002) Multiple dispatching rule based heuristic for multi-objective scheduling of job shops using tabu search. In: Proceedings of the 5th international conference on managing innovations in manufacturing, Milwaukee, USA, pp 396–402
11. Loukil T, Teghem J, Tuyttens D (2005) Solving multi-objective production scheduling problems using metaheuristics. Eur J Oper Res 161:42–61
12. Sbalzarini LF, Müller S, Koumoutskos P (2000) Multiobjective optimization using evolutionary algorithms. In: Proceedings of the summer program 2000, Center of Turbulence Research, pp 63–74
13. Kulcsár Gy, Kulcsárné Forrai M (2009) Solving multi-objective production scheduling problems using a new approach. Prod Syst Inf Eng 5:81–94
14. Kulcsár Gy, Kulcsárné Forrai M (2013) Detailed production scheduling based on multi-objective search and simulation. Prod Syst Inf Eng 6:41–56

The Context Between the Shift of Average Demand and the Safety Stock of Purchased Parts

János Korponai, Ágota Bányainé Tóth and Béla Illés

Abstract In order to optimize the stockpile management costs, the suppliers of car manufacturers answer by building up a safety stock of different extents to avoid uncertainties arising from the fluctuation of demands and to minimize their impact. During the calculation of the safety stock in the case of both the periodic and the continuous review models, the relations proceed from the system of conditions, according to which the average level of forecasted demands does not change with the progress of time. In practice, however, we can see a certain shift, thus, during the definition of safety stocks and the order of purchased parts, the historical data can only be used by knowing the direction and the extent of the shift. During our analysis, we examine the impact of shifts of different directions and extents on the stock of purchased parts, and we build up a model that predicts the probability of the occurrence of a stock shortage of an unplanned extent and of an overstocking.

1 Introduction

Unlike in several industrial sectors, the automobile industry is characterized by the fact that in the supply chain the demands of suppliers in direct contact with the car manufacturer can be well planned. This means that car companies provide long-term forecasts for their suppliers, and the demands can change only to the agreed extent for a short term. These demands, however, do not reflect completely the market mechanisms; predictability and plannability are in many cases related to the capacity of the car manufacturers, thus the demands themselves imply stochastic

J. Korponai (✉) · Á.B. Tóth · B. Illés
University of Miskolc, Miskolc, Hungary
e-mail: janos.korponai@gmail.com

Á.B. Tóth
e-mail: altagota@uni-miskolc.hu

B. Illés
e-mail: altilles@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_26

features. The fluctuation of demands influences the supplier's production demand and, implicitly, the stocks of purchased parts [1–3].

The stock level of a company's purchased parts is the consequence of the joint action of several factors. It is influenced by the quantity and temporal scalability of each phase of the supply chain, the features of the continuity of production and supply, and to a great extent, by the frequency of purchases and the individual order quantities as well. The objective of the logistics management is to guarantee the stock level required for the adequate handling of production at the lowest possible level of costs [4–6].

2 The Analysis of Planning Reliability by Comparing the Planned and Actual Ordering for the Past Period

The calculation of the safety stock takes into account the fluctuation of previous demands, the forecasted demands and the stock replenishment time of a length agreed upon with the suppliers. In practice, however, we can see a certain shift, e.g. in case the customer demands are under-forecasted or if there is a significant deviation in the production process from the planned percentage of rejects. The result of these impacts is that the previous use and the future demands cannot be interpreted as a single continuous series of data, thus, during the definition of safety stocks and the order of purchased parts, the historical data can only be used by knowing the direction and the extent of the shift [7, 8].

By knowing the historical data, we can describe the observable regularities in the data line and we can forecast the future development of the phenomenon. The development of the future values of a time series is usually the result of the joint action of several factors, thus the possible outputs can be defined with random variables, from which only one value will be realized.

The fact that the demands can be well forecasted is characteristic of the automobile industry. On the short term, major changes take place only at certain actors of the supply chain due to unexpected disturbances. However, it is a common phenomenon that the forecasted demand continuously differs from the actual utilization. This reverberates through the production demand towards the supplier through the issued order and the forecast. The purchase demand for purchased parts is influenced not only by the deviation derivable from the customer; it is also influenced by the deviations observable in the production processes within the company. These external and internal effects—e.g. the deviation in one direction of the percentage of production scrapes, or the conservative order forecast received from the customer—can jointly generate a deviation between the demand forecasted toward the supplier and the actual order. The deviation can be bidirectional, and underplanning and overplanning of a certain extent can occur as well.

We can see in practice that the forecasted customer demand in relation to time can be divided into three successive periods:

- Short-term scheduling
- Medium-term forecast
- Long-term forecast [9, 10].

In practice, the question could arise as to which periods and which forecast related to which date to compare when comparing the planned and actual utilization, since several forecasts were valid for the same period. Accordingly, the question is what to consider as being the basis of reference. On the basis of the fact that the tracing of the eventual change in demands after the placement of the order can happen with the rescheduling of the planned time of the arrival of goods and the modification of the ordered quantity, the focus of our analysis is the examination of the stock replenishment time between the placement of the order and the arrival.

First, we must define the relation between the forecast and the actual utilization, thus the reliability of the forecast. By drawing up for the given article the demand forecasted for the stock replenishment time for the different order periods, and by assigning the actual utilization for the same time, we have a developing pattern (Fig. 1). By considering a longer period, this pattern can reflect the constant underplanning and overplanning, and the quasi-stationary and non-stationary fluctuations. The ad hoc changes of signs can be smoothened out by comparing the cumulated value lines of the planned and the actual utilization demand, the real tendencies can be represented, and the rate of under- and overplanned periods can be quantified. Our analysis looks for the solution guaranteeing a safe planning when the actual utilization systematically exceeds the forecasted demand.

When the demand is overplanned, or the actual utilization decreases after the definition of the purchase demand of purchased parts and the placement of the order, a larger stock than the actual demand is purchased. In case the order quantity is not reduced and the planned arrival of goods is not rescheduled parallel to the decrease of demand after the placement of the order, the stock level will grow at the end of the period, and the production supply is guaranteed. The subject of our

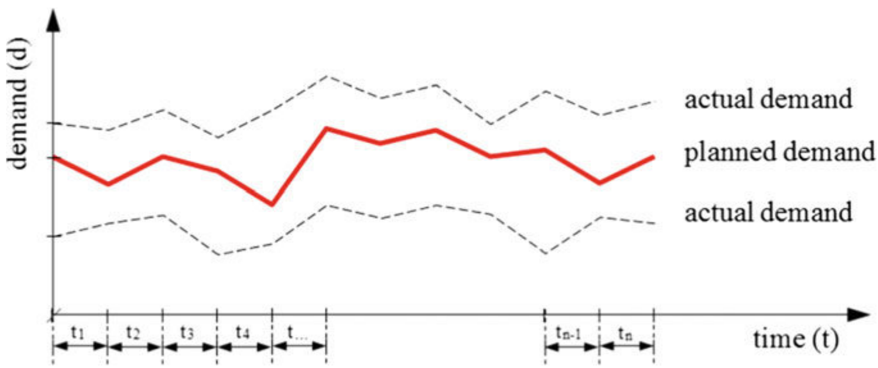


Fig. 1 Demand changes, development of planned and actual demand in relation to time [personal editing]

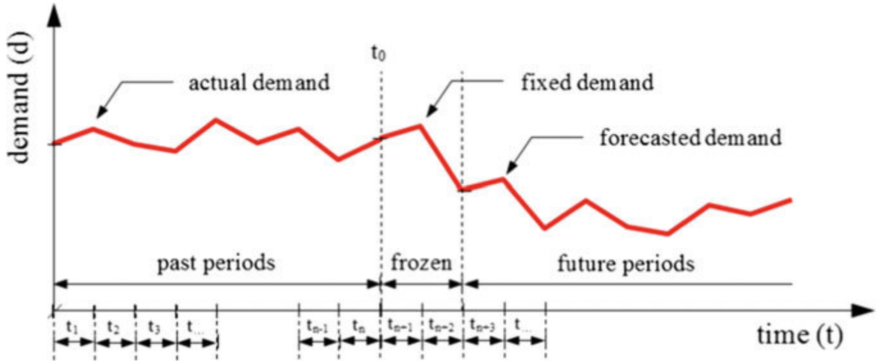


Fig. 2 Past actual utilization and the development of future expected demand in relation to time [personal editing]

analysis is the examination of the cases, when the satisfaction of production could be in danger, thus the case of overplanned demand is not part of the analysis.

After comparing the planned and actual utilization for the past periods, we shall proceed with the representation of past and future data lines. Figure 2 shows the actual utilization from the past and the expected demand for the future; the planned utilization can be represented as two further sections. In the period closer in time, the utilization demand can be regarded as fixed and frozen, and the demand after that can vary to different extents. The length of the frozen period can vary in practice, however, for the analysis, we separated two versions. In the first case, the frozen period is longer than or identical with the stock replenishment time, which implies that we can completely calculate with deterministic values when defining the required quantity of purchased parts, thus the stock coverage is guaranteed throughout the whole stockpiling period. In case the frozen period is shorter than the stock replenishment time, the stockpiling mechanism carries uncertainty. The cause of this uncertainty is that in a part of the planning period, thus the period after the frozen period, the demand is probably underplanned. In case there is an underplanning, the stocks would drop to zero before the end of the stockpiling period or before the arrival of the ordered quantity as a consequence of a demand increase to a realistic level.

3 Definition of the Trend Function that Gives the Most Accurate Estimate with an Alignment Analysis

After presenting several past stockpiling periods and the planned demands, quasi-stationary and non-stationary fluctuations can occur. In order to define the extent of the underplanning to calculate with, comparable values must be obtained from the time series. The average and the standard deviation can be the simplest features of the time series, thus the average value and the standard deviation of the actual utilization of the examined past period as well as the average and standard deviation of the expected demand can be defined. However, these features do not carry any information about the trends within these time series. Continuous increase, decrease and hidden trends can occur during a long period. Accordingly, if we characterize a time series only with the average and the standard deviation, we may draw wrong conclusions [11].

For a more exact description, we must define the basic trend of the time series. The aim of trend calculation is the presentation of the basic trend by evening the time series, and by eliminating the periodical fluctuations and the accidental factors. We can even the time series with the method of the moving average, analytic trend calculation and graphic representation. Since the examined time series becomes shorter as a result of the moving average method, and the graphic representation for more articles is time-consuming and provides only approximate information, the analytic trend calculation has become the focus of our analysis.

As a first step of trend calculation, we must decide the type of the function, we wish to use to estimate the trend of the examined time series. The regression function describing the development tendency of the time series the best can be defined as a result of graphic representation. However, in case of more articles, this is not a viable option in practice, thus we analyse the basic trend for past data lines with different functions instead of one function type, and we define the type of the function with the best estimate by deriving from the result of the goodness of fit used for the different functions. The most common functions for trend calculation are the linear, the exponential and the parabolic trend function. We will present the first two function types.

The equation of the linear trend function can be defined by the following formula [12–15]:

$$\hat{y}_{t(\lambda)} = \lambda_0 + \lambda_1 \cdot t, \tag{1}$$

where

- $\hat{y}_{t(\lambda)}$ the trend value of the t-th element in case of a linear trend function,
- λ_0 the value of the basic trend at a t_0 time in case of a linear trend function,
- λ_1 the slope of the linear trend function, thus the extent of the average growth for a time unit per one period,
- t the series of periods at equal distances from each other expressing the time variable.

The following relations can express the standard equations necessary for the initial value of the basic trend and the slope of the trend function [13]:

$$\sum_{t=1}^{n} y_t = n \cdot \lambda_0 + \lambda_1 \cdot \sum_{t=1}^{n} t, \tag{2}$$

$$\sum_{t=1}^{n} t \cdot y_t = \lambda_0 \cdot \sum_{t=1}^{n} t + \lambda_1 \cdot \sum_{t=1}^{n} t^2, \tag{3}$$

where

y_t the value belonging to the t-th period of the analyzed series,
n the number of analyzed periods.

Since from the factors of the system of equations we know the values $y_{1(\lambda)}, y_{2(\lambda)}, \ldots, y_{n(\lambda)}$, and the dates t and the value of n, the unknowns λ_0 and λ_1 can be defined by solving the system of equations as a first degree system of equation with one unknown.

The equation of the exponential trend function can be defined by the following formula [12–15]:

$$\hat{y}_{t(\varepsilon)} = \varepsilon_0 \cdot \varepsilon_1^t \tag{4}$$

where

$\hat{y}_{t(\varepsilon)}$ the trend value of the t-th element in case of an exponential trend function,
ε_0 the value of the basic trend at a t_0 time in case of an exponential trend function,
ε_1^t the average relative change per time unit of the exponential trend function.

In case of an exponential trend function, the normal equations can be expressed with the following relations [13]:

$$\sum_{t=1}^{n} \log y_t = n \cdot \log \varepsilon_0 + \log \varepsilon_1 \cdot \sum_{t=1}^{n} t, \tag{5}$$

$$\sum_{t=1}^{n} t \cdot \log y_t = \log \varepsilon_0 \cdot \sum_{t=1}^{n} t + \log \varepsilon_1 \cdot \sum_{t=1}^{n} t^2. \tag{6}$$

Unknowns ε_0 and ε_1 can be defined by solving the system of equations as a first degree system of equation with two unknowns.

By carrying out the trend calculated separately for the times series of past actual utilizations and future expected demands, we must define the most accurate regression function describing the alteration of the demand of the examined article in relation to time. The demand calculable as previously recorded and due during the frozen period can be regarded as actual utilization, thus the time series can be extended until the end of the frozen period during the quantification of the past data (Fig. 3).

When defining the deviations from the actual trend, the residual standard deviation and the relative residual standard deviation must be quantified for both

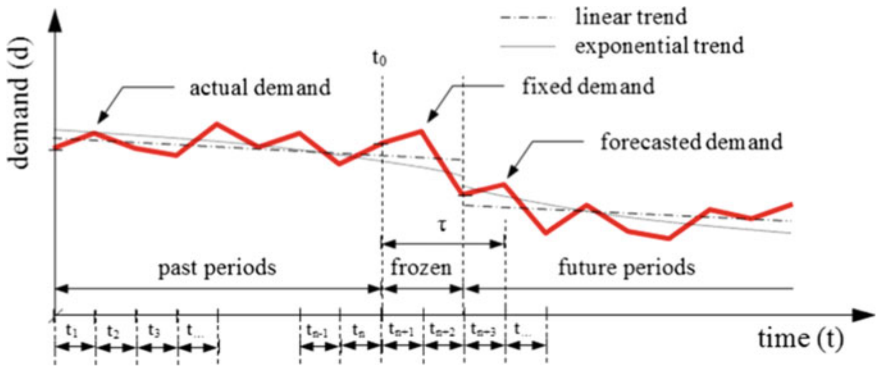


Fig. 3 Past actual utilization and the development of future expected demand by presenting the linear and exponential trends in relation to time [personal editing]

regression functions. The residual standard deviation shows the quadratic sum of deviations from the values according to the trend of the time series values [12–15]:

$$s_{e(\lambda)} = \sqrt{\frac{\sum_{t=1}^{n} \left(y_t - \hat{y}_{t(\lambda)}\right)^2}{n}}, \tag{7}$$

where

$s_{e(\lambda)}$ value of residual standard deviation in case of a linear trend function.

The relative residual standard deviation shows the deviation of the value estimated with the defined trend function from the actual value. By comparing indicators quantified by applying several types of functions, we can decide which trend function is the most accurate for the examined time series [13]:

$$V_{e(\lambda)} = \frac{s_{e(\lambda)}}{\bar{y}_t}, \tag{8}$$

where

$V_{e(\lambda)}$ value of the relative residual standard deviation in case of a linear trend function,

$\bar{y}_t$ mean value of the examined series.

The value of the residual standard deviation in case of an exponential trend function can be defined with the following relation [12–15]:

$$s_{e(\varepsilon)} = \sqrt{\frac{\sum_{t=1}^{n} \left(y_t - \hat{y}_{t(\varepsilon)}\right)^2}{n}}, \tag{9}$$

where

$s_{e(\varepsilon)}$ value of residual standard deviation in case of an exponential trend function.

The value of the relative residual standard deviation in case of an exponential trend function can be defined as [13]:

$$V_{e(\varepsilon)} = \frac{s_{e(\varepsilon)}}{\bar{y}_t}, \tag{10}$$

where

$V_{e(\varepsilon)}$ value of the relative residual standard deviation in case of an exponential trend function.

The function type aligning better to the examined time series is the one, where the residual standard deviation is smaller, thus where the quadratic mean of the deviations of the time series values from the trend values shows a more favourable value. Similarly, the alignment with the relative residual standard deviation can be examined as well, in case of which the more accurate trend function is the one showing a lower value [16, 17].

4 Forecasting the Future Expected Demand Based on the Past Actual Utilization, the Definition of Safety Stocks

After the function type was selected as a result of the alignment analysis of the time series and the actual trend, the trend value and the equation system of the selected function can be given to every single period of the time series. During our analysis, we aim to define the extent of underplanning, thus we wish to quantify the demand that is not predicted at the time of the order of purchased parts, but which is expected to occur in the knowledge of past data. This requires the anticipation of the trend calculated from the actual utilization demand, thus the definition of the expected data line that assume the future continuation of the discovered regularities; and during the extrapolation, we calculate the expectable trend values for the future periods as well (Fig. 4).

We can outline the extent of the underplanning in the knowledge of the projected trend values and the trend values calculated from the planned demand. The trend values resulting from extrapolation can be compared with the actual planned demand of the different periods, and with the trend values calculated for the planned demands. Figure 4 shows both versions, Δd shows the comparison of trends, $\Delta d_{t_{n+m}}$ shows the difference of trend values belonging to a t_{n+m} time period.

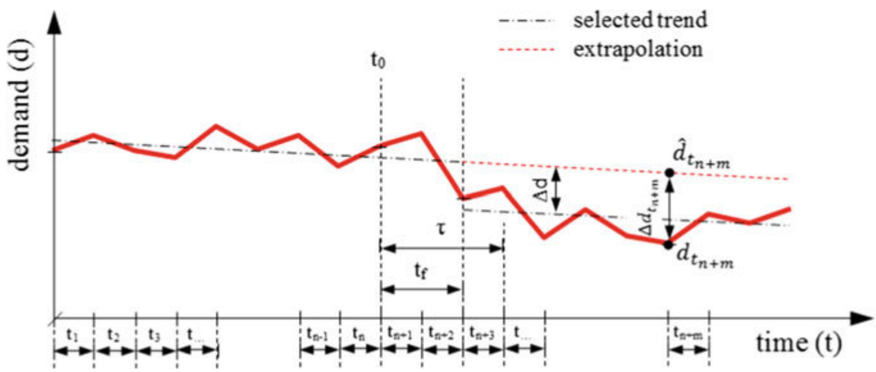


Fig. 4 Forecasting the future expected demand based on the past actual utilization by applying the most accurately aligning trend function [personal editing]

The formulae can be substituted with concrete demands:

$$y_t = d_t \quad \text{and} \quad \hat{y}_t = \hat{d}_t, \tag{11}$$

where

d_t the demand belonging to the t-th period,
$\hat{d}_t$ the trend value of the demand for the t-th period in case of the chosen trend function.

We can define the value of the underplanned demand for the τ stock replenishment time by considering the length of the frozen period (t_f) and the extent of the related deterministic demand (d_{t_f}):

$$\Delta d_\tau = \hat{d}_\tau - d_{t_f} - d_{\tau - t_f}, \tag{12}$$

where

Δd_τ demand for the stock replenishment time,
$\hat{d}_\tau$ the value of the demand for the stock replenishment time forecasted with the trend function,
d_{t_f} the value of the demand for the frozen period,
$d_{\tau - t_f}$ the planned value of the demand belonging to the stock replenishment time period exceeding the frozen period.

There are two possible directions to reduce the risk caused by underplanning:

- Ordering an increased quantity in the knowledge of the forecasted trend
- Rescheduling the planned quantity for an earlier arrival date.

Given that the conclusion from a forecast carries a certain extent of uncertainty, it can happen that the underplanned demand actually remains on a low level, thus a potentially increased order quantity can lead to a durable increase of the stocks' level or to the formation of obsolete stocks. Therefore, we do not recommend this version in practice. However, we must underline its advantage, since the arrival date of the ordered quantity does not change and the delivery is more plannable.

On the other hand, the risk defined by considering the trend values can be avoided by rescheduling the arrival date of the planned quantity. In case the forecasted demand is realized at a lower level, the rescheduling of the arrival date of the ordered quantity to an earlier date will only cause a temporary increase of the stock level, thus the stock is available at an earlier date due to the arrival of the originally planned quantity. The advantage of rescheduling is that the forecasted demand was not increased, thus the increase of the durable stock level and the formation of obsolete stocks can be avoided. A disadvantage of the model is that an arrival of goods scheduled for an earlier date can mean the dispensing of the supply chain on the one hand, and an auxiliary delivery demand that differs from the scheduled delivery frequency and the generation of additional charges on the other hand.

5 Conclusions

The risk caused by underplanning can lead to supply problems in the satisfaction of production. In practice, the covering of the production demand for the stock can be guaranteed through the following steps:

- The analysis of planning reliability by comparing the planned and actual ordering for the past period.
- In case of insecure planning, the presentation of the past and planned demands of the examined article, as well as of the frozen period and the stock replenishment time.
- The definition of trend values related to the time series of past and planned demand by applying different regression functions.
- Definition of the trend function that gives the most accurate estimate with an alignment analysis.
- With the selected trend function, forecasting the past time series for a future period.
- Quantification of the quantity between the planned demand and the demand forecasted with extrapolation regarding the stock replenishment time, by considering the utilization demand of the frozen period.
- Rescheduling the order date and the arrival date in order for the stock to be able to cover the demand to be realized at an increased level and on time.

Beside linear and exponential trend functions, other function types can be used to analyze the demands. Furthermore, the model can be further refined by considering seasonal fluctuations and accidental factors, so that the impacts of planning uncertainties can be further moderated.

Acknowledgements The described study was carried out as part of the EFOP-3.6.1-16-00011 “Younger and Renewing University – Innovative Knowledge City – institutional development of the University of Miskolc aiming at intelligent specialisation” project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund.

References

1. Krampe H, Lucke H-J, Schenk M (2012) Basics of logistics (in German). Huss-Verlag GmbH, München
2. Kummer S, Grün O, Jammernegg W (2009) Basics of purchasing, production, logistics (in German). Pearson Studium, München
3. Vörös J (2010) Production and service management (in Hungarian). Akadémia Kiadó, Budapest
4. Chopra S, Meindl P (2007) Supply chain management: strategy, planning, and operation. Pearson Prentice-Hall Publishers, New York
5. Juxiang W, Lijuan R, Changju C, Guozheng L (2006) Influence of different forecasting modes on automobile manufacturing supply chain. In: Proceedings of the 2006 IEEE international conference on service operations, logistics and informatics, SOLI 2006, pp 775–779
6. Kása R, Gubán Á (2015) Business process amelioration methods, techniques, and their service orientation: a review of literature. In: Vastag G (ed) Research in the decision sciences for global business: best papers from the 2013 annual conference of the European Decision Sciences Institute. Pearson Education Limited, New Jersey, pp 219–238
7. Yamazaki T, Shida K, Kanazawa T (2016) An approach to establishing a method for calculating inventory. Int J Prod Res 54(8):2320–2331
8. van der Veen B (1986) Safety stocks and the order quantity that leads to the minimal stock. Eur J Oper Res 27(1):34–49
9. Koltai T (2009) Production Management (in Hungarian). Typotex Kiadó, Budapest
10. Stock JR, Lambert DM (2001) Strategic logistics management. McGraw-Hill Higher Education, Boston
11. Cselényi J, Illés B (2006) Planning and controlling of material flow systems (in Hungarian: Anyagáramlási rendszerek tervezése és irányítása). Miskolci Egyetemi Kiadó, Miskolc
12. Shao J (2005) Mathematical statistics: exercises and solutions. Springer Science + Business Media, Inc.
13. Korpás A (ed) (1997) Elementary statistics II (in Hungarian: Általános statisztika II.). Nemzeti Tankönyvkiadó, Budapest
14. Browder A (1996) Mathematical analysis—an introduction. Springer, Berlin
15. Triola MF (2013) Elementary statistics. Pearson Education
16. Triola MF (2015) Essentials of statistics. Pearson Education Limited, Essex
17. Devore JL (2016) Probability and statistics for engineering and the sciences. Cengage Learning, Boston

An Overview of Autonomous Intelligent Vehicle Systems

Daniel Tokody, Imre János Mezei and György Schuster

Abstract Vehicles, whose functions are enriched with attributes to increase safety, environmental awareness, effectiveness, comfort level and prestige, so that they can play a key role in creating optimal mobility, are now being invented, planned and manufactured for general use. Throughout the full spectrum of transport, vehicles will soon exempt people from the routine of driving. If people do not need to drive their cars, will their driving skills deteriorate or will they entirely fail to develop this skill later on? Is this threatening us in the near future? What are the latest research results and regulations on autonomous vehicles? What are the actual advantages of vehicle automation? We are trying to find the answers to these questions in our article, while analysing and systematizing information from the national and international literature on the development of intelligent vehicles by examining the interaction between various ground transport vehicles, and the related developments on the subject. Our goal is to create automatic intelligent vehicle systems, within the concept of intelligent infrastructures and smart cities. The paper provides an FMEA analysis of intelligent vehicles. To decrease the explored deficiencies in the present system, applicable proposals are formulated about development areas, such as forming a communication between vehicular traffic and railed vehicles. We feel that such developments are important steps in increasing traffic safety, and we regard them as elements of intelligent transport.

Keywords Automation · Self-driving cars · Levels of automation · Automatic train operation · Grades of automation · Intelligent vehicles · Smart sustainable city · Intelligent infrastructure · V2X

D. Tokody (✉) · I.J. Mezei · G. Schuster
Doctoral School on Safety and Security Sciences, Óbuda University, Budapest, Hungary
e-mail: tokody.daniel@mav.hu

I.J. Mezei
e-mail: mezei.janos.imre@gmail.com

G. Schuster
e-mail: schuster.gyorgy@kvk.uni-obuda.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_27

1 Introduction

According to the 2014 traffic accident statistics, the primary causes of road accidents are the drivers' faults (Drivers' fault total 14 616). This means that 14 161 out of 15 847 accidents are caused by the drivers, which makes 92% of all accidents arise directly from human error. In less than 0.5% of all cases, the primary cause is the vehicle's technical problem [1].

The facts are the same regarding the accidents between road vehicles, locomotives and rolling stock. There were 19 collisions of road vehicles and railway vehicles, suburban trains and trams in 2014 (Collisions of public road vehicles with rail vehicles) [1].

Based on the data found in the statistics, we can increase traffic safety, if we aim to develop a system that supports the driver in the transport process. Such systems could be driver assistant solutions, different awareness and attention maintenance tools or even a collision avoiding emergency brake system. The development of such systems is already in progress, but in 2010 the importance of developing an intelligent transport system was stated by a European Directive.

"Intelligent transport systems are such advanced applications that aim to offer innovative services to a variety of transport modes and in relation to traffic management, without the actual incarnation of intelligence, as well as enabling different users to get better information on how to use the transportation network safer, more coordinated and in a smart way" [2]. This definition has changed in its essential elements in the past five years. According to the latest research, in many ways we need to use artificial intelligence in order to develop fully autonomous cars. Therefore, it is not entirely about creating a fully integrated co-operative system. We think that in the future the fully autonomous transport system will create a living, intelligent technosphere in the whole world.

Implementing cooperation between the various participants in traffic is one of the main elements of the intelligent system. This was the pioneering concept of the cooperative intelligent transport system.

In our view, ground transport development does not follow a single direction. This means that the development of road and rail transport systems do not have common sections. It is also uncommon to put a well working practice (for example communication protocols, management systems, etc.) from one transport sector to the other. Therefore a cooperative collaboration—best case scenario—can only be realised through an interface that provides simplified connection.

The road–rail level crossing point can be a significant threat to vehicles. Still researching on the subject does not cover common system development, while it would be critical regarding the cooperative intelligent transport system's traffic safety. There is no better proof for the separation of disciplines than the standards that only deal with communication between road vehicles and road infrastructure. They do not include the possibility of communicating with other transport operators. In a few cases, though, they mention tangentially the importance of developing the communication between different vehicles. Regarding the potential global goals and long term solutions, they do not really bring any solution in the topic of road and rail transport. With our article, we aim to show the possibility of achieving a complex, multi-vehicle operation in the intelligent traffic systems, not forgetting the two main participants in ground transport, that are the road and rail vehicles.

2 The Situation of Intelligent Vehicles in Europe

Today's research in the field of robotics is dealing with systems that are capable of autonomous operation, are equipped with intelligent sensors and are able to adapt to environmental changes. Autonomous intelligent vehicles are basically robots like these systems, and while developing them, we need to use the results of many different disciplines. Robotics is a discipline that includes other disciplines, and it is the most advanced field of automation. In other words, vehicle automation is being in progress with great efforts [3].

Today's modern robotic systems, such as autonomous cars are capable of progressing in convoy or even achieving an objective on their own. Driverless, autonomous vehicles can be categorized into three groups: UGV (Unmanned Ground Vehicles), UMV (Unmanned Marine Vehicles) and UAV (Unmanned Aerial Vehicles) [3]. In our article we are dealing with the possibilities of creating a complex transport system consisting of UGV vehicles.

2.1 Autonomous Unmanned Ground Vehicle

In the case of ground-level vehicles, our research involves autonomous automobiles and fixed-track vehicles, capable of automatic operation.

Among fixed-track vehicles, not only heavy railways, subways, trams, suburban railways and trolley buses should be considered.

In Sweden, Scania has just recently introduced (22nd June 2016) their first true parallel hybrid trucks (using electricity and biofuel), which can be powered on the motorway by an overhead line [4]. Naturally, development has been ongoing for a few years.

Self-driving cars—Road Transport

In the United States, as of now, five states have introduced legislation and regulatory actions regarding autonomous transport and 15 more have pending legislative regulations concerning the technology, according to recent news.

In Hungary, on 19th May 2016, Günther Oettinger, the European Commissioner for Digital Economy and Society in the European Commission stated that Europe lags behind the US regarding software research and development. The strengthening of cooperation of the digital market, vehicle industry and information technology has to be facilitated by creating a unified digital policy in the fields of cyber security and new traffic regulations, and this has also become necessary with the development of autonomous vehicles. With his proposal, the development and production of autonomous vehicles in Europe could get closer to its competitors' achievements. According to plans, a test track will be built in Zalaegerszeg for testing autonomous vehicles [5].

Legislators of the European Union (http://eur-lex.europa.eu/) mention the term of "self-driving cars" in only three documents. All of the dates of these documents are of 2016. This clearly shows the situation of introducing autonomous road vehicles in Europe.

15th January 2016—the opinion of the European Economic and Social Committee on the document entitled "The effects of digitalisation on the service industry and employment within the framework of industrial restructuring" was issued.

"Recent years have seen major breakthroughs in the development of digital technologies. The new and extremely powerful capabilities of digital technologies allow for the automation of ever more tasks and occupations (e.g. the self-driving car, the internet of Things, Industry 4.0). Additionally, digital technologies are leading to drastic reductions in search and transaction costs and thus enabling the development of entirely new and highly scalable business models in services (e.g. online marketplaces and platforms including the so-called sharing economy, Uber, Airbnb). This is encouraging the digitalisation of business models and processes in a wide range of economic sectors. Some of them have already been addressed in previous EESC opinions (3)" [6].

19th April 2016—The announcement of the European Committee directed towards the European Parliament, Council, the European Economic and Social Committee and the Committee of the Regions, concerning the digitalisation of Europe's industry and the advantages of fully exploiting a unified digital market:

"Stakeholders expressed a need to examine the regulatory framework for digital innovations with a view to provide further clarity on the following: … Autonomously acting systems such as self-driving cars or drones pose a challenge to current safety and liability rules where a legal person is ultimately responsible. Legal implications of the roll-out of IoT are wider than the allocation of liability as recognised in the DSM strategy and also need to be addressed" [7].

25th May 2016—In the European Committee's document entitled "COMMISSION STAFF WORKING DOCUMENT Online Platforms Accompanying the document Communication on Online Platforms and the Digital Single Market":

"Furthermore, online platforms are continuously changing and developing in new directions. They are at the forefront of developments in the emerging platforms of the future: virtual reality, connected and self-driving cars, Internet of Things, Big Data and Artificial Intelligence" [8].

According to the documents above, it is noticeable that although the legislation is not final, it is kept on the agenda, and based on the results of the latest research, the digital evolution of vehicles is an important economic issue in Europe, which triggered a new industrial revolution in the world. The building of a society based on digital and cyber-physical systems is in progress.

Research of intelligent vehicles in the European Union

Autonomous vehicles use intelligent systems in order to become cost-efficiently operated. Their environmental impact is smaller than that of traditional vehicles, since for optimal fuel consumption during their run accelerating and decelerating profiles are used in accordance with traffic ratios [9].

Creating an active security system is a vital issue when it comes to self-driving vehicles. Examples of this would be the implementation of systems designed to avoid head-on collisions and reduce risks [10–12].

Regarding vehicle development, further significant roles are played by the systems, supervised by man, within which various driving-assistance systems are being developed. A kind of system could be one which can adapt to the driver's behaviour, maintaining travel speed and signalling possible collisions (cruise control and forward collision warning) [13].

Multiple high-volume cutting edge European projects are in progress. Such is the AdaptIVe project, which is a 25 million Euro budget project coordinated by the Volkswagen Group Research from 2014 to 2017. The goal of the research is to create a highly automated vehicle. Taking into account and utilising the human factor, a new type of cognitive integrated structure is formed by developing the functions (28) and applications (65) of intelligent vehicles. During the research, the handling of the systems' complexity is a crucial aspect [14].

Several organizations are currently working on the elaboration of intelligent transportation systems and the technological solutions related to them. This spring in Munich at the "Trends in Automotive Radar and Impact on System Architecture Workshop" convention, an expert in radiospectrum management, Thomas Weber at the European Communication Office drew attention to the following:

"Regulators consider other transportation applications should share the ITS frequencies such as Urban Rail Systems (based on an analysis in progress—looks like it could be possible)—also perhaps relevant for future ITS applications—such as the driverless vehicles or autonomous systems" [15].

"New ECC Decision (16)02 on rotorcraft anti-collision radars operating in 76–77 GHz approved for publication on 4 March 2016. Be aware that 76–77 GHz, includes automotive radars, infrastructure radars in the current regulation; it is also used at railway level crossings or airplanes during taxiing" [15].

This claim also leads to the point of our article, which is the possibility of interactions between railway vehicles and road vehicles at railway crossings.

Vehicle automation from conventional to intelligent

The first table shows the levels of vehicle automation, defined by the international SAE association. From level 0 to level 5, we gradually get from a vehicle driven by a person to a fully autonomous vehicle. With LoA 0, the driver controls the vehicle in terms of braking, accelerating, speed, change of direction, parking, maintaining distance and changing lanes. With LoA 1, a function-specific automation is implemented, which helps the driver, e.g. adapting speed-maintaining automation. With LoA 2, the vehicle driving procedures are partially automated, which further assists driving. With LoA3, specific automated systems are cooperating with each other. With LoA 4, the vehicle is capable of executing complex driving functions, such as parking itself. And LoA 5 means the total automation of the vehicle, in other words, every operating function of the vehicle is executed autonomously without the driver.

Considering the functional description of these levels, a few important systems from the possibilities of potential automation are listed here:

LoA 0 (driver only): Advanced Driver Assistance Systems (AEBS), Anti-lock Braking System (ABS), Electronic Stability Control (ESC), Forward Collision Warning (FCW), Lane Departure Warning (LDW) [14]. LoA 1 (assisted—Automated vehicle): Adaptive Cruise Control (ACC), Park Steer Assistance, Lane Keeping Assistance [14]. LoA 2 (partial automation—Automated vehicle): Traffic Jam Assistant (TJA), Parking assistant [14]. LoA 3 (conditional automation—Automated vehicle): Highway traffic jam system [14]. LoA 4 (high automation—Automated vehicle): Parking garage pilot [14]. LoA 5 (full automation—Autonomous vehicle): Robot Taxi [14] (Table 1).

Automatic Train Operation—Guided Transport

With the automation of the operation of vehicles [17], concerning fixed-track transportation, significant results have been achieved, e.g. in automatic train driving, based on iterative learning [18]. In the line of the system's automatic functions manipulating the vehicle's speed (using a speed profile) [19] is technologically an already solved problem, as well as vehicle starting, stopping, standby [20], the traffic of vehicles based on a genetic algorithm [21], collision and catch-up protection [22] and the prevention of opposite traffic situations. On the other hand,

Table 1 Levels of motor vehicle automation adapted from SAE Standard J3016 [16]

Levels of automation	Type of motor vehicle operation	Steering, acceleration, deceleration	Monitoring driving environment	Fallback performance of dynamic driving task	System capability (driving modes)
LoA 0	No automated, traditional car	Driver	Driver	Driver	n/a
LoA 1	Drive assistance	Driver/Automatic	Driver	Driver	Some driving modes
LoA 2	Partial automation	Automatic	Driver	Driver	Some driving modes
LoA 3	Conditional automation	Automatic	Automatic	Driver	Some driving modes
LoA 4	High automation	Automatic	Automatic	Automatic	Some driving modes
LoA 5	Full automation	Automatic	Automatic	Automatic	All driving modes

heavy railway automation is lagging behind its public road counterpart, since, most probably, the application of newer technologies takes significantly more time, because of the infrastructural costs and the calculated lifespan of the vehicles and the infrastructure.

The average calculated lifespan of railway vehicles is 30 years or 10 million km, of signalling and interlocking systems it is 15–50 years, while structures and platforms must be usable for up to 100 years. In the case of automobiles, there is a huge dispersion of lifespan periods from 300 000 up to 1 million km, and for trucks and lorries, lifespans of 1.5–4 million km are not unusual [23–25].

There are roughly 50 (increasing) automated subway lines in the world, including one in Budapest, Hungary. With the help of automatic operation, timetables can reach an almost 100% accuracy. Vehicles can automatically maintain departure and arrival times stated in the timetable, using different driving characteristics. If a bigger crowd at a station would withhold the vehicle, it could work off the delay and affect the following vehicles' traffic itself within certain limits. Automatic operation significantly increases the safety and energy-efficiency of vehicle traffic. The vehicle can achieve this by using automatic speed profiles, exceptionally accurate positioning and automatic breaking curve calculation. The position of the vehicle is calculated by an odometer, fixed balises and using the map of the infrastructure. The vehicles are controlled by the instructions of the interlocking system with the help of the CBTC system (Communications-based train control). Traffic operators supervise the system through ATS (Automatic Train Supervision), but the operation is fully automatic [26, 27].

IEC 62290-1 standard defines the main concepts of Urban Guided Transport Management and Command/Control Systems. This standard includes the definition of the levels of automation concerning fixed-track transport. These levels are shown in the second table. GoA 1 (Manual Protected Operation, MPO: Automatic Train Protection (ATP) with driver) ensures in a non-automatic railway operation that a safe track path and distance between vehicles are guaranteed, and the vehicles' partial supervision of speed is solved. GoA 2 (Semi-automated Operation Mode, STO: ATP and Automatic Train Operation (ATO) with driver), in addition to the previous level, is in total control of the train's speed, braking and accelerating. GoA 3 (Driverless Train Operation, DTO) adds the ability to detect and avoid obstacles and people on the railway. GoA 4 (Unattended Train Operation, UTO) makes a fully automated operation possible without the need of a train attendant. This is the operational level in which the Nr.4 subway line in Budapest operates. This level is currently the most advanced with regard to both functionality and safety. Full automation means that the safe traffic of trains is guaranteed, the driving of the train is automated in accordance with the timetable and there is an active track management system detecting and avoiding objects and people falling in the path of the train. Moreover, supervision of the transport of people, automated closing of the passenger doors and the guarantee of safe start-up conditions are all ensured [28, 29] (Table 2).

Automation considerably improves the competitiveness of the railway sector. The development of heavy railway automation and the installation of systems, like

Table 2 Grades of automation by UITP [30]

Grades of automation	Type of train operation	Setting train in motion	Stopping train	Door closure	Operation in event of disruption
GoA 1	ATP with driver	Driver	Driver	Driver	Driver
GoA 2	ATP and ATO with driver	Automatic	Automatic	Driver	Driver
GoA 3	Driverless	Automatic	Automatic	Train attendant	Train attendant
GoA 4	UTO	Automatic	Automatic	Automatic	Automatic

ETCS (European Train Control System), ERTMS (European Rail Traffic Management System), GSM-R is an important step in the sector. At the same time, establishing ETCS Level 4 is inevitable, in which trains (not only subway lines) can operate automatically without a driver. Fixed-track, ground-level transport could be a good alternative to road traffic, when you consider the differences in capacity in both passenger and cargo transport, or the possibility of energy-saving provided by automated operation [31].

3 Technological Overview of the Development of Intelligent Vehicle Systems

3.1 Definitions/Terminology of Intelligent Vehicle Technology

Co-operative intelligent transport system (C-ITS): "In the previous section a new concept has emerged, namely the concept of 'Intelligent Transport Systems' (ITS). It refers to the endeavour to develop the integrated operation of various transport structures by applying the research results of interdisciplinary fields, for example by using the infocommunication technologies. Logistic hubs are also in the focus of this endeavour. Main aims of Intelligent Transport Systems are to find more environmentally friendly methods of transport, to implement an efficient transport system and to enhance transportation safety" [32].

Intelligent infrastructure: "Intelligent infrastructure can be defined as an integrated system which includes the complete traditional intra- and interurban infrastructures, collects data of them (with the help of sensors), and, by evaluating these data, improves the efficiency and safety of operation. It also optimizes and maintains the operation of the city, and supports environmental protection efforts. Based on the collected data, this system is able to provide help for humans to prevent accidents. This can be achieved by carrying out an early diagnosis of the problems in the system based on the evaluation of the Big Data collected and analysed about the system. Furthermore, the ‘learning ability’ of the system helps to recognise such schemes which can lead to abnormal operation. The intelligent infrastructure provides the basis for the smart city. The intelligent infrastructure is not limited to the city centre only. This system also connects different smart cities” [33].

Smart sustainable city: “A smart sustainable city (SSC) is an innovative city that uses information and communication technologies (ICTs) and other means to improve quality of life, efficiency of urban operation and services, and competitiveness, while ensuring that it meets the needs of present and future generations with respect to economic, social and environmental aspects” [34].

Smart mobility: “The concept of mobility is used for the moving and travelling of people.” “Smart mobility is one of the main characteristics of the smart city.” “Mobility: Safe and green mobility is possible because of intelligent vehicles and coordinated traffic management systems with the help of distributed real-time situation awareness and solution finding” [35].

Intelligent vehicle: A system of robotic applications to collect information on the position, kinematics and dynamics of the vehicle, the state of the environment and the state of the driver and the passenger, to assess such information and make decisions based on it. It is capable of duplex communication with roadside infrastructure and other vehicles, to use digital map applications and satellite positioning systems, it has an active internet connection and its own physical address [36].

Self-driving car (Autonomous car): A vehicle which is able to perform the driving tasks without the intervention of the driver, with a high degree of safety, to recognise any obstacles in its environment, and to stop before or go round such obstacles. It can communicate with the surrounding infrastructure and with various other vehicles. It has internet access and route design capable of online change management. It has special driving patterns and ensures its energy supply in an autonomous way. It can park autonomously [37–39].

V2V communication (VANET—Vehicular ad hoc network): It is an ad hoc mesh network-based wireless data transmission method for inter-vehicle communication. Its purpose is to enable vehicles to share information about their position, speed, direction or about any dangers in traffic, based on which they can take the necessary preventive measures, i.e. they can slow down or stop in time. It is one of the dedicated short-range communications technologies [40–43].

V2I and I2V communication: It ensures information share between the infrastructure and the vehicles approaching a crossroads. It can also provide data for a parking vehicle to start its route, or send speed instructions to vehicles. It can also be used for communication between railway level crossings and vehicles [44]. Further applications: “Red Light Violation Warning, Curve Speed Warning, Stop Sign Gap Assist, Reduced Speed Zone Warning, Spot Weather Information Warning, Stop Sign Violation Warning, Railroad Crossing Violation Warning, and Oversize Vehicle Warning” [44].

Vehicle-to-Everything (V2X): A system including several communication methods: Vehicle-to-Infrastructure (V2I), Vehicle-to-vehicle (V2V), Vehicle-to-Pedestrian (V2P), Vehicle-to-device (V2D), Vehicle-to-grid (V2G) [45].

GSM-R communication: The communication system of fixed-track vehicles for infrastructure-to-vehicle communication. Global System for Mobile Communications—Railway [46].

IoT based vehicle technology (vehicle-to-cloud communication, cloud technology for vehicles system): The technology of the Internet of Things connects various devices in the virtual space. It can transfer a large amount of data to extra-vehicular applications. It takes advantage of cloud technology to connect vehicles to the infrastructure, other vehicles, intelligent buildings and the smart grid for a more efficient resource management [47]. It can provide the basis for a toll system [48].

Smart grid and electric vehicles: The growing number of electric vehicles causes a considerable increase in electric energy consumption in the world. Meanwhile, some research proves that, besides their consumption, the energy producing and storing capacity of vehicles can also change the characteristics and operating philosophy of the supply network in a considerable way. It is based on the fact that the parking and charging time of these vehicles are certainly longer than their daily running time [49–51].

Network of everything, networked society: "Cities in themselves can also be described as networks, or the aggregate of systems and subsystems. The individual elements comprise a continuously changing artificial system, which functions like a living organism, and its technological complexity has significantly increased over the last decades" [52]. The network elements are the intelligent vehicles too.

Cyber-physical systems (CPS): "Cyber-Physical Systems are such systems which connect the virtual world and the physical world with the help of embedded systems and infocommunication systems." It is a key element of vehicle automation, facilitating the digitalisation of vehicles [52].

Micro-electromechanical systems for passenger safety in automotive vehicles: As a result of the technological revolution of the 21st century, the size of the devices used in robotic technology has been reduced considerably. The use of silicon-based microelectronic and michromechanic constructions in the vehicle industry has now become everyday practice. Micro-sized electro-mechanical devices are capable of detection, control, intervention and influence at a micro level. MEMS technology is used in the following systems: automotive airbag sensors, internal navigation sensors, air conditioning compressor sensors, brake force sensors and suspension control accelerometers, fuel level and vapour pressure sensors, "Intelligent" tyres [53].

Multi-agent system (MAS): Its sensor agents are able to detect the environment, while its intervention agents are able to change it. Multi-agent systems can model the complex processes of nature, as its individual agents act in co-operation with each other in a dynamically changing environment, while performing their specific tasks in the system and creating common knowledge through interactive information-based communication [54, 55].

Inter-Vehicle Communication (IVC): "Intervehicle communication (IVC) networks, a subclass of mobile ad hoc networks (MANETs), have no fixed infrastructure and instead rely on the nodes themselves to provide network functionality. However, due to mobility constraints, driver behaviour, and high mobility, IVC networks exhibit characteristics that are dramatically different from many generic MANETs. This paper elicits these differences through simulations and mathematical models and then explores the impact of the differences on the IVC communication architecture, including important security implications" [56].

Automated Transport Systems: "An Automated Transport System (ATS) is an innovative holistic mobility concept, where all its different elements (i.e. vehicle, travellers, public transport, infrastructure, operations and control) are capable of self-organizing and operating at an "automated" manner, addressing in real time the needs of all and each participant of a specific traffic scenario, applying different levels of automation and supporting all transport modes for both passenger and freight" [57], © Original definition, ICCS, 2013.

In addition to their definitions, Fig. 1 aims to show the relationships between the above concepts revealing the interconnections that can be important from the point of our research. As it has been illustrated in Fig. 1, a smart, sustainable city has six primary elements [58], and one of these elements is smarty mobility. Smart mobility requires the creation of an intelligent cyber-physical infrastructure, which involves the technologies of telemetrics, cybernetics, artificial intelligence, telecommunication and infocommunication. From a holistic aspect, the development of intelligent transport systems also includes the various methods of transport in the whole spectrum of transportation.

The autonomous systems which have already been developed are characterised by a communication-based structure. Autonomous intelligent cars communicate with each other through short-range communication (several metres) or long-range communication (50–80 m). These communication solutions are enough, for example, for the communication of a tram or a car. However, they are not suitable for the communication of a high-speed train. Further technological developments are needed to find a solution to this problem.

4 Failure Mode Effects Analysis

Definitions:

- Failure mode: the type of failure.
- Failure effect: Effects and consequences caused by the failure.
- Failure cause: Defining the exact cause of the failure.
- Failure extent: The gravity of the failure's consequences.
- Frequency: The frequency of the failure's occurrence.
- Recognisability: The efficiency with which the failure can be detected.
- Measures: Actions preventing the recurrence of a failure.

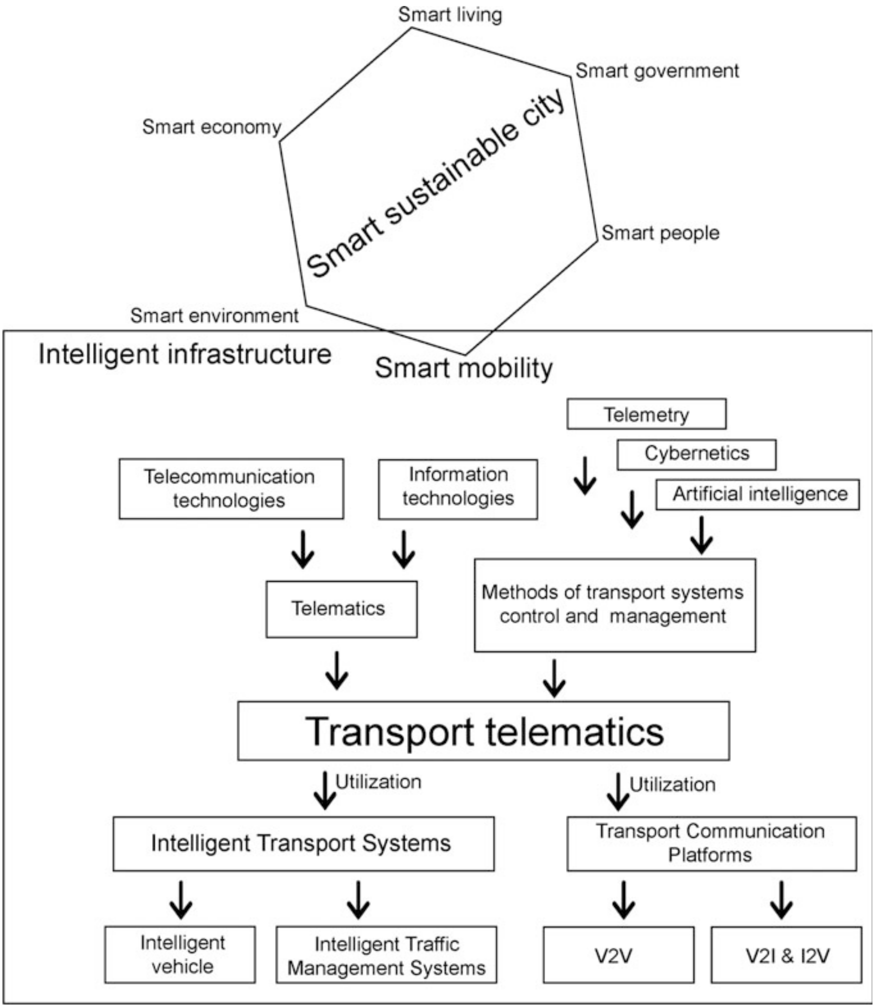


Fig. 1 The relationship between Intelligent Transportation Systems and intelligent vehicles (revised figure) [59]

The aim of Sect. 4 is to realise complex, intelligent vehicle systems, which can collectively and automatically handle vehicles equipped with rubber wheels (cars, trucks, motorcycle, etc.) or with iron wheels (heavy railway, subway, suburban railway, tram, etc.). These systems could still be called “utopian”, but a future goal is to make them reality. It can be clearly seen that this is an incredibly complex system, therefore, regarding the FMEA analysis, we mainly concentrated on the main hazards, system components and errors, and less on the errors of sub-systems (Table 3).

Table 3 Failure mode effects analysis

Number	Failure mode	Failure effect	Sever	Cause	Occur	Detect	RPN	Provisions
1	Turning off human intervention	Speeding irregular direction	10	Free provision over the system	10	10	1000	Restricted overwrite options
2	Hacker attack	Taking over the vehicle's control/incorrect data communication	10	System is not sufficiently encrypted	6	10	600	Testing
3	Faulty code	Speeding collision hazard	10	Human mistake	5	10	500	Multiple quality controls and tests
4	Weather anomalies	Change in braking distance data loss	10	–	5	10	500	Preparing the system for extreme weather
5	Wheeled vehicle signal loss	Positioning loss	10	Technical failure, shielding	4	9	360	In case of signal loss, change to manual control
6	Iron wheeled vehicle signal loss	Positioning loss	10	Technical failure, shielding	4	9	360	In case of signal loss, change to manual control
7	Distinguished vehicle handling	Distinguished vehicles are not integrated to the system, therefore increased accident risk	10	Incorrect handling of distinguished vehicle	7	6	420	Indicating distinguished vehicles to the system
8	Oversized vehicle handling	Oversized vehicle badly positioned delimitation thus creating a dangerous situation	10	Incorrect handling of oversized vehicle's specialties	4	8	320	Data handling of oversized vehicles

(continued)

Table 3 (continued)

Number	Failure mode	Failure effect	Sever	Cause	Occur	Detect	RPN	Provisions
9	Wheeled vehicle's delayed data sending/receiving	Incorrect positioning speed setting	10	Technical failure, shielding	5	9	450	Using a proper time limit, after which the data loss is considered as delayed sending/receiving
10	Iron wheeled vehicle's delayed data sending/receiving	Incorrect positioning speed setting	10	Technical failure, shielding	5	9	450	Using a proper time limit, after which the data loss is considered as delayed sending/receiving
11	Sending incorrect data to vehicles	Incorrect positioning and inaccurate speed setting	10	–	2	8	160	–
12	Wheeled vehicle sensor failure	Road conditions, traffic barriers and loss of sensing other vehicles	10	–	3	9	270	–
13	Iron wheeled vehicle sensor failure	Traffic barriers and loss of sensing other vehicles	10	–	3	9	270	–
14	Technical failure	Termination of the operation of sub-systems, system shutdown	10	Poorly developed physical parameters	1	10	100	Strict in-process quality control
15	Unexpected effects of passive components	Unexpected appearance of pedestrians, animals, coercion to emergency manoeuvre	8	Unexpected effects of passive components	5	10	400	–

5 Possibilities to Increase Safety by Using Intelligent Systems

Active vehicle safety systems, which are based on intelligent communication and used at road and railway level crossing and in railway infrastructure, have the following elements:

- Intelligent infrastructure agent (traffic environment, traffic lights, cameras, signallers, gates, switches and crossings, sensor networks, etc.).
- Intelligent railway vehicle agent.
- Intelligent road vehicle agent.

The agent is able to detect its environment and interact with it for its own interests through its interventive capacity. The autonomous behaviour of intelligent agents depends on their freedom to make decisions based their accumulated knowledge.

The agent-based conception of intelligent transport systems means that an ambient sensor system continuously monitors the state of the transport system by sensing the various transport infrastructure elements and transport system units (vehicles, participants in traffic). The types of sensors include radars, infra gates, opening sensors, pressure sensors, stretch sensors, acceleration sensors, etc. The agents are continuously capable of intervening, as without intervention or reaction none of the autonomous elements could perform their tasks. Not only the transport infrastructure itself, but the whole environment could be intelligent.

Vehicle and fixed-track vehicle, or intelligent vehicle agents include the following: communication module, vehicle intervention/control/braking module, information assessing and decision making unit, interface for on-board systems, module for collecting and storing information from the authorities, black box, etc.

Adaptive running properties and timetable subsystem: fixed-track vehicles run according to a well-defined timetable, but various traffic situations can dynamically modify it. This dynamic timetable is continuously revised and corrected by the system, and it is sent to the road vehicles within a certain zone. The running properties of road vehicles will also be sent to fixed-track vehicles through a wireless communication system. Each vehicle agent has their own running properties and timetable.

The key element of the system is the wireless communication network which helps to transfer the information created by the agents. Possible types of the wireless communication system by the parties of communication:

- Train to train communication,
- Train to infrastructure and infrastructure to train communication,
- Train to road vehicle and vehicle to train communication,
- Vehicle to infrastructure and infrastructure to vehicle communication.

In order to achieve autonomous transport, a database of intelligent infrastructure maps (map of railway lines, road maps) is required, which can be partly found in all vehicle agents, or retrieved from fix infrastructure agents. Maps are modified and corrected by the agents.

The above-described system is demonstrated by a system composition in Fig. 2, which includes the following numbered elements: 1 traditional and high-speed passenger trains, 2 freight trains, 3 road vehicles transporting goods, 4 cars and vehicles using distinguishing signals, 5 railway gate equipment, 6 communication, 7 railway tracks.

Messages of inter-vehicle communication: speed, exact position, direction of movement and the priority of the vehicle. Communication messages of the fixed infrastructure: operability and state (e.g. gate is open or closed), while the infrastructure can also give speed restrictions, brake warning, or immediate stop instruction.

The messages will be received by all vehicles within the defined zone around the railway gate. Zones will be defined by considering the longest general braking distance of the vehicles, both in the case of the road and the railway vehicles. Long-distance communication, within 50–80 m, has already been realised between cars.

The system recognising adaptive traffic situations gives greater priority to fixed-track vehicles. It also has an emergency plan for unavoidable collisions. The virtual protected environment surrounding the vehicle depends on the speed and the direction of approaching or distancing. The system makes it possible to use the radar and the camera together. The images of the cameras can be seen in the vehicles, but intervention may only happen by the automatic assessment of other sensors and image processing. All vehicles contain a data recorder unit given by the authorities (a black box—EDR—event data recorder).

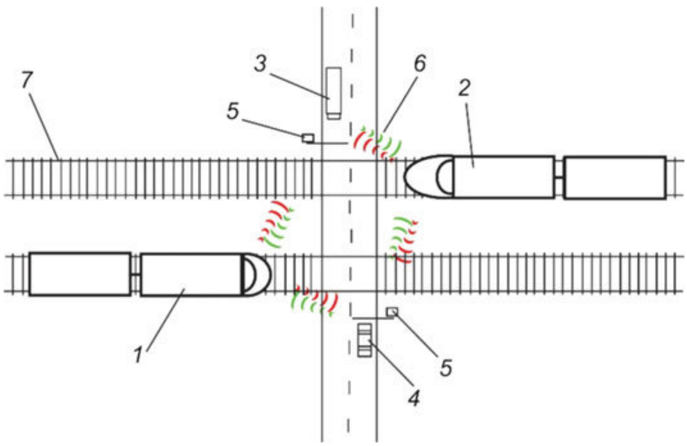


Fig. 2 Intelligent communication-based active vehicle protection system used at road and railway level crossings and in railway infrastructure (figure by author)

6 Conclusion

In conclusion, more than 92% of accidents can be directly attributed to human errors. This is also true for the accidents involving road vehicles and fixed-track vehicles. There is hardly any intersection between road transport and fixed-track transport developments. The automation endeavours presented in our article envisage a uniform autonomous intelligent operation for both road transport and railway vehicles. We see the opportunity to increase safety in the application of intelligent systems. Road and railway level crossings are the most obvious hazards. According to our suggestion, it is possible to implement an active vehicle safety system used in railway infrastructure at road and rail level crossings based on intelligent communication. It can also be part of intelligent infrastructure as we defined it, and it relates to smart mobility, one of the basic pillars of smart city.

Acknowledgement The research on which the publication is based has been carried out within the framework of the project entitled “The Development of Intelligent Railway Information and Safety Systems”. This research has been realised by using the resources of the National Talent Programme, Grant Scheme for the Nation’s Young Talents (Application number: NTP-NFTÖ-16-0582) and the support of the Human Resource Support Office and the Ministry of Human Resources.

References

1. Statistical Yearbook of Traffic Accidents (2014) [Online]. https://www.ksh.hu/docs/hun/xftp/idoszaki/baleset/baleset14.pdf. Accessed 10 Aug 2016
2. Directive 2010/40/EU, Intelligent Transport Systems in the field of road transport, as well as framework for interfaces with other transport modes [Online]. http://eur-lex.europa.eu/legal-content/HU/TXT/PDF/?uri=CELEX:32010L0040&from=EN. Accessed 10 Aug 2016
3. Mester Gy (2010) Service robots, VMTT Konferencia, Novi Sad, Serbia. ISBN 978-86-88077-02-6 (in Hungarian)
4. World’s first electric road opens in Sweden [Online]. https://www.scania.com/group/en/worlds-first-electric-road-opens-in-sweden/. Accessed 10 Aug 2016
5. Előd F (2016) According to Orbán and Varga the autonomous car-manufacturing in the future of the Hungarian industry [Online]. http://index.hu/gazdasag/2016/05/19/orban_es_varga_szerint_az_onvezeto_autogyartas_a_magyar_ipar_jovoje/. Accessed 12 Aug 2016 (in Hungarian)
6. Opinion of the European Economic and Social Committee on ‘The effects of digitalisation on the services sector and employment in relation to industrial change’(exploratory opinion) (2016/C 013/24) [Online]. http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1470258424676&uri=CELEX:52015AE0765. Accessed 12 Aug 2016
7. Communication from the commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions Digitising European Industry Reaping the full benefits of a Digital Single Market {SWD(2016) 110 final}, Brussels [Online]. http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1470258424676&uri=CELEX:52016DC0180. Accessed 12 Aug 2016
8. Commission staff working document, Online Platforms Accompanying the document Communication on Online Platforms and the Digital Single Market {COM(2016) 288 final},

Brussels [Online]. http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1470258424676&uri=CELEX:52016SC0172. Accessed 12 Aug 2016
9. Urmson C, Whittaker W (2008) Self-driving cars and the urban challenge. IEEE Intell Syst 23 (2):66–68. doi:10.1109/MIS.2008.34. ISSN 1541-1672, URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4475861&isnumber=4475846
10. Kim T, Jeong HY (2010) Crash probability and error rates for head-on collisions based on stochastic analyses. IEEE Trans Intell Transp Syst 11(4):896–904. doi:10.1109/TITS.2010.2053536. ISSN 1524-9050. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5504226&isnumber=5635426
11. Sugimoto M et al (2000) Realization of head-on collision warning system at intersections-DSSS: driving safety support systems. In: Intelligent vehicles symposium, 2000. IV 2000. Proceedings of the IEEE, Dearborn, MI, 2000, pp 731–735. doi:10.1109/IVS.2000.898436. ISBN 0-7803-6363-9. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=898436&isnumber=19432
12. El-Shamy EF (2010) Head-on collisions of two electrostatic solitary waves in electron–positron–ion plasmas. IEEE Trans Plasma Sci 38(4):909–914. doi:10.1109/TPS.2010.2042465. ISSN 0093-3813. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5433006&isnumber=5444580
13. Wang J et al (2013) An adaptive longitudinal driving assistance system based on driver characteristics. IEEE Trans Intell Transp Syst 14(1):1–12. doi:10.1109/TITS.2012.2205143. ISSN 1524-9050. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6236181&isnumber=6470717
14. Kelsch J (2016) Human-Vehicle Integration in EU-AdaptIVe, ITS World Congress, Bordeaux, France [Online]. https://www.adaptive-ip.eu/files/adaptive/content/downloads/Deliverables%20%26%20papers/Kelsch_AdaptIVe%20SP3%202015%20-%20ITSCongress.pdf. Accessed 25 Aug 2016
15. Weber T (2016) ECO, spectrum management: activities in CEPT (European Spectrum Regulation) for applications in transport and traffic telematics including sensors, trends in automotive radar and impact on system architecture workshop, March 14–17. Munich, Germany
16. SAE International (2016) Automated driving levels of driving automation are defined in New Sae International Standard J3016 [Online]. http://www.sae.org/misc/pdfs/automated_driving.pdf. Accessed 25 Aug 2016
17. Venkateswaran KG et al (2015) Impact of automation on the capacity of a mainline railway: a preliminary hypothesis and methodology. In: 2015 IEEE 18th international conference on intelligent transportation systems, Las Palmas, 2015, pp 2097–2102. doi:10.1109/ITSC.2015.339. ISSN 2153-0009. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7313431&isnumber=7312804
18. Wang Y et al (2008) A novel automatic train operation algorithm based on iterative learning control theory. In: IEEE international conference on service operations and logistics, and informatics, 2008. IEEE/SOLI 2008, Beijing, pp 1766–1770. doi:10.1109/SOLI.2008.4682815. ISBN 978-1-4244-2012-4. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4682815&isnumber=4682741
19. Dominguez M et al (2012) Energy savings in metropolitan railway substations through regenerative energy recovery and optimal design of ATO speed profiles. IEEE Trans Autom Sci Eng 9(3):496–504. doi:10.1109/TASE.2012.2201148. ISSN 1545-5955. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6213573&isnumber=6226882
20. Zuo J et al (2010) Feedback control of pneumatic brake of urban railway train under ATO mode. In: 2010 international conference on E-Product E-Service and E-Entertainment (ICEEE), Henan, 2010, pp 1–4. doi:10.1109/ICEEE.2010.5660567. ISBN 978-1-4244-7159-1. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5660567&isnumber=5660092
21. Chang CS, Sim SS (1997) Optimising train movements through coast control using genetic algorithms. In: IEE proceedings—electric power applications, vol 144, no 1, pp 65–73.

doi:10.1049/ip-epa:19970797. ISSN 1350-2352. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=583371&isnumber=12657
22. Liu J et al (2013) An analysis of BeiDou Navigation Satellite System (BDS) based positioning for train collision early warning. In: Intelligent vehicles symposium (IV), 2013 IEEE, Gold Coast, QLD, 2013, pp 1065–1070. doi:10.1109/IVS.2013.6629607. ISSN 1931-0587. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6629607&isnumber=6629437
23. Skillingberg M, Green J (2007) Aluminum applications in the rail industry. Light Metal Age-Chicago 65(5):8
24. Jun HK, Kim JH (2007) Life cycle cost modeling for railway vehicle. In: ICEMS. international conference on electrical machines and systems, Seoul, 2007, pp 1989–1994. ISBN 978-89-86510-07-2. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4412054&isnumber=4411938
25. Baumgartner JP (2001) Prices and costs in the railway sector. Laboratoire d'Intermodalité des Tansports Et de Planification, École Polytechnique Fédérale De Lausanne
26. Dominguez M et al (2012) Energy savings in metropolitan railway substations through regenerative energy recovery and optimal design of ATO speed profiles. IEEE Trans Autom Sci Eng 9(3):496–504. doi:10.1109/TASE.2012.2201148. ISSN 1545-5955. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6213573&isnumber=6226882
27. Zhu L et al (2012) Handoff performance improvements in MIMO-enabled communication-based train control systems. IEEE Trans Intell Transp Syst 13(2):582–593. doi:10.1109/TITS.2012.2188288. ISSN 1524-9050. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6183515&isnumber=6205705
28. Darai L (2013) Metrók, metró biztonsága, oktatási vázlat, slide show [Online]. http://www.kjit.bme.hu/images/stories/targyak/metro/automatiz_goa_5r.pdf. Accessed 10 Jan 2016
29. Oh S, Yoon Y, Kim Y (2012) Automatic train protection simulation for radio-based train control system. In: 2012 international conference on information science and applications, Suwon, 2012, pp 1–4. doi:10.1109/ICISA.2012.6220965. ISBN 978-1-4673-1402-2. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6220965&isnumber=6220919
30. International association of public transport: metro automation facts, figures and trends, p 1, 2012, Brüsszel [Online]. http://metroautomation.org/wp-content/uploads/2012/12/Automated_metros_Atlas_General_Public_2012.pdf. Accessed 26 Aug 2016
31. Rail automation—reaching out across Europe, 6th August 2016 [Online]. http://www.railengineer.uk/2016/08/06/rail-automation-reaching-out-across-europe/. Accessed 26 Aug 2016
32. Tokodi D, Schuster G, Ihász J (2014) SMART Rail technológiák lehetőségei, az intelligens vasúti hálózatok kialakításának kérdései. Vezetékek Világa: magyar vasúttechnikai szemle 19 (2):11–15
33. Tokody D, Schuster Gy (2015) I2—Intelligent Infrastructure. In: Schmidt P et al (ed) Reviewed proceedings fifth international scientific videoconference of scientists and PhD. students or candidates: trends and innovations in E-business, education and security, 129 pp, Bratislava, Szlovákia. University of Economics in Bratislava, pp 121–128. ISBN 978-80-225-4191-6
34. International Telecommunication Union, smart sustainable cities: an analysis of definitions, p 21 [Online]. https://www.itu.int/en/ITU-T/focusgroups/ssc/Documents/Approved_Deliverables/TR-Definitions.docx. Accessed 21 Dec 2015
35. Tokody D, Holicza P, Schuster G (2016) The smart mobility aspects of intelligent railway. In: 2016 IEEE 11th international symposium on applied computational intelligence and informatics (SACI), pp 323–326. IEEE
36. Trebuňa F et al (2012) Modelling of mechanical and mechatronics systems intelligent vehicles as the robotic applications. Procedia Eng 48:105–114. ISSN 1877-7058. http://dx.doi.org/10.1016/j.proeng.2012.09.492
37. Milanés V et al (2010) Clavileño: Evolution of an autonomous car. In: 2010 13th international IEEE conference on intelligent transportation systems (ITSC), Funchal, 2010, pp 1129–1134.

doi:10.1109/ITSC.2010.5624971. ISSN 2153-0009. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5624971&isnumber=5624963
38. Jo K, Sunwoo M (2014) Generation of a precise roadway map for autonomous cars. IEEE Trans Intell Transp Syst 15(3):925–937. doi:10.1109/TITS.2013.2291395. ISSN 1524-9050. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6680774&isnumber=6823199
39. Pozna C, Antonya C (2016) Issues about autonomous cars. In: 2016 IEEE 11th international symposium on applied computational intelligence and informatics (SACI), Timisoara, Romania, 2016, pp 13–18. doi:10.1109/SACI.2016.7507360. ISBN 978-1-5090-2379-0. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7507360&isnumber=7507333
40. Uhlemann E (2015) Introducing connected vehicles [connected vehicles]. IEEE Vehicular Technol Mag 10(1):23–31. doi:10.1109/MVT.2015.2390920, ISSN 1556-6072. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7047294&isnumber=7047264
41. Kaempchen N et al (2009) Situation assessment of an autonomous emergency brake for arbitrary vehicle-to-vehicle collision scenarios. IEEE Trans Intell Transp Syst 10(4):678–687. doi:10.1109/TITS.2009.2026452. ISSN 1524-9050. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5161310&isnumber=5338639
42. Tokody D (2016) Smart mobility. In: Enikő B (ed) Proceedings of the XXI-th international scientific conference of young engineers, 452 pp, Kolozsvár, Románia. Erdélyi Múzeum-Egyesület (EME), pp 401–404
43. Sukuvaara T, Nurmi P (2009) Wireless traffic service platform for combined vehicle-to-vehicle and vehicle-to-infrastructure communications. IEEE Wireless Commun 16(6):54–61. doi:10.1109/MWC.2009.5361179. ISSN 1536-1284. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5361179&isnumber=5361167
44. Harding J et al (2014) Vehicle-to-vehicle communications: readiness of V2V technology for application (Report No. DOT HS 812 014). National Highway Traffic Safety Administration, Washington, DC
45. Seo H et al (2016) LTE evolution for vehicle-to-everything services. IEEE Commun Mag 54 (6):22–28. doi:10.1109/MCOM.2016.7497762. ISSN 0163-6804. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7497762&isnumber=7497754
46. Tokody D, Maros D, Schuster Gy, Tiszavölgyi Zs. (2016) Communication-based intelligent railway—implementation of GSM-R system in Hungary. In: 2016 IEEE 14th international symposium on applied machine intelligence and informatics (SAMI), Herlany, pp 99–104. doi:10.1109/SAMI.2016.7422989
47. Hank P et al (2013) Automotive ethernet: in-vehicle networking and smart mobility. In: Design, automation & test in Europe conference & exhibition (DATE), Grenoble, France, 2013, pp 1735–1739. doi:10.7873/DATE.2013.349. ISSN 1530-1591. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6513795&isnumber=6513446
48. Cvijić B et al (2016) Cloud based web application supporting vehicle toll payment system. In: 2016 5th Mediterranean conference on embedded computing (MECO), Bar, Montenegro, 2016, pp 489–492. doi:10.1109/MECO.2016.7525700. ISBN 978-1-5090-2220-5. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7525700&isnumber=7525673
49. Ipakchi A, Albuyeh F (2009) Grid of the future. IEEE Power Energy Mag 7(2):52–62. doi:10.1109/MPE.2008.931384. ISSN 1540-7977. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4787536&isnumber=4787510
50. Deilami S et al (2011) Real-time coordination of plug-in electric vehicle charging in smart grids to minimize power losses and improve voltage profile. IEEE Trans Smart Grid 2 (3):456–467. doi:10.1109/TSG.2011.2159816. ISSN 1949-3053. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5986769&isnumber=5993785
51. Sortomme E et al (2011) Coordinated charging of plug-in hybrid electric vehicles to minimize distribution system losses. IEEE Trans Smart Grid 2(1):198–205. doi:10.1109/TSG.2010.2090913. ISSN 1949-3053. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5664815&isnumber=5715606

52. Tokody D, Schuster G (2016) Driving forces behind smart city implementations—The Next Smart Revolution, manuscript
53. Prime Faraday Technology Watch—January 2002: an introduction to MEMS. ISBN 1-84402-020-7. http://www.lboro.ac.uk/microsites/mechman/research/ipm-ktn/pdf/Technology_review/an-introduction-to-mems.pdf
54. Tokody D, Papp J, Schuster Gy (2015) The challenges of the intelligent railway network implementation: Initial thoughts from Hungary. In: Gogolák L, Fürstner I (eds) Proceedings of the 3rd international conference and workshop mechatronics in practice and education—MECHEDU 2015, Szabadka, Szerbia, 2015.05.14–2015.05.16. Subotica Technical College of Applied Sciences, pp 179–185. ISBN 978-86-918815-0-4
55. Mesterséges Intelligencia Almanach [Online]. http://mialmanach.mit.bme.hu/aima/ch01. Accessed 21 Dec 2015
56. Blum JJ et al (2004) Challenges of intervehicle ad hoc networks. IEEE Trans Intell Transp Syst 5(4):347–351. doi:10.1109/TITS.2004.838218. ISSN 1524-9050. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1364012&isnumber=29884
57. Lytrivis P (2015) A holistic approach for automated transport systems. In: iMobility forum plenary meeting, 21 October 2015, Brussels
58. Cohen B (2015) The smartest cities in the world 2015: methodology [Online]. http://www.fastcoexist.com/3038818/the-smartest-cities-in-the-world-2015-methodology. Accessed 20 Dec 2015
59. Siergiejczyk M (2015) Communication architecture in the chosen telematics transport systems. [Online]. http://cdn.intechopen.com/pdfs-wm/37575.pdf. Accessed 21 Dec 2015

Software Reliability of Complex Systems Focus for Intelligent Vehicles

György Schuster, Daniel Tokody and Imre János Mezei

Abstract Using software became a part of our everyday life, in the last few decades. Software is widely used in areas, such as national defence, aeronautics and astronautics, medicine or even transport. There are 100 million lines of codes in a modern high-end car's engine control unit. In comparison, the Space Shuttle needs 400 000, the F22 fighter jet needs less than 2 million, the Boeing 787 airplane needs 14 million and the Facebook needs more than 60 million lines of codes to function. Even a smaller error can lead to devastating consequences in safety-critical systems, such as those operating in vehicles. There have been several examples in recent years, when an automotive recall was necessary due to dangerous software, and there were cases when these errors presumably caused fatal accidents. Definition of software reliability is the error-free working probability of software for a specified period of time under well-defined environment. Usage of software is inevitable. It can be found in every vehicle to control almost everything. Therefore software can be considered as a critical success factor and it has a strong effect on the reliability of the whole system. The software systems are getting more and more complex. Known fact is a more complex system has more possibility to have errors. The most difficult problem is that the traditional methods of reliability cannot be used. For example fatigue and wearing of mechanical parts or features of lubricant systems can be calculated quite well, since we have enough prior knowledge on their features. Unfortunately, in case of software systems this knowledge is missing. This paper deals with the question of software reliability. In the first part it lists the

G. Schuster (✉) · D. Tokody · I.J. Mezei
Doctoral School on Safety and Security Sciences, Óbuda University,
Budapest, Hungary
e-mail: schuster.gyorgy@kvk.uni-obuda.hu

D. Tokody
e-mail: tokodi.daniel@bgok.hu

I.J. Mezei
e-mail: mezei.janos.imre@gmail.com

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_28

problems and the second part gives some mathematical issues to calculate working probability.

Keywords Software reliability · Complex systems · Safety critical system · Intelligent vehicles

1 Introduction

1.1 Cases

On 9 May 2015, an Airbus A400M Atlas cargo plane on a test flight crashed near Seville, Spain. Four of the six aircraft crew were killed and the remaining two were seriously injured. Three of the aircraft's four engines failed during the A400M's departure from Seville. The crash was caused by software [1].

The US air safety authority issued a warning and maintenance order 1 May, 2015 over a software bug that causes a complete electrical shutdown of Boeing's 787 Dreamliner. The software bug was found in the plane's generator-control units [2, 3].

Non-professionals would think once a software works, it will not become faulty. Reality, however, differs. Since software—like human controllers—makes decisions; its operation can also be confused by environmental effects. A software, which worked well for the Ariane 4 rocket, caused the explosion of Ariane 5 on its first voyage. The cause is simple; the overflow of a variable [4].

1.2 Definitions

Software: the whole of programs and data. The part of the system that is not tangible.

Software reliability: the probability of failure-free software operation for a specified period of time in a specified environment [5, 6].

Failure: unexpected software behaviour perceived by the user.

Fault: the software characteristic caused by failure.

1.3 Circumstances that Have an Effect on the Reliability of the Software

- Human factor: Experience and ability to faultless work of people writing the program and providing the data. The skilfulness of those doing the tests is equally important [5, 6].

- Software development model: Agile and Scrum software development practices should be avoided in the development of safety-critical systems [5, 6].
- Aptitude of management: unfit management is a recurrent problem. Managers are often not competent in the given field (with some exceptions), and what they expect is many times senseless or unachievable [5, 6].
- Faultless development environment: In many cases a faultless source code does not work properly because of the faults of the compiler or the software library [5, 6].
- Faultless test environment: testing is an essential part of the life cycle of the software, thus it strongly affects reliability. The majority of tests needs to be automated, especially endurance and time tests. If the test environment is faulty or incomplete, it evidently influences the results [5, 6].
- Choice of hardware environment: Software does not exist without hardware. If the hardware provides an adequate platform for the execution of tasks with sufficient reserves, the software can work well. There have been cases when the software configuration worked well in the chosen simulation environment and in real life for years, but then the hardware ran out of its reserves, and the system stopped, or worked improperly [5, 6].

1.4 Indices of Software Reliability

Factors influencing software reliability need to be quantified. However, whereas choosing these factors is straightforward and easy in case of other engineering fields, finding objective influencing factors is uncertain in this field [7].

Influencing factors include, but are not limited to:

- complexity of the software: It seems easy, but it is not. Indices can be various:
  - LOC (Line of Code) and KLOC (Kilo Line Of Code). This data is good as a first approximation, but it is not definitive. For example the complexity of a relatively long program with one thread does not match the complexity of another one with multiple threads.
  - Number of the files of the software: Experienced programmers take the system apart into processes to avoid great and unnecessary complexity and to make the testing of parts possible.
  - Number of tasks running parallel: it is a useful index, but it is not enough in itself, either.
  - Number of IPC elements: These, together with LOC provide a good approximation, but still not a complete description of complexity. The number of IPC elements shows the probability of the occurrence of different special faults (Dead Lock).
- Experience of programmers. This is an unquantifiable index. Great parts of the software are not written by one programmer, not even one group. Thus, this important index cannot be estimated.

- Indices of the project management. It is a well-known fact a well-managed project is likely to have better results. If project management is supportive and evaluate rationally, development work is done in a more relaxed and effective way than a work where programmers are asked to report too frequently and in unproductive ways. The problem is the same here, too: there is no index that can be given [8].
- Quality aspects: Quality assurance methods and standards can serve as a base for quality, as well as the CMM certification of the development company.
- The extent of the error: This is the most important index, since it provides an exact number that shows reliability. Reliability depends on time. The following types of tests can be applied:
  - Stress endurance test. The software is tested in an environment that influences its operation heavily, e.g. it reduces the resources.
  - Load testing and stability test. The software is tested over a long period of time under peak load conditions.
  - Reliability test. It is similar to the previous one, but testing is done under normal load conditions.
  - Regression test. It examines whether the software remains stable if new elements appear in the system, or if some part of the system starts to work improperly.

The most problematic feature of these tests is that they take a long time, as they are statistical procedures. Results are expressed in the following units of measurement: [5, 6].

- Probability of Failure on Demand (POFOD). The probability that the system will fail when a service request is made.
- Rate of Change of Frequency (ROCOF). It is given for a certain time period.
- Mean Time Between Failures/Mean Time To Failure (MTBF or MTTF).
- Availability: how long a system is available in a given time period and how long maintenance lasts.

Another problem is the adequateness of the test environment. Obviously testing cannot be done by humans in the long run, it has to be automated. If the test environment is incomplete, some parts of the software are not tested, or the automated tester does not notice differences.

1.5 Faults and Failures

As the definition shows, failures do not always cause faults. It is possible that an expert notices that the operation of the system operates is slightly different from the specifications, but it is not certain that the failure manifests itself in the operational properties of the system. Failures that are apparent on a system level are faults.

Faults can be: [9]

- insignificant (unimportant),
- significant (critical),
- known,
- unknown.

Insignificant faults—known or unknown—do not cause serious functional disorders. If, however, the fault is detected, it needs to be corrected soon. In systems that are not safety-critical faults are sometimes not corrected after a risk assessment. Minor colouring faults are a good example.

The generator problem of a Boeing 787 is somewhere between failure and fault. It is a very serious problem, but to have this problem the generator needs to operate for an improbable length of time. This fault needs to be repaired, but the 787 fleet does not have to be called back immediately [2, 3].

Critical faults need to be corrected without delay. Unknown critical faults are the most dangerous. Both the circumstances and their cause are unknown (see the catastrophe in Seville).

Though prompt correction of faults seems apparent, correction itself carries the possibility of faults. Validation of the corrected software is essential [9].

It is not apparent why the detection of faults is difficult. The reasons can be:

- the fault only occurs in the joint occurrence of certain circumstances,
- the fault does not occur in normal circumstances/operation, or its effect remains hidden,
- the fault develops continuously until at some point it causes a system-level problem.

The first case is typical for concurrent systems. In case of imperfect development the whole system or parts of it can become blocked in a given constellation.

In the second case the system to which it is not prepared and it may even give an answer to it. An unfortunate situation like this can be e.g. a dead code.

The third fault can be very banal. The system wastes some resource continuously and unnoticed. A typical example for this is which leads to the full consumption of the memory, another example is a memory leak.

2 Increasing Reliability

Increasing reliability is a primary goal. Possible methods are the following:

- The use of appropriate development methods. Observing security and coding procedures, e.g. MISRA, and complying with the relating standards (e.g. IEC 61508) [9].

- Testing the produced data and programs most exhaustively, even at the expense of doing more testing projects at the same time. It is even more important in case of long-term, stress and regression tests.
- The code should be, it should be able to correct small faults so that they do not lead to serious failures [5, 6].
- The use of software units that have already been used many times and proved perfect. This is not an ultimate solution, either, because these units can get into circumstances in which they do not work well.
- The code should contain which can help shed light on more serious faults, and help prevent the emergence of these faults.

The problem is still persistent, since all redundancies increase complexity, which in turn is a source of further faults. There have been examples for a security software rendering a good system unserviceable.

3 Reliability Calculation

Reliability can be practically expressed by the probability of failure, since probability of well-functioning P_W (1) is the complement of probability of failure P_f. Thus:

$$P_w = 1 - P_f. \tag{1}$$

In many cases the mathematical analysis of a problem is easier this way.

The first example is a one-thread program, which does not occur in reality, but it will be of significance in the following examples.

Let us suppose that the statistical data of the program examined are known. In this case the easiest mathematical approximation of failure (2) is Poisson distribution:

$$P_f(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \tag{2}$$

$\lambda > 0$, where is the expected value of failure for the given time period, and is the $k = 1, 2, 3, \ldots$ number of failures.

Problem: the probability of failure can be computed in this case. Another question is how we obtain reliable information on the parameter.

Explanation: software packages are well designed and implemented constructions. Thus, conventional observation techniques do not yield useful results. In these cases people can be examined on the basis of their active program. If there is a software described above, already in use, with tens of thousands of them running, conclusions can be drawn concerning parameter λ examining collected failure data.

Example: let us suppose that 1000 copies of a software are running. The mean failure rate is one failure every 100 h. In case of one copy, it means one failure for every 100000 h.

However, it would be preferable for software designers if this value could be determined before the release of the software, in the design phase.

Question: Can this method be applied to concurrent systems?

The answer is yes, in case of individual program sections, but only to the given tasks and threads. This problem is further complicated by Inter Process Communication (IPC).

If we suppose that the failure of an individual software section leads to the faulty operation of the whole system, the probability of failure (3) is:

$$P_{f\Sigma} = \sum_{i=1}^{n} P_f. \tag{3}$$

Let us suppose a program with four threads (IPC excluded), all the four threads of which have the same probability of failure, and a failure leads to the faulty operation of the whole system.

In this case the probability of failure increases fourfold.[1]

In case of IPC the relationship between processes constitutes another possibility of failure. A common occurrence is deadlock between two processes. It can happen with semaphores and MUTEXes.

3.1 Example

There are two processes: A and B and two MUTEXes: X and y. The priority of B is higher than that of A. The problem:

- The process is running, and gets MUTEX X.
- The scheduler stops process A, because process B is ready, and starts B.
- Process B gets MUTEX Y.
- Process B is blocked on MUTEX X.
- Process A is running, since B is in blocked state.
- Process A is blocked on Y MUTEX

Process A cannot run, because it is waiting for the release of MUTEX Y, and for this reason it cannot release MUTEX X, on which process B is waiting, and so B cannot release MUTEX Y.

[1]Increase in complexity leads to decrease in reliability.

This is the simplest example; it is easy to notice this problem and handle it. A more difficult problem is when this phenomenon does not occur in pairs, but more processes get into the cyclic blocking state.

The probability of this state can be estimated. The following method is suitable for two processes: Let us suppose that both A and B processes use X and Y MUTEXes with normal distribution. Based on that:

- In case of normal running, the probability of process A being in a field protected by MUTEX X is P_X.
- In case of normal running, the probability of process B being in a field protected by MUTEX Y is P_Y.
- The probability of running of process A is P_A.
- The probability of running of process B is P_B.
- The probability of processes A and B running at the same time is

$$P_{AB} = P_A\, P_B. \tag{4}$$

- The probability of deadlock (5) of the two MUTEXes is

$$P_{dead-lock} = P_A\, P_B\, P_X\, P_Y. \tag{5}$$

The solution does not seem complicated, but the result is unfavourable from the point of view of running reliability. From the point of view of testing, it is advantageous, since the error is detected quickly.

The above example is perfunctory. With tens or hundreds of processes the probability of an error like this decreases significantly, but so does the chance of its detection, too.

Results are similar, if the hardware is taken into consideration as well. Processing a state originating from the hardware takes time, too.

A typical case is the processing of network allocation, when a device finds the network device available, but by the time it would start to use it, another device has already started to use it. This time interval is called a soft time slot.

Unfortunately, this problem also happens when this task is performed by hardware. It can be solved by different methods, like CSMA/CD or CSMA/CA. Depending on the usage of the network device, intolerable delays may occur in firm and hard real time cases.

It happens when the resources of the system are not sufficient in the critical time period. If the hardware and software resources suffice, the system can handle the most disadvantageous state, and does not experience time related disturbances. Our method of computing is the following:

ρ_i is the relative frequency of resource usage (6), where

$$\rho_i = \frac{\sum_j \tau_{ij}}{T_i}, \tag{6}$$

τ_{ij} is the time span of the critical phase, during which the problem may occur (regarding the ith process). If another process wants to access the resource, it blocks access

T_i the time span examined (regarding the ith process)

The limit at T_i infinity of p_i is the probability of the occurrence of the critical phase.

There is an error or deviation, when a certain number of processes clash.

If p_i probabilities are the same for the access part resource, the binomial distribution (7).

$$p_{ik} = \binom{n}{k} p_i^k (1 - p_i)^{n-k}. \tag{7}$$

For the error or deviation to happen, k processes out of n must happen during the given time span. However, there is an error, if more than k processes enter at the given time.

$$p_f = \sum_{j=k}^{n} \binom{n}{j} p_i^j (1 - p_i)^{n-j}, \tag{8}$$

where p_f is the probability of the error or deviation (8).

To avoid it, sufficient resources need to be allocated during the designing of the system, so that no problems can arise even with maximum load.

Memory leaking is the next case to be examined. The cause of the deviation—and later the error—is that the process fails to release memory it has been allocated, thus the available memory is diminishing.

The error model is a simple subprogram, which has many exit points. This subprogram exits at one of the exit points—where the memory is not released—with p_i probability, thus reducing available memory, p_i can also be estimated using relative frequency, supposing uniform distribution.

Let us suppose that:

M_a is the memory available,

Δm is the allocated, but not released memory,

k the number of runnings,

τ the average "cycle" time of runnings,

T_f the expected interval of the error

The available memory must come to an end for the error or deviation to occur.

$$0 > M_a - k\Delta\Delta, \text{ where } T_f > \tau \frac{M_a}{\Delta m}. \tag{9}$$

Memory leaking is an error. However, software developers often do not solve the problem, but advise the restart of the system from time to time.

The next issue is the examination of software as a Markov process.

Markov process is a stochastic process, the future behaviour of which depends only on the present state of the system [10]. In short, it is described as "memory less".

This boundary condition is not met by the software. A simple example for this is the resource usage of the system for one moment of time.

Let us look at the following scenario: there is a process that has considerable memory access. The data storage device becomes fragmented due to high load, this way its access time gets longer. It is, without doubt, an effect from the past.

If the system has layered structure, the "upper" layers have supervisory and correcting functions—especially in case of SIL3 and SIL4 [11]. It is not suitable to treat the system as a Markov process in this case, either.

These were all examples of problems which lead to deviations or failures due of some designing error. In all cases the probability of the occurrence of the sub process was attempted to be established. This probability value is, however, more difficult to determine, since the probability of the occurrence of these critical phases is so small that it would take years to observe the given process, or large number of elements need to be monitored in order to determine the desired value.

A further problem is that software packages are without an exception concurrent system. Thus, observing them is a far more complex task than observing any other traditional technical systems.

Software manufacturers and users would like to determine risk elements before the release of the software. The methods described above are not suitable for this task.

The solution is the monitoring of the producer of the software. The producer most certainly has many software products, several of which run under real conditions. Collecting these data allows for a statistical sample that can be analysed.

Our study aims to examine how arbitrary software errors are, and what kind of trends can be observed considering the software manufacturer.

If we consider the software development procedure as a time series, we can apply Hurst analysis [12].

Note: Hurst analysis has been designed for time series, but since software development is not time dependent enough to examine time spans, time is substituted by the number of code lines.

The number of occurrences of errors and deviations has been determined regarding a given number of code lines.

E_i the number of errors and deviations regarding a given number of code lines, $i = 1, 2, \ldots, m$
m the number of time series,
N the number of code lines of samples,
E_N average number of errors (10)

$$E_N = \frac{1}{m}\sum_{i=1}^{m} E_i. \tag{10}$$

Let us calculate the cumulated averages of the given periods (11).

$$X(i, N) = \sum_{i=1}^{m} E_i - E_N. \tag{11}$$

Then let us calculate the difference of the maximum and minimum value of $X(i, N)$.

$$R = \max X(i, N) - \min X(i, N). \tag{12}$$

This value is normalised with the standard deviation of the whole series:

$$\frac{R}{S_N} = N^H \tag{13}$$

H is the Hurst exponent, the value of which is characteristic of the errors and trends.

$H = 0.5$ indicates a completely uncorrelated series; that errors occur arbitrarily.

$H > 0.5$ indicates a time series with long-term positive autocorrelation, i.e. if errors tend to decrease in a given time period, the same is expected in the next period of time.

$H < 0.5$ predicts a switching between high and low values in adjacent pairs. That is, a single high value will probably be followed by a low value, and that the value after that will tend to be high.

This is method is promising. However, it has two problems. The producers of the software tend to keep smaller errors for themselves, since this way their reputation will not be damaged. The other problem is that users may not become aware of deviations, or may not deal with the problem. In both cases the statistical sample becomes less accurate.

4 Consequences and Conclusion

Software reliability can be expressed with the help of probability attribute. It is difficult to find a model that enables us to estimate the value fairly. Having examined the question from a multitude of viewpoints it can be stated that there is no one perfect theory. It is difficult to find a model that enables us to estimate the value fairly. Having examined the question from a multitude of viewpoints it can be stated that there is no one perfect theory.

The application of the mathematical solutions explained in this article requires some preliminary knowledge, which cannot be obtained with 100% certainty.

The method that we were studying aims at the producer of the software, since the statistical examination of its product yields results that can form the basis of further research on the software.

Reliability can be increased with different methods, rules and testing, but experience shows that just like other man-made systems, reliability of software packages cannot be 100%. Reliability is usually unchanging after the phase of design, but the fact that codes depend on humans is a serious problem in itself.

The reliability of a software is probability. It is difficult to find a probability model that gives a good estimate of the software's reliability. Having examined this topic, we can assert that it is impossible to name one perfect theory. Different methods, rules and testing can increase reliability, but experience shows that—similar to other artefacts made by humans—software packages cannot be 100% reliable.

References

1. Flottau J, Osborne T (2015) Software cut off fuel supply in stricken A400M. http://aviationweek.com/defense/software-cut-fuel-supply-stricken-a400m. Accessed 19 May 2015
2. The Aviation Safety Network, Accident description. http://aviation-safety.net/database/record.php?id=20150509-0. Accessed 19 May 2015
3. Mouawad J, FAA (2015) Orders fix for possible power loss in boeing 787. http://www.nytimes.com/2015/05/01/business/faa-orders-fix-for-possible-power-loss-in-boeing-787.html. Accessed 19 May 2015
4. Lions JL (1996) ARIANE 5 Flight 501 Failure. https://www.ima.umn.edu/~arnold/disasters/ariane5rep.html. Accessed 19 July 2015
5. Pan J (1999) Software reliability. https://users.ece.cmu.edu/~koopman/des_s99/sw_reliability/
6. Schneidewind FN (1997) Reliability modeling for safety critical software. IEEE Trans Reliab 46(1)
7. Van Solingen R, Berghout E (1999) The goal/question/metric method: a practical guide for quality improvement and software development. McGraw-Hill International
8. Long J (ed) (2008) Metrics data program, national aeronautics and space administration. http://mdp.ivv.nasa.gov/index.htm

9. Tokody D, Schuster G, Papp J (2015) Study of how to implement an intelligent railway system in Hungary. In: 2015 IEEE 13th international symposium on intelligent systems and informatics (SISY), Subotica, pp 199–204. doi:10.1109/SISY.2015.7325379
10. Markov process (mathematics)—Britannica Online Encyclopedia
11. Marszal EM, Scharpf EW (2002) Safety integrity level selection—systematic methods including layer of protection analysis. The Instrumentation, Systems, and Automation Society, Research Triangle Park, NC, USA. ISBN: 9781556177774
12. Szilágyi GyA (2016) The fractal dimension of the air accidents. Repüléstudományi Közlemények XXVIII(2):41–48. HU ISSN 1417-0604

Usage of an Optical Flow Sensor in Robotics to Define Orientation

Máté Koba, Roland Bartók and László Czap

Abstract In autonomous robotic control it is necessary to measure the exact tracks of the robot and then calculate the best route for it. This article is about to introduce a method which can be used to measure the route of an Unmanned Ground Vehicle to help the controller software to guide the vehicle. For now, the most popular way to measure and track of a vehicle or a robot's movement is the incremental or absolute encoder installed on the vehicle's wheels. This method measures movement in an indirect way because there can be gear system on its motor or in omnidirectional systems, there can be a slip between the wheel and the ground so the data of these sensors could be easily tampered. With this method we install an optical flow sensor (or sensors) to the undercarriage of the robot. This way the system measures the movement in a fully direct way because the sensors detect the movement between the robot and the ground and there is no mechanical or electrical connection between the sensor and the reference.

1 Introduction

This article explains the usage of an optical flow sensor through an example of a custom-built robot. The robot built for a Hungarian robot-building contest called "Magyar Alkalmazott Mérnöki Tudományok Versenye". The software of the robot is running ROS (Robot Operating System), it is controlled over wireless communication and it has autonomous functions too. In the following, there will be written about the robot's structure and the optical flow sensor based navigation system will be explained thoroughly.

M. Koba (✉) · R. Bartók · L. Czap
Institute of Automation and Infocommunication, University of Miskolc,
Miskolc, Hungary
e-mail: autkmate@uni-miskolc.hu

R. Bartók
e-mail: bartokroland@hotmail.hu

L. Czap
e-mail: czap@mazsola.iit.uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_29

2 Introducing the Robot

The task of the robot is manoeuvring on a track which has 15 pillars made of PVC and it has to occupy as many pillars as it can by putting a beacon into them. The robot is equipped with 3 omnidirectional wheels (omniwheel) so it is capable of complex movements. There are 3 main stages of the robot. At the lowest level there are the 3 motors equipped with omniwheels, 2 optical flow sensors and the battery-pack, shown in Fig. 1. On the middle level there are the computing units, an STM32F4 Discovery board and a Raspberry PI 3. The top level contains the beacon on a mechanism which is responsible for putting the beacon in the right position.

2.1 *Why Optical Flow Sensors?*

For a precise navigation it is very important to know the exact orientation of the robot. In robotics there are a lot of methods of defining the position and orientation of a robot such as using inertial sensors, GPS navigation based applications but these are quite expensive and easily jamming appliances. Using optical flow sensor is much cheaper than the above mentioned ones and it is less sensitive to electrical interference still it is an accurate sensor. With these sensors it is possible to measure velocity of multiple-driving systems and make measurements of conveyor systems like in [1].

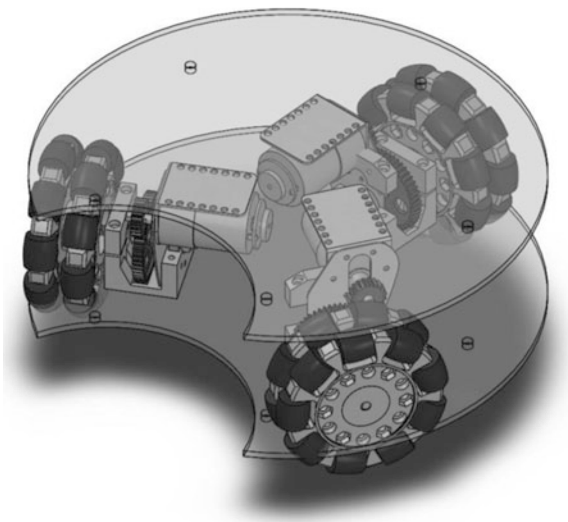


Fig. 1 Bottom level sketch without battery and sensors

3 Optical Flow in Theory and Presenting an Actual Sensor

3.1 Optical Flow Phenomenon

Optic flow is a visual event which can be noticed during our everyday life. Optic flow is the visual motion we can experience during movements. It can be introduced to a simple situation. Suppose that you are sitting in a car or a train and looking out the window. The objects outside like trees, buildings, etc., seem to move backwards. The motion you experience is optic flow. From this motion, you also could tell your distance from these objects. Distant objects appear still and closer objects appear to move backwards, faster than the objects far away from you. There are mathematical relationships between the magnitude of the optic flow and the chosen object's relation from the spectator. When the speed of travel doubles, the optic flow you see will also double. If an object is twice as close as it was, the optic flow will double again. The magnitude of optic flow also depends on the angle between the direction of travel and the direction of the inspected object. Let's see the case of travelling forward on Fig. 2 [2].

The optic flow is the largest when the object is to the inspector's side by 90° or directly above or below from it. In front of the inspector the optic flow will be zero. So the object in front of the inspector will appear to be still.

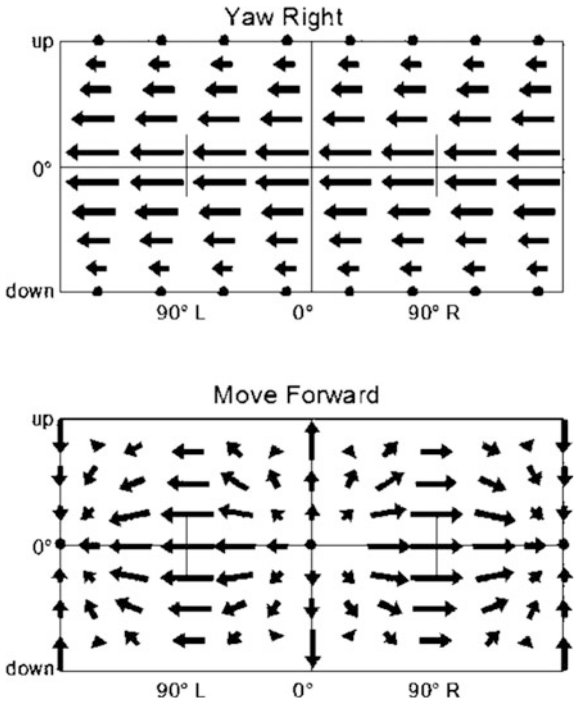


Fig. 2 Optical flow in case of yaw right and moving forward [2]

3.2 Operation of an Optical Flow Sensor

Optical flow sensors could find in optical mice. These sensors could be called intelligent sensors, they process the texture of the surface the sensor is inspecting and after that they define the relative movement of the sensor and the surface at the x, y plane. The camera makes pictures sequentially and compares the data in the order to the sequence. According to the texture and the time between the two pictures it is possible to calculate the amount and the direction of the movement the object travelled. With further calculations, we could define current speed, object orientation from start point, turning angle [3, 4].

3.3 Benefits of Using These Sensors

With flow sensors, the calculations of speed and movement are much more punctual than using encoders on the motors or using GPS even on hard terrain. In our case, with the omnidirectional wheels to measure speed and distance is more simple and accurate using optical flow sensors rather than encoders. With omniwheels the direction of travel is not clearly definable from the rotation of the motors because in this case the robot is not heading to the exact direction the wheels are spinning. Also during each round of the competition there are four robots on track and it is possible that robots collide or push each other. When the robot is pushed or there is any slip between the wheel and the track the encoders provide wrong data for the speed and orientation calculations. It is also possible to use GPS but commercial GPS provides weak signal inside buildings. It is also possible to use sensor-fusion to define movement, but it requires a lot of different types of sensors which are very sensitive and easily jammable. The best way to measure speed and travelled distance is to use optical flow sensors [5].

3.4 ADNS-3080

The sensor is based on Optical Navigation Technology, which measures changes in position by optically acquiring sequential surface images (frames) and mathematically determining the direction and magnitude of movement. It contains an Image Acquisition System (IAS), a Digital Signal Processor (DSP), and a four-wire serial port. IAS acquires surface images via a lens and illumination system. The images are processed by the DSP to determine the direction and distance of motion. The DSP calculates Δx and dy relative displacement values. With an external microcontroller it is possible to read the Δx and dy information via the sensor serial port [6]. The ADNS-3080 block diagram is shown in Fig. 3.

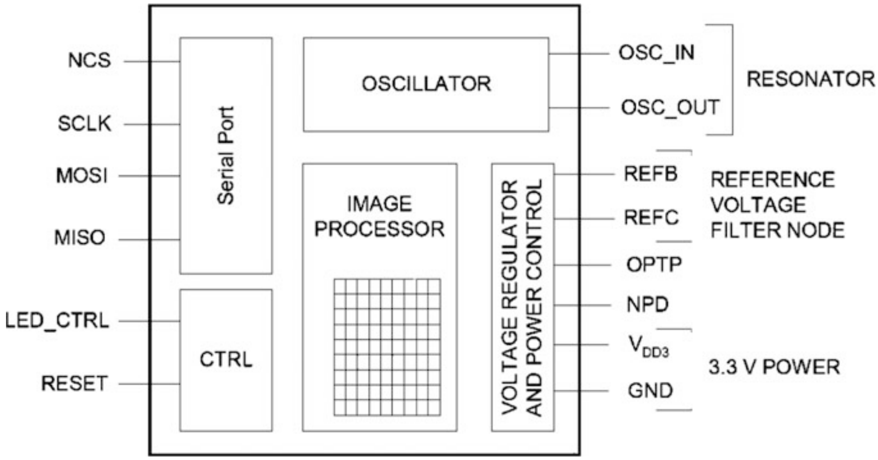


Fig. 3 Block diagram [6]

<u>The sensor has 5 main parts:</u>

- serial port
- ctrl logic
- image processor
- power circuit
- oscillator

<u>The sensor structure built up by 3 parts:</u>

- Image Acquisition System (IAS)
- Digital Signal Processor (DSP)
- 4 wire serial port (SPI)

After the reflected light from the illuminated surface reaches the sensor the IAS acquires the images and then the DSP processes the image and the processor calculates the magnitude and direction of the movement and after that it is possible to calculate the relative movement [6].

3.5 Sensor Sensitivity

The sensor is not equally sensitive at the entire light spectrum. The figure below shows the sensor sensitivity curve related to light wavelength. The sensor is most sensitive in the 600–700 nm range. The sensor sensitivity is shown in Fig. 4 [6].

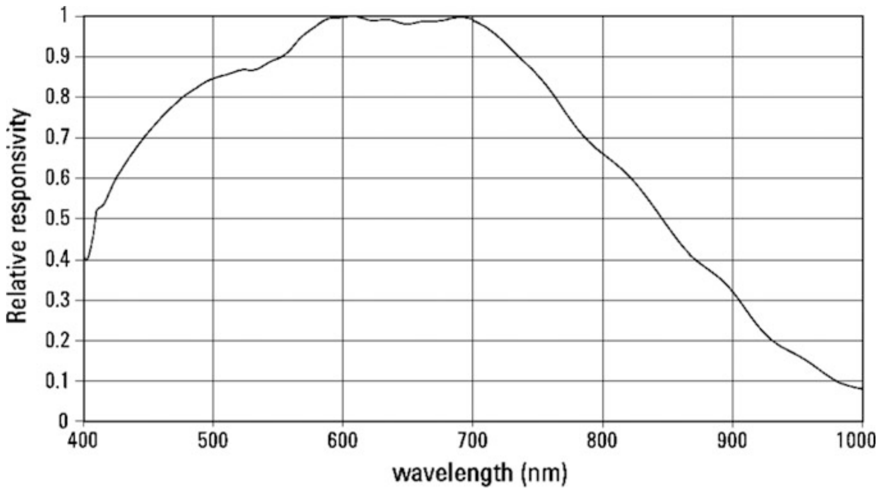


Fig. 4 Sensor sensitivity [6]

4 Using the Sensor

4.1 Sensor and Lens

The sensor is accessible with a PCB mount so the main circuits and the camera are attached and the user could change lenses and connect the sensor to a microprocessor. The mounted sensor is shown in Fig. 5.

There are a wide variety of lenses compatible with this mount. The lens image shown in Fig. 6 at the farthest and nearest focal point, the grid is 1 mm × 1 mm.

4.2 Guarantee Appropriate Light Conditions for the Sensor

Since the sensor is very sensitive to the light conditions it is necessary to provide eligible amount of light for it. The light is provided by 4 LEDs each has 2180–4200 mcd intensity and 60° radiating angle and their wavelength is 619–629 nm. Every 4 pieces are mounted in a 3D printed case which provides the optimal angle for each LED. It is necessary to notice that every lens needs a different placement according to its optical specification. The correct LED angle is calculated by the 3D designing program. Figure 7 shows the 3D model of the LED case.

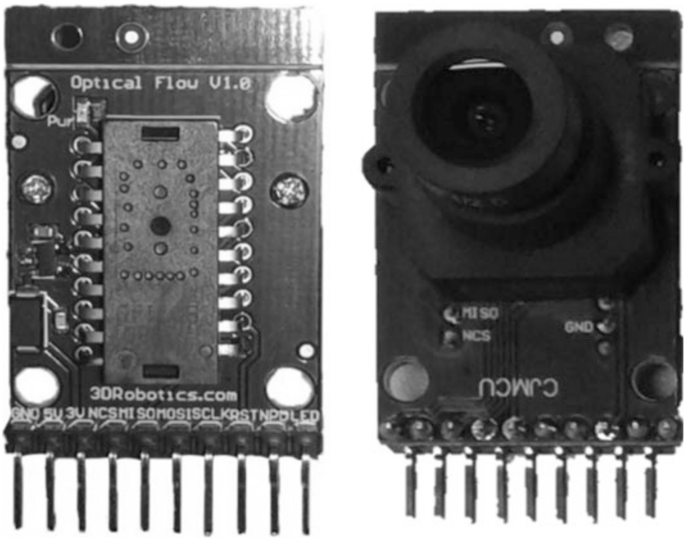


Fig. 5 Sensor PCB

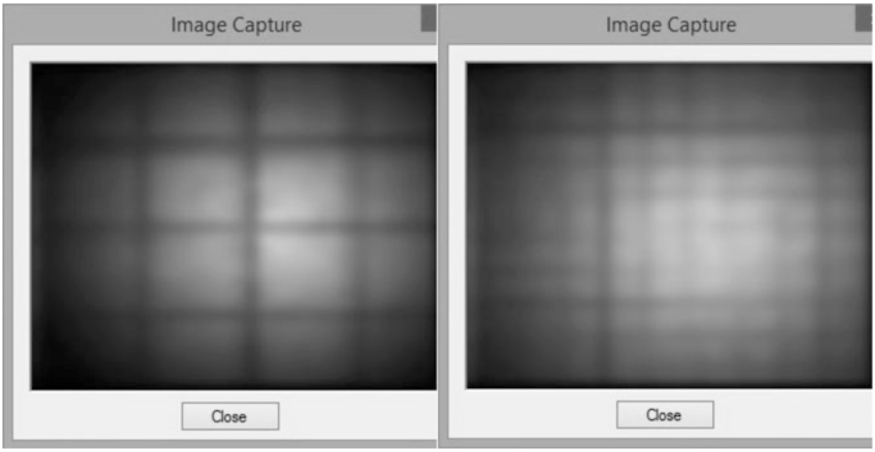


Fig. 6 Selected lens image over test grid

4.3 Sensor Placement

There are 2 sensors on the robot. The sensors are on the x-axis of the robot, rotated by 90° to each other. The result of the rotation is that the given data is compensated and this way it is possible to calculate rotating angle besides x and y distance. Figure 8 shows one of the attached sensors.

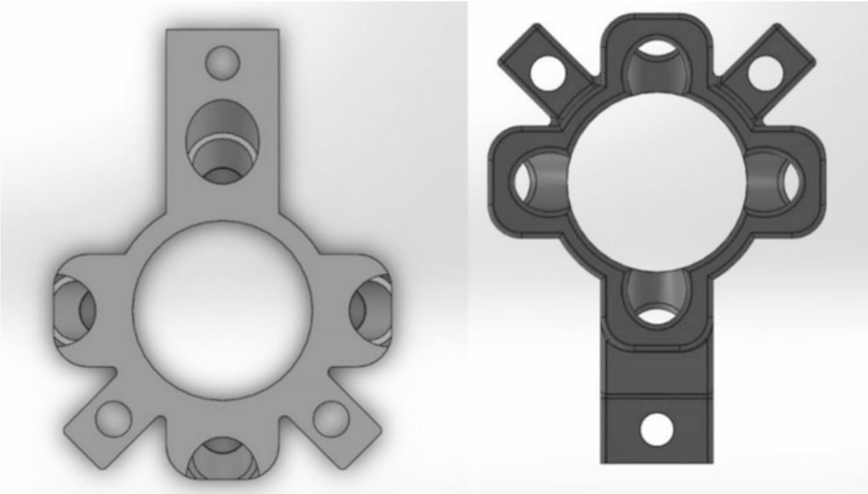


Fig. 7 LED case

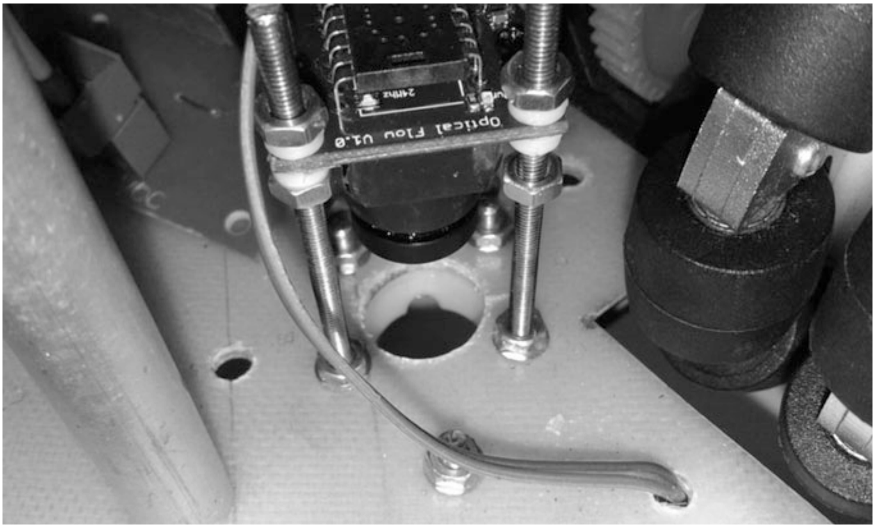


Fig. 8 Sensor placement

Applying lenses with longer focal distance is very good when the quality of the surface is not good or there is some stain or some sort of roughness on it. With long focal length the depth of field and the field of view is bigger, the sensor could see larger area and less sensitive for stains.

5 Mathematics

To define the exact distance, speed and angle values the microcontroller has to calculate them from just the two sensor's Δx and Δy data. After placing the sensors, we could define coordinate systems referring to the sensor and robot orientation (Fig. 9). Below is shown the references [7].

From the coordinate system we could make equations referring to x (1) and y (2) movement, the following calculations are from [7].

$$x_i = \Delta X_R \cdot \sin(\theta_i + \Phi_i) - \Delta Y_R \cdot \cos(\theta_i + \Phi_i) - \Delta\omega \cdot r_i \cdot \cos(\theta_i + \Phi_i) \quad (1)$$

$$y_i = \Delta X_R \cdot \cos(\theta_i + \Phi_i) + \Delta Y_R \cdot \sin(\theta_i + \Phi_i) + \Delta\omega \cdot r_i \cdot \sin(\theta_i + \Phi_i) \quad (2)$$

For the computing we should redistribute the equations and make a vector form for the constants as in (3).

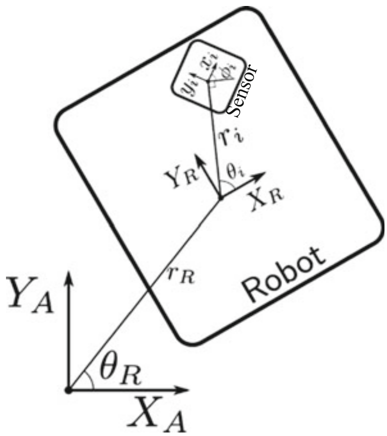
$$\begin{bmatrix} x_i \\ y_i \\ x_i \\ y_i \end{bmatrix} = \Delta X_R \begin{bmatrix} \sin(\theta_i + \phi_i) \\ \cos(\theta_i + \phi)_i \\ \sin(\theta_i + \phi_i) \\ \cos(\theta_i + \phi_i) \end{bmatrix} + \Delta Y_R \begin{bmatrix} -\cos(\theta_i + \phi_i) \\ \sin(\theta_i + \phi_i) \\ -\cos(\theta_i + \phi_i) \\ \sin(\theta_i + \phi_i) \end{bmatrix} + \Delta\omega \begin{bmatrix} r\cos\phi_i \\ r\sin\phi_i \\ r\cos\phi_i \\ r\sin\phi_i \end{bmatrix} \quad (3)$$

After that we separate the constants and the variables (4).

$$A = \begin{bmatrix} \sin(\theta_i + \phi_i) & -\cos(\theta_i + \phi_i) & r\cos\phi_i \\ \cos(\theta_i + \phi_i) & \sin(\theta_i + \phi_i) & r\sin\phi_i \\ \sin(\theta_i + \phi_i) & -\cos(\theta_i + \phi_i) & r\cos\phi_i \\ \cos(\theta_i + \phi_i) & \sin(\theta_i + \phi_i) & r\sin\phi_i \end{bmatrix} \quad (4)$$

For the absolute distances we need to multiply the sensor values with the Moore-Penrose inverse of matrix A (5).

Fig. 9 Coordinate systems



$$A^{+} = A^{T}(AA^{T})^{-1} \tag{5}$$

The final equation defines the absolute distances and rotating angles from each iteration when the microcontroller's program measures the movement (6).

$$\begin{bmatrix} \Delta X_R \\ \Delta Y_R \\ \Delta \omega \end{bmatrix} = A^{+} \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{bmatrix} \tag{6}$$

6 Conclusion

The system we designed is working well, the provided data is good for measuring movement. The provided data is more reliable than the encoder applications. The system is sensitive to ambient light so it is needed to provide proper light conditions for the sensor, especially in dark environment. For best operation, using light sources with the right wavelength and measure and set focal length precisely is necessary. For more precise operation, more sensors in different orientation could be used.

Acknowledgements This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 691942. This research was partially carried out in the framework of the Center of Excellence of Mechatronics and Logistics at the University of Miskolc.

References

1. Németh J, Illés B (2015) Determination of the ratio of centripetal forces in the friction drive used at more places. XXIX. In: microCAD international multidisciplinary scientific conference, University of Miskolc, Miskolc (in Hungarian)
2. Centeye (2013) [Online] http://www.centeye.com/technology/optical-flow/
3. Sorensen DK (2004) Texas A&M University, on-line optical flow feedback for mobile robot localization/navigation
4. Sekimori D, Miyazaki F (2007) Precise dead-reckoning for mobile robots using multiple optical mouse sensors. Informatics in control, automation and robotics II. Springer, pp 145–151
5. Tresanchez M, Pallejà T, Teixidó M, Palacín J (2009) The optical mouse sensor as an incremental rotary encoder. Sens Actuators A 155(1):73–81
6. Avago, ADNS-3080 Datasheet (2007)
7. Bell S (2011) High-precision robot odometry using an array of optical mice

Pose Determination for Autonomous Vehicle Control

Ahmed Bouzid, József Vásárhelyi, Roland Bartók and László Czap

Abstract For the purpose of determining the position and orientation of a moving robot and autonomous vehicle, inertial sensors and magnetometer data are computed in order to enhance GNSS (Global Navigation Satellite System) data accuracy. This paper presents a method called hybrid localization that combines absolute localization, using exteroceptive data, and dead reckoning technique, using proprioceptive data. A positioning method based on dead reckoning technique is developed in this paper.

1 Introduction

Nowadays, objects in movement (autonomous robots in particular) have to know in real time their navigation parameters for better trajectory control. The parameters of navigation intended are: orientation, position, velocity and acceleration. The process of calculating these parameters is called pose determination. Autonomous vehicles should consider also the dragging force, which acts in the drivetrain of the system [1].

In the case of the navigation, IMU (Inertial Measurement Unit) is used for data acquisition to estimate pose information (all those parameters above mentioned parameters) of a moving object [2]. The first application of an inertial unit was a directing system for a torpedo. The system is based on a gyroscope patented by Ludovico Obry in 1894. The usage of inertial units was limited due to high prices

A. Bouzid (✉) · J. Vásárhelyi · R. Bartók · L. Czap
University of Miskolc, Miskolc, Hungary
e-mail: qgebouzid@uni-miskolc.hu

J. Vásárhelyi
e-mail: vajo@uni-miskolc.hu

R. Bartók
e-mail: qgeroli5@uni-miskolc.hu

L. Czap
e-mail: czap@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_30

until MEMS (Micro-Electromechanical Systems) technology appeared. MEMS devices are very tiny and inexpensive, but they provide some systematic and random noises and generate many errors. To solve this problem we need some tools and implement adequate software/hardware architecture.

For a mobile robot or an autonomous vehicle the collision avoidance is very important. A wide range of sensors are used for pose determination, distance to objects, walls and collision avoidance. Better results are obtained from different type of sensors (for example infrared, ultrasonic and laser distance sensors). Using many sensors allows sensor fusion to obtain higher precision results.

After pose determination, the robot behaviour control is relatively easy by using the FRI based Behaviour Description Language or the Bayes-classifier [3].

This paper is organized as follows: the first section gives details about the proposed method for the pose determination and control. Then theoretical background is presented, which is followed by the presentation of the simulation results. Finally, some conclusions are presented. In the paper the expressions autonomous vehicle and autonomous robot are considered equivalent expressions.

2 Proposed Method for Pose Determination

Pose determination is a complex process; it is divided into 4 subsystems. It requires and IMU (Inertial Measurement Unit) of 9 sensors (3 accelerometers, 3 gyroscopes and 3 magnetometers) and GNSS (Global Navigation Satellite System) receiver. In order to process the navigation parameters, the processing unit needs hybridization of sensors data to get accurate estimation. The subsystems are (see Fig. 1) [4]:

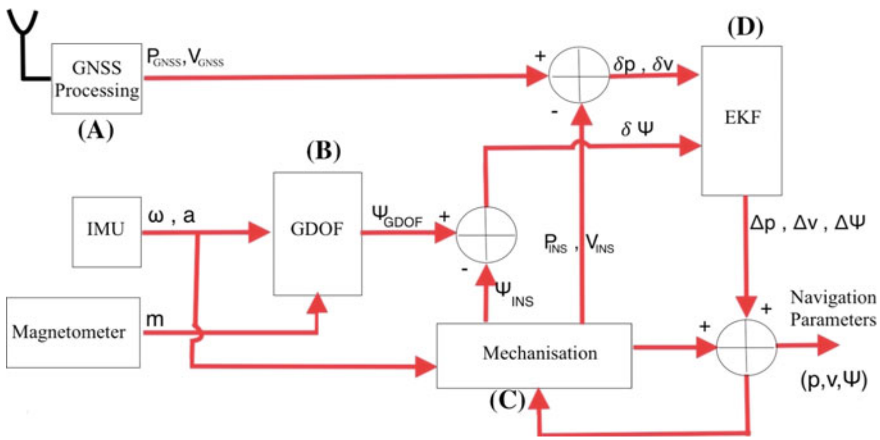


Fig. 1 Block diagram of the pose estimation solution

- GNSS processing **(A)**: uses a GNSS receiver to process the position and the velocity;
- GDOF processing **(B)**: the Gradient-Descent Orientation Filter is an algorithm to estimate the orientation using the 9 sensors above-mentioned [5].
- Mechanization **(C)**: is the step intended to estimate the position and the orientation using 3 accelerometers and 3 gyroscopes.
- EKF processing **(D)**: Extended Kalman Filtering is a robust algorithm that estimates the states of the system in a noisy environment.

In Fig. 1 the following notations were used: P_{GNSS}, V_{GNSS} are respectively the position and the velocity after GNSS processing; P_{INS}, V_{INS} are respectively the position and velocity given by the mechanisation; δp and δv are the difference between the results of GNSS processing and Mechanisation; ω, a are respectively the angular velocity and the acceleration given by the 6 DoF IMU; m is the data of the magnetometer. Ψ is the orientation vector. Δp, Δv and Δ Ψ are the errors computed by the EKF.

In the paper 2D positioning is treated in the plan map of the autonomous vehicle. In real time, the robot calculates his position and stores it in order to reconstitute its trajectory. The block diagram of the proposed system is given in Fig. 2. The processing unit uses raw data of the accelerometer sensors to measure the acceleration of the robot. The used sensor is ADXL335 that is a set of 3 accelerometers (one for each degree of freedom) orthogonally placed on a single chip providing analog data. A secondary processing unit (ZYNQ device) is responsible for the acquisition of the analog data (analog to digital conversion) from sensors and the position estimation. Then the processing unit puts the processed results in a serial frame and sends via the UART to the main processing unit (STM32). For more accurate data, the main processor sends an interrupt to the secondary processing unit each time the wheels stop to reinitialize the dead reckoning equations and avoid divergence.

The primary processing the STM32—ARM Cortex M4 based board, while the secondary boar is a Zynq System on Chip contain a dual ARM Cortex A9 MP Core and Neon Single precision Floating point unit for each processor unit. The Zynq also contains Programmable Logic (PL) for user based algorithm implementation. The Zynq based board is Zedboard, which finally will do all the computation for the robot.

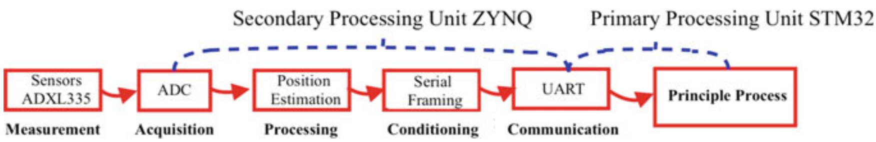


Fig. 2 Block diagram of the system

3 Mathematical Background of the Dead Reckoning System

Knowing that the calculations are iterative, the equations for dead reckoning are:

$$\begin{cases} x = \frac{1}{2} a_x dt^2 + v_x^- dt + x^- \dots \\ y = \frac{1}{2} a_y dt^2 + v_y^- dt + y^- \dots \end{cases} \qquad (1,2)$$

where *x, y* are the coordinates (position) of the material point (the Robot); a_x, a_y are the accelerations measured by the sensors, *dt* is the sampling time; v_x^-, v_y^- are the velocities calculated during the previous iteration, x^-, y^- finally are the coordinates of the previous position (Eqs. (3) and (4)). The velocities from Eqs. (3) and (4) must be injected in the Eqs. (1) and (2) on each iteration. The velocities are measured as follows:

$$\begin{cases} v_x = a_x dt + v_x^- \dots \\ v_y = a_y dt + v_y^- \dots \end{cases} \qquad (3,4)$$

4 Simulations

In order to validate the algorithm, the system has been modelled in Matlab Simulink and simulated considering ideal conditions (non-noisy environment). The robot moves on a 2D plane (X and Y axis) and starting from the origin (0, 0).

In the simulation process, some conditions have been supposed and they were considered three scenarios (*with the sampling rate of dt = 0.1 s during 100 iterations*):

4.1 First Scenario

Figure 3 describes the trajectory of the robot with the following conditions: initial velocity for X axis is $v_{xo} = 1\,\text{m/s}$ and the acceleration on the *y* axis is $a_y = 9.8\,\text{m}\,\text{s}^{-2}$ (with $v_{yo} = 0$, $a_x = 0$). The simulation results are given below with the already mentioned scenarios.

The graphic looks like the trajectory of a horizontally launched projectile.

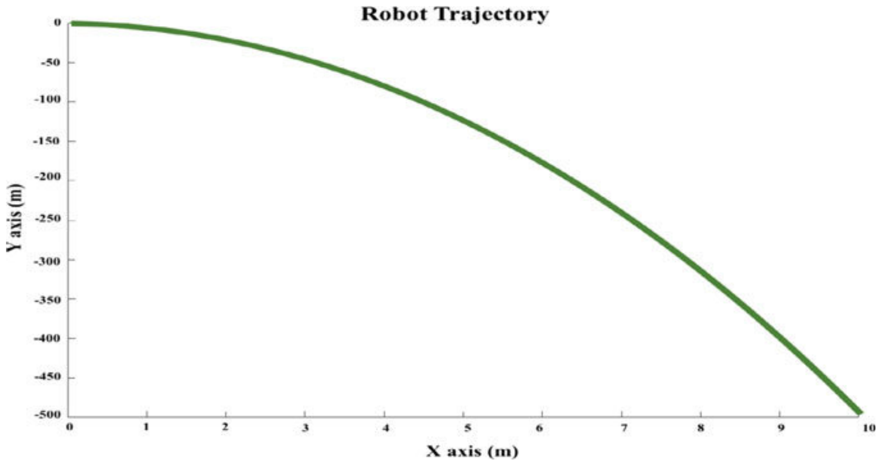


Fig. 3 Trajectory of the robot on the 1st scenario

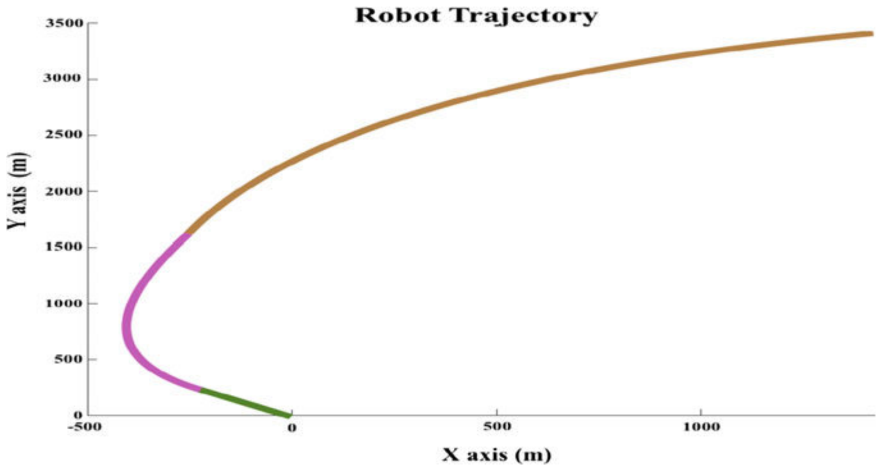


Fig. 4 Trajectory 2 of the robot on the 2nd scenario

4.2 Second Scenario

The simulation is done without initial velocities. The variations on the accelerations are given in Fig. 6 (left part).

(*2nd Scenario:* $v_{xo} = v_{yo} = 0$, $a_{x1} = -1\ \mathrm{m\,s^{-2}}$, $a_{x2} = 1$, $a_{y1} = 1$, $a_{y2} = -1$)

Figures 4 and 5 describes the trajectory of the robot with the conditions described in Fig. 6 without initial velocities $v_{xo} = v_{yo} = 0$.

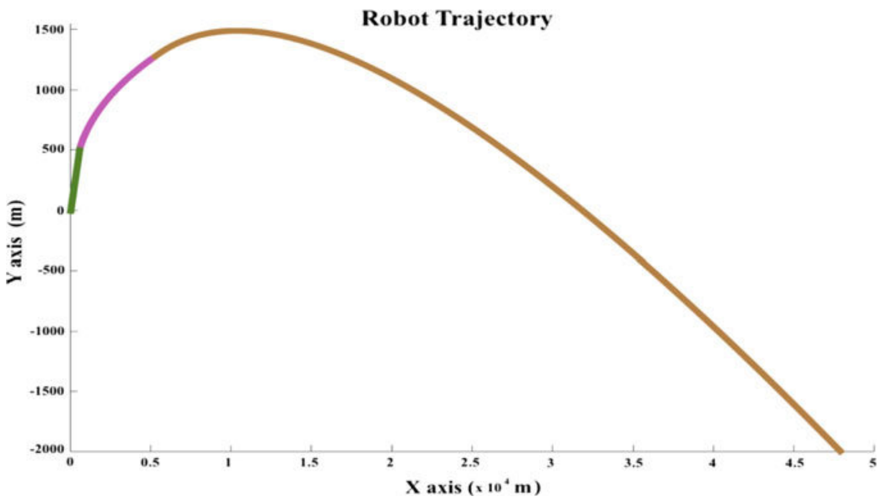


Fig. 5 Trajectory of the robot on the 3rd scenario

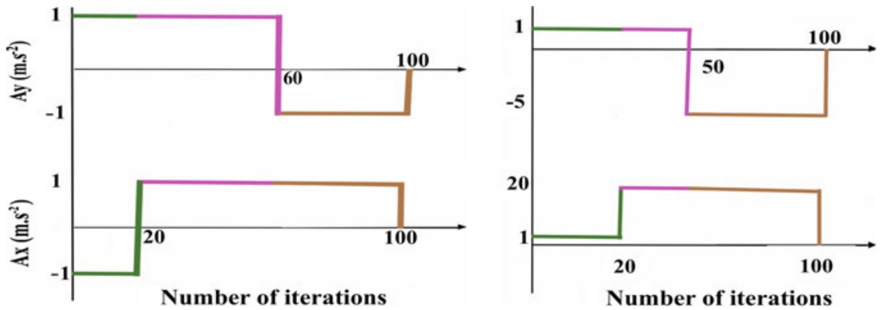


Fig. 6 Accelerations versus number of iterations (*Left* 2nd scenario. *Right* 3rd scenario)

4.3 Third Scenario

The simulation is done without initial velocities. The variations on the accelerations are given in Fig. 6 (right part).

(*3rd Scenario:* $v_{xo} = v_{yo} = 0$, $a_{x1} = 1\ \mathrm{m\,s^{-2}}$, $a_{x2} = 20$, $a_{y1} = 1$, $a_{y2} = -5$)

5 Conclusions

The simulation results confirmed our expectations. The dead reckoning system algorithm gives good results in ideal environment. There were presented three scenarios for robot movement in three different conditions that were studied for the evolution of acceleration versus the number of iterations.

The next step will be to model the accelerometers and introduce noise in the system to be close to real conditions. Then implement the solution in ZYNQ device in order to experimentally validate the proposed model.

Acknowledgements The research work was (partially) supported by the Hungarian Scientific Research Found grants OTKA 29326 and Fund for the Development of Higher Education FKFP 8/2000 project. This research was (partially) carried out in the framework of the Center of Excellence of Mechatronics and Logistics at the University of Miskolc.

References

1. Németh J, Illés B (2015) Determination of the ratio of centripetal forces in the friction drive used at more places. XXIX. In: microCAD international multidisciplinary scientific conference, Miskolc, University of Miskolc (in Hungarian)
2. Noureldin A et al (2009) Performance enhancement of MEMS-based INS/GPS integration for low-cost navigation applications. IEEE Trans Veh Technol 58(53):1077–1096
3. Bartók R et al (2016) Embedded behavioral model implementation. In: Proceedings of the 17th international carpathian control conference (ICCC), Slovak Republic. IEEE
4. Bouzid A et al (2016) Implementation of INS/MAG/GNSS hybridisation technique for pose determination based on SoC and low cost sensors: theoretical approach and synthesis. In: 17th Carpathian control conference (ICCC), Slovak Republic. IEEE
5. Madgwick SOH (2010) An efficient orientation filter for inertial and inertial/magnetic sensor arrays. University of Bristol, UK

Description of a Method for the Handling of Customer Needs in Logistics

Béla Illés, Róbert Skapinyecz and György Wagner

Abstract The paper describes the application of the QFD method, a technique used for the evaluation and proper realization of the different customer expectations, in the quality management of logistics systems. Both the theoretical basics of the method, as well as the main steps of its implementation are introduced. The implementation itself is presented with the help of a practical example that is strongly related to both the logistics and the automotive industries, as the latter especially relies on complex supply chains that require the extensive utilization of quality management tools. Besides the previous, the paper also provides an overview of all the possible areas of utilization for the QFD in the logistics industry. Therefore, the described method can have a great value from both the academic and the industrial perspectives.

1 Introduction

By the present day, logistics has become a common concept. The word "logistics" can be seen on notices on the vehicles of transport companies. The excellent "logistical quality" is advertised on the stationery posters of facilities. But really, what is in the background of the concept of "logistics"?

Logistical problems and tasks arise in all areas of the industry and economics. This puts a lot of requirements on the comprehensive field that is today called as logistics. In the meantime, logistics is under constant change due to the technical innovations and the changing social and political boundary conditions.

B. Illés (✉) · R. Skapinyecz · G. Wagner
University of Miskolc, Miskolc, Hungary
e-mail: altilles@uni-miskolc.hu

R. Skapinyecz
e-mail: altskapi@uni-miskolc.hu

G. Wagner
e-mail: wagner@iit.uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_31

The diverseness and the dynamics present a great challenge for the logistics professionals of the future, as intelligent and economical solutions are awaited from them.

Logistics assures,

- the flow of materials and the connected information,
- the flow of waste materials and the connected information,
- the change of the location of objects (persons, animals, things) in order to avoid losses (quantitative differences), to avoid changes in the quality of objects (damage), while in the meantime new objects are not created. Logistics can be distinguished from manufacturing or processing procedures which are generally aimed at the production or modification of new products.

As it can be seen from the previous, the quality of logistics processes affects the overall performance of a logistics system in multiple ways, therefore it is an integral part of the value creation process in logistics [1]. Quality itself can be basically measured through customer expectations. For these reasons, the proper handling of the customers is essential for all companies. In the long term, the company will only be successful if it perfectly knows its customers and their expectations, and it shapes its products, processes and systems in such a way that they properly and effectively fulfil these expectations.

Many companies already applied special software for the registration and maintenance of their client data. The purpose of using such systems is to attain the expected benefits and income, thereby assuring the success of the company. The variability of such software is very large. Its capabilities span from the simple customer data base to data mining tools which generate new implications from the customer data, while also including work processes which realize automatic information distribution. These techniques became especially important today with the rapid deployment of big-data applications.

One of the tools that can utilize the results of the previous techniques is the QFD (Quality Function Deployment), a method that is also known in the automotive industry. It basically provides a controllable way to effectively transform the expectations of the customers into measurable technical data. QFD relates to the first principle of Lean thinking, which is identifying the customer and the customer's value [2, 3]. In the paper, the most important theoretical basics of this method will be described, together with a practical example which is related to both the logistics and automotive industries.

2 QFD—A Method for the Realization of Customer Expectations

QFD is an abbreviation standing for Quality Function Deployment. It originally came from Japan, the year of its appearance was 1966. In Japanese, the term means Hin Shitsu, Ki No, Ten Kai, which means the "Presentation of Characteristic

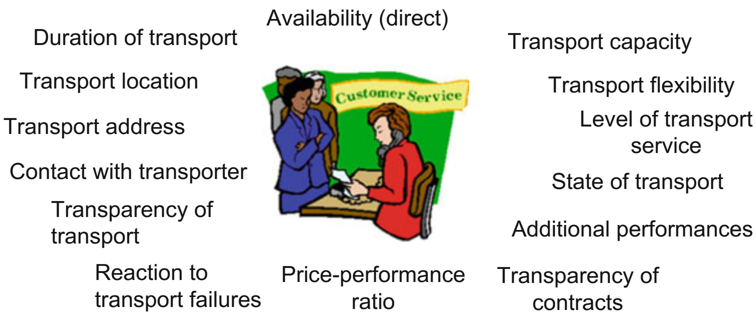


Fig. 1 Customer's expectations as the starting points of QFD

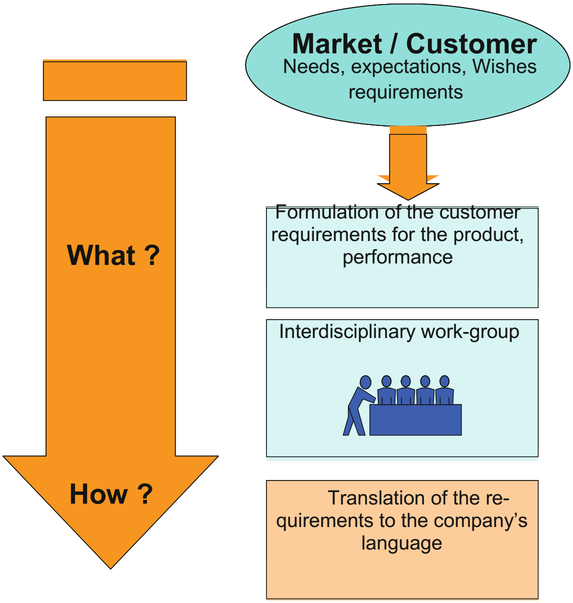


Fig. 2 QFD—basic philosophy

Functions". The founder of the QFD was Yoji Akao. His book on the QFD was presented in Japan in 1978 (see [4]). A German edition of this book was also published, first in 1992 [5].

The goal of the QFD method is to properly select the customer's expectations and reformulate them into technical or organizational solutions. Therefore, the starting point for the QFD are the customer's expectations themselves (Fig. 1). However, the knowledge of the customer's expectations alone is not enough, as they consistently have to be realized as well.

The QFD method has a systematic, multi-step procedure. Generally, it can be stated that with the help of the QFD, the question "What should be done?" is reformulated into the question "How it should be done?" (Fig. 2).

Table 1 Advantages and disadvantages of the QFD

Advantages	Disadvantages
Assures the general customer orientation at the company	Significant time expenditure
Preventive disposition	Personal expenditure
Sector dependent	The method is less known outside quality management
Systematic, proven procedure	The results are dependent on the competencies and the collaboration skills of the participants
Applicable in multiple manner	The purpose oriented usage requires the special knowledge of the method
Interactions are consciously investigated	It partially requires significant abstraction skills
Needs team work and demands communication over different areas	It reduces the investigation to a few relevant parameters

The QFD aims to achieve both the consistency and completeness of the goal. For this reason the best is to utilize team work with the participation of all the concerned partners. For instance in case of the development of a new product, the team needs to develop all of the followings: the marketing, the construction, the work preparation, the production, the maintenance and the service background. The usage of the QFD in logistics also covers the marketing, the logistics planning and the involved logistics areas, among them controlling.

In the literature [4], three starting points are differentiated in case of the QFD:

- the excluded start according to Akao (comprehensive QFD),
- the 4-phase model according to the American Supplier Institute (ASI),
- the King matrices [6].

At this place we do not wish to compare these different approximations. Table 1 lists and summarizes the advantages and the disadvantages of the QFD.

In the followings, the ASI 4-phase model will be described, which is often used in industrial environments and for that reason it is especially applicable in the field of logistics as well.

3 General Procedure of the QFD

The individual phases of the ASI 4-phase model, together with their inputs and results are shown in Table 2.

From Table 2 it can be seen that the most important outputs of the individual phases are the inputs of the following ones. Thereby the continuous stream of arrows is realized.

For transparent documentation, the usage of a worksheet has proven to be the best practical solution. The structure of the worksheet can be seen in Fig. 3.

Table 2 ASI 4-phase model

Phase	Name of the phase	Input = What	Output = How
Phase 1	Product-planning	Purchaser expectation	Product parameters (performance list)
Phase 2	Component-planning	Product parameters	Component parameters (given product)
Phase 3	Process-planning	Component parameters	Production prescriptions (process data)
Phase 4	Production-planning	Process data	Production instructions (task and checking prescriptions)

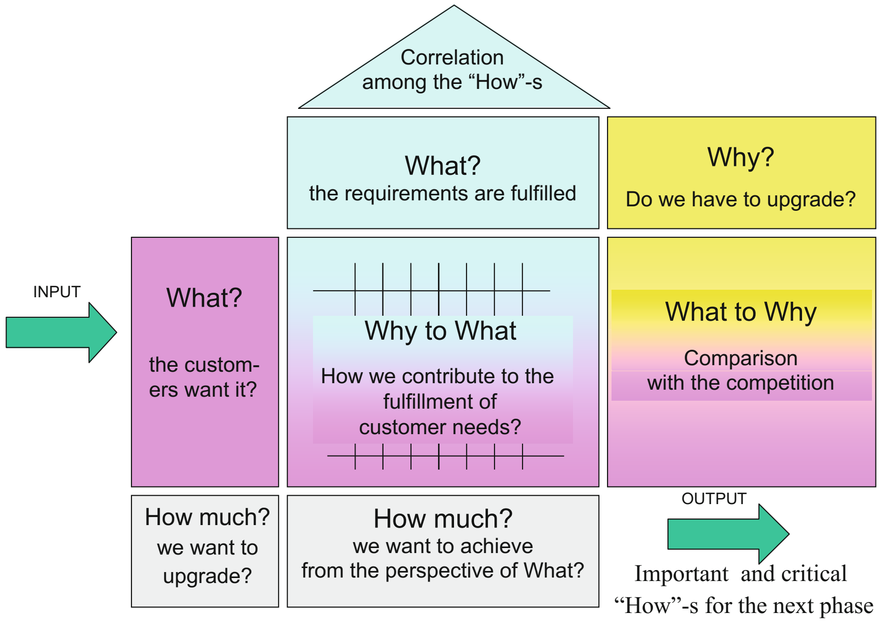


Fig. 3 Basic structure of a QFD work sheet

4 Example for an Automotive Industry Application

The methodology of the QFD will be shown through an example. The task is to develop a new transport vehicle under the codename "Citysprinter" for an urban distribution center, in order to supply various shops in the city. In the followings,

Table 3 Customer expectations for the Citysprinter

Characteristics	Expression
Good safety	• Safety of the driver • Safety of the cargo (refrigeration, theft, safety against impairment)
Very good travel characteristics	• Turning radius <3 m
Good acceleration	• From 0 to 100 km/h in 10 s
Low fuel cost	• Less than 3.500 euro/year
Lower initial cost	• Less than 25.000 euro
Simple storage	• Up to 1500 kg accessory cargo • Loading/unloading should be possible from three sides

the individual steps of this task will be presented in order, according to the procedure of the QFD method.

4.1 Registration of the Customer Expectations

The customer expectations represent the first "What to do" question. Above all, this "What to do" describes the expectations of the customer. A lot of useful information is derived from the CRM (Customer Relationship Management) system. In principle, the historical data is stored in the CRM system, while the consideration of the actual trends is also necessary.

Important questions regarding the registration of the customer expectations:

- Who are my customers?
- What is the importance of these customers for my company?
- What kind of wishes and expectations my customers have?
- How important these wishes and expectations are for my customers?

The potential customer of the "Citysprinter" would like to get the vehicle with the characteristics shown in Table 3.

4.2 Evaluation of the Customer Expectations According to Kano

The evaluation of the customer expectations will first be realized according to the triple classification introduced by Kano [7]. Kano [7] distinguishes:

- The basic expectations of the customer (basic demands),
- The function related expectations of the customer (performance demands) and
- Innovative properties (inspirational demands).

Table 4 Characterization of the customer expectations according to Kano

Characteristics	Expression	Basic expectation	Functional expectation	Innovative properties
Safety	Safety of the driver Safety of the cargo (refrigeration, theft, safety against impairment)	X		
Travel characteristics	Turning radius <3 m		X	
Acceleration	From 0 up to 100 km/h in 10 s		X	
Operational cost	Less than 3.500 euro/year		X	
Price	Less than 25.000 euro		X	
Loading	Up to 1500 kg accessory cargo Loading/unloading should be possible from three sides		X	

The lastly specified points are often the decisive demands from the point of the actual purchase. In case of the Citysprinter, the evaluation of the customer demands give the results shown in Table 4. From the table, it can be seen that most of the customer expectations are directly named by the customers themselves. It can be concluded that the innovative characteristics are missing. Under such conditions, the “Citysprinter” would probably not be an outstanding hit among the customers.

4.3 Weighting of the Customer Expectations

Not all of the customers’ expectations are equally important; they have to be weighted against each other. Any kind of known weighting procedure can be applied. In our example the method of pairwise comparison is utilized (see Table 5). The principle of the algorithm is the following: each time, two characteristics are compared with each other. If one of them is more important than the other, then it gets 2 points, while the other gets 0. If both of them are equally important, then they both get 1 point. After that, the normalization is carried out, where the highest value will be 10 (Table 5).

4.4 Derivation from the Technical Expectations (HOW)

In this step, the “WHAT” purchaser expectations are translated into the “HOW” for the “Citysprinter” product. The “HOW” means the desired characteristics of the product. We decided that the product should fulfil the state of the art technical

Table 5 Weighting of the customers' expectations with pairwise comparison and normalization on a 1–10 scale

Pairwise comparison matrix 2—more important 1—equally important 0—not important	Travel characteristics	Acceleration	Operational cost	Price	Loading	Summary	Normalization on 1–10
Travel characteristics		1	0	2	2	5	7.1
Acceleration	1		0	1	1	3	4.3
Operational cost	2	2		2	1	7	10.0
Price	0	1	0		0	1	1.4
Loading	0	1	1	2		4	5.7

requirements in the case of the vehicle's dimensions, the motorization, the hybrid engine, the production costs and the payload.

4.5 Quantitative Expression of the Technical Requirements

Now the objective values can be determined according to our technical requirements, but it is more preferable to evaluate the technical significance and the competitiveness first. This is the step where the connections between the product characteristics (HOW) and the customer expectations (WHAT) are determined. Here it is asked that how strongly each unique product criteria affect the fulfilment of the customer expectations (correlation). The strength of the connection between the HOW and the WHAT is evaluated with the help of a previously fixed scale.

In our example, the evaluation is carried out in four steps:

0—no connection,
1—narrow/weak connection,
2—medium connection,
3—high level/strong connection

4.6 Calculation of the Technical Meanings of the "HOW"-s

For the evaluation of the Technical meaning, the sum should be formed by taking into account the importance and the suitable demand in relation to all products. Based on the values, the priorities among the characteristics of our technical solution for the "Citysprinter" is decided.

In our example, the following priorities are revealed:

1—the motorization has the highest priority,
2—the manufacturing costs has the second priority,
3—the hybrid engine has the third priority,
4—the vehicle dimensions have the fourth priority,
5—the useful loading has the lowest priority.

4.7 Comparison with the Concurrent Products (Comparison of the "WHAT")

Now the comparison between the own and the concurrent products is carried out. By doing this, the own products and the similar products of the competitors are determined and listed. This creates a documentation regarding the market image of the own and the concurrent products, while it also gives the prestige of the different products in relation to each customer expectation. The data which are the basis of the evaluation has to be specifically produced, or it could be provided by the sales department based on the systematic study of the market.

4.8 Investigation of the Interdependencies Among the Characteristics (Correlation of the HOW to the HOW) and the Determination of the Direction of Optimization

It should be emphasized, that the parameters are mutually strengthen (positive correlation) or weaken each other (negative correlation). Based on the correlation, critical constellations are revealed which should be resolved. This can be done for example through the modification of the prescribed values ("how much").

Now the House of Quality in phase 1 can be fully completed. The result is shown below in Fig. 4.

4.9 Difficulty at the Realization and Technical Comparison with the Competition

In the next step, the difficulty of the realization of the "HOW" is evaluated. Among others, the evaluation contains the statistics, the tests, the planning, the experiences regarding the improvements, the reclamations, the terms of the guaranteed performances and for the waste disposal, the legal terms and the mode of realization of all the previous. As can be seen from the example, the "House of Quality" provides a very good support as a worksheet, and also for the documentation.

For the further development of our "Citysprinter", at least three extra Houses of Quality are needed. The first one is required for the deduction of the characteristics of the constituents of the concrete product, the second one is required for the manufacturing prescriptions and the third one is required for the deduction of the process instructions (services, checks etc.).

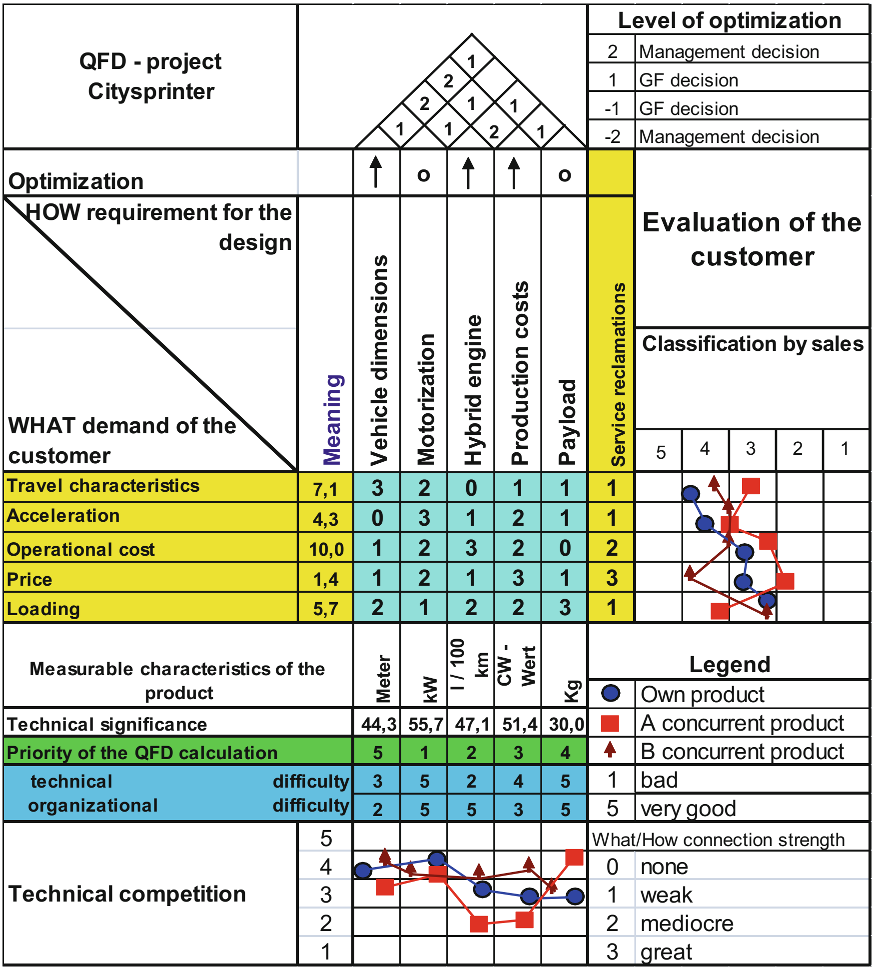


Fig. 4 The accomplished House of Quality after the close of phase 1

5 Important Matrices

The method uses two important matrices. The first is the component matrix and the second is the process matrix. We have to examine these matrices.

5.1 *Component Matrix*

In the component matrix, the determination and the specification of the critical product components is carried out. The selected product characteristics (How), their (vertical) and prescribed values (“How much”) are taken over to a new “House of Quality” (component matrix).

For all products (all of the own products and all of the concurrent products) an own matrix is formulated. The product functions such as stability, driving, etc. are put vertically into this matrix (these can be regarded as “What”) and are supplemented by such aspects as transportation, selling and delivery. Horizontally, the product components and aspects such as ordering, accounting and transportation are inserted as “How”. In addition, the estimated costs of each component are numbered. The function costs of each component for the “From which to How” are given in percentage, in order to have an overview of the cost structure and its distribution. From the representation, it will be visible if for example the concurrent products are able to satisfy the same functions with fewer components.

5.2 Process Matrix

The harmonization of the procedures and the processing of the critical process elements, the optimal values of the process parameters, the costs and the reliability are done in the third phase of the QFD for further follow-up. The process during the process planning has to be determined in such a way, that the specifications of the components in the manufacturing could be kept in a reproducible manner. Through the optimization of the process (e.g. by statistical experimental methodology) often principal improvements can be gained without significant extra investment.

6 Other Areas of Utilization of the QFD Method in Logistics

As it can be seen, the QFD is a general methodology at the company in order to maintain the continuous customer orientation. It is easy to apply it for a new task, at a specific company and at a given sector. Inside logistics, its utilization is possible for example in the areas shown in Fig. 5.

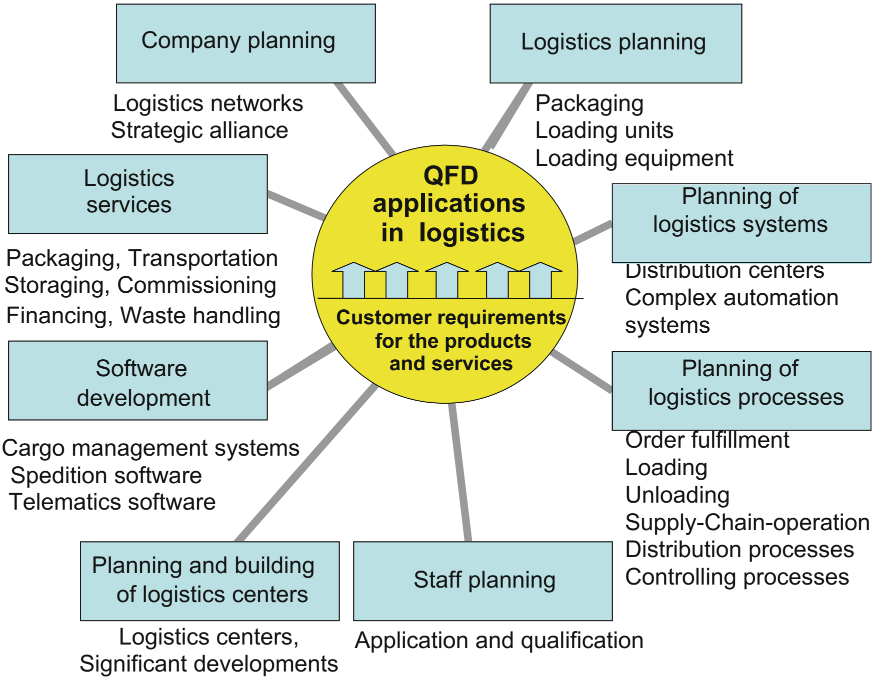


Fig. 5 Examples in the fields of application of QFD in logistics

7 Conclusions

In the paper, we presented the QFD method, which is a well refined tool for the evaluation and quantification of the customer needs and expectations. The method is becoming increasingly applied in the field of logistics, while it is also known in the automotive industry. For this reason, we chose to describe the QFD through a practical application that relates to both industries, which represents very well the strong connection between these two important fields of engineering.

Besides the presented example, the QFD can be applied in various other fields of logistics as well, areas that are also crucial in supporting the modern automotive manufacturing processes. These areas include the planning of logistics systems, the planning of logistics processes, the proper development of cargo management systems and software, the efficient design and operation of complex supply chains, the proper utilization of controlling processes and many other essential areas.

It is important to see that the QFD is just one tool among the multiple methods which are utilized in the quality management of modern logistics systems. Other important tools are for example the prevention methods utilized for the avoidance and reduction of failures (methods such as FMEA, Fault Tree Analysis, Poke-yoke and others), the methods of Benchmarking and Business Process Reengineering, the crucial field of Statistical Process Control, or the very well-known and widely applied methodology of Kaizen with its various techniques. The latter is especially important, as it forms one of the backbones of the Lean philosophy, which again plays a vital role in the efficient operation of the modern automotive manufacturing system [1, 8]. All of these examples show that the continuous development of the quality management methods used in logistics has a wide and positive effect on many other related areas, out of which one outstanding beneficiary is the automotive industry.

Finally, it should be noted that such modern trends like the extensive use of big data analysis and the rapid spread of the internet of things (trends that together define the concept of the so called "industry 4.0") also support the wider use of such data-intensive techniques as the QFD. As the accumulated data related to customer expectations grows with an almost exponential rate, the importance of these quality management tools will grow accordingly in both the logistics and automotive industries.

Acknowledgements "This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 691942". "This research was (partially) carried out in the framework of the Centre of Excellence of Mechatronics and Logistics at the University of Miskolc".

References

1. Tamás P (2016) Application of value stream mapping at flexible manufacturing systems. Key Eng Mater 686:168–173
2. Kovács Gy (2012) Productivity improvement by lean manufacturing philosophy. Adv Logistic Syst: Theor Pract 6(1):9–16
3. Kovács Gy, Illés B (2011) Productivity improvement by application of lean manufacturing, conference proceeding, international scientific conference (MASXXI 2011), ISBN:978-959-250-693-0, pp 1–6
4. QFD—quality function deployment/ausgearb. von der Arbeitsgruppe 132 "Quality Function Deployment". Hrsg.: Deutsche Gesellschaft für Qualität e.V. (DGQ).—Berlin; Wien; Zürich: Beuth, 2001. DGQ-Band; 13–21 ISBN 3-410-32899-8
5. Akao Y (1992) QFD—Quality Function Deployment: Wie die Japaner Kundenwünsche in Qualitätsprodukte umsetzen. Verlag Moderne Industrie, Landsberg. ISBN 3-478-91020-6
6. King B (1994) Doppelt so schnell wie die Konkurrenz; dt. Übersetzung: Kossmann; Hofstetter; Lange; Grohn; St. Gallen; gfmt
7. Kano N, Seraku N, Takuhashi F, Tsuji S (1984) Attractive quality and must-be-quality. Hinshitsu: J Japan Soc Qual Control 39–48
8. Tamás P (2016) Application of simulation modeling for formation of pull-principled production control system. J Prod Eng 19(1):99–102

Sensorless Determination of Load Current of an Automotive Generator Applying Neuro-Fuzzy Methods

Csaba Blága

Abstract This paper presents a sensorless method for determination of the load current of an automotive generator applying a neural-fuzzy implementation. We developed a simulation model of the automotive generator and its voltage regulator in order to get information about its behaviour at different operating conditions. The model takes into consideration the nonlinearity caused by the saturation of the magnetic flux and the effect of the shaft speed on to the internal impedance. The simulated results are compared to those that are available in the literature and to the results gained from real system measurements. A laboratory test rig was developed to study the operation of the automotive generator and voltage regulator in different conditions. In contrast to general spread methods in this area the measurement results are plotted in 3D to emphasize the hidden operation fields. In the operation of the system the parameters of the DFM—Digital Field Monitoring—signal have important roles. Both the frequency and even more the duty of the DFM signal carry on important information about the condition of operation of the automotive generator and its voltage regulator, especially about the load current. This has an important influence on the whole system of the electric circuit of a car starting from battery to end consumers, different ECUs—Electronic Control Units—fuel consumption and emission of the ICE—Internal Combustion Engine. Applying neural-fuzzy theory, we could realize a sensorless method for this aim.

Keywords Automotive generator · Voltage regulator · Sensorless current determination · Neural-fuzzy implementation

1 Introduction

This paper presents a simulation model of a voltage regulator of an automotive generator (or alternator) in order to get information about its behavior at different operating conditions. The role of the voltage regulator is to keep the voltage of the

C. Blága (✉)
University of Miskolc, Miskolc, Hungary
e-mail: elkblaga@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_32

alternator of the car at a constant value that it is necessary to charge the battery and to supply the consumers. Nowadays the most spread passenger cars have an electric circuit of “12 VDC”. It is known that even the starter (lead-acid) battery has a greater voltage than 12 VDC. If we take into consideration that the battery has to be charged, then we can assume that it is necessary to produce at least 14 VDC quite at low operation speed of the internal combustion engine (for example at idling) to avoid the discharge of the battery especially in urban traffic. The speed of an Otto-engine can increase from about 800 rpm to 6500 rpm. If the induced voltage is somehow proportional to the rotational speed of the shaft, then it means the voltage would increase over 100 V and all the consumers would blow out. The voltage regulator is set in such a way that the voltage is 14.5 VDC. So the voltage regulator has an important role, especially nowadays, because there are many electronic equipment and apparatus in a car. Also there are new strategies concerning to the charging methods and operation of batteries that requires important interventions from the voltage regulator. The voltage regulator is also controlled by the central electronic control unit of the car through different communication lines as LIN, CAN or FlexRay. Integration of voltage regulator into the diagnostic system of vehicle has become a common service, also.

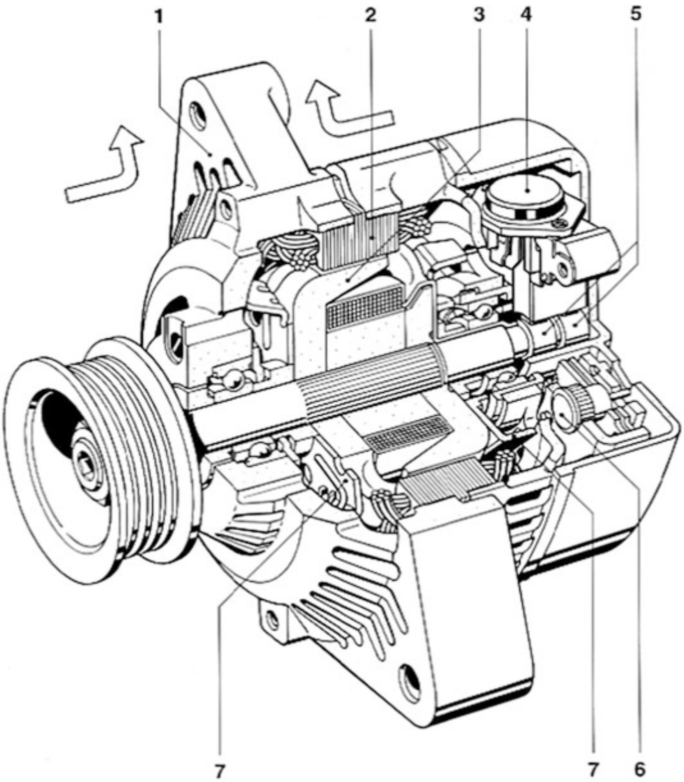


Fig. 1 Construction of an alternator of a car: *1* housing, *2* stator, *3* rotor, *4* transistor voltage regulator with brush holders, *5* collector rings, *6* rectifier, *7* fan [2]

Nowadays, special generators have already appeared on the market and built into cars, the so called starter-generators, as it is described in paper [1]. These are suitable to start the internal combustion engine as a starter motor and afterwards to operate as a generator.

In order to be able to simulate the behaviour of the voltage regulator we should know the construction and the operation of the entire system. The alternator basically is a synchronous generator as it is presented in Fig. 1 and described in many references e.g. [2]. It can be observed that the housing contains a rectifier and a voltage regulator, nearby other elements as brushes, fan, etc.

2 Simulation Model of an Alternator

To create the simulation model of the alternator we started from a simplified version. An improvement of this model has been carried out by introducing into the model the followings:

- the saturation of magnetic flux and
- the effect of rotational speed of the shaft on the internal impedance of the synchronous generator.

The induced voltage depends on the magnetic flux. The saturation of magnetic flux has a great influence on the operation of an alternator [3, 4]:

$$\Phi = \Phi_{sat}\sin\left(\mathrm{arctg}\left(\frac{F_f I_{ex}}{\Phi_{sat}}\right)\right), \tag{1}$$

where Φ_{sat} = saturated value of flux, F_f = flux factor, $F_f = \Phi/I_{ex}$, and I_{ex} = the excitation current.

The induced voltage can be calculated as follows:

$$V_i = \mathrm{k}\Phi\,\omega, \tag{2}$$

where k = machine constant, ω = angular velocity [rad/s].

Moreover the frequency depends on the rpm of the shaft, so the internal impedance also depends on the rotational speed of the shaft.

$$X = \omega L = 2\pi f L = 2\pi\left(\frac{pn_a}{60}\right)L = \frac{\pi\, pniL}{30}, \tag{3}$$

where L = inductance [H], f = frequency [1/s], p = number of pole pairs, n_a = rotational speed of the shaft of the alternator [RPM], n = rotational speed of

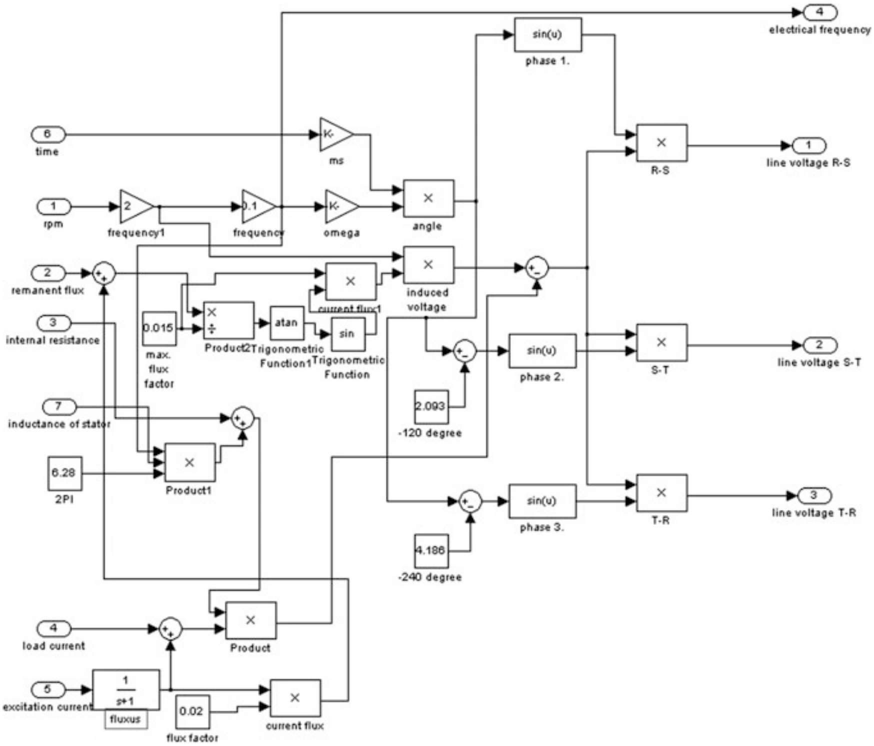


Fig. 2 Simulation model of an automotive generator. Inputs: *1* rpm, *2* remanent flux, *3* internal resistance, *4* load current, *5* excitation current, *6* time, *7* inductance of stator. Outputs: *1* line voltage R-S, *2* line voltage S-T, *3* line voltage T-R, *4* frequency

the internal combustion engine [RPM], i = rate of the mechanical transmission between the internal combustion engine and the alternator.

We had to build in a feedback from the output current. The output current is the load of the alternator. There will be a voltage drop inside of the alternator and the output voltage will decrease if the load current increases (Fig. 2).

Concerning to the external behaviour of the alternator we can make two affirmations:

- the higher the speed is, the higher the output voltage is (taking into consideration the nonlinearity caused by the saturation of the magnetic flux), and
- the higher the load current is, the lower the output voltage is.

The rectifier model is presented in Fig. 3.

A simulation model of the battery is presented in Fig. 4.

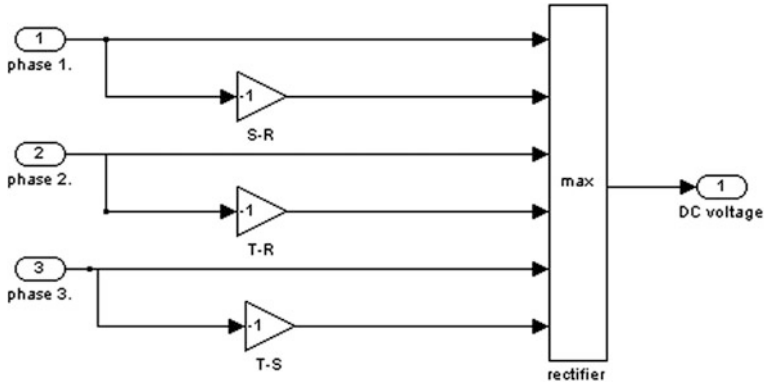


Fig. 3 Model of the rectifier. Inputs: *1* phase 1, *2* phase 2, *3* phase 3. Output: DC voltage

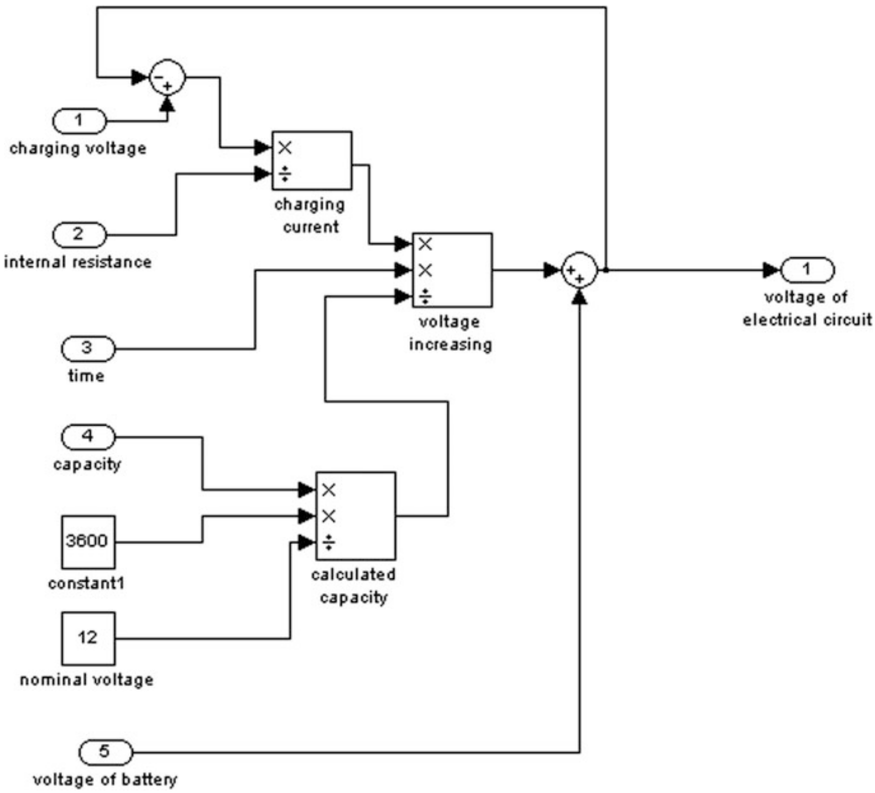


Fig. 4 Model of the battery. Inputs: 1—charging voltage, 2—internal resistance of battery, 3—time, 4—capacity of the battery, 6—internal voltage of battery. Output: 1—network voltage

3 Performance Curve

In order to have the performance curve, the alternator has to operate in such a condition that the regulator should not have any influence on its operation. So this would give an unregulated characteristic: the load current is changed at different rpm in order to have constant voltage. In real cases the alternator has its regulator built in. Maybe it is difficult, but it would be necessary to separate its influence. Similar happens on a test bench where in the first step the level of regulated voltage should be found. The test voltage should be set below this value, that means the voltage regulator will not control the excitation field. In this case we have an easier situation because the voltage regulator can be eliminated from the system. Also the battery is removed because its internal voltage and charging current would modify the behaviour of the alternator.

It can be observed (Fig. 5) that the output voltage (2) of the alternator was kept constant, but the load current (1) increases and there is no any change in the excitation current (3). The performance curve is presented in the Fig. 6.

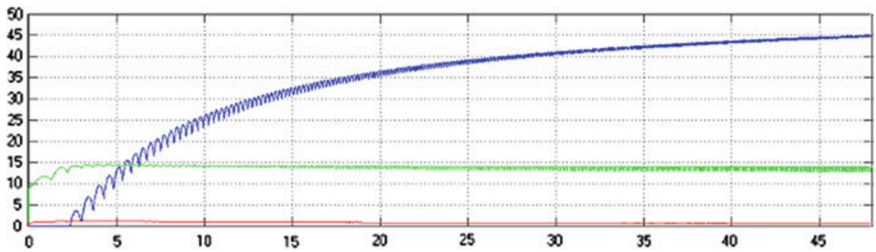


Fig. 5 Time domain analysis of the performance curve: *horizontal axes* represents the time [s]; on the *vertical axes* there are presented the following quantities: *1* the load current [A], *2* output voltage [V], *3* excitation current [A]. (The scale should read as it is.)

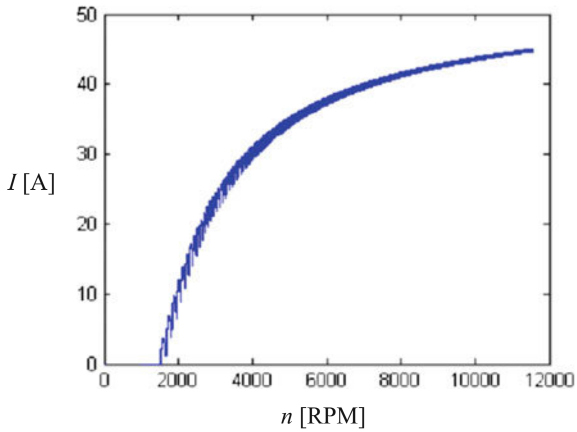


Fig. 6 The performance curve of an alternator by simulation: $I(n)$, V = const. On horizontal axes it presented the rotational speed [RPM] of the shaft of the alternator and on the vertical axes it is presented the load current [A]

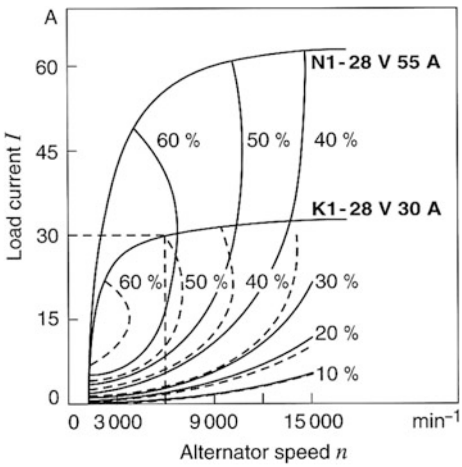


Fig. 7 The performance curve $I(n)$ that can be found in literature [2]

The simulated characteristic and the curves known from the Ref. [2] are identical if same circumstances are assumed (Fig. 7).

4 Simulation Model of a Voltage Regulator

The voltage regulator has passed a nice trajectory of evolution since the hero era of car driving. It started from the electro-mechanical regulator, over through-hole technology and surface-mounted technology until to power modules and specialised integrated circuits (Fig. 8).

At the base of each regulator stays a switching element that is a contactor or a transistor. It works mainly as a simple delta modulator. Each time the voltage reaches the upper and lower limits of the tolerance band around the reference value of the voltage (14.5 V) the switch changes its state.

It results that the excitation current rises and falls alternatively. The ratio of the switch-on and switch-off times will differ from one operating point to other. The steepness of the slope depends on the level of the current. The switching frequency is changing also. So the pulse width and the pulse frequency change in the same time (Fig. 9).

On Fig. 10 we can observe the electric circuit of alternator, voltage regulator and battery. Inside the voltage regulator there are two main components: a switching element (1) and a comparator (2). If we would like to build the model of a voltage regulator we have to use such kind of components. The simulation model of a voltage regulator is presented in Fig. 11.

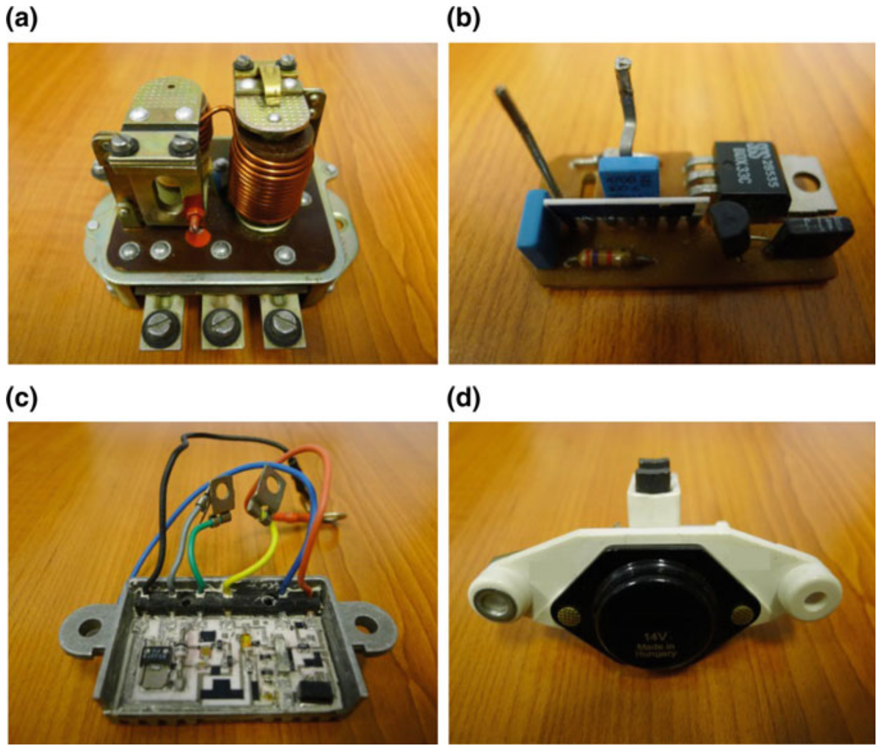


Fig. 8 Evolution of technology of voltage regulators: **a** electro-mechanical, **b** through-hole mounted, **c** surface mounted, **d** power modules (smart power)

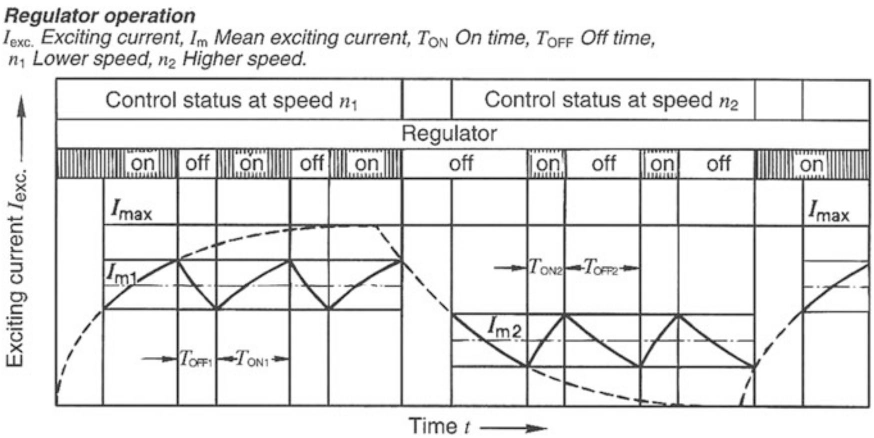


Fig. 9 Operation of voltage regulator [2]

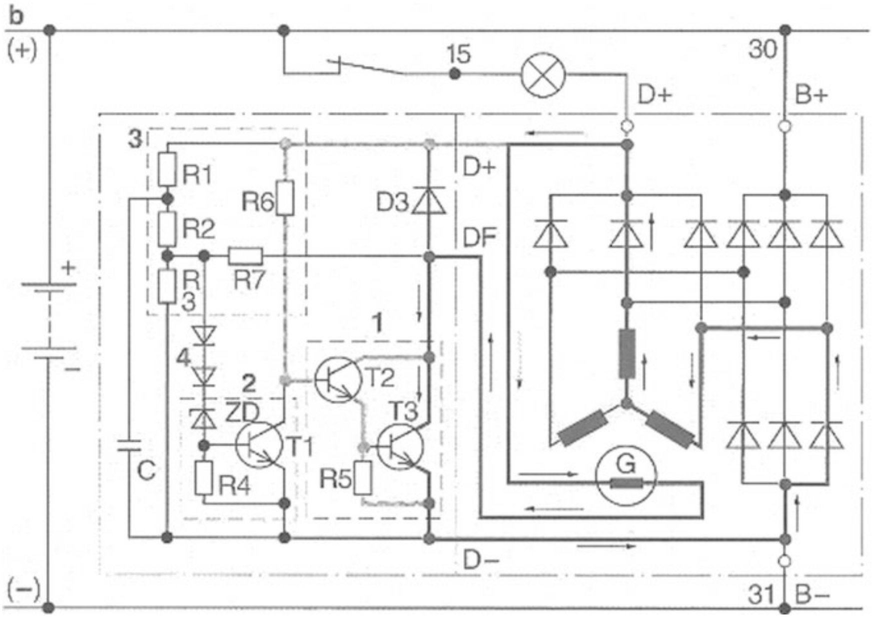


Fig. 10 Electric circuit of an alternator, voltage regulator and battery

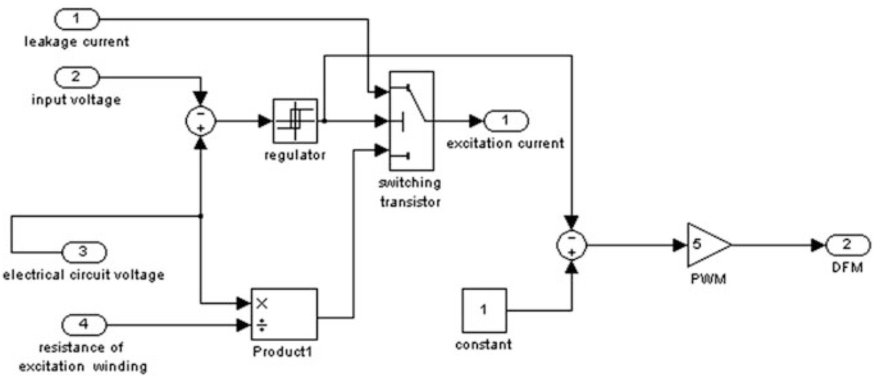


Fig. 11 Model of the voltage regulator. Inputs: *1* leakage current, *2* voltage base signal, *3* network voltage, *4* resistance of excitation, *5* frequency, *6* inductance of excitation. Outputs: *1* excitation current, *2* DFM = Digital Field Monitoring

5 Simulation Model of Entire System

Based on the previously described elements of the system a nonlinear model of the alternator including the rectifier and the voltage regulator can be created. Setting the parameters of the inputs the simulation can be carried out. The rpm of the internal combustion engine is set to change linearly from 1000 to 6000 rpm in every 50 s, and the load resistance changes in steps from 5.5 to 0.5 Ω in every 10 s (Fig. 12).

The results of the simulation are presented in Fig. 13. It can be seen that the induced voltage (1) has ripples due to the rectification and due to the voltage regulator. The network voltage (2) is highly influenced by the charging level of the battery and the state of the load current (3). The DFM (Digital Field Monitoring) signal gives information about the operation mode of the voltage regulator and so about the excitation current. A value of 5 V of DFM signal means that the switch state is ON: so the transistor conducts. If it is zero that means the switch is OFF: so the transistor does not conduct. The higher the duty is, the higher the excitation is.

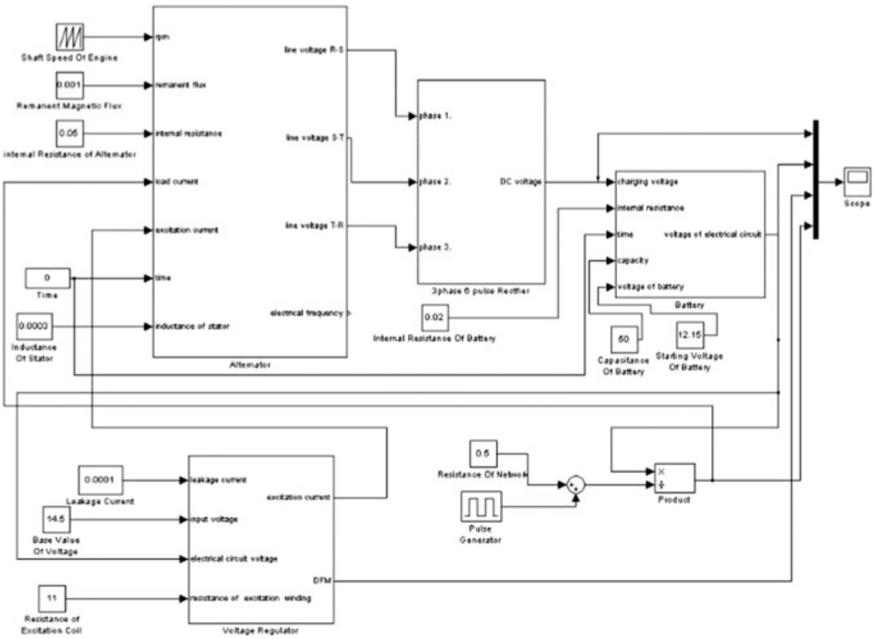


Fig. 12 The entire model of system

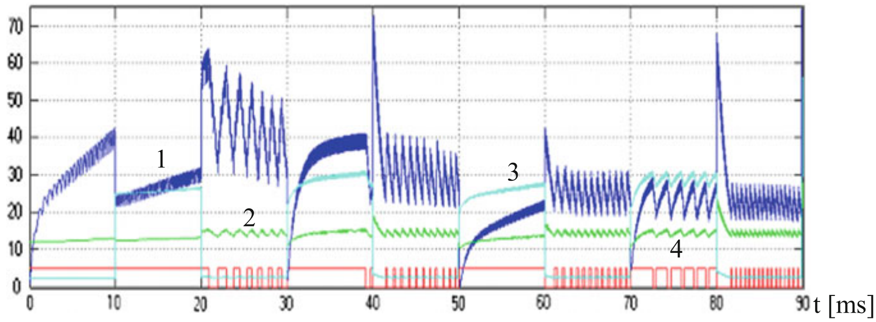
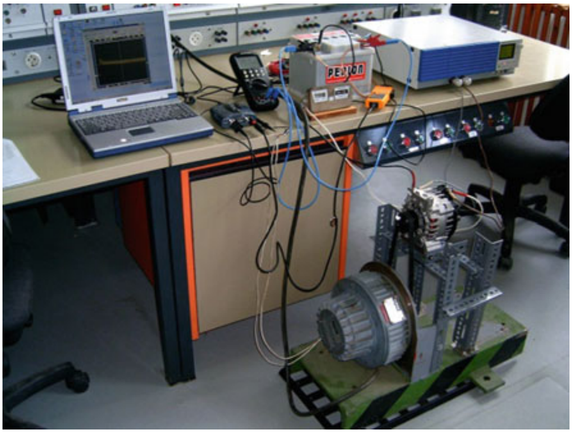


Fig. 13 The result of the simulation: horizontal axes it presents the time [s]; on the vertical axes are presented the following quantities: 1—the induced voltage in the stator winding [V], 2—voltage of the network [V], 3—load current [A], 4—DFM signal [V]

Fig. 14 Laboratory measurement of the alternator



6 Laboratory Measurements

A special test bench (Fig. 14) was built to validate the simulated model with real alternator. Because there was not possible to use an internal combustion engine, we applied a DC motor that was able to drive the alternator with suitable velocity. The speed could be changed by changing the voltage of the DC power supply. The load current of the alternator could be set by an electronic (artificial) load.

A series of measurements were carried out using equipment presented in Table 1.

In the first step we set the value of the load current, then changing the supply voltage of DC motor we set certain values of the speed and finally we measured the output voltage of the alternator. In this way we obtained the following characteristic surfaces as you can see in Fig. 15.

Table 1 Measurement equipment

Item	Model	Specifications
Electronic load	KIKUSUI PLZ1004 W	1000 W, 15 V 0–200 A
Current clamp	Fluke i30	DC to 20 kHz, 5 mA–30 A conversion 100 mV/A
Digital multimeter	Volcraft	
Battery	Perion	44 Ah, 12 V
DC motor	CEM	2000 W, 169 V, 3.6 A 3000 rpm
DC power supply	Built in laboratory power	0–500 V, 0–10 A
Digital oscilloscope	SDS 200 A	200 MHz, 5 GS/s, max. 42 Vpk
PC	Portocom	Freestar

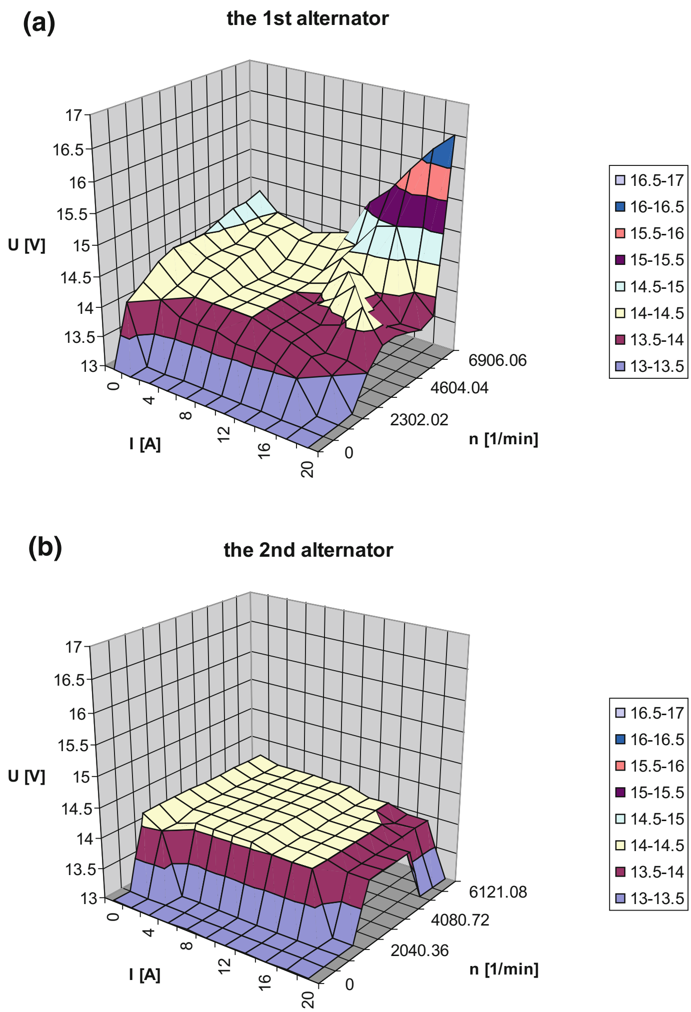


Fig. 15 Characteristic surface of the voltage regulator in function of the load current and the shaft speed **a** bed regulator, **b** good regulator

In the case of the first alternator we could observe that at high speed and at high load the voltage regulator had an interesting behaviour: the voltage increased up to value approx. 16.5 V. This would result in a malfunction of several consumers or even a broken off some consumers of the electric circuit of the car. In the case of the second generator the voltage regulator worked properly.

Moreover, we could find some interesting correlations between the duty and the frequency of the exciting current and the load current of the generator, as well as the rotation speed of the shaft of the generator (RPM) presented in Figs. 16 and 17.

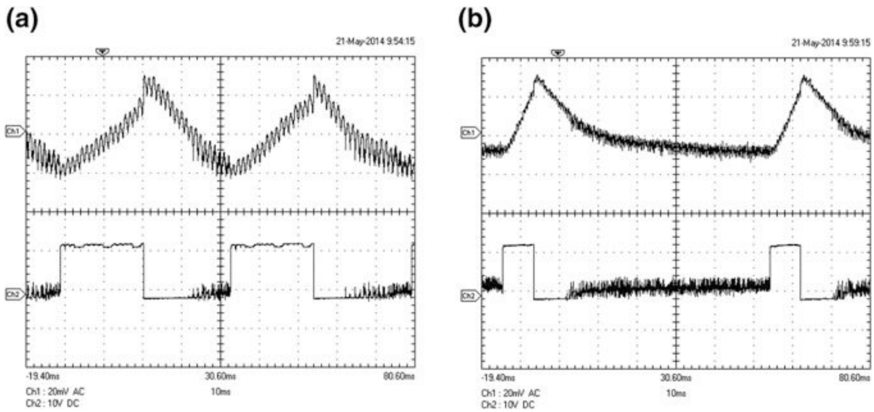


Fig. 16 Measurement results of exciting current versus time: correlations between **a** duty and current, **b** frequency and RPM

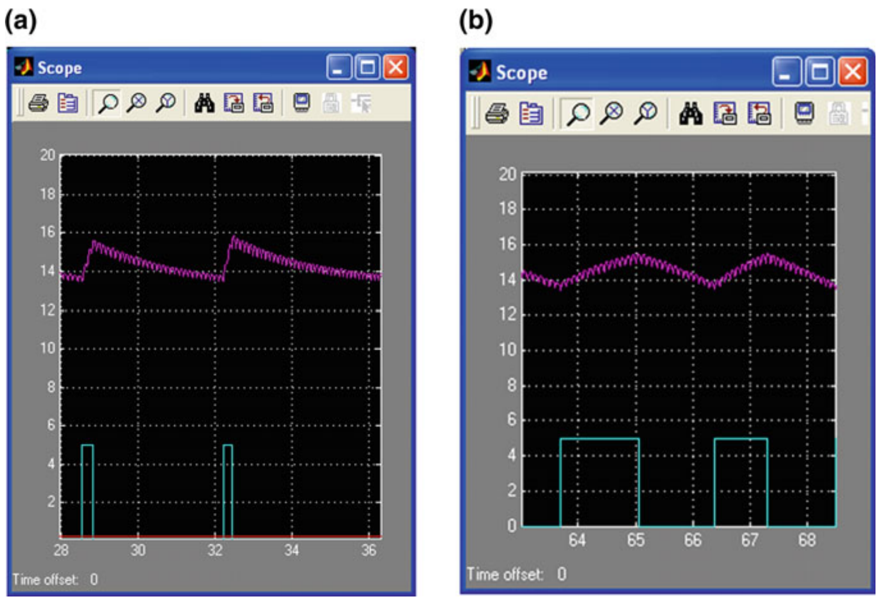


Fig. 17 Simulation results of exciting current versus time: correlations between **a** duty and current, **b** frequency and RPM

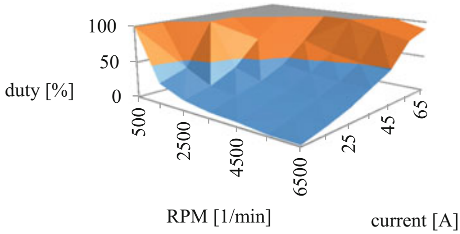


Fig. 18 Influence of current and RPM on duty got from simulations: duty = $f(I, \text{RPM})$, where duty [%] = 100 · TON · f

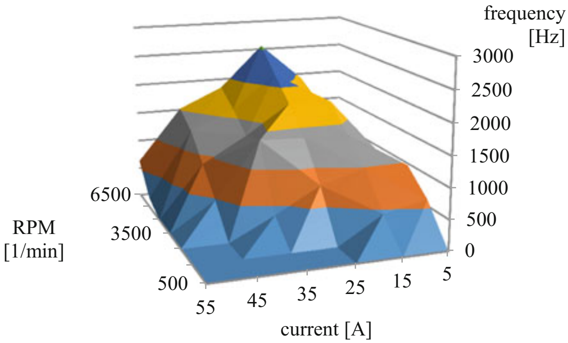


Fig. 19 Influence of current and RPM on frequency [Hz] got from simulations: $f = g(I, \text{RPM})$

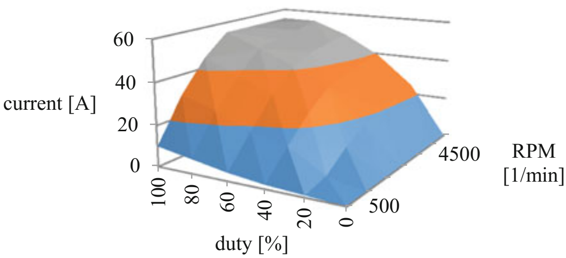


Fig. 20 Unequivocal dependency of current from duty and frequency: $I = h(\text{duty, RPM})$

For us Fig. 19 is less important, but changing the variables of surface presented in Fig. 18 we get the following diagram presented in Fig. 20.

7 Adaptive Neuro-Fuzzy Implementation

The previous characteristic surface can be implemented in a neuro-fuzzy application as follows in Figs. 21, 22, 23, 24 and 25.

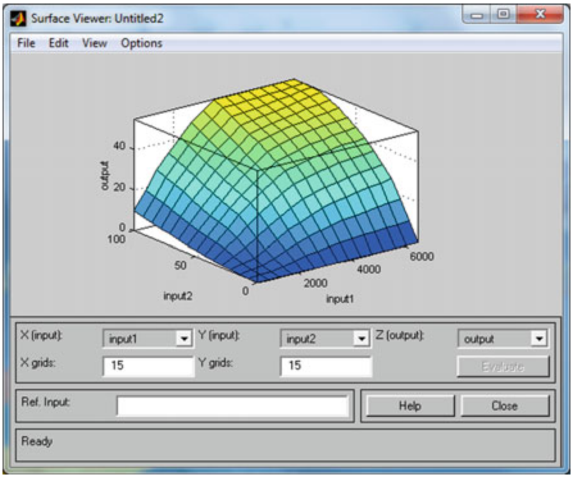


Fig. 21 Surface of current versus duty and frequency—$I = h$(duty, RPM)—implemented in a neuro-fuzzy application

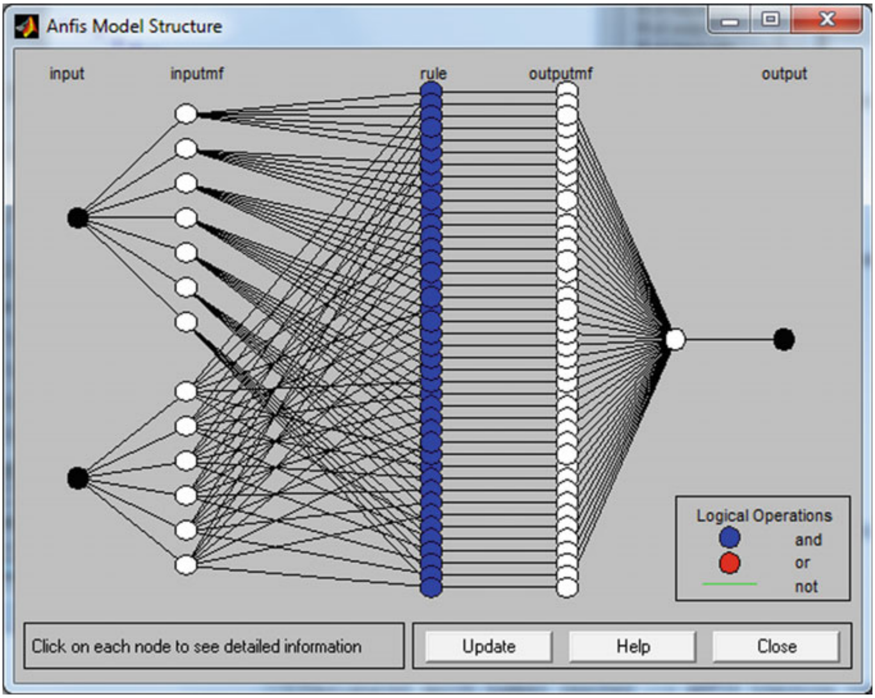


Fig. 22 Neural network of dependency of current versus duty and frequency $I = h$(duty, RPM)

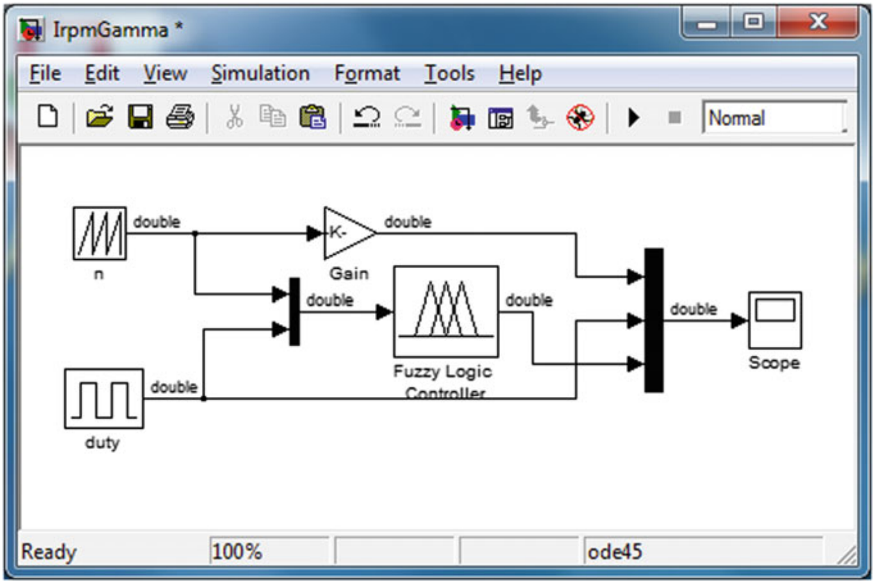


Fig. 23 Block diagram for determination of current versus duty and frequency—$I = h(\text{duty}, \text{RPM})$—based on a neuro-fuzzy implementation

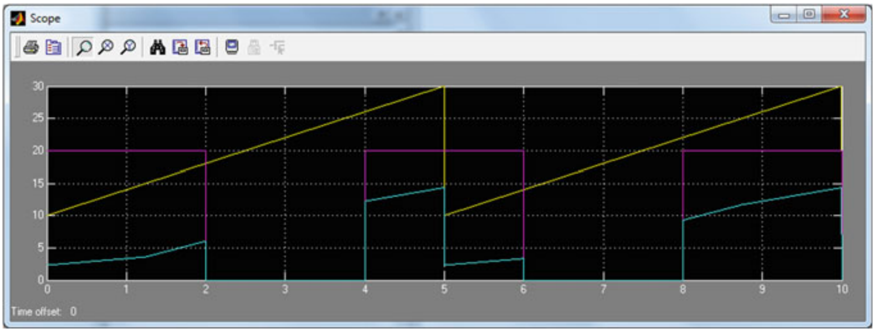


Fig. 24 Simulation results for determination of current versus duty and frequency—$I = h(\text{duty}, \text{RPM})$—based on a neuro-fuzzy implementation

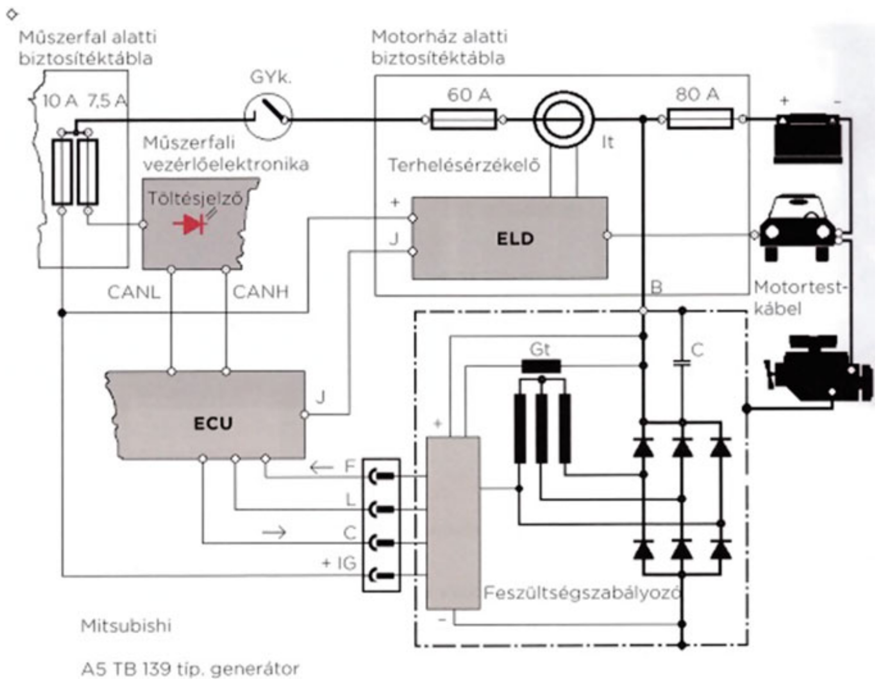


Fig. 25 Possibility of elimination of the electric load detector (ELD) [5] as a practical application of determination of current versus duty and frequency—$I = h(\text{duty, RPM})$—based on a neuro-fuzzy implementation

8 Sensorless Determination of RPM

If we measure the voltage of electrical circuit of the car we can detect pulses coming from the rectifier of the generator. The frequency of the pulses depends on the RPM of the generator that is driven by the internal combustion engine. So, both the RPM of the generator and internal combustion engine can be determinate from the frequency of the pulses. Only we to build an electronic circuit to detect the commutation point between pulses, for example. This method and circuit was developed and built as it is presented in the literature [6]. The measurement results are presented in Fig. 26.

Determination of the RPM can be done with the following relation:

$$n_{ICE}[\mathrm{min}^{-1}] = \frac{n_{ALT}}{i} = \frac{60 f_{el}}{p \cdot i} = \frac{60 \cdot \frac{f_p}{P}}{p \cdot i} = \frac{5 \cdot f_p\,[\mathrm{Hz}]}{3 \cdot i}, \tag{5}$$

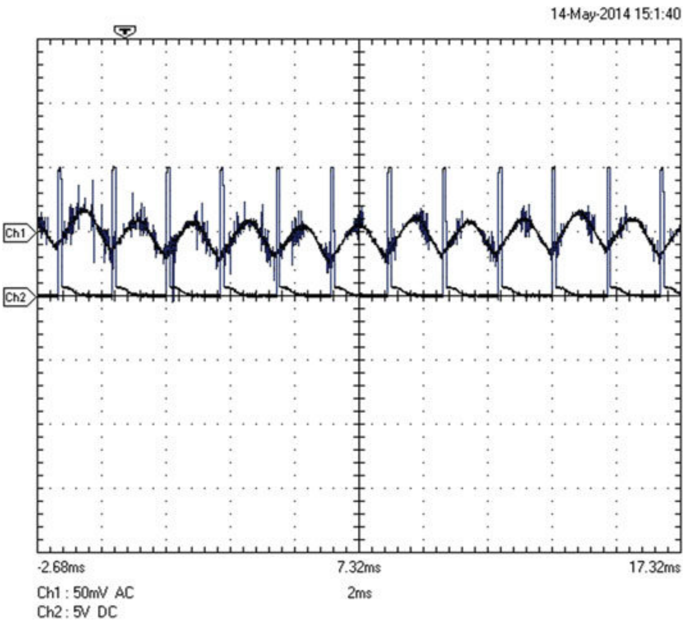


Fig. 26 Detection of commutation point of pulses of voltage of electric circuit of car

where
n_{ICE} = RPM of internal combustion engine, n_{ALT} = RPM of alternator,
i = ratio of the belt transmission, f_{el} = frequency,
p = pole pairs of alternator (= 6), f_p = frequency of pulses,
P = number of pulses (6-pulse Graetz-bridge).

9 Conclusions

This paper contains a lot of theoretical and special knowledge on the field of electrical machines and power electronics, mathematical deductions, simulation of linear and non-linear systems, building of test equipment and interpretations of many measurement results. The main aims have been achieved: the simulation model has been built and measurements were done. The simulation model of the voltage regulator has been created. It can be stated that the simulated behavior of the voltage regulator corresponds to the expectations and to the results of the laboratory measurements. We can simulate the operating conditions of the power electronic semiconductor device that switches the excitation current of a car alternator. As a continuation of this work we can imagine to build simulation models based on soft-computing methods as fuzzy-logic, neural network or quite use an adaptive neuro-fuzzy inherence system.

Acknowledgements This research was carried out in the framework of the Center of Excellence of Mechatronics and Logistics at the University of Miskolc.

References

1. Viorel A, Szabó L, Löwenstein L, Şteţ C (2004) Integrated starter-generators for automotive applications. Acta Electrotehnica 44(3):255–260
2. Bauer H (2003) Alternators and starter motors. Robert Bosch GmbH, Stuttgart
3. Blága Cs, Kovács E (2003) Solenoid based actuators, EDPE'03. In: Proceedings of the international conference on electrical drives and power electronics. The high tatras, Slovakia, 24–26 Sept 2003, pp 207–210
4. Blága Cs, Kovács E (2011) Modeling and measurement of an alternator. In: Proceedings of the 17th international conference on electrical drives and power electronics. The high tatras, Stara Lesna, Slovakia, 28–30 Sept 2011, pp 210–214. ISBN 978-80-553-0734-3
5. Veres L, Analysis of intelligent generators, Autotechnika, X-Meditor Kft., Győr, 2014/4. (in Hungarian)
6. Domonkos D (2014) Measuring of the internal combustion engine speed using electrical interference, final thesis, Miskolc (Hungary), 2014. (in Hungarian)

Distributed CAN-Bus Based Driving Assistance System on Autonomous Vehicle

Gergely Kovács and László Czap

Abstract This article focuses on CompactRIO based driver assistance system and the goal was a completely autonomous operation. By the development our vehicle had to be able to realize several intelligent driving assistance functions, such as adaptive cruise control, lane keeping, predictive emergency braking, brake energy regeneration, automated parallel- and cross parking and GPS navigation, and we also had to design hybrid drive on the go-kart. For implementing the intelligent functions mentioned above, we chose NI cRIO during the development. By taking the advantages provided by the developing environment and the modularity of the system, we could solve the scheduled tasks.

1 Introduction

The Hungarian Bosch Group organized a series of competitions in 2013 for the engineering universities of Hungary. The goal was to design and create a driving assistant system for a small racing car, for a go-kart which was provided by the company itself. The competitors were also forced to use several and existed Bosch components, such as an MPC (Multi Purpose Camera) camera, a ParkPilot ECU (Electronic Control Unit) with its ultrasonic sensors and a Bosch Mid-Range radar (MRR). During the preparation phase the individual teams had to fulfil special milestones in order to achieve minor goals and allowed the Bosch Company to monitor and support the teams. The final aim of the competition is to develop a smart driving system not just for go-karts but for passenger car, to create a basis for an autonomously driving vehicle.

G. Kovács (✉) · L. Czap
University of Miskolc, Miskolc, Hungary
e-mail: kovacs.gergely17@gmail.com

L. Czap
e-mail: czap@mazsola.iit.uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_33

2 The RIO Architecture

The mechanical impacts affected on the go-kart, such as vibrations or knocks, the thermal stress and the problems come up, especially from outside, the environmental damages require an industrial designed platform. CompactRIO's modular I/O enables users to flexibly implement a wide array of sensor types and industrial connectivity. The CompactRIO real-time controller and reconfigurable FPGA chassis provide a customizable inherent ability to implement many parallel loops on the FPGA (Field Programmable Gate Array) and real-time controller, its rich set of complex math functions makes it powerful software to control all aspects of a smart machine. Additionally, the open and flexible nature of LabVIEW helps implement reliable communication architectures for local operator interfaces and centralized resource management, diagnostics, and remote updates. These functions are essential once the machines are deployed around the world and need maintenance.

2.1 Real Time Controller

The LabVIEW Real-Time Module (Fig. 1) combines LabVIEW graphical programming with the power of a real-time operating system, enabling you to build deterministic real-time applications. National Instruments designed the LabVIEW Real-Time Module to execute VIs on three different real-time operating systems. The LabVIEW Real-Time Module can execute VIs on hardware targets running the real-time operating system of the NI Embedded Tool Suite (ETS), Wind River VxWorks or NI Linux Real-Time. The dataflow programming model of LabVIEW frees you from the sequential architecture of text-based programming languages. Additionally, LabVIEW makes it easy to assign thread priorities with the Timed Loop structure. Each loop can contain a separate task with a configurable timing source, period, priority, and more. LabVIEW Real-Time supports multicore processing and automatically maps parallel sections of your code to individual OS threads, which eliminates the need to manually create and manage them. By default, these threads are also automatically balanced across the CPUs available on your real-time hardware [1].

We have chosen the Linux Real-Time operating system, because it provides a wide range of opportunities to connect other devices to our system. NI Linux Real-Time contains embedded user interface and Mini Display Port to implement a local HMI (Human Machine Interface).

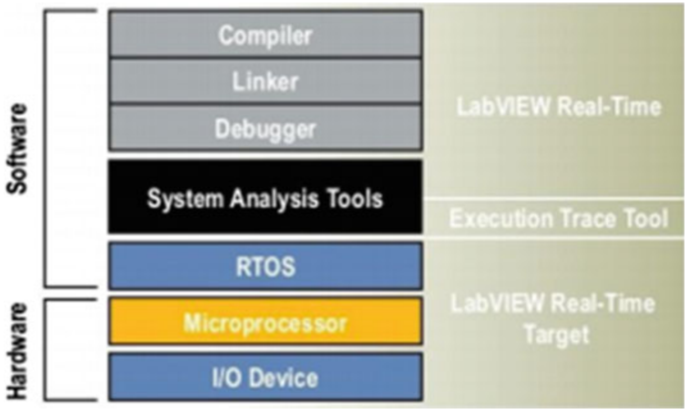


Fig. 1 Real-time development tools

2.2 *FPGA Module*

The basic components of a LabVIEW FPGA system are:

- the LabVIEW FPGA Module,
- the NI-RIO driver,
- and an NI-RIO device, also known as RIO target.

The LabVIEW FPGA Module allows LabVIEW to target FPGAs (Fig. 2) on NI Reconfigurable I/O (RIO) hardware, so scientists and engineers can take advantage of the performance and flexibility of FPGAs without the needing to learn low-level design tools. The LabVIEW FPGA Module compiles the LabVIEW VI to FPGA hardware, the graphical code is translated to text based VHDL code. LabVIEW is

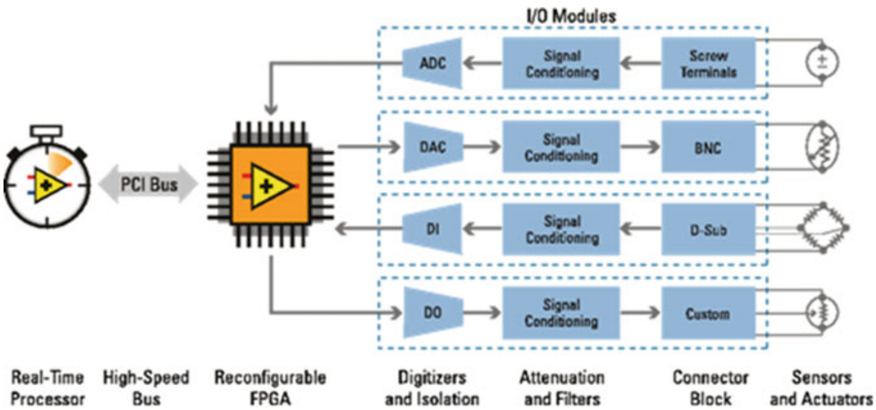


Fig. 2 NI FPGA architecture

well suited for FPGA programming because it clearly represents parallelism and dataflow, so users who are both experienced and inexperienced in traditional FPGA design can productively apply the power of reconfigurable hardware. The Xilinx compiler tools synthesize the VHDL code into a hardware circuit realization of the LabVIEW design. The result is a bit-file that is loaded to the FPGA chip before running the application [2].

3 The Control System

By choosing our control unit we had to take the external environmental impacts and the software development into consideration. During the two years we had to implement numerous driver assistances, so we needed a control unit which not only has the suitable industrial environment, and the algorithms can be implemented on a high level. Because of this development time can be drastically reduced, compared to lower level programming languages. During the operation of the go-kart, mechanical effects like vibrations, heat effects and other environmental effects occur more often, so industrial protection is a requirement. After taking these into consideration, we used the CompactRIO platform (Fig. 3) from National Instruments from the beginning, which are fully suitable for us. Its compact and provides robust protection against the environmental effects, the operating system running on the controller provides the time critical development.

In the first year we used the 9022 controller with the 9012 FPGA provided by National Instruments. Because of the complexity of the task and the enormous number of control tasks, it was necessary to use a more powerful CompactRIO platform, because we needed a larger FPGA and better computing capacity of the control system. Our other problem was the lack of space of electric components and because of the lack of ventilation we needed a platform, whose heat production is lower. This is the reason why we chose National Instruments 9031 CompactRIO, where the FPGA was given. It also played a role, that the controller is using the

Fig. 3 NI 9031 CompactRIO

Real-Time Linux operating system, so other devices can be easily connected to the CompactRIO.

3.1 I/O and Communication Modules

We had to choose the modules we would like to use during the development. It was important that we only had 4 slots available, so we preferred multifunction modules. Two multifunction modules were selected: a digital I/O module with 32 channels and a universal analog input module. We chose the following modules:

NI 9403 Bidirectional Digital I/O

- 32 digital I/O
- 5 V TTL level

NI 9219 Universal Analog Input

- 4 channels universal module
- 100 S/s per channel simultaneous inputs

NI 9263 Analog Output

- ±10 V output range,
- 16-bit resolution

NI 9853 High-Speed CAN Module for NI CompactRIO

- 2 ports

4 Distributed CAN-Bus System

CAN communication implements with the NI 9853 High Speed module. The main reasons why we chose this module are,

- High Speed CAN.
- Two ports (CAN0, CAN1), two separated CAN networks with different speeds.
- Support of 11 and 29 bit arbitration field.

The only difference between the two ports is that only one of them needs external power supply, which we can provide with 5–30 V DC power. Each port of the NI 9853 has an NXP SJA1000 CAN controller that is CAN 2.0B-compatible and fully supports both 11-bit and 29-bit identifiers. Each port also has an NXP TJA1041 High-Speed CAN transceiver that is fully compatible with the ISO 11898 standard and supports baud rates up to 1 Mbps.

The most important factor of our choice was that the communication functions can be easily accessible from the FPGA [3].

4.1 Developing of the CAN-Bus System

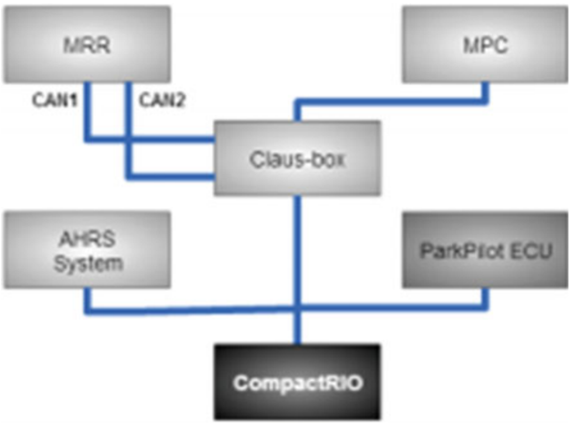
The first step was to decide what and how fast the devices are going to communicate with CAN.

This was a problem, because other devices which are not using this timing won't be able to properly communicate on that network. For this reason the Park Pilot ECU was moved to CAN1. The camera and the radar are communicating with a transfer rate of 500 kb/s (Fig. 4). These rates are not changeable, so the AHRS system and the microcontroller in the steering wheel are going to communicate on that rate of the CAN0 port. We can set the speed of our network, and we can create a so called "white list", where we can define what kind of messages we would like to receive [4, 5].

4.2 Communication with the Park Pilot ECU

The Park Pilot ECU was the first Bosch component we had to fit on the CAN network. Bosch gave the specifications about the network and the messages. The Park Pilot ECU is the only device on a separate network as mentioned above. There are twelve ultrasonic sensors connected to the Park Pilot ECU, which signals are

Fig. 4 CAN-BUS system



fitted with temperature and humidity compensation. The sensor data arrive in four messages, so three sensors send their data in one message:

- eight of the sensors can see up to 2.5 m,
- four sensors can see up to 5 m.

We decided that we position the four longer range sensors on the corners of the go-kart which was important because of the parallel parking and the home environment recognition [4, 6].

4.3 Communication with the Radar and the Camera

In the middle of the second year Bosch gave two other devices: a mid-range radar and an intelligent lane and object information providing camera which transmits its data through the CAN bus. Bosch gave a high speed CAN gateway for the devices, which directly connects to our system. For proper functioning it needs the steering angle, the speed in [m/s] and the velocity of the go-kart along the vertical axis (Yaw rate). Both devices are communicating through the CAN0 port [7–9].

5 Summary

We closed all three years very successfully. Every year our team was the absolute winner of the completion and in the first year we won the first prize in the best technical solution (Fig. 5).

Fig. 5 Awards

In the second year we were the absolute winner of the competition and we won the “Best Teamwork” and the “Smartest Go-Kart” awards too. In the third year beside the absolute first prize, we have got the first prize in the self-driving category.

Acknowledgements This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 691942. This research was partially carried out in the framework of the Center of Excellence of Mechatronics and Logistics at the University of Miskolc.

References

1. National Instruments: LabVIEW Real-Time 1 Course Manual, Sept 2012
2. National Instruments: LabVIEW FPGA Course Manual, Aug 2012
3. Trohák A, Végh J, Kovács G, Forgács Zs (2015) Implementations of driving assistance systems using labview based platform. Proceedings 16th international carpathian control conference (ICCC), pp 548–551
4. Trohák A, Illés B, Bíró Z (2013) Can message filter algorithms for remote diagnostics of vehicles. Appl Mech Mater 309:213–220
5. Ajtonyi I (2008) Industrial communication systems I. AUT-INFO Kft., Miskolc, (Hungary), p 431. ISBN: 978-963-06-5813-3, HU ISSN: 1789-5456 (in Hungarian)
6. Trohák A, Kolozsi-Tóth M, Rádi P, Méhes L, Biró Z (2011) The development of a remote diagnostic system for vehicles. Adv Logistic Syst 5:216–220. HU ISSN 1789-2198
7. Schoettle B, Sivak M (2015) Road safety with self-driving vehicles: general limitations and road sharing with conventional vehicles, Report No. UMTRI-2015-2. Michigan, USA
8. Schoettle B, Sivak M (2015) Potential impact of self-driving vehicles on household vehicle demand and usage, Report No. UMTRI-2015-3. Michigan, USA
9. Litman T (2015) Autonomous vehicle implementation predictions: implications for transport planning, transportation research board 94th annual meeting, Washington DC, USA

Wall and Object Detection with FRI and Bayes-Classifier for Autonomous Robot

Roland Bartók, Ahmed Bouzid, József Vásárhelyi and Márton L. Kiss

Abstract To detecting the nearly objects and walls is very important for a mobile robot or an autonomous vehicle to collision avoidance. It needs many of sensors for wide range detection. The results are better with more type of sensors for example infrared sensors, ultrasonic sensors, laser distance sensors. Because of the number of sensors it needed to use sensor fusion. It is a relatively easy way by using the FRI based Behaviour Description Language or the Bayes-classifier.

1 Introduction

There are many expressions in human language which are difficult to translate (compile) in computer code. These expressions make the human communication easy. To describe the distance between two objects one can use expressions like: “The two objects are far from each other”; “The two objects are relatively close to each other”. Or another example is how one expresses the speed of a moving object: one can use the words: “fast”, “slow” and “much slower”.

To translate these features, which can describe a moving autonomous robot fuzzy logic and fuzzy sets are used. In fuzzy logic to describe these features is used the interval [0, 1], instead of using the logic values 0 (false) and 1 (true). Fuzzy sets use this interval to characterise the membership functions. The members of sets have a membership function which shows what extent is part of a sets. Fuzzy sets were first described by L. A. Zadeh in its publication “Fuzzy sets” (see [1, 2]).

R. Bartók (✉) · A. Bouzid · J. Vásárhelyi · M. L. Kiss
University of Miskolc, Miskolc, Hungary
e-mail: qgeroli5@uni-miskolc.hu

A. Bouzid
e-mail: qgebouzid@uni-miskolc.hu

J. Vásárhelyi
e-mail: vajo@uni-miskolc.hu

M. L. Kiss
e-mail: autlkiss@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_34

1.1 Fuzzy Rule Interpolation

In fuzzy the knowledge is represented by the rulebase. Rulebases are a set of rule collection, i.e. they are "If … then" type sentences, which are sets of observations and conclusions.

In classical fuzzy logic one needs a rule for every observation. This can result in the description of a control system with a large number of rules which implies a rulebase with great dimension. A complete description of a system, which contains all the rules is called full rulebase.

There are some observations which cannot described by a rule (have no rule). In this case there is no conclusion for that observation. The missing conclusions can be computed from their neighbourhood using existing rules. This method is called fuzzy interpolation, based on the description of the system behaviour. In [3] was described the Behaviour Description Language, which uses the FIVE (Fuzzy Interpolation in Vague Environment) method described in [4]. The FIVE method makes unnecessary the classical steps of a fuzzy control system design, such as fuzzification and defuzzification step. In this method defuzzification became unnecessary.

1.2 Behaviour Description Language

Complicated behaviour models are created in a simple way with Fuzzy Rule Interpolation (FRI) based fuzzy automaton. The fuzzy automaton has got its own declarative description language. This language is based on ethology behaviour description languages, which already exist, most of them are built from primary declared samples by observing the reactions and behaviours of living creatures.

The simplicity was important in the design of the behaviour description language [3]. The language is easy to use with minimal programming knowledge.

The behaviour engine is a fuzzy automaton where the state is a vector of membership values. The state changes are controlled by fuzzy rulebase. The observations and conclusions are continuous values. The knowledge representation is based on rules for modelling behaviours. A sequence of events can be described with continuous values and continuous states. This automaton is computed in discrete time, but its model is continuous in states.

The Description Language contains keywords and user-defined tags, which helps to design an easy to read and easy to modify behaviour model of a given system. The list of keywords is as follows [3]:

- *rulebase*: The main base elements of behaviour descriptions are the rulebases. It contains the antecedents (observes) and consequents.
- *universe*: Defines the symbols and value range. Contains the user-defined symbols and value pairs. The universe can define the antecedents and consequents too.

- *description*: Optional keyword for comments and documenting.
- *rule, when, is, and*: Keywords for compose a rule. The syntax is the follow: *rule* "output symbol" *when* "antecedent symbol" *is* "antecedent value" *end*.

2 Naive Bayes-Classifier

The Bayes-classifier is a simple, scalable classifier to determine class label for attributes. The Bayesian classifiers can be used for text categorization such as spam filter or in medical diagnostics. The naive Bayes–classifier supposes that the attributes are independent. The method is based on the conditional probability. The Bayes-theorem is:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)}, \tag{1}$$

where i is the number of classes, B_i, A is a vector which contains the observed attributes. In this paper is considered that the classes show *"where are the walls around the mobile robot"*. The observed attributes are the measured data received from infrared sensors. The selected classifier will be the one where the probability $P(b|A)$ is maximal for the A attribute vector [5].

3 The Mobile Robot and Its Environment

The mobile robot is equipped with two wheels with differential drive. The autonomous robot is equipped with several different sensors, such as: incremental phase encoder for each shaft, six-axis (Gyro + Accelerometer) MEMS, infrared diode-phototransistor pairs for distance measurement. The robot shape and sensor layout are shown on the Fig. 1.

The robot moves in maze-like corridors and creates maps and also navigating. The shape of maze is square, build from square shaped cells (Fig. 2). The cells dimension is 90 * 90 mm^2, wall height is 25 mm, and the size of maze is 5 * 5 cells. The robot knows the size of the maze, but not where the walls are situated.

3.1 Data Collection from the Environment

There were defined 7 cases in the robot movement. Using Bayes-classifier, a lot of training data were collected for every case. Some movements of data collecting method are shown on Fig. 2. The mobile robot moved backward and forward in the

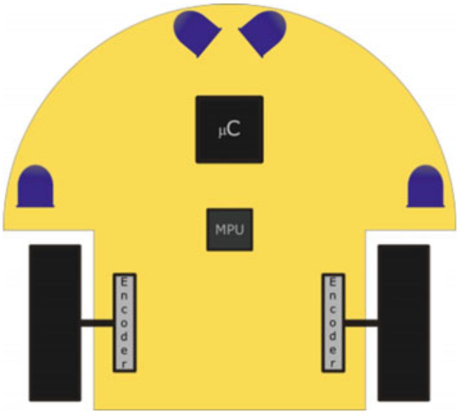


Fig. 1 Mobile robot. Infrared LEDs are represented in *blue*. Measurement area of sensors is 10°

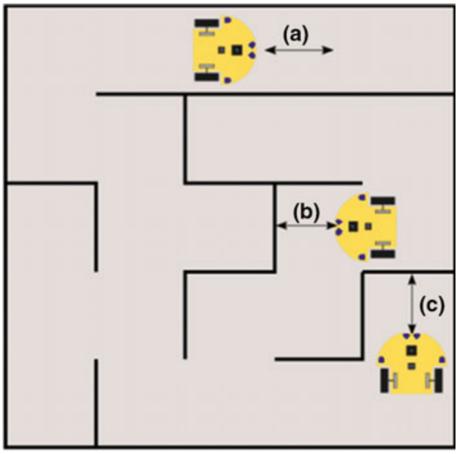


Fig. 2 Data collection method. For example **a** corridor, **b** front-right, **c** left-front-right

target area while collecting the measured values. The defined cases were as follows with their probability in the train maze:

a. wall on front (P = 0.08)
b. wall on left (P = 0.12)
c. wall on right (P = 0.12)
d. wall on left and right (P = 0.24)
e. wall on front and left (P = 0.36)
f. wall on front and right (P = 0.36)
g. wall on front and left and right (P = 0.24)

The probability for the above situations is calculated with the formula:

$$P = (\text{number of possible cases})/(\text{number of all cases}) \tag{2}$$

where the number of all cases means the robot can enter in one cell from 4 directions; the number of possible cases were defined above (a–g).

The mobile robot was moving in each case, while collecting the measured data. The listed cases become classes for Bayes-classifier and the first 3 cases become classes in behaviour description.

4 Behaviour Description for Wall Detection

The Behaviour description gives 3 classes: wall detected on the left, front or right. The given value range is [0, 1] for each class. If the returned value for each above situations is higher than 0.6 this represents an obstacle (wall) in front of the moving robot. The binary input value range is [0.4096] according to analogue digital converter, where the 0 means the wall is very far and the 4096 means the wall is very close.

The universe defined for all sensors are as follows. The given example below is for wall detected on the left back IR sensor:

universe "*leftBack*"
"*far*" 0 0
"*middle*" 100 0.7
"*close*" 720 1
"*very close*" 4096 1
end

The first parameter of symbol is an element of input range; the second parameter is the membership value. This means the input characteristics runs as described, 0 when the wall distance is far and the close and very close gives 1. The output universe is as follows:

universe "*isWallOnLeft*"
"*no*" 0 0
"*yes*" 1 1
end

This is the consequence of rulebase for wall detected on the left side of the robot. It gives 0 if no wall and gives a value larger than 0 when wall detected is approaching. If the calculated value returned by the rulebase is converted into Boolean value when True if obstacle occurs, false if no obstacle. This value and the values in the input universes results of the experimental measurements.

The example rulebase for wall detected on the left is the following:

rulebase "*isWallOnLeft*"
Rule
"*No*" **when** "*leftBack*" **is** "*far*" **and** "*left Front*" **is** "*far*"
End
Rule
"*Yes*" **when** "*leftBack*" **is** "*middle*" **and** "*lefFfront*" **is** "*middle*"
End
End

The right sensor data is not used, because generated errors in the cases d–g. The number of rules must be equal with the number of user-defined symbols in the universe.

5 Bayes-Classifier for Wall Detection

The number of collected data for listed classes was enough for statistical analysis. The occurrence of each value from 0 to 4096 were counted to the defined ranges, for example from 0 to 30, from 31 to 60, this was needed for noise reduction. Then probability was computed for every range. This training data were used for Bayes-classifier. A sample of distortion calculated is shown in Fig. 3.

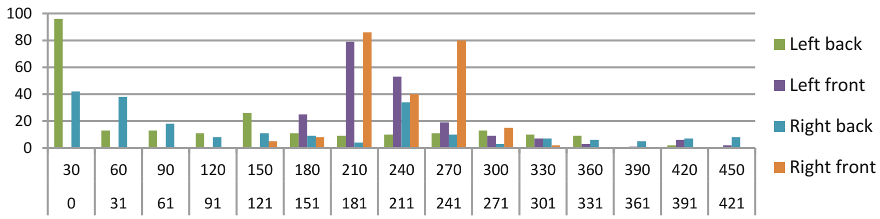


Fig. 3 Calculated probabilities for cluster values resulted from the different sensors for "wall in front and *left* and *right*" class (sensors represented on: *green*, *violet*, *blue*, *orange*)

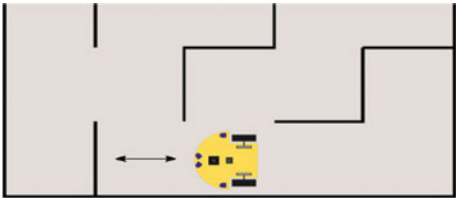


Fig. 4 The situation when the Bayes-classifier gives better result. Thin wall caused a false detection

6 Conclusions

Both methods succeeded to detect the walls around the mobile robot. The FRI method was not tuned. The Bayes-classifier was trained with collected data. The Bayes-classifier could detect the walls earlier.

The difference of distance between the FRI method and Bayes-classifier was around 5–8 mm from the centre of the cell. When the mobile robot moved too close (5–10 mm) to wall both of methods generates false detection. The Bayes-classifier gave better result in the situation shown in Fig. 4. Further research is needed to implement the FRI and Bayes-classifier in FPGA using VHDL.

Acknowledgements The research work was (partially) supported by the Hungarian Scientific Research Found grants OTKA 29326 and Fund for the Development of Higher Education FKFP 8/2000 project. This research was (partially) carried out in the framework of the Center of Excellence of Mechatronics and Logistics at the University of Miskolc.

References

1. Kovács S (1993) Fuzzy logic control, M.Phil. theses, Technical university of Budapest, Faculty of informatics and electrical engineering, Branch of Computer Science, Budapest (Hungary), p 116
2. Zadeh LA (1965) Fuzzy Sets. Inf Control 8(3):338–353
3. Piller I, Vincze D, Kovács S (2015) Declarative Language for Behaviour Description. Emergent Trends Robot Intell Syst 316:103–112
4. Kovács S (2006) “Extending the fuzzy rule interpolation “FIVE” by fuzzy observation”, Computational intelligence, Theory and applications: 9th international conference on Dortmund Fuzzy Days. 802 p, Dortmund (Germany), 2006.09.18-2006.09.20. Berlin; Heidelberg: Springer-Verlag, pp 485–497. ISBN:978-3-540-34780-4
5. Downey AB (2014) Bayesian statistics made simple. Green Tea Press, p 194. ISBN: 13-978-1449370787

Optimal Formation of Logistics Networks

György Kovács, Zoltán Varga and Márk Mihalik

Abstract The growing market globalization, increasing global competition, more and more complex products require new technologies, methods and processes. The shorter life cycle and the complexity of final products and new customer requirements require efficient operation of supply chains (SC). Optimization of supply chains results new models, concepts of value chains and new organization and cooperation forms of members. Reduction of total cost and lead time of the chains, and higher customer service level can be the objectives of the supply chain optimization.

1 Supply Chain, Virtual Organization, Virtual Enterprise

Stevens [1] defined the *supply chain* as a system whose constituent parts include material suppliers, production facilities, distribution services and customers linked together via a feed forward flow of materials and feedback flow of information.

Due to fast changing market environment, globalization and global competition, supply chains have become more and more complex networks. Reduction of total cost and lead time of the chains, and higher customer service level becomes the most common used objectives for operation of supply chain networks.

A *virtual organization* [2] is a short-term form of cooperation among legally independent co-producers in a logistics network of long-term duration of potential business partners for the development and manufacturing of a product. This is true for procurement and production, as well as for product and process innovation. Co-producers produce the service on the basis of mutual values and act towards the

G. Kovács (✉) · Z. Varga · M. Mihalik
University of Miskolc, Miskolc, Hungary
e-mail: altkovac@uni-miskolc.hu

Z. Varga
e-mail: altvarga@uni-miskolc.hu

M. Mihalik
e-mail: miha02@freemail.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_35

third party as a single organization. Each co-producer is active within the area of its core competence. The choice of a co-producer depends on the co-producer's innovative power and its flexibility to act as a partner in the logistics network [2]. The strength of virtual organizations is their ability to form quickly and gain competitive advantages.

Camarinha-Matos [3] interpreted the *virtual enterprise* (VE) as a temporary alliance of enterprises that come together to share their skills, core competencies, and resources in order to better respond to business opportunities, and whose cooperation is supported by computer networks.

Gunasekarana et al. [4] defined that virtual enterprises are characterized by several strategic objectives:

- maximizing flexibility and adaptability to environmental changes,
- developing a pool of competencies and resources,
- reaching a critical size to be in accordance with market constraints, and
- optimizing the global supply chain.

Virtual organizations are used in more and more industries, e.g. fashion industry, food industry, automotive industry, etc.

Figure 1 shows supply chain networks and a virtual enterprise as a temporary alliance of enterprises. The network includes a large number of customers, production companies and service providers. Customers can be consumers, end-users, etc. Production companies are the final assemblers, primary-, secondary- …suppliers and raw material suppliers.

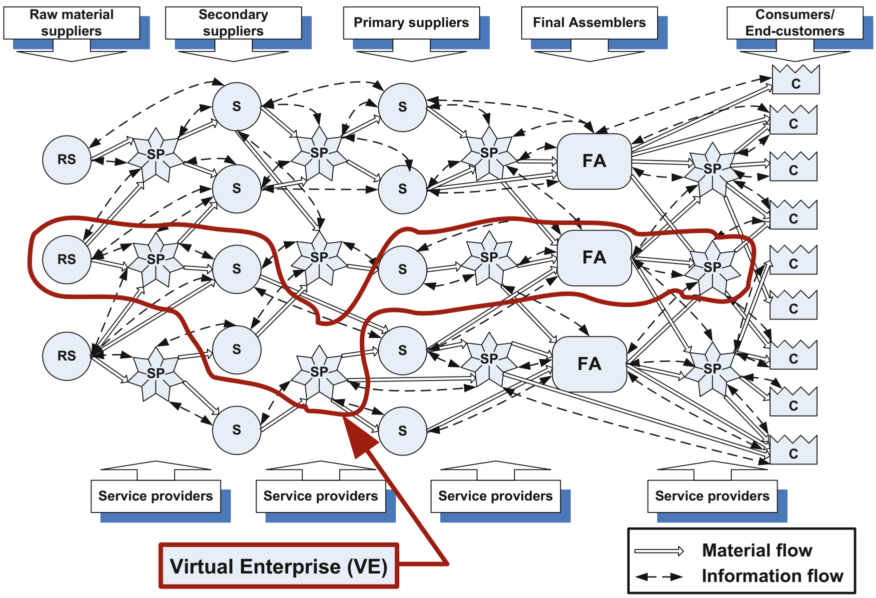


Fig. 1 Supply chain networks and a virtual enterprise

Possible service providers in the supply chains can be the followings:

- Logistics service providers (forwarding comp., warehousing comp., SC managing comp., …),
- Information service providers (IT companies, telecommunication companies, …),
- Research and Development service providers (Research institutes, Universities, Consulting enterprises, …),
- Financial service providers (Banks, Lease brokers, Consulting enterprises, …).

2 Network Optimization in Case of a Virtual Enterprise

The objective in the optimization of a supply chain network is to produce and deliver finished products to end consumers in the most cost effective and timely manner and to achieve higher customer satisfaction. Supply chain management and supply chain optimization appear frequently in the literature. Large number of publications has been also presented in the literature on topic of virtual enterprise concepts and characteristics [5–10].

This study shows the optimization method for the operation of a virtual network. Objective function and constraints will be elaborated.

Total cost should be minimized, taking into consideration constraints relating to production and service capacities, inventories and flexibility in the chain.

In our theoretical problem (Fig. 1) the followings are given: products types and quantities ordered by customers, possible suppliers of components, possible final assemblers, possible service providers, list of required services fulfill the customer satisfactions, manufacturing capacities of production plants, capacities at service providers, distances between the possible objects of the network, cost and time of activities required for manufacturing of components and final products to fulfill the customer orders.

The optimal formation of a virtual enterprise (optimal variation of members of the supply chain) is searched.

Indices are used in mathematical formulation are the followings: i—products; j—possible suppliers (including raw material suppliers, primary-, secondary-, etc. suppliers); k—final assembler companies; l—customers; m—possible service providers; t—time periods.

2.1 Objective Function—Total Cost

Total cost is including the raw material and component costs, production costs, transportation costs, inventory costs, cost of activities of service providers and operation cost of the virtual enterprise.

$$O_1 = C_{mat} + C_{prod} + C_{tran} + C_{inv} + C_{serv} + C_{oper}\,[Euro]. \tag{1}$$

2.1.1 Raw Material and Component Cost

The total material cost is the sum of material costs at suppliers and material cost at final assembler:

$$c_{mat} = cm_{ij} \cdot V_{ijt} + cm_{ik} \cdot V_{ikt}\,[\text{Euro}], \tag{2}$$

where: cm_{ij}—unit material cost of raw materials and components of product i at supplier j [Euro/piece]; cm_{ik}—material cost of product i at final assembler k [Euro/piece]; V_{ijt}—production volume of raw materials and components of product i at supplier j in each t time period [piece]; V_{ikt}—production volume of product i at final assembler k in each t time period [piece].

2.1.2 Production Cost

The production cost (including additional activities joining to manufacturing activity, e.g. machine setup, material supply, loading unit formation, …) is the sum of production costs at suppliers and production cost at final assembler company:

$$c_{prod} = cp_{ij} \cdot V_{ijt} + cp_{ik} \cdot V_{ikt}, \tag{3}$$

where: cp_{ij}—unit production cost of raw materials and components of product i at supplier j [Euro/piece]; cp_{ik}—unit production cost of product i at final assembler k [Euro/piece]; V_{ijt}—production volume of components of product i at supplier j in each t time period [piece]; V_{ikt}—production volume of product i at final assembly k in each t time period [piece].

2.1.3 Transportation Cost

The transportation cost (including additional activities joining to transport activity, e.g. loading in, loading out, loading unit formation, etc.) is the sum of transportation of each product in each relation of the network, between suppliers and final assembler and between final assembler and customers:

$$c_{tran} = ct_{ijk} \cdot V_{ijkt} + ct_{ikl} \cdot V_{iklt} \tag{4}$$

where: ct_{ijk}—unit transportation cost of raw materials and components of product i from supplier j to assembler k [Euro/piece]; ct_{ikl}—unit transportation cost of product i from final assembler k to customer l [Euro/piece]; V_{ijkt}—flow of components of product i from supplier j to assembler k in each t time period [piece]; V_{iklt}—flow of components of product i from final assembler k to customer l in each t time period [piece].

2.1.4 Inventory Cost

The inventory cost is including the holding and storage costs of stocks at suppliers, at assembler, at customers and in some cases at service providers (warehousing service providers):

$$c_{inv} = ci_{ij} \cdot I_{ijt} + ci_{ik} \cdot I_{ikt} + ci_{il} \cdot I_{ilt} \tag{5}$$

where: ci_{ij}—unit inventory cost of raw materials and components of product i at supplier j [Euro/piece]; ci_{ik}—unit inventory cost of raw materials and components of product i at assembler k; ci_{il}—unit inventory cost of product i at customer l [Euro/piece]; I_{ijt}—inventory of components of product i at supplier j in each t time period; I_{ikt}—inventory of components of product i at assembler k in each t time period [piece]; I_{ilt}—inventory of product i at customer l in each t time period [piece].

2.1.5 Cost of Activities at Service Providers

Activities of service providers are including the processes required for financing, manufacturing, warehousing, packaging, etc. of product i.

$$C_{serv} = Csp_{imt} \tag{6}$$

where: Csp_{imt}—cost of activities of service providers required for manufacturing and delivery of product i in each t time period.

2.1.6 Operational Cost of the Virtual Enterprise

Operational cost of the VE (C_{oper}) is including the management cost of the network, cost of the information and communication technology and all of additional costs which are required for the optimal operation of the virtual network.

$$C_{oper} = C_{oper}(n_{comp}, p) \tag{7}$$

This cost component is depending on the size of the network (n_{comp}) and profile of member companies (p).

2.2 Constrains

2.2.1 Production and Service Capacity Constraints

If the product i is allocated to supplier j during time period t, the production volume (V_{ijt}) should be limited by minimum and maximum volume at suppliers:

$$V_{ijt}^{min} \leq V_{ijt} \leq V_{ijt}^{max} \tag{8}$$

If the product i is allocated to final assembler k during time period t, the production volume (V_{ikt}) should be limited by minimum and maximum volume at final assembler:

$$V_{ikt}^{min} \leq V_{ikt} \leq V_{ikt}^{max} \tag{9}$$

If the services required for production of product i is allocated to service provider m during time period t, the activity capacity (C_{imt}) should be limited by minimum and maximum volume at service providers:

$$C_{imt}^{min} \leq C_{imt} \leq C_{imt}^{max} \tag{10}$$

2.2.2 Inventory Constraint

Based on lean philosophy the inventory is one of the biggest waste in the supply chain, but it provides the flexibility of the chain. Depending on the inventory strategy of the supply chain, the volume of inventories at manufactures and service providers should be limited:

$$I_{ijt}^{min} \leq I_{ijt} \leq I_{ijt}^{max}; I_{ikt}^{min} \leq I_{ikt} \leq I_{ikt}^{max}; I_{imt}^{min} \leq I_{imt} \leq I_{imt}^{max} \tag{11}$$

2.2.3 Flexibility Constraint

The virtual enterprise is characterized by a dynamic and flexible network, so the following flexibility constraints are defined for the chain members:

- flexibility of the manufacturing system at the manufacturing companies of the chain,
- flexibility of service providers of the chain,
- flexibility of the IT infrastructure at the companies of the chain,
- liquidity of the chain members,
- flexibility due to the organizational structure of the chain members.

These constraints are not easy to defined, but these can be given by a value in a given interval (1–5). Number 5 represents the most flexible value.

3 Software Conception for Optimal Network Design

Based on the above mentioned optimization concept our research team developed a software application for the optimization of a virtual enterprise network.

The software is written in Java programming language. Java is a general-purpose computer programming language that is concurrent, object-oriented, and class-based. Java is a platform free programming language. This is a big advantage comparing other languages, it means our written software is independent to operating systems [11].

We used Eclipse software framework to develop the software. Eclipse is an open source integrated development environment (IDE), one of the most widely used Java IDE. It contains a base workspace and an extensible plug-in system for customizing the environment. Eclipse is written mostly in Java and it is primary used for developing Java applications, but it may also be used to develop applications in other programming languages [12].

3.1 *Numerical Example for Formation of Optimal Supply Chains*

In this chapter we would like to show the application and operation of a software developed by our research team.

In our example (Fig. 2) the supply chain includes one final assembler (FA), four primary suppliers (S_{11}, S_{12}, S_{13}, S_{14}) and five possible secondary suppliers (S_{21}, S_{22}, S_{23}, S_{24}, S_{25}).

Relations of possible suppliers can be seen in Table 1, the distances of different members can be found in Table 2.

We assume that the unit production cost (cp_i) of our fictive product (product A) is 5 Eur/pieces in Europe and in America and 2.85 Eur/pieces in Asia. We also

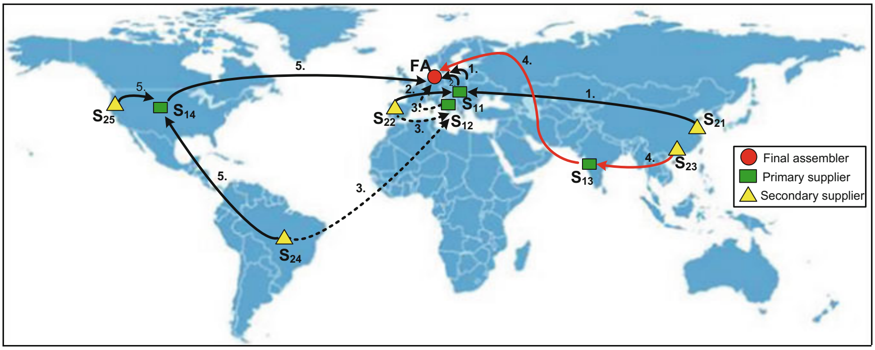


Fig. 2 Possible supply chain combinations

Table 1 Relation matrix

	FA	S_{11}	S_{12}	S_{13}	S_{14}	S_{21}	S_{22}	S_{23}	S_{24}	S_{25}
FA	0	1	1	1	1	0	0	0	0	0
S_{11}	0	0	0	0	0	1	1	0	0	0
S_{12}	0	0	0	0	0	0	1	0	1	0
S_{13}	0	0	0	0	0	0	0	1	0	0
S_{14}	0	0	0	0	0	0	0	0	1	1
S_{21}	0	0	0	0	0	0	0	0	0	0
S_{22}	0	0	0	0	0	0	0	0	0	0
S_{23}	0	0	0	0	0	0	0	0	0	0
S_{24}	0	0	0	0	0	0	0	0	0	0
S_{25}	0	0	0	0	0	0	0	0	0	0

Table 2 Distance matrix (km)

	FA	S_{11}	S_{12}	S_{13}	S_{14}	S_{21}	S_{22}	S_{23}	S_{24}	S_{25}
FA	0	810	960	6570	7600	8830	1440	9170	9180	9390
S_{11}	810	0	800	5170	8370	8350	1970	8580	9670	9800
S_{12}	960	800	0	6180	8400	9120	1360	9290	8910	10,050
S_{13}	6570	5170	6180	0	13,430	5040	7540	4300	13,760	13,510
S_{14}	7600	8370	8400	13,430	0	11,360	7670	12,570	7800	2420
S_{21}	8830	8350	9120	5040	11,360	0	10,260	1230	18,000	9890
S_{22}	1440	1970	1360	7540	7670	10,260	0	10,550	7740	9320
S_{23}	9170	8580	9290	4300	12,570	1230	10,550	0	18,020	11,110
S_{24}	9180	9670	8910	13,760	7800	18,000	7740	18,020	0	9780
S_{25}	9390	9800	10,050	13,510	2420	9890	9320	11,110	9780	0

assumed in the calculation that the material cost (cm_i) is 5 Eur/pieces all over the world. The specific transportation cost (ct_i) inside Europe is 0.00024 Eur/Km (road transport), between Europe and Asia and between Europe and America is 0.00012 Eur/Km (possibility of water transport).

The objective function in the optimization is the total cost (Eq. 7), which including production costs at the final assembler and at the suppliers and transportation costs between the members (in our example the inventory cost, costs of service providers, operation cost of the virtual enterprise and time consumptions of activities are not taken into consideration).

Table 3 shows the most important data relating to the final assembler (FA), primary suppliers (S_{1i}) and secondary suppliers (S_{2i}). Flexibility parameters and production capacities are also defined in Table 3.

Table 3 Specific production- and material costs and flexibility parameters

	cpi (Euro/piece)	*cmi* (Euro/piece)	Production capacity (pieces)	Flexibility of the manufacturing system	Flexibility of the IT infrastructure	Liquidity	Organizational structure
FA	5	5	450	5	4	3	4
S_{11}	5	5	250	3	3	3	3
S_{12}	5	5	250	3	5	3	3
S_{13}	2.8	5	250	3	3	5	3
S_{14}	5	5	350	3	4	3	3
S_{21}	2.8	5	200	3	3	3	3
S_{22}	5	5	100	3	3	3	3
S_{23}	2.8	5	250	3	3	3	3
S_{24}	5	5	150	3	3	3	3
S_{25}	5	5	350	3	3	3	3

3.2 Operation and Main Screens of the Software Application

In the next part of the paper we show the most important print screens of the developed software.

The software provides the possibility of determination of:

- data for the products to be produced,
- for data potential members of the supply chain and,
- relations for members of the supply chains (Fig. 3).

In the menu "Data for the products to be produced" we can define the parameters of the product to be manufactured.

In the menu "Data for potential members of the supply chain" we can define the most important parameters relating to the final assembler, suppliers and forwarding service providers (Fig. 4).

In the menu "Relations for potential members of the supply chain" the relation matrix, distance matrix and transport modes can be defined (Fig. 5).

In the menu "Results of the optimization" we can select the objective function of the optimization (Fig. 6). In this paper we show the cost optimization procedure, the time optimization and the multi objective optimization is under development.

Flexibility constraints relating to the chain members can be also set (Fig. 6). Recently the optimization in case of single objective optimization is performed by systematic search. This method is absolutely good in case of a small network, but for a huge size network we will apply more robust optimization algorithm. The multi objective optimization algorithm is also under development.

The result of the cost optimization is also can be seen in this screen. The possible chain combinations—which fulfill the constraints—can be listed in the screen.

As it can be seen in Fig. 6, the optimal formation of the supply chain in our example is **FA–S_{13}–S_{23}**, when the total specific cost of the supply chain will be minimal.

Fig. 3 Main menus of the program

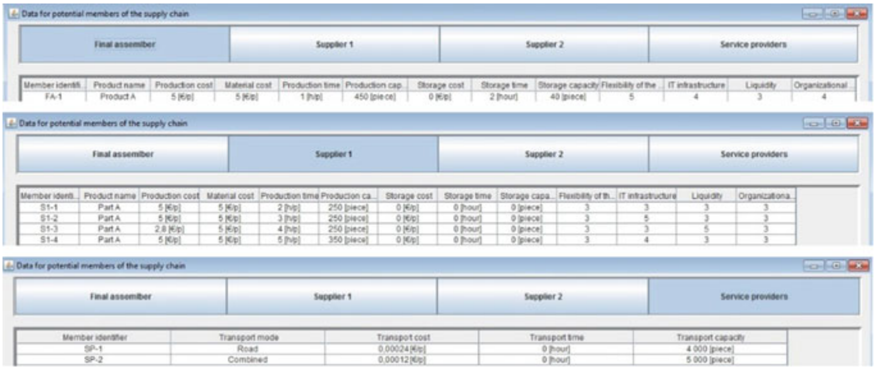


Fig. 4 Screens for parameter setting of supply chain members

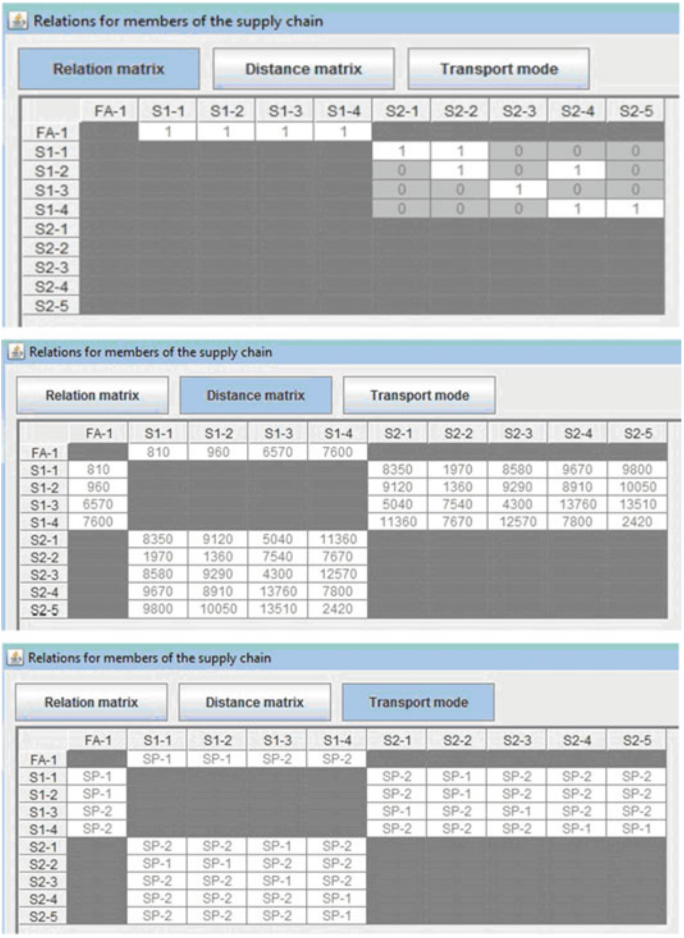


Fig. 5 Screens for parameter setting of relation matrix, distance matrix and transport modes

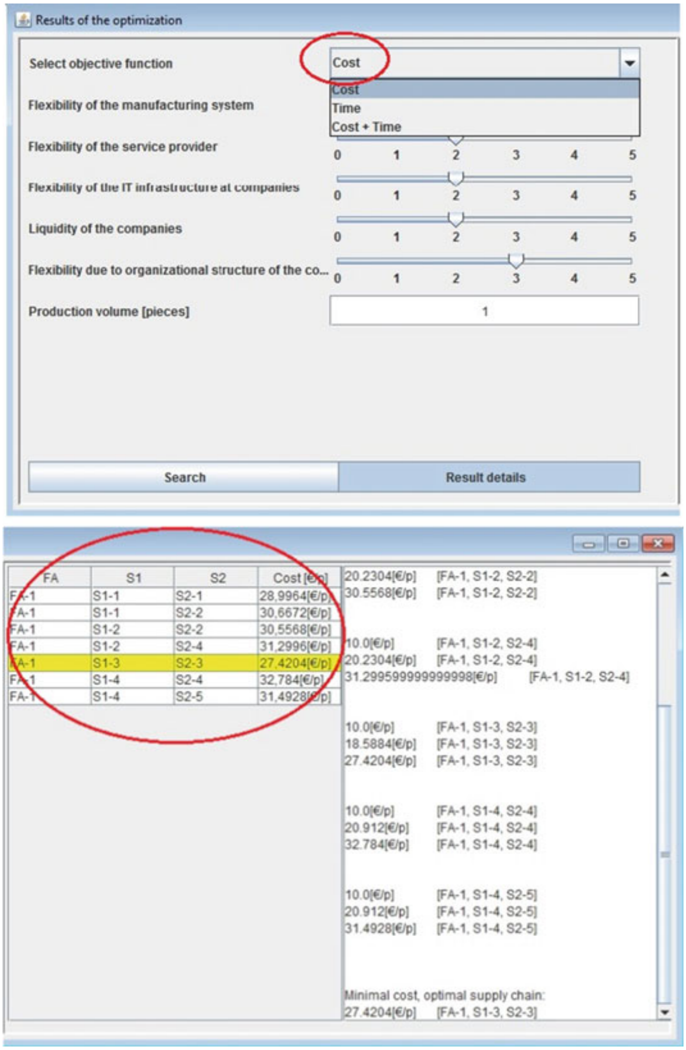


Fig. 6 Screens for selection of objective function, parameter setting of flexibility constraints and results of the optimization

4 Conclusion

In this paper the definition and characteristics of supply chains and virtual enterprises were introduced.

Authors elaborated the objective functions and constraints for optimization of virtual enterprises. Possible objective functions can be the minimization of total cost and lead time of the chain, possible constraints can be also defined e.g. for the production- and service capacities, inventories and flexibility of the chain members.

The research group developed a software application for formation of optimal supply chains which was introduced in this article.

Acknowledgements This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 691942. This research was partially carried out in the framework of the Centre of Excellence of Mechatronics and Logistics at the University of Miskolc.

References

1. Graham CS (1989) Integrating the supply chain., Int J Phys Distrib Mater Manag 19(8):3–8
2. Schönsleben P (2000) With agility and adequate partnership strategies towards effective logistics networks. Comput Ind 42(1):33–42
3. Camarinha-Matos LM (2001) Execution system for distributed business processes in a virtual enterprise. Future Gener Comput Syst 17:1009–1021
4. Gunasekaran A, Lai K, Edwin Cheng TC (2008) Responsive supply chain: a competitive strategy in a networked economy. Omega 36(4):549–564 (2008)
5. Tamás P, Illés B (2015) The concept of a virtual logistics center for a Hungarian Region. J Prod Eng 18(2):107–110
6. Esposito E, Evangelista P (2014) Investigating virtual enterprise models: literature review and empirical findings. Int J Prod Econ 148:145–157
7. Skapinyecz R, Illés B (2014) Presenting a logistics oriented research project in the field of E-marketplace integrated virtual enterprises. Appl Inf Sci Eng Technol Selected Topics Field of Prod Inf Eng IT Manuf Theory Pract 197–211
8. Illés B, Buczkó K (2008) Virtual logistics network for supporting the supplier selection. In: Proceedings 22th MicroCAD International Science Conference Miskolc (Hungary), 20.03.2008.–21.03.2008, 7–12
9. Telek P (2010) Operation strategies for delivery models of regional logistic networks. Adv Logis Syst Theory Pract 4:33–40
10. Gubán M (2011) Non-linear programming model and solution method of ordering controlled virtual assembly plants. Proc Logistics—the Eurasian bridge: materials of V. International scientifically-practical, Krasnoyarsk (Russia), pp 49–58
11. Java Programming, https://en.wikibooks.org/wiki/Java_Programming
12. Simon K (2009) Object oriented programming using Java. ISBN 978-87-7681-501-1, 1st edn. Simon Kendal&bookboon.com

Part IV
Welding

Development of Complex Spot Welding Technologies for Automotive DP Steels with FEM Support

László Prém, Zoltán Bézi and András Balogh

Abstract In this paper spot-welding technology of sheets made of Dual-Phase (DP) steel, which are one of the most important materials in the automotive industry, has been analyzed with experimental research and finite element modelling. Besides conventional mild steels that have been dominant for decades in automobile production, the application of advanced high-strength steels (AHSS) comes into focus. Among the first generation AHSS, DP steels are of the utmost importance concerning their automotive use. In Hungary a lot of small and medium sized enterprises function as the suppliers of the big automakers settled in our country. Nowadays these enterprises, besides the welding of conventional mild steel parts, often have the task of joining ferrite-martensitic DP steels with spot welding. During the spot welding of DP steels, when using the technology familiar with mild steels, the risk of hardening, the unfavourable failure of the joint or even cracking during operation must be taken into account, therefore welding technology must be planned on different theoretical bases. The traditional, continuous energy input and the symmetric double pulse as the non-continuous energy input were compared for resistance spot welding, focused on the advantages of pulsed energy input. For numerical analysis an axisymmetric coupled finite element model (FEM) is developed to study the effect of welding time and current intensity on nugget size and thermal history in resistance spot welding process using MSC.Marc software package. The cross-section macrostructures of the welded specimens are examined and compared to the predicted size of the weld nugget and heat affected zone (HAZ). Shear testing, cross tension testing, peel testing and hardness mea-

L. Prém (✉) · A. Balogh
University of Miskolc, Miskolc, Hungary
e-mail: metpreml@uni-miskolc.hu

A. Balogh
e-mail: balogh.andras@uni-miskolc.hu

Z. Bézi
Bay Zoltán Nonprofit Ltd, Miskolc, Hungary
e-mail: zoltan.bezi@bayzoltan.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_36

surements are also performed on specimens for each welding process to examine the effect of the changes in different welding parameters on the load bearing capacity of the joint.

1 Introduction

The automotive industry is one of the propulsive industries of Hungarian economy. In Hungary several automotive clusters exist which represent hundreds of small and medium sized enterprises which function as the suppliers of the big automakers settled in our country. Thanks to the high-expenditure automotive material developments in the world, engineers working at these automotive suppliers today may encounter steel grades with different strength and deformability during their work. It follows that the processing of these materials with the various forming and welding technologies can be very different, which could mean a great challenge in certain cases for the engineers designing these technologies. For decades the low strength unalloyed (so called mild) steels were decisive in the automotive industry because of their good deformability and excellent weldability. Due to the need for self-weight reduction of passenger cars, the ever increasing safety standards and gradually tightening environmental specifications the application of conventional and advanced high-strength steels came into view in recent years. Among the first-generation advanced high-strength steels (AHSS) the DP steels with a ferrite-martensite microstructure play a leading role concerning the application ratio and their processing with spot welding [1]. However, DP steels possess prominently high strength and significantly lower deformation ability compared to mild steels which are otherwise excellent for spot welding. Therefore the engineers planning the spot welding technology must take into consideration that the proven technology of the spot welding of mild steels might cause local hardness peaks, hardening, decreasing of load bearing capacity, as well as an unfavourable failure mode in the joints of DP steels [2].

The automotive sheets are mostly joined by some kind of resistance and arc welding process, or in some special cases (e.g. in the case of poorly weldable or non-weldable materials) by joining methods belonging to the welding related technologies. As a result of innovations and developments in recent years, among automotive joining technologies we can observe the application and slow expansion of processes such as adhesive bonding, soldering and various mechanical methods (riveting, clinching, etc.). However, it is an indisputable fact that today resistance spot welding is still the leading joining method of automotive steel sheets. Compared to the above listed joining methods resistance spot welding is the most advantageous from engineering and economic point of view, so currently this is the determinative method in car production and it can be safely predicted that it will

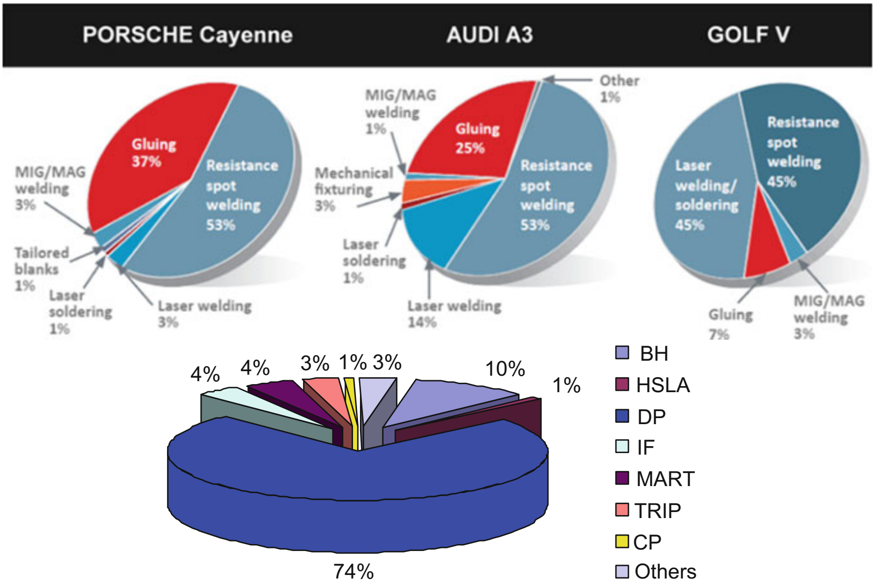


Fig. 1 Application ratio of welding and other joining methods applied in the production of the body in white unit of luxury cars [5]. Application ratio of different type of automotive steels in a modern passenger car [6]

keep its dominance in the near future. It can be clearly seen in Fig. 1 as well that in the production of modern, luxurious passenger cars the application ratio of spot welding is approximately equal to the total ratio of all the other joining methods applied. The plastic shaping technologies, the CAD/CAM systems, as well as the continuous development and evolution of finite element software made it possible for the car manufacturers and their suppliers to carry out more accurate strength calculations and produce larger compound parts with a complex geometry. Consequently the number of parts to be joined by welding and this way the number of welding spots decreased. The foregoing is proved by the fact that while some years ago an average passenger car's integral body and frame had 4000…5000 spot welds [3], this number was reduced to 2800…3500 [4] in the case of cars produced today (eg. Audi, VW, BMW, Porsche).

2 Dual-Phase Steels in Automotive Industry

The development of DP (Dual Phase) steels were started in the early 1970s with the aim of creating a type of steel with much better deformability within the strength range of HSLA steels [7]. Since the main user was the automotive industry, measures characterizing cold-workability (such as strain hardening exponent, hole expansion characteristics, minimum bending radius, spring-back etc.) gained significant role besides strength and elongation after fracture.

The main reason for the development of DP steels was the rising of crude oil prices due to the oil crisis caused by the 1973 Arab-Israeli war (Yom Kippur) and the response measure taken to mitigate the effects of this crisis, namely to produce passenger cars with a low fuel consumption. In favour of self-weight reduction (from among many other possibilities) the reduction of wall thickness of steel parts seemed to be the most efficient and have the most favourable cost implications, but this required steels with higher strength. On the other hand reducing production time and costs required such an upgrading of cold forming technology which needed the approximation of the high deformability of deep-drawable mild steels besides a twice-three times higher strength.

The expression 'Dual-Phase steel' was first used by Hayami and Furukawa (Japan) to emphasise the heterogeneous microstructure compared to mild steels thought to be Single-Phase (LSS) [8]. Despite its inaccuracy, this expression quickly spread and became generally used.

DP steels contain fine, dispersedly distributed, hard martensite islands embedded in a ferrite matrix. The soft ferrite ensures high deformability, whereas the hard martensite increases the strength of steel. By controlling the chemical composition, more precisely the carbon content and cooling rate, the ratio of martensite is variable between 20…70% [9]. This heterogeneous microstructure is usually achieved by continuous heat treatment (in a continuous annealing furnace) or by controlled cooling after hot rolling [4]. The essence of these heat treatments is that the desired ferrite/austenite structure is formed within the $A_1 \ldots A_3$ temperature range, then the steel is cooled down at high rate from the intercritical temperature, this way martensite is produced from the austenite during the process. The greatest difference of the production of DP steels compared to the normal hardening of steels is that the austenitization is only partial, therefore the austenite crystallites are surrounded by ferrite [10]. DP steels possess very high tensile strength (R_m = 600…1000 MPa), high toughness and relatively high deformability (A_{80} = 10…18%), as well as an excellent energy absorbing capacity in the case of a high-speed deformation. It follows from these very advantageous properties that the bumpers, A, B, C pillars and various reinforcing elements of the body (doors, roof etc.) of a passenger car could be primarily produced from the DP steel type and through their significant energy absorbing capacity these elements are intended to serve the safety of passengers during an incidental collision. Figure 2 demonstrates the application possibilities of DP steels in a modern car's body in white unit manufactured today [11].

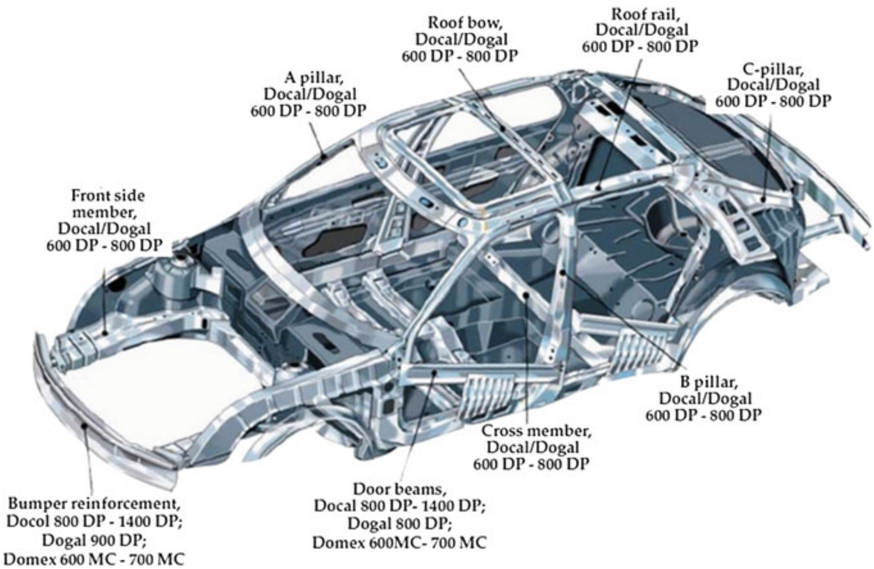


Fig. 2 Application of DP type Docol steels in a modern passenger car's body in white unit [11]

3 Resistance Spot Welding of Automotive Dual-Phase (DP) Steel Sheets

From among the available types of DP steels we have carried out experiments on the DP 1000 material which possesses the highest strength and lowest deformability, being the most challenging from a welding engineering point of view. The nominal thickness of the Docol DP 1000 steel sheets purchased from Swedish manufacturer SSAB was 1.0 mm. During the examination of the spot weldability of the experimental steel we examined the microstructure of the base materials (ferrite-martensite ratio), we determined the mechanical properties with destructive testing and the chemical composition of the sheets was checked with a spectrometer. The results of the tests are summarized in Tables 1 and 2.

During welding it has to be taken into consideration at any case that the base material (BM) contains so called 'soft' martensite with low carbon content, and that

Table 1 Mechanical properties of DP 1000 steel

Steel grade	Mechanical properties					
	R_m (MPa)	$R_{p0.2}$ (MPa)	$R_{p0.2}$ / R_m (−)	A_{80} (%)	Ratio of martensite (%)	Average hardness (HV0.2)
DP 1000	1047	721	0.69	11.33	55	324

Table 2 The chemical composition of experimental DP 1000 steel given in mass percent

Steel grade	Chemical composition							
	C (%)	Si (%)	Mn (%)	P (%)	S (%)	Nb (%)	V (%)	B(%)
DP 1000	0.148	0.49	1.50	0.010	0.002	0.015	0.01	0.0004

the microalloying elements can facilitate the hardening of the austenitized material volumes [12].

It is less known among welding experts that resistance spot welding also has its own weldability conditions just as fusion welding processes. It is an expectation for spot welding that the spot welds with the specified spot diameter should be created reproducibly, without cracks and with a load capacity characteristic of the base material and the joint type. Weldability strongly differs from that of arc welding processes due to the characteristics of spot welding (fast heating, small weld pool, compressive stress and crystallization during the intense heat-removal caused by the copper electrodes), belonging to pressure welding processes. The test criteria of spot weldability are usually given as the maximum hardness of joints and the appearance of unfavourable failure modes in welds determined by a qualification method.

During the examination of spot weldability the effect of the chemical composition of the base material is expressed by the carbon equivalent, similar to fusion welding. Equation (1) shows such a carbon equivalent, which was introduced by Japanese researchers for the weldability qualification of automotive AHSS steels. The 0.24% boundary value marks the limit where the strength (as well as the ratio of cross tension strength and tensile shear strength) of the joints begins to decrease as a function of the increasing strength of the base material [13].

$$CE_{RSW} = C + \frac{Si}{30} + \frac{Mn}{20} + 2 \cdot P + 4 \cdot S \leq 0.24\% \tag{1}$$

If the value $CE_{RSW} \leq 0.24\%$, the spot welded joint will predictably suffer plug failure, however, if $CE_{RSW} > 0.24\%$ the joint will show partial or interface failure, or will not suffer plug failure at all [14]. We have calculated the carbon equivalents of Docol DP 1000 from the chemical composition of experimental sheet metal with the help of Eq. (1), the values are summarized in Table 2.

The carbon equivalents of DP 1000 steel sheet exceed the 0.24% boundary value, the exact value is 0.27%. This means that in the case of using the conventional continuous energy input spot welding technology for the joints of DP 1000 materials cracking and/or unfavourable fracture modes can appear, or even the premature fatigue failure or brittle fracture of joints might occur during operation.

Therefore, the appliance of a complex working programme with non-continuous energy input, where the cooling rate can be controlled, is favourable in the case of spot welding DP steels. However, even with such a complex working programme, the cooling rate is greater than the upper critical cooling rate so the microstructure of the spot welded joint will be 100% martensitic.

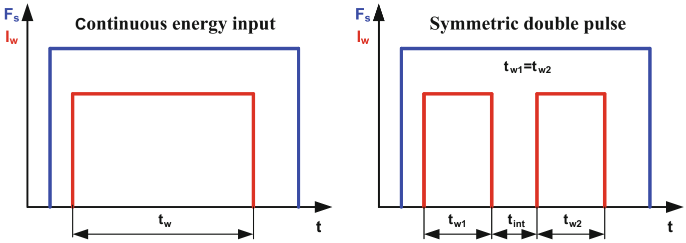


Fig. 3 The continuous energy input and the symmetric double pulse as the non-continuous energy input during RSW

We compared the traditional, continuous energy input and the symmetric double pulse as the non-continuous energy input during our experiments (Fig. 3). The effect of the intermediate time between the first and second pulses on the hardness distribution, susceptibility to hardening, load bearing capacity and failure modes of high strength steel joints, in case of spot welding with symmetric double pulse current, have been studied by experimental and numerical method. Based on the results of preliminary spot welding experiments conducted on DP steels we managed to define such parameter combinations (Table 3) which make it possible with both energy input methods to create spot welds with the highest possible and approximately equal tensile shear strength [12].

Then we prepared tensile shear for checking, cross tension, peel test specimens with the parameters of Table 3, because we assumed that the tensile shear test is less sensitive to technological variants than the other two test methods and it does not provide accurate information on the quality of spot welds in high-strength steels, primarily on their susceptibility to cracking.

The tests were performed on a sample containing 11 specimens. Design tensile shear strength, cross-tension strength and peel strength values were determined at 95% confidence level by using Student's t-test. Macroetching specimens and hardness distributions of joints examined during the comparison analysis are shown in Fig. 4.

The microscopic examination and the hardness measurement of welds gave very exciting results. We found that in the case of applying 14…15 cycles of intermediate time the outer part of the weld nugget starts to become recrystallized in a ring shape. This ring increasingly widens from the outside towards the inside with the increase of the intermediate time. The recrystallized ring is a fairly mild part of the spot weld with a good deformability, its hardness is approximately equal to the hardness of the base material, while the hardness of the non-recrystallized internal core with a coarse dendritic structure stays around 475…500 HV. In Figs. 4 and 5 the recrystallized ring and the internal core can be well observed [15].

The design shear strength of joints welded by symmetric double pulse is lower by 5…6%, than the same of joints welded by continuous energy input. This value approximately equals the rate of the differences in spot diameters. Despite the smaller spot diameters, more favourable failure modes and better load bearing

Table 3 Welding parameters of resistance spot welding experiments

Nomination	Continuous energy input	Pulsed energy input
Number of pulses:	1	2
Intermediate time: t_{int} (cycle)	0	15
Weld time: t_w (cycle)	12	6
Weld current: I_w (kA)	5.75	6.35
Electrode pressure: F_e (bar)	3	3

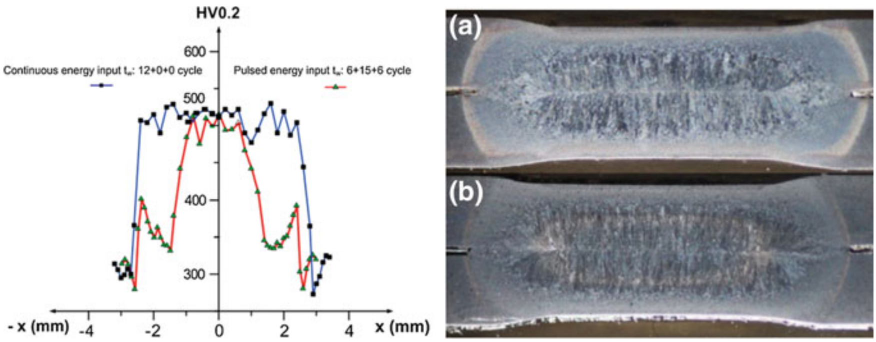


Fig. 4 Hardness distribution and spot weld of the joints of DP 1000 steel examined in the experiment, welded **a** with a continuous energy input, **b** with a symmetric double pulse. Magnification: 16×. Etching: Nital

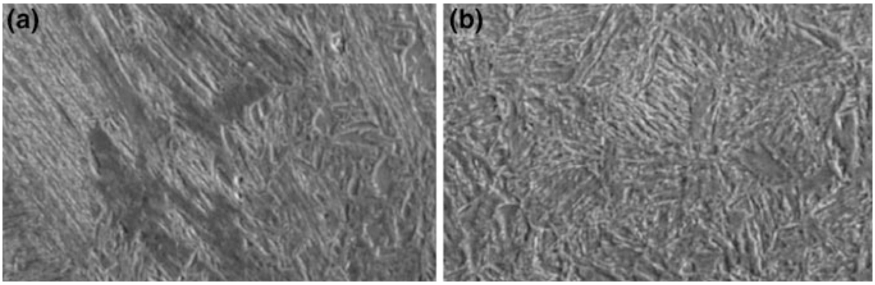


Fig. 5 SEM pictures from the structure of nuggets **a** dendrite structure with continuous energy input, **b** tempered ring with symmetric double pulse. Magnification: 1000×

capacities of joints welded by non-continuous energy input were observed during cross-tensile and peel tests (Figs. 6 and 7).

The results of the tension tests were proved that using an appropriately increased intermediate time between pulses the outer part of weld nugget can be recrystallized and tempered with the second current pulse, the microstructure of spot welds in DP steels can be fined, their hardness distribution and the conventional static mechanical properties (summarized in Table 4) and the failure mode of the joints can be significantly improved [12, 15].

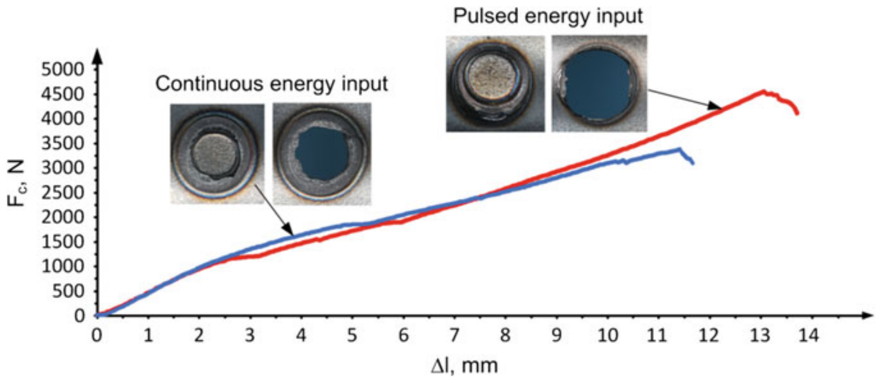


Fig. 6 Comparison the cross tension test results

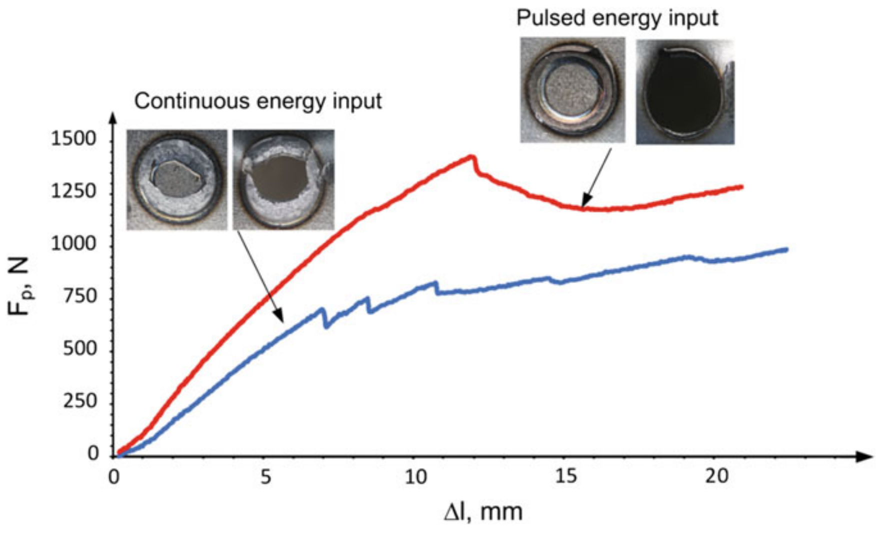


Fig. 7 Comparison the peel test results

Table 4 Summary of the results of tensile shear, cross tension, and peel tests

Nomination	Energy input	
	Continuous energy input	Pulsed energy input
Tensile shear force: F_s (kN)	12.38	11.93
Cross tension force: F_c (kN)	3.44	4.08
Peel force: F_p (kN)	0.95	1.24

In order to appreciate the effects of microstructure on the dynamic load capacity and hardness distribution of the joints, it is important to recognize the various weld regions. In the Fig. 4 the typical welding zones can be identified clearly. The fusion zone (FZ) is created by heating above the melting point, while the surrounding heat affected zone (HAZ) material itself consists of several regions which experience thermal cycles with progressively decreasing peak temperature from the fusion boundary. Outside the fusion zone, near the fusion boundary, the peak temperature exceeds the A_{c3} upon which 100% austenite forms. This region is called the supercritical HAZ. Further from the fusion boundary, the peak temperature is within the intercritical ($A_{c1} - A_{c3}$) range, which results in austenitization of the carbon-rich martensite phases while large islands of undissolved ferrite remain unchanged. The cooling rate is typically high enough to result in the decomposition of austenite to hard metastable phases such as martensite, bainite and carbides. Thus, the room temperature microstructures of the fusion zone, the supercritical and intercritical regions consist of a higher volume fraction of hard phases than the base metal. In the subcritical HAZ, the peak temperature is below the A_{c1} resulting in tempering of metastable martensite. In general, the subcritical area will soften during the welding of DP steels. The soft zone is partly attributed to the disappearance of martensite and partly due to tempering of the remaining martensite pre-existing in the base metals during RSW [16].

In order to analyse the effect of RSW processes on the microstructure of the spot welds and the welding metallurgy of the DP 1000 steel numerical models must be used. The microstructure of the welded joint depends heavily on the base metal structure. So it is important to know the thermal history and the effects of the different regions on the resulting microstructure.

4 Finite Element Modelling

FE modelling of the spot welding process can be difficult for most of the modelling tools including finite element based software, as RSW is governed by electrical-thermal, mechanical and metallurgical phenomena. It is difficult to simulate the RSW process because three different physical phenomena are interacting with each other. The model takes the following physical and metallurgical interactions into consideration in the simulations: interaction between the electro-kinetics and heat transfer via the Joule effect, heat transfer and phase transformations through latent heat and heat transfer, electro-kinetics, and mechanical behaviour via contact conditions.

The welding process starts with analysing the squeeze cycle in which electrode force is applied to the electrodes. The results of this mechanical analysis include initial deformations and contact area, which serve in electro-thermal analysis. In this stage, the temperature distribution by Joule heating is calculated for an increment from the fully coupled electrical thermal FEA. In the electrically thermally coupled analysis the electrical and thermal boundary conditions are applied to the model in a house. The calculations of Joule heating at the sheet-sheet and electrode-sheet faying surfaces, as well as in the base material and electrode are then performed. As a result, the temperature distributions are obtained in the first increment and sent to the mechanical analysis as a nodal thermal load. Contact pressure and deformations are the results of mechanical analysis that obtains a new contact condition. So, the mechanical results are transferred to the electro-thermal analysis to update contact conditions for the next increment analysis. This loop continues until the welding time is finished.

Numerical calculations were performed with commercial finite element (FE) software package MSC.Marc and Simufact.welding solver. The problem was solved numerically in a cylindrical coordinate system due to axial symmetry of the geometry. The actual geometry of the electrode was measured and replicated in the model. The associated element mesh is shown in Fig. 8. The finite element mesh contains 4095 linear quadrilateral elements and 3822 nodes.

Thermal-electric and mechanical boundary conditions were applied to the FE model. Heat transfer to the surrounding air, using convection and 20 W/m^{2}K convective heat transfer coefficient was used. Temperature at the far end of the steel sheet and the water temperature inside the copper-chromium-zirconium (Cu–Cr–Zr alloy) electrodes were both assumed to be 20 °C. At the upper electrode, current and mechanical pressure was applied uniformly to the top of the electrode.

The elastic behaviour is modelled using the isotropic Hooke's rule with temperature-dependent Young's modulus. The thermal strain is considered using thermal expansion coefficient. The yield criterion is the Von Mises yield surface. In the model, the strain hardening is taken into account using the isotropic Hooke's law [12].

4.1 Material Properties

In order to capture the correct microstructure evolution, all of the relevant mechanical, thermal and electrical properties of the sheets and electrodes must be

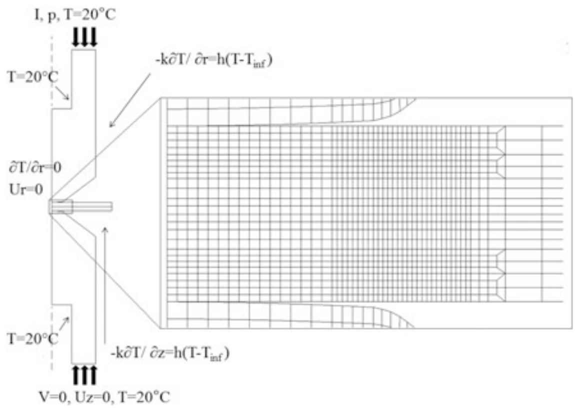


Fig. 8 Associated finite element model

Table 5 Electrical properties

	Electrical resistivity DP 1000 (Ωm*10^{-7})	Electrical resistivity electrode (Ωm*10^{-7})	E. contact conductance sheet/sheet (1/Ωm*10^{9})	E. contact conductance electrode/sheet (1/Ωm*10^{9})
20	2.432	0.257	0.203	1.075
100	2.975	0.315	0.298	2.174
200	3.756	0.397	0.401	3.334
300	4.656	0.487	0.505	4.348
500	6.985	0.675	0.714	6.668
700	10.03	0.845	0.909	8.621
1000	14.44	1.001	1.222	12.05
1300	15.87	–	1.538	16.67
1600	17.51	–	1.852	18.52
2000	17.62	–	3.025	26.13

known. Since electrical and physical properties vary with temperature and are not readily available, many of these values were generated with JMatPro software and assumed to be homogeneous. Contact resistivity of the sheet-to-sheet and electrode-to-sheet interfaces were also assumed to vary with temperature (Table 5).

Material properties (i.e. thermal conductivity, specific heat, thermal expansion coefficient, density, Young's modulus and yield strength) for steel having nominal tensile strength of 1000 MPa were considered to be temperature dependent. Example of properties at room temperature has been listed in Tables 1 and 6.

For numerical simulation of RSW, the accuracy of critical austenite transformation temperatures is crucial. The start temperature for the austenite transformation depends strongly on three essential parameters, namely heating rate, initial microstructure and chemical composition (Table 2). Ferrite-austenite transformations at high heating rates are characterized by JMatPro welding-cycle interface. The temperature and strain rate depend flow curves are calculated by JMatPro Simufact.premap interface with 30 μm initial grain size starting at 1300 °C for each phase. A similar calculation has been done in [17].

Cu–Cr–Zr alloy, which has high thermal conductivity performance, is selected for electrodes. Non-linear time dependency of thermal and electrical material properties as well as convection coefficient rate for water, air and gas flow are all obtained from literature and paper [17].

Table 6 Physical and mechanical properties at room temperature

	Thermal conductivity (W/mK)	Specific heat (J/kgK)	Thermal expansion coefficient (1/K*10^{-6})	Young modulus (GPa)	Density (kg/m^3)
DP 1000	55	448	12.8	218	7850
Cu–Cr–Zr alloy	450	380	18.5	129	8840

5 Comparisons Between the Numerical Calculation and Experiment Results

A simulation model has been developed and extensive numerical calculations were carried out to find out the spot diameter and hardness distribution of resistance spot welded DP 1000 joints. Figure 9 shows the thermal cycle curves in different locations calculated by MSC.Marc software according to the welding parameters listed in Table 3.

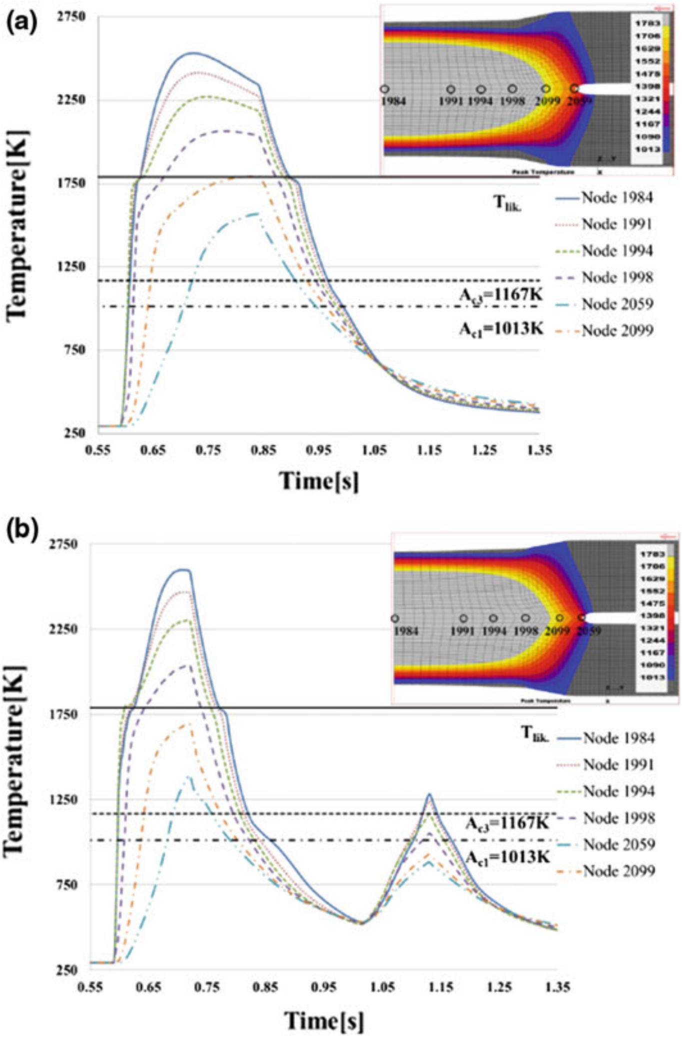


Fig. 9 Temperature history: **a** continuous energy input, **b** symmetric double pulse

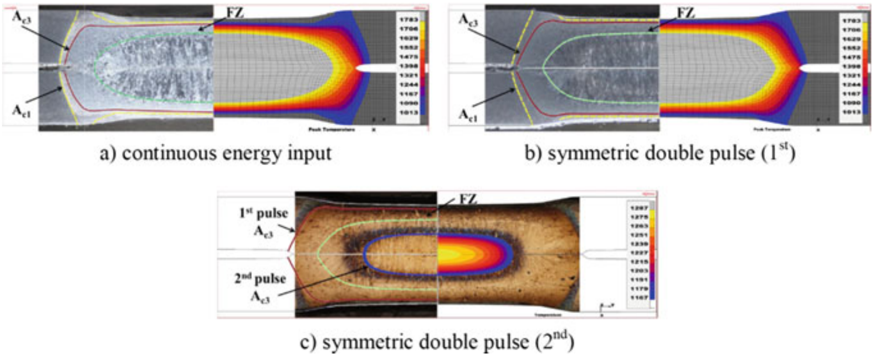


Fig. 10 Comparison of simulated peak temperature map to experimental micrograph

Predefined mesh nodes are identified in Fig. 9. The location of these nodes was used to compare the thermal history within the sub-regions for different RSW parameters. During the single-pulse RSW process (Fig. 9a), the DP steel was heated to the peak temperature of 2530 K at an average rate of about 8000 K/s. After reaching the peak temperature, the steel was quenched to about 300 K with 4500 K/s during where the martensitic transformation can occur. Figure 9b shows the modelled temperature profile at the reference nodes for double-pulse RSW. After the heating of the first pulse the specimen was cooled to about 550 K, where partial austenite may transform into martensite, and then it was reheated by adding the second pulse. The reheating effect of second pulse was predicted to occur in all weld regions. For the inner-part of the FZ region the temperature range is typically 1200–1300 K at the reference nodes 1984–1991. The outer-part of the FZ and intercritical (IC) region peak temperature was predicted (node 1994–1998) to be 1100 K and typical range from A_{c1} and A_{c3}. The peak temperatures near the node 1994 was exceed A_{c3}, resulting in a fully austenitized local structure. The short time above A_{c3} was limited grain growth, producing an ultra-fine structure upon cooling in this region.

Figure 10 compares the peak temperature contours with experimental cross-section micrographs. The FZ, HAZ and BM region can be clearly observed. The grey region mostly represents the FZ and the black grey region mostly represents the BM that is not affected by heat. The regions with other colours are the HAZ, where the material is partially transformed into austenite. The size of FZ and HAZ is again in good agreement with the experimental observations.

The microindentation hardness profiles (measured and simulated) of welded DP 1000 steel joints are shown in Fig. 11. Significantly higher hardness values, approximately 1.5 times higher than that in the base metal were observed in FZ zone. The profile was relatively flat across the fusion zone with an average FZ hardness of 490HV. In the IC HAZ of joints near the fusion boundary also exhibited high hardness, dramatically decreasing toward the unaffected base metal

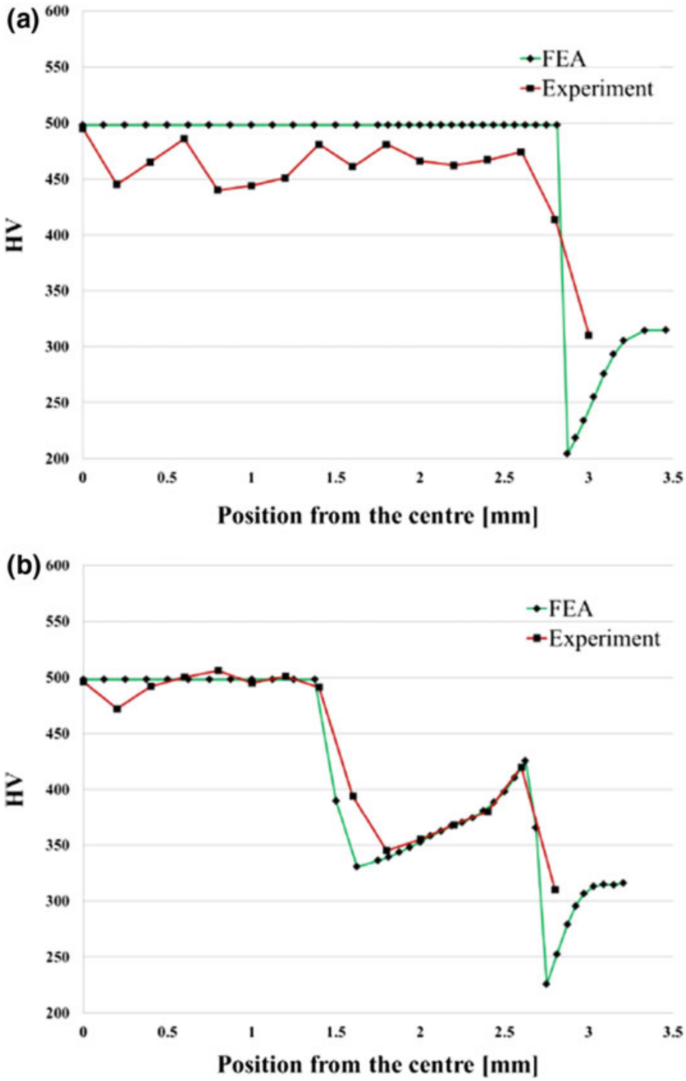


Fig. 11 Comparison of calculated and measured hardness data **a** continuous energy input, **b** symmetric double pulse

(324HV). A softened zone was observed in the outer-HAZ where the hardness was significantly lower (230HV) than the base metal.

Figure 11b shows how the second pulse affects the measured hardness values. For a second pulse softening of the FZ was a result of low cooling rates. This suggests an increase in the transformation time and auto-tempering of the martensitic structure during cooling [18]. The results of the numerical simulation support this experimental result. So as to fully understand softening kinetics and the relationship between microstructure and mechanical properties a more extensive study should be done on this type of DP steel with various welding parameters.

6 Conclusions

Concerning the experiments and numerical modelling of resistance spot welding of automotive DP steel sheets, the following conclusions can be drawn:

1. The results of the tests show that the failure mode of joints, the plastic strain reserves and the conventional static load bearing capacity can be improved by applying symmetric double pulsed energy input. Because of its many advantages it is more favourable to use pulsed energy input for the spot-welding of such automotive parts and load bearing elements produced from advanced high strength DP steels, that are important for life and property safety (e.g. B-pillars).
2. The simulation shows that the second pulse can improve the weld strength of the welded joint. The results of the FE simulations can be used for in process planning of RSW in modern manufacturing. The results show that the accuracy between physical experiments and numerical simulations is comparable to the variations of weld output in practical welding. Thus, the prediction accuracy of simulations is useful for engineers using process planning.

MSC.Marc simulations can be used to predict post weld hardness, adding some metallurgical capability to the software. Additional investigations are required to determine which the most suitable combination of the welding parameters for the achieving of the ideal hardness distribution.

Acknowledgements The research work presented in this paper based on the results achieved within the TÁMOP-4.2.1.B-10/2/KONV-2010-0001 project and carried out as part of the TÁMOP-4.2.2.A-11/1/KONV-2012-0029 project in the framework of the New Széchenyi Plan. The realization of this project is supported by the European Union, and co-financed by the European Social Fund.

References

1. H. World Steel Association (2009) Advanced high strength steel (AHSS) application guidelines, Version 4.1, pp 1–16
2. Prém L (2014) Spot welding experiments of automotive Dual-Phase steel sheets. Publications of the MultiScience. In: Proceedings 28th microCAD, international multidisciplinary scientific conference, University of Miskolc (Hungary)
3. Weman K (2003) Welding process handbook. Woodhead Publishing Ltd and CRC Press LLC
4. Janota M, Neumann H (2008) Share of spot welding and other joining methods in automotive production. Welding in the World

5. IIW White Paper (2012) Improving global quality of life through optimum use and innovation of welding and joining technologies, p 36
6. Advances in high strength steels for automotive applications, www.wordautosteel.org
7. ASM Handbook (2005) Properties and selection: irons, steels, and high performance alloys, 10th edn., vol 1, p 697
8. Tsipouridis P (2006) Mechanical properties of dual phase steels, PhD dissertation, Technische Universität, München (Germany)
9. Dziedzic M, Turczyn S (2010) Experimental and numerical investigation of strip rolling from dual phase steel. Arch Civil Mech Eng 10(4):21–30
10. Tisza M (2015) Material and technological developments in sheet metal forming with special regards to the needs of the automotive industry. Arch Mater Sci Eng 71(1):36–45
11. DOCOL advanced high strength steels for automotive industry, www.ssab.com
12. Bézi Z, Prém L, Balogh A (2016) Development of resistant spot welding technology for automotive ferrite-martensitic dual-phase steels with joint application of finite element modelling and experimental research. Adv Mater Res 1138:43–48
13. Oikawa H, Sakiyama T, Ishikawa T, Murayama G, Takahashi Y (2007) Resistance spot weldability of high strength steel (HSS) sheets for automobiles. Nippon Steel Technical Report, No. 95
14. SSAB: welding of AHSS/UHSS steel. A guide for the automotive industry
15. Prém L, Balogh A (2015) Symmetric double pulse with increased intermediate time for resistance spot welding of ferrite-martensitic DP steel series. In: Young welding professionals international conference, YPIC 2015, Budapest (Hungary)
16. Khan MI, Kuntz ML, Biro E, Zhou Y (2008) Microstructure and mechanical properties of resistance spot welded advanced high strength steels. Mater Trans 49(7):1629–1637
17. Bézi Z, Baptiszta B, Szávai Sz (2014) Experimental and numerical analysis of resistance spot welded joints on DP600 sheets. BID-ISIM welding and material testing 23(4):7–12
18. Khan I, Kuntz M, Zhou Y, Chan K (2007) Monitoring the effect of RSW pulsing on AHSS using FEA (SORPAS) software. SAE Technical Paper 2007-01-1370

A Lightweight Design Approach for Welded Railway Vehicle Structures of Modern Passenger Coach

István Borhy and László Kovács

Abstract The design of railway vehicle structures is a complex task requiring absolute care, well-founded theoretical knowledge and significant practical experience. The requirements related to the vehicle structures make it necessary to manufacture structures that have the right amount of stiffness and are adequate also in terms of fatigue resistance, but have the lowest weight possible and can economically be manufactured, and at the same time are simple and easy to maintain. The main aspects of designing for modern welded vehicle structures are, as follows: suitable load-bearing capability and safety, the correct construction, according the welding technology, and the same, the cost-effectiveness, as well. We can achieve significant reduction of weight by increasing of the strength of elements and by reduction of the thickness of plates and construction elements. The minimum of the necessary thickness of plates can be calculated by optimization of sizes. In our lecture we will present the weight reduction design approach, and new innovative design methods and technologies for lightweight welded railway structures of the modern passenger coach by the examples of new IC+type railway passenger couch family, developed by MÁV-START Ltd.

1 Introduction

The development of railway transportations is the priority for the European Union and for Hungarian National Transport Policy [1, 2]. The Hungarian government specified the share incise and prove of competitiveness of railway transportation, as part of National Railway Development Conception [3] in its government decision about the Strategy of National Transportation Infrastructure Development,

I. Borhy (✉)
TÜV Rheinland InterCert Kft, Budapest, Hungary
e-mail: iborhy@hu.tuv.com

L. Kovács
MÁV-START Vasúti Személyszállító Zrt, Budapest, Hungary
e-mail: kovacs.laszlo13@mav-start.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_37

in August 2014. To achieve these objectives, necessary to decry the average age of rolling stock, it means refurbishment (main repair) or changing of weathered railway fleet [4]. This could be a good possibility to revivify the traditional, domestic manufacturing of railway stock, the same via innovative development it could contribute the income generating capacity and the employment, improve domestic small- and medium-sized enterprises in the field of sub-suppliers. As the first step of this process, the MÁV-GÉPÉSZET Ltd. in 2012—on the base of EU support and co-financing of European Regional Development Found—started to design the first members of IC+type railway passenger couch family and the manufacturing of the first example [5].

This lecture describes the designing, manufacturing and conformity assessment of the body structure as a welded vehicle structure, via presentation of examples of engineering, manufacturing and conformity assessment of 2nd class coach.

2 Presentation of IC+Type Railway Passenger Coach Family

MÁV-GÉPÉSZET Ltd. in 2011 got owner's authorization to start own manufacturing of railway vehicles, as the "spine" of national railway passenger transport, to develop national inventions, domestic manufacturing potential and to keep and develop domestic employment. The goal of this project to meet the high demands of passengers to develop a vehicle type Z1, which is economically manufactured and maintainable, and which is able to travel with 200 km/h speed in domestic and international traffic. The main elements of the family:

- 2nd class saloon type coach;
- Shared 1st class and bistro coach;
- Multipurpose grow wagon, for person with reduced mobility with wheel-chair and transportation of bicycles;
- Driving car with 2nd class compartment interior.

As the first phase of the project, the developing and preparing of manufacturing of two pieces 2nd class grow wagon was done. Vehicles have an air-conditioned passenger area (Figs. 1 and 2), closed-circuit vacuum toilet, GPS controlled, audio-visual passenger information system. Vehicles have an energy supply at each seats and internet connection via free Wi-Fi [6].

The European Parliament and Council of European Union defined rules of Interoperability of the rail systems Directive [7] for international mobility of railway vehicles. This Directive defines rules interoperability of safety designing, manufacturing, operation and maintenance of elements, the same, the roles conformity assessment and licensing of these elements. The adoption of these Directives of the domestic rule of law is done [8], so the licensing by the National Safety Authority is done, the type approval is issued (the number of type approval is:

Fig. 1 The prototype of IC+2nd class coach (Photo: Sándor Czeglédi)

Fig. 2 Saloon type interior arrangement of the 2nd class coach (Photo: Sándor Czeglédi)

UVH/VF/320/9/2015 (1st June 2015)). The precondition of the licensing is the approval of the full compliance of requirements by a Notified Body. The process of conformity assessment is done by experts TÜV Rheinland InterCert Kft, in cooperation with Luxcontrol Nederland B.V. and TÜV Rheinland InterTraffic GmbH., belonging to the TÜV Rheinland Group, according requirements of SB+SD audits defined by resolution No.: 2010/713/EU [9].

On the base of traffic experiences of prototype vehicles, the Hungarian Government issued Government Decision No. 1167/2016 (IV.6.) [10] about support of 100 pieces IC+type passenger coach. The serial manufacturing starts in 2017 in Szolnok.

3 The Carbody Structures of Railway Vehicles

The main construction part railways vehicles is the carbody structure, it task is to ensure the strength and stiffness against the outsider dynamical stresses, the same the reliable protection of passengers and intern area [11]. The forming of the carbody structure—as the load-bearing element, is done according the principle of lightweight construction. This lightweight construction decreases the weight of the coach, via this fact it has positive influence to operational costs, decreases the load of landing gears, which is important in the case of vehicles, running over 160 km/h speed [12].

The engineering of railway vehicles is a complicated task, it demands a high degree of care, a solid theoretical knowledge. Engineers have to accept and evaluate a lot of viewpoints of—each other often conflicting—requirements for the aim to meet the requirements. To meet all requirements is a hard task, necessary to accept requirements of strength, functionality, reliability, aesthetics, etc., necessary to accept the requirements of manufacturability, controllability, and to minimize the costs of full lifetime (LCC). The typical failure mode if welded structures of vehicles—in addition of the corrosion—appearing and developing of cracks in result of the exhausting stress, thus for these effects have to be accepted during the full process of engineering, and make all necessary measures to avoid these.

Over these, due attention has to be paid to the development of bright and spacious interiors, heat and sound deadening and to ensure the flow and smooth internal temperature during the elaboration of engineering concept of carbody structure.

The requirements of the vehicle structure with sufficient rigidity and the wearing stress is also suitable, but at the same time the lowest possible weight, economical to manufacture and easy respectively. The development to maintain a low cost structure is required. The carbody structures of passenger coach are characterized by rib stiffened hull plate that bears the chassis together with all the load acting on the vehicle. The individual structural elements plating themselves participate in the rigidity of the structure and design of the vehicle's longitudinal axis parallel forces uptake. In case of a differentiated method of construction for each strengthen elements and panels are made out separately and are compiled into an assembly unit. For integrated moulded method of construction elements, beam-profiles, which at the same time provide the structural and cladding functions. The development of modern passenger rail vehicles uses both structures carbody structures construction methods characterized by: IC+passenger coach carbody structure was constructed on the basis of differentiated method of construction (Fig. 3).

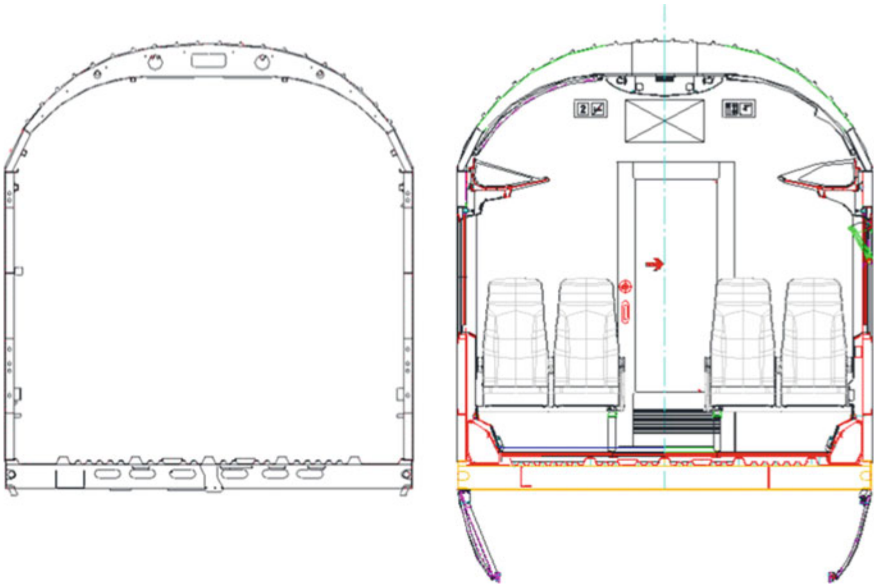


Fig. 3 Cross-section of the carbody structure and the welded structure

Fig. 4 Welded side wall elements (Photo: István Borhy)

The railway structures of the high-powered building will be designed according to the principle by which the pre-built components (chassis, side panels, front, roof, etc.) installed in the cabinet assembler machine welded together (Figs. 4 and 5). The main structural elements of the carbody structures:

Fig. 5 Welded carbody structure in the assembling machine (Photo: István Borhy)

- Corrugated floors covered chassis;
- Sidewalls;
- End walls;
- Roof.

4 Design of the Carbody Structures of IC+Type Railway Passenger Coach

The car body structure of the IC+type coach is developed according requirements, mentioned above. The full life-time of coach is 30 years. During engineering was paid attention to increase the quality and reliability of vehicles structure, and the same to decry manufacturing costs. The elements of family meet requirements via unified solution of platform-conception (Figs. 6 and 7).

The 3D design of IC+coach done with Autodesk Inventor 2012 software, the cooperation between engineering groups is supported with Autodesk Vault centralized controller. The requirements of strength are detailed in standard MSZ EN 12663-1:2010 [13]. The standard categorizes the several vehicles into several groups, the same, the kind and volume of test loads are categorized into groups, too. The carbody structure if IC+type coach belongs to the category P-I.

The results of engineering have to be proved (verificated and validated) by calculations and measuring to order assessment. The finite element calculations are made by pre- and post-processing MSC Nastran 2012 with MSC SimXpert 2012 (Fig. 8).

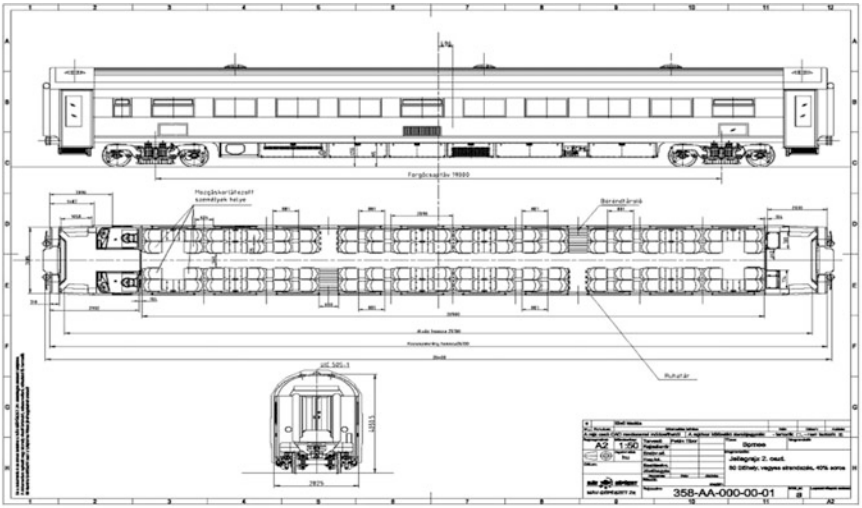


Fig. 6 Layout drawing of IC+2nd class coach

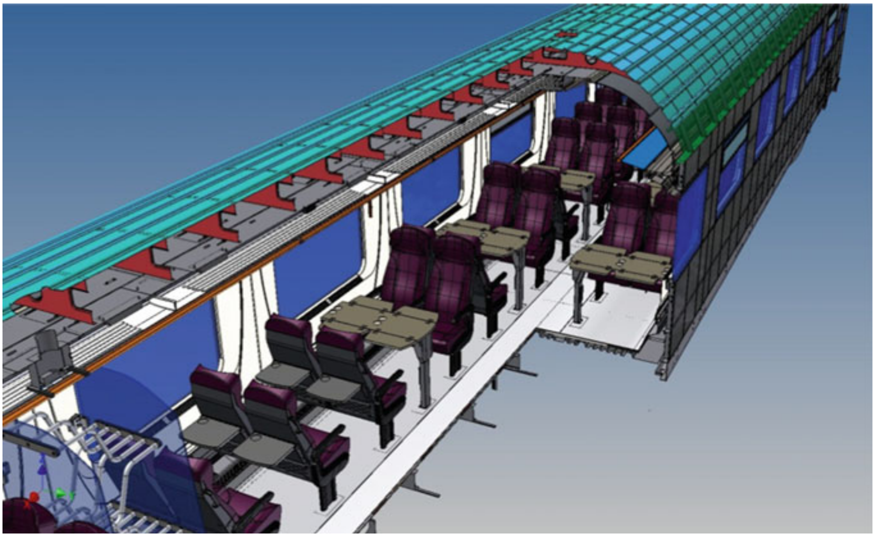


Fig. 7 Section view of carbody 3D model

The strength measurements are made by MÁV Ltd. Vasúti Mérnöki és Mérésügyi Szolgáltató Központ (VMMSZK) in the III. quarter of 2012 (Fig. 9).

The calculation and measurements proved strength of carbody structures. The necessary constructional changings are made on the base of measurements results.

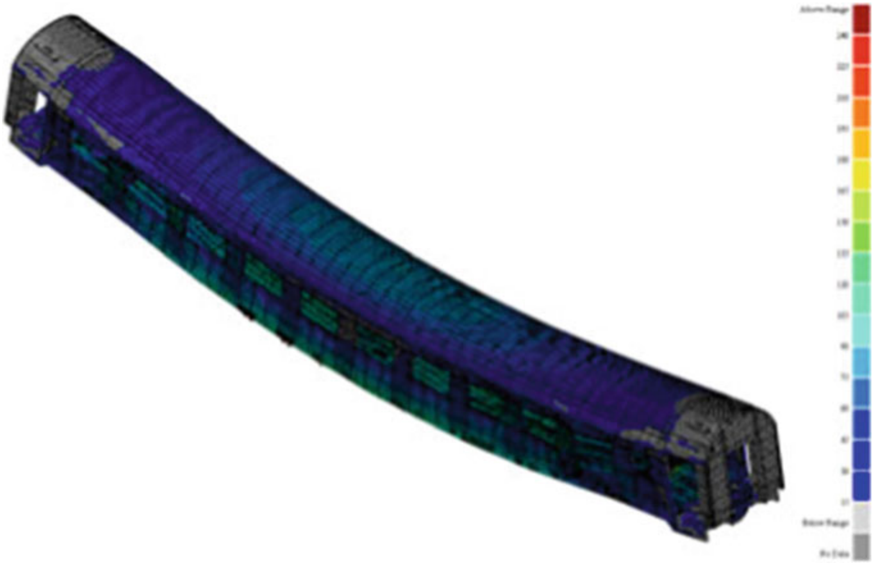


Fig. 8 3D FE model of carbody (Stress distribution and deformation due to standard loads)

Fig. 9 Strength test of carbody (Photo: Sándor Czeglédi)

5 Developing of Innovative Methods and Joining Technologies

When we prepared the serial production of IC+coaches, the secondary aim was to elaborate and develop innovative methods of design and production technologies. The first target was to elaborate and prove the ability to use during the engineering of construction and technology of welded railway structures, a multiple objective function optimization procedure.

The environmental awareness developed some changing in manufacturing of railway vehicles at the late '90. The main method to reduce the mass of vehicles is to use materials with increased strength. The main material for carbody structure manufacturing are the unalloyed and low alloyed steels, but in result of the material development, the percentage of strength/mass of used materials is better and better, which has a positive influence on the load-bearing capacity and mass of vehicles.

The use of strengthen materials and decreased weight, demand to use less energy, so it decries the operational cost. The moving of vehicles is possible with less powered drives, or in the case of similar performance of the drive, the dynamical characteristics of vehicles can be higher (speed, acceleration). The use of high strengthen steels is provided by stricter regulations, which allow less and less exhaust emissions.

The more effective method reduced the mass and cost is to make thicker the plates and beams. The result is the thick-walled welded body construction. The engineering of these constructions demands to accept a lot of questions. The shrinkage of welded connection makes permanent tensions and warps, thin plates and shells reformate and needs to damp vibration, these structures are sensitive to inhibit warping. To avoid these effects the thin plate walls should be wrapped. The optimal thickness of plates should be calculated. The main target was to reduce or minimize the local and global deformation of vehicle structures.

One of the preferred technologies of vehicle structure manufacturing is the resistance spot welding, which is proved by a lot of facts, for example the consistent quality, effective manufacturing, easy to automate, etc. The engineering, economic, environmental factors, the permanent developing if technology (non-continuous energy input, decried the specific heat input) ensure, that this technology will be a decisive technology. Our target is to define optimal technological parameters (welding sequence, parameters), to explore the sensitivity of manufacturing against these parameters.

6 Optimization of Spot Welded Sidewalls

Our presentation is part of our earlier research activity, which was reported for you [14–17]. The sidewalls—as welded structures—have to meet requirements of strength and stiffness, and the same have to meet aesthetical requirements, and the manufacturing has to be economical, too. The economic manufacturing in this case means to find the minimal manufacturing cost; the target described by the next function:

$$k = k_{material} + k_{fabrication} = k_{material}\rho V + k_{fabrication}\sum_{i=1}^{n} T_i \quad (1)$$

where: $k_{material}$ and $k_{fabrication}$ are manufactural costs, ρ is the material density, V is the volume of construction, n is the number of spot welding phases (incl. positioning of electrodes and factual realization of welding), T_i is the time required for the ith welding phase.

The engineering of the vehicle carbody meets requirements of MSZ EN 15085-3 [18], the annex F details requirements for quality of spot welded junctions. The characteristics of weld performance class (CP C1-CP D), and requirements of surface strength are fixed too. During the calculation of sidewalls, the requirements of MSZ EN 12663-1:2010 are accepted, the determining the value of the fatigue limit stress is important too. The fatigue life than limiting condition can be used in a vehicle related to the welded constructions of optimization of construction [19] (Figs. 10, 11 and 12).

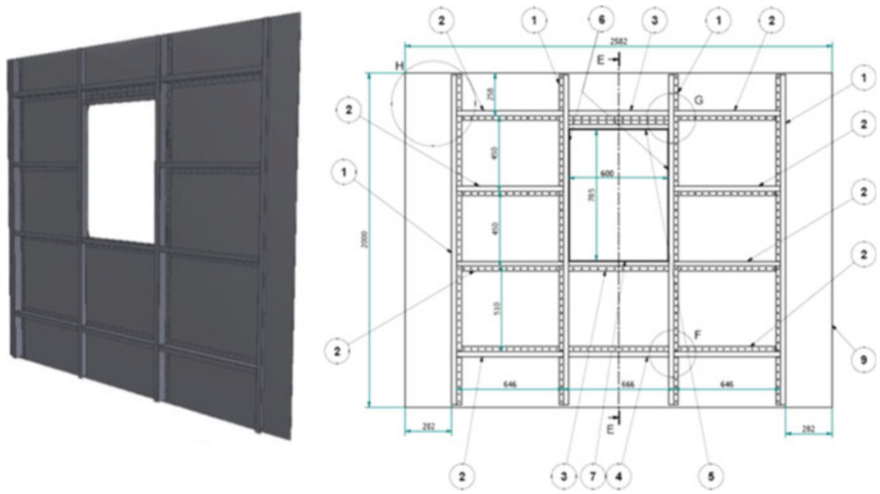


Fig. 10 Compilation of 3D sidewall model and its drawing

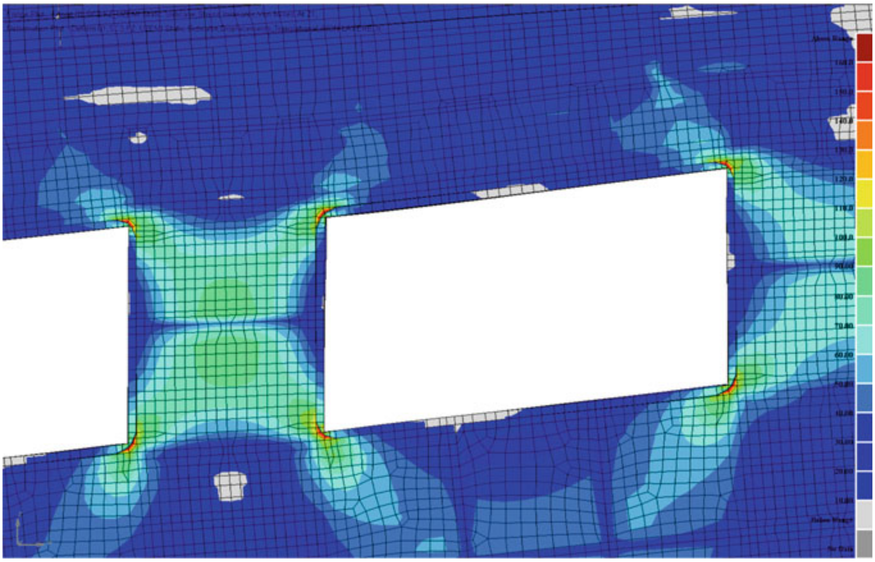


Fig. 11 Stress and deformation on sidewall near window

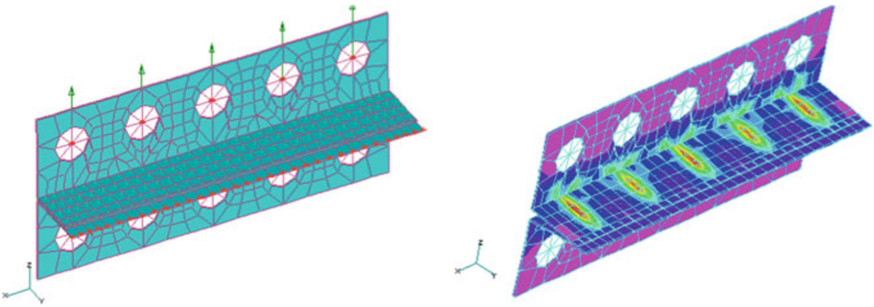


Fig. 12 Finite-element model, equivalent stresses and deformations in the case of the HP-type specimen (F_i = 100 N)

7 Conclusions

In this lecture we demonstrated questions of manufacturing, conformity assessment of 2nd class IC+type passenger coach, developed by MÁV-START Ltd. We turned the attention to the preferences of resistance spot welded constructions as a technological possibility in the in serial production. The actuality of this topic is proved by the fact, the results of research are close to use in the serial manufacturing. Given that the family development allows for further tests, so the development of applied research with our work in the welded plate structures designing and manufacturing optimization methods we intend to contribute.

Acknowledgements This presentation was conducted as part of the Project IC-PR01-GOP131/A, on the basis of contract No.: K2378-11, between MÁV-GÉPÉSZET Ltd. and TÜV Rheinland InterCert GmbH. The realization of this project is supported by the European Union, and co-financed by the European Regional Development Fund.

References

1. White paper: roadmap to a single european transport area—towards a competitive and resource efficient transport system, http://ec.europa.eu/transport/themes/strategies/doc/2011_white_paper/white_paper_com(2011)_144_en.pdf, COM(2011) 144, Európai Bizottság, Brüsszel, 2011
2. Government Decision: 1486/2014 (VIII.28.) Korm. határozat: National Transport Strategy http://www.3k.gov.hu/remos_downloads/NTS_Strategic%20document_version%20for%20public%20discussion_EN_2013.70.pdf, Hungarian Transport Administration, Budapest, 2013
3. National Transport Strategy (2013) National railway development concept. Hungarian Transport Administration, Budapest
4. The business strategy of MÁV-START Ltd. (period from 2014 to 2023) for improve the level of public transport (in Hungarian), Founders Decision: 10/280/2014 (20 Mar 2014)
5. Kovács K (ed) (2011) The MÁV-GÉPÉSZET Ltd. started the development of new railway passenger coaches (in Hungarian), Railway Engineering (ISSN 0230-0605), No. 3, p 11
6. Mezei I (ed) (2014) Design and development of IC+type domestic InterCity passenger coaches (in Hungarian) (ISBN 978-963-9852-20-4). MÁV-START Zrt, Budapest
7. Directive 2008/57/EC of the European Parliament and of the Council on the interoperability of the rail system within the Community http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:191:0001:0045:EN:PDF, Official Journal of the European Union (17 June 2008)
8. Regulation of Ministry of National Development: 31/2010 (XII.23.) NFM rendelet: Process of authorisation for placing in service, periodic inspection and the national vehicle registers of the railway vehicles (in Hungarian) http://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=A1000031.NFM, Hungarian Official Journal, No. 197, pp 30412–30500
9. Commission Decision 2010/713/EC on the modules for the procedures for assessment of conformity, suitability for use and EC verification to be used in the technical specifications for interoperability adopted under Directive 2008/57/EC http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32010D0713&from=EN, Official Journal of the European Union (9 November 2010)
10. Government Decision (2016) 1167/2016 (IV.6.) Korm. határozat: Development of the rail vehicle industry in Szolnok and financial support of the production of IC+type passenger coach (in Hungarian). Hungarian Off J 47:4295
11. Zobory I (ed) (2006) Handbook of railway engineering (in Hungarian) (ISBN 963 204 127 5). Magyar Államvasutak Zrt, Budapest
12. Radnai G (1976) Lightweight constructions in the vehicle and machinery industry (in Hungarian) (ISBN 963 17 0765 2). Tankönyvkiadó, Budapest
13. MSZ EN 12663-1:2010—Railway applications. Structural requirements of railway vehicle bodies. Part 1: Locomotives and passenger rolling stock
14. Borhy I, Schwartz I (2002) Practical application of FEM in the technology design of resistance spot welding in railway production. In: Proceedings 3rd conference on mechanical engineering (ISBN 963 699 1650), Budapest, (Hungary), pp 479–484

15. Borhy I, Szabó P (2005) Topical issues in the development of expert systems for use in the process planning of resistance spot welded railway vehicle structures. Math Modell Weld Phenomena 7 (ISBN 3-901351-99-X), Verlag der Technischen Universität Graz, pp 1099–1109
16. Borhy I, Belső L (2012) Design and optimization of spot welded railway vehicle structures (in Hungarian), Proc. 26th Hegesztési Konferencia (ISBN 978-615-5018-28-2), Óbudai Egyetem (Hungary), pp 91–96
17. Borhy I, Belső L (2014) Design and conformity assessment of welded railway vehicle structures, demonstrated via IC+project (in Hungarian). In: Proceedings 27th Hegesztési Konferencia (ISBN 978-963-08-8585-0), Óbudai Egyetem (Hungary), pp 49–58
18. MSZ EN 15085-3:2008—railway applications. Welding of railway vehicles and components. Part 3: Design requirements
19. Borhy I, Szabó P (2008) Possibilities of predicting the fatigue life of resistance spot welded joints. In: Proceedings international conference on design, fabrication and economy of welded structures (ISBN 978-1-904275-28-2), Horwood Publishing, Chichester (UK), pp 201–210

Challenges and Solutions in Resistance Welding of Aluminium Alloys—Dealing with Non Predictable Conditions

Jörg Eggers, Ralf Bothfeld and Thomas Jansen

Abstract For welding of aluminium alloys with Harms and Wende medium frequency inverters (1000 Hz) we developed a special control mode AMC (Aluminium Mode Classic) and its extension AMF to handle aluminium alloys (mainly 5000th and 6000th alloys). These modes have been made to handle the alloy groups. The new methods cover every day issues as undefined oxide layers or other coatings. From the commercial side a standard inverter is used to reduce investments. Internal monitoring while welding is another important topic of this mode as well as constant weld time. The last item is the key topic for high volume production customers. Examples are shown for several surface conditions. Methods are suggested to preserve against crackles and cracks. In advance first tendencies in quality monitoring are presented using the force signal during the current phase of the welding process. In this paper we give the solution first before showing the background.

Keywords Resistance welding · Adaptive control · MFDC inverter · Aluminium alloys

1 Introduction

When dealing with aluminium joints of the system has to handle conditions between the tips which are not known. Go one step ahead these conditions vary from weld spot to weld spot.

J. Eggers (✉) · R. Bothfeld · T. Jansen
Harms & Wende GmbH & Co KG, Hamburg, Germany
e-mail: joerg.eggers@harms-wende.de

R. Bothfeld
e-mail: ralf.bothfeld@harms-wende.de

T. Jansen
e-mail: thomas.jansen@harms-wende.de

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_38

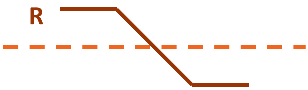


Fig. 1 Resistance drop, as described the first time in [1]

On the following pages solutions re provided which run out in the field and have shown that such work stable.

1.1 Pre-conditioning

In the first step the uncertain situation has to be removed. When oxide layers are on the surface of the material this causes a higher resistance as the base material, the aluminium itself. Such coatings can be thicker or thinner depending how long the material has been sitting on the shelf and how it was treated before reaching the factory.

Hence what has to be done is to break through this layer to touch the base material. Since the base parameter of the base material does not chance the first phase is called a pre-conditioning phase. Figure 1 shows the measurable resistance drop.

1.2 Force-Profile

A force profile can be used in the AMC case. A high electrode force is used during the pre-conditioning phase and a lower during the weld and a higher force again during the post welding phase. The lowered force results in a lower energy demand due to the hither contact resistance.

A type of advanced process control suggested by DVS Merkblatt [2] (Fig. 2) and before by [3].

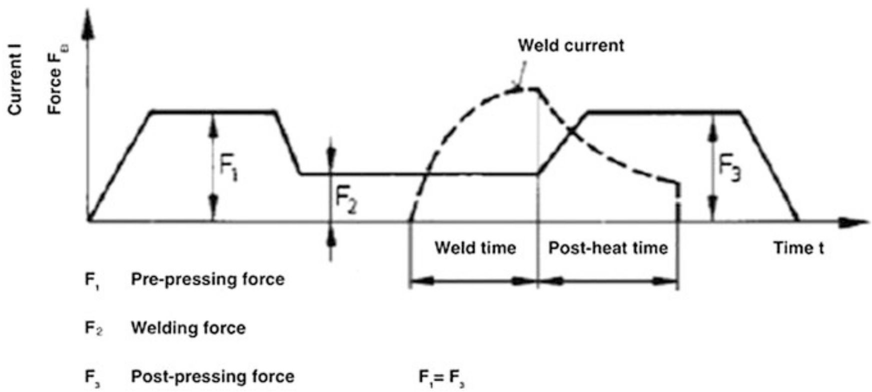


Fig. 2 Suggested force and current profile from [2]

2 Aluminium Mode Classic AMC

The aluminium mode classic combines the pre-conditioning phase and weld phase and if needed the force profile.

Inside the inverter the weld is divided into the two sections as shown in Fig. 3. Since the AMC is intended to use, for example, in automotive business the cycle time becomes very important. One parameter to stop the weld process is, when the pre-conditioning phase is not finished to a certain time segment. Then it is assumed the conditions cannot be stabilized. The following weld phase is a constant weld time segment. The sum of the maximum allowed conditioning phase plus weld time is the maximum cycle time for the weld itself. Other times, as pre-hold time and so forth have to be added.

2.1 *User Interface XPegasus*

The whole process control is assembled for the user in one XPegasus page. The challenge in the design was to make the user feel familiar with a standard weld schedule to keep the learning phase short.

This user interface is independent from the tools (Servo gun, pneumatic gun, or 7th Axis robot system). The user interface as in Fig. 4 is designed in the way that is reminds to a normal Constant Current Regulated (CCR) weld. A graphical window shows the entered process.

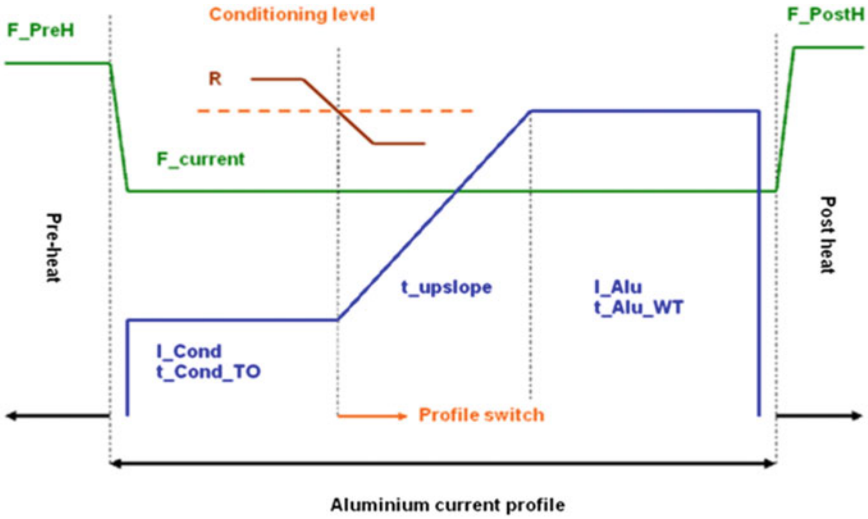


Fig. 3 Schematic diagram of AMC (Alu Mode Classic)

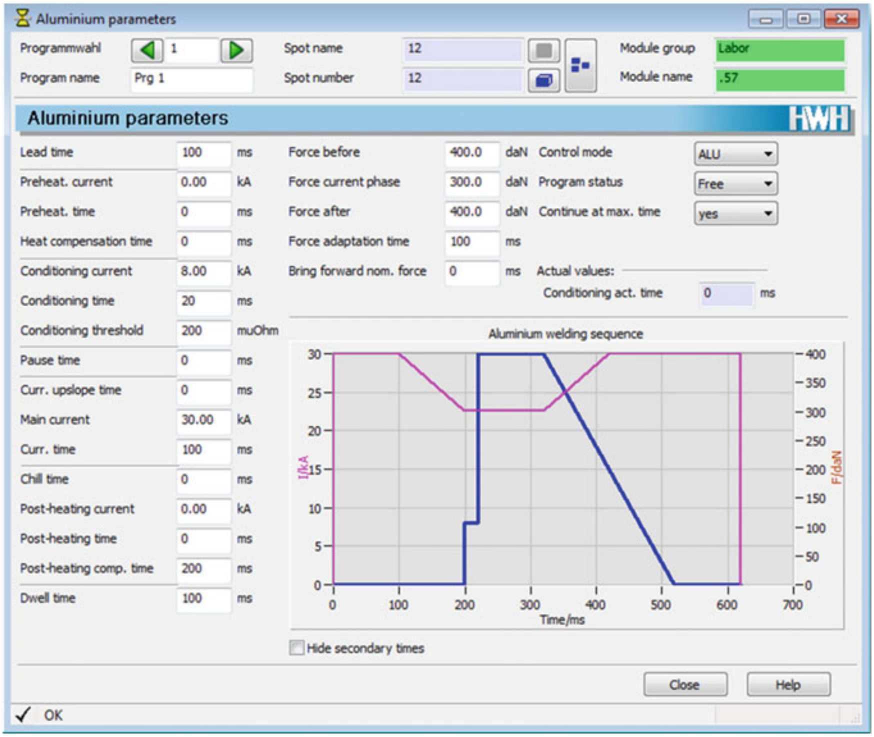


Fig. 4 XPegasus page for aluminium process control

As difference from the standard constant current schedule the break-through resistance can be entered and the force profile can be programmed.

2.2 Weld System Set-Up

Figure 5 shows the gun the trials in Sect. 2.2 have been made with. It's a servo driven C-gun. This gun can keep up a pressure of 4000 N (899.23 lbf). There is a second gun in the lab with a larger transformer and a maximum force of 7000 N (1573.66 lbf). The inverter used in this system is a GeniusHWI424 (1600 A max output current) with AMC installed. For power requirements, please refer to Sect. 5 later in this paper.

Almost flat D16 electrodes in standard alloy CoCrMo have been used. It's important to notice that no special electrodes have been used. These electrodes show a large radius on both gun arms. However, one flat electrode can be used single sided for better surface quality.

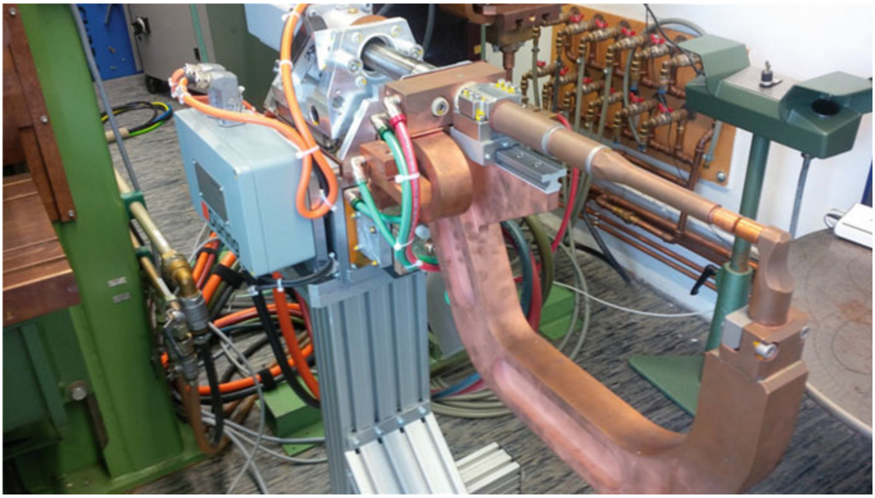


Fig. 5 Servo-electric C-gun

2.3 *Development of Process Factors*

Using such a device you end up with following process diagrams:

Figure 6 shows a typical aluminium weld process using the AMC. The cure moment of the process is the time frame between 900 and 1300 ms. The pink curve shows the force profile. The green curve shows the resistance and the blue the current.

When the current is applied at 900 ms the resistance jumps to a high level due to the oxide layers and other resistance between the tips. After a short moment the resistance curve drops. This is the moment when the tips break through the oxide and touch the real aluminium.

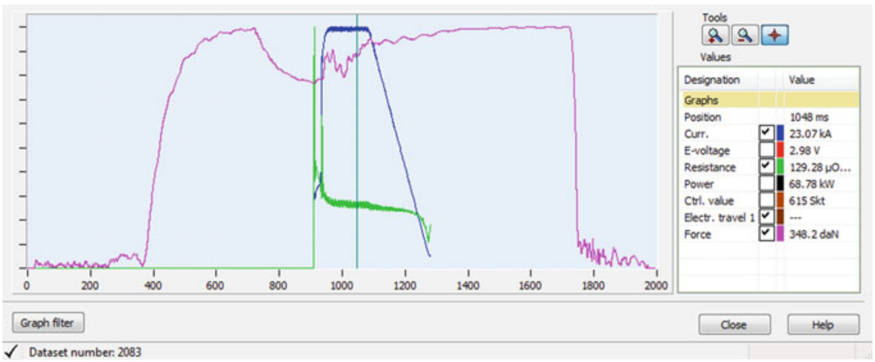


Fig. 6 Process diagram resulted for AMC usage (XPegasus)

Direct after the current starts to raise the force value is moving slightly up and then down again just left to the vertical line. This is the nugget build-up or in other words the material expansion.

In this moment the inverter raises the current to the welding level of 23 kA.

For this weld we used a gun, which is usually intended for steel welds. The reason is to show the weld process curves very clearly. Usually the gun would be much stiffer and the drop in the force will be much less.

The force measurement shows a significant force drop when the material starts to soften and a quite steep peak value when the nugget starts to grow. This peak is related to the expansion material factor. Aluminium expands twice as fast as steel (factor 24–12 on steel).

3 Pre-conditioning

The phase of pre-conditioning is combined by the force increase, followed by a fading phase to the welding force. Then the welding force is reached the control starts with a pre-conditioning current. In Fig. 7 is the case shown, that the resistance threshold is reached relative lately.

3.1 Example with High Resistance Sn Coating

One example is shown here with high resistance Sn coating. Normally with the standard welding method for aluminium (one pulse with no pre-conditioning) you will get an expulsion because of the high start resistance.

With pre-conditioning this effect is reduced.

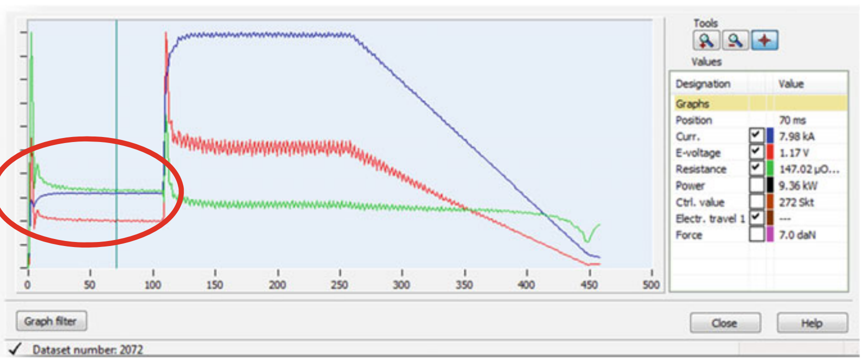


Fig. 7 Example for a long conditioning time of an old charge (*red circle*)

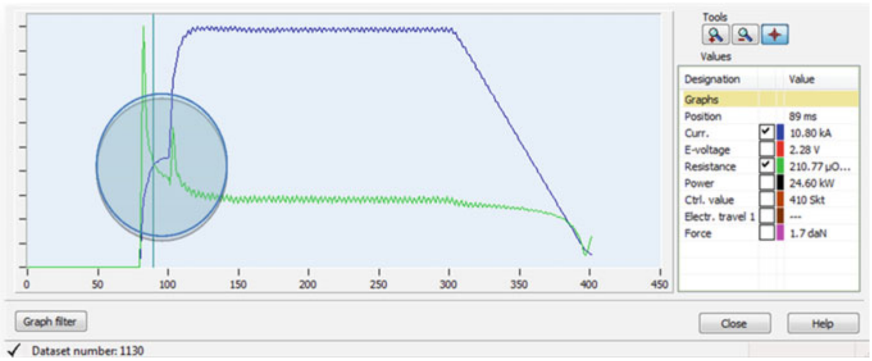


Fig. 8 Example with high resistance Sn coating

When material is tin coated the initial resistance is higher (see circle in Fig. 8) as on uncoated material. This increases the splatter risk by 20% especially on short term welds.

4 Typical Welding Quality Until Tip Dressing

There is no hope to reach higher values at 20–50 welds. The main effect is electrode pick up of aluminium. There every millisecond with high current power reduces the lifetime.

The hardest effect is on the anode electrode (Fig. 9).

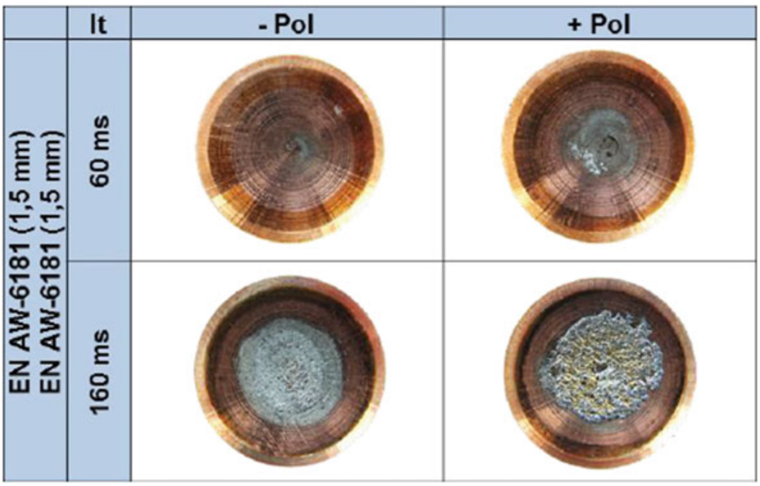


Fig. 9 Electrode erosion after 50 welds with different weld times of the high current main pulse [4]

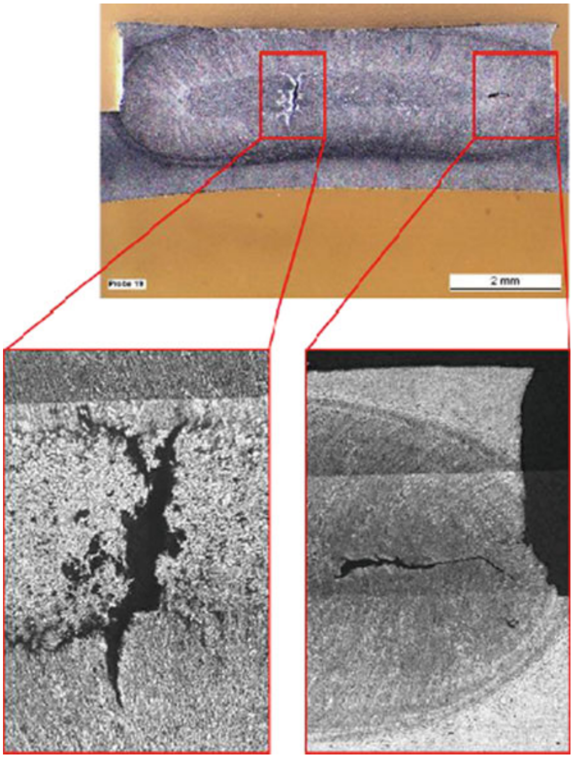


Fig. 10 Imperfections increased excessive the limits of tip dressing cycles [4]

If you try to increase the number of welds between tip dressing you will get an increase of surface cracks, strongly erode of surfaces and at the end the electrode sticks on the aluminium plates (Fig. 10).

4.1 Methods to Decrease the Number of Imperfections

The main method is the usage of a longer downslope in current after the main pulse combined with an increase in the electrode force. This will control the cooling down of the spot. The volumetric change with time of the spot is reduced.

A typical control profile looks like Fig. 11.

The result is a nice looking spot like in Fig. 12.

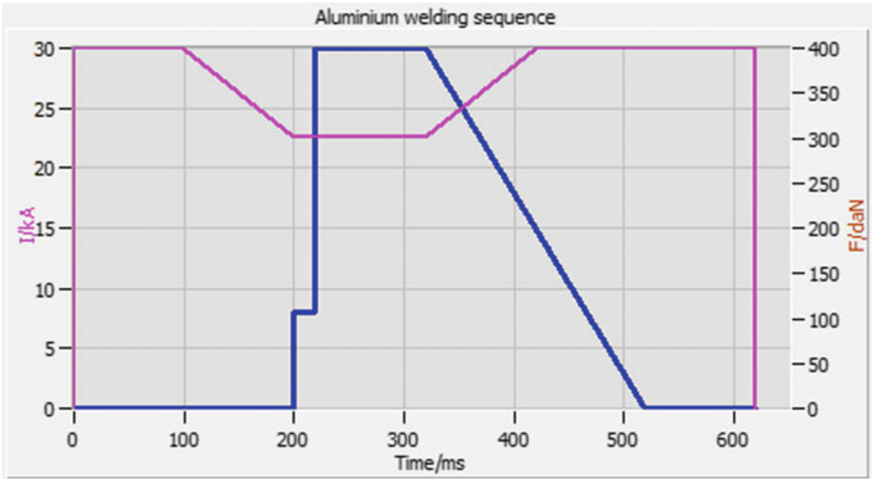


Fig. 11 Control profile to reduce the number of imperfections

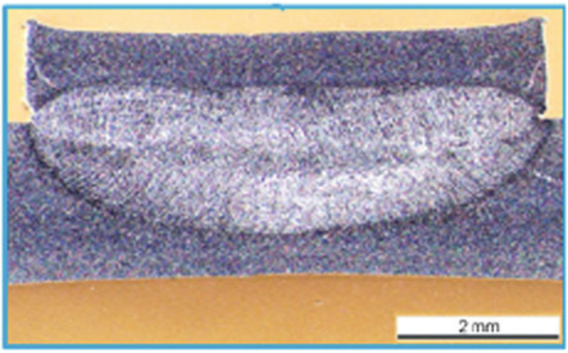


Fig. 12 A spot without imperfections maintained by control profile (etchant side view)

5 Price of High Currents and Welding Times

An important factor of a welding system is the power requirements. The most common inverter in spot welding is the 80 kVA (max. 600 A output current) which is widely used for steel welds of any kind.

In case of aluminium welds the inverter power class differs. Typical inverter power classes used for such welds are 1600 A (i.e. GeniusHWI 2824) or 2400 A (i. e. GeniusHWI2836) inverters. This has an impact not only on the fuse and on the supply lines as well as on the power consumption of a plant. Table 1 below gives an overview about technical key data of different inverter classes.

Table 1 Different inverter classes

Type	HWI	2813 L	2816 L	2816 L	2816 W	2824 W	2836 W
Input							
Nom. voltage	V	400–480, 690[a]	400–480, 690	400–480, 690	400–480, 690	400–480, 690	400–480, 690
Nom. current 100% DC	A	66	153	74	229	305	430
Phases		3, PE, (N)	3, PE, (N)	3, PE, (N)	3, PE, (N)	3, PE, (N)	3, PE, (N)
Frequency	Hz	50/60	50/60	50/60	50/60	50/60	50/60
Output							
Voltage @ 400 V	V	500	500	500	500	500	500
Voltage @ 480 V	V	600	600	600	600	600	600
Nom. power @ 20% DC	kVA	130	300	145	450	600	845
Max. current	A	800	800	1200	1200	1600	1600
Cooling qty.	l/min	Air	4	Air	4	4	4
Weight	kg (lb)	21 (46)	21 (46)	21 (46)	21 (46)	24 (53)	24 (53)

[a]690 V mains supply for certain volkswagen plants only

Hence, when a production is designed with a larger number of aluminium weld cells, the infrastructure must be considered as well. An intermediate way is to limit the load of the power system by a power divider or similar. This solution will have an impact on the cycle time since the cells have to wait when the maximum load allowed is reached.

Another factor on aluminium weld systems is the cooling. While cells welding steel materials are typically equipped with air cooled inverters require aluminium weld cells water cooled inverters. Hence, apart from the power consideration and water network must be considered as well.

A third subject is the nominal voltage a system is operating in. In Europe, it's generally 400 V and in the US its 480 V. The difference in the voltage implies that 480 V systems have 20% mode energy as 400 V systems. This fact becomes very important when designing a system.

Since power is calculated as $P = U \cdot I$ the current flow in 480 V systems is lower as in 400 V systems. This becomes even more important when systems are for example designed in Europe for operation in the USA or the other direction. Systems running well under 480 V may fail below 400 V since 20% of the energy is lost and the system reaches its maximum.

6 Extension with Quality Monitoring (AMF)

With an ideal electric motor servo gun it is very easy to construct a quality monitoring system. You can simply take a tolerance band around a force reference curve and from there you get a monitoring signal during current time.

This can be easily managed by several third party systems. In this chapter we want to show the case for an older 'used-up' gun with a non-ideal force signal. It will also show tendencies in future monitoring systems to build up robust algorithms.

The force examples shown here are for non-ideal welding equipment. Especially equipment with lower capabilities beyond 4 kN. This welding gun here has a larger amount of lateral movement during force build up.

Here 1 mm × 1 mm 6000 aluminium alloy plates are welded. We used 16 mm electrodes of type A.

With a current of 30 kA we get a spot diameter of around 7 mm.

In Fig. 13 we see an increase in force of around 25–30 % during main current. But the force signal is distorted by the lateral movement of the electrodes by larger oscillations on the signal. The movement is not smooth and the mechanical system is oscillating.

If we produce an expulsion with a higher current we see it in the force signal clearly (Fig. 13). In the resistance signal no occurrence is observed.

One future requirement for expulsion detection (red circle in Fig. 14) is the distinction between expulsion and lateral movement. The other requirement is the measurement of the total amplitude by a robust filter mechanism.

Figure 15 shows a weld with a 2 mm spot diameter. So simple by signal amplitude a differentiation might be possible. Here we reach only 10–15% increase of the force during main weld current.

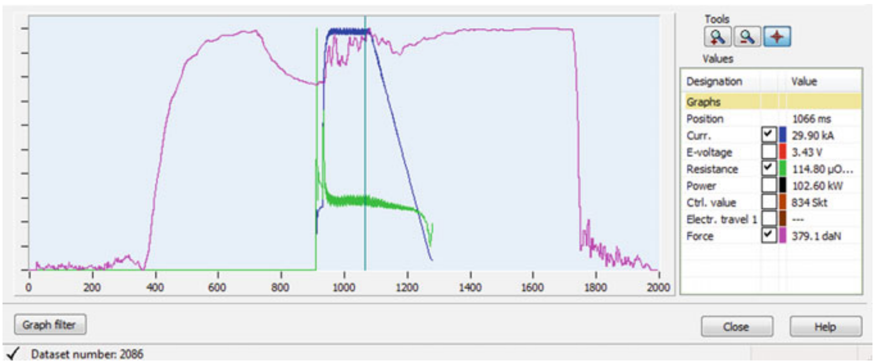


Fig. 13 Welding is OK. Reached 7 mm spot diameter

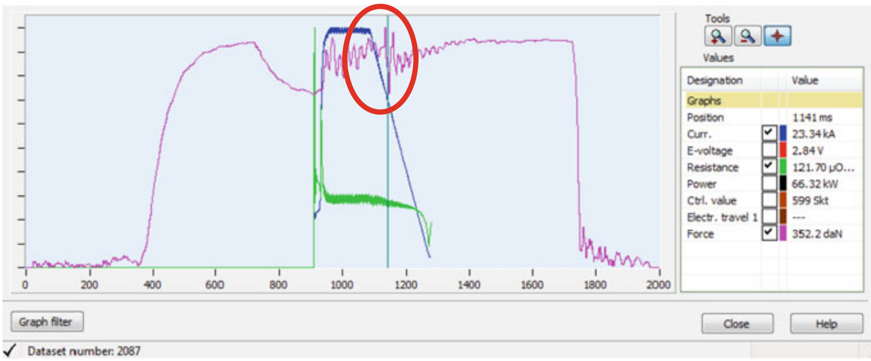


Fig. 14 Welding with expulsion in the late phase of the weld

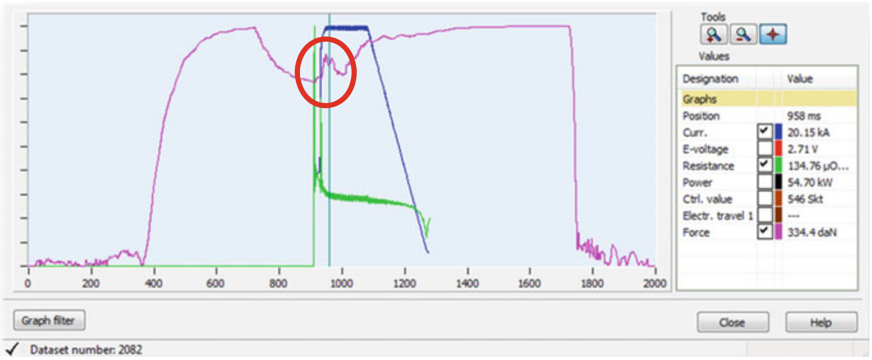


Fig. 15 Welding for a not ok (*red circle*) welding with 2 mm spot diameter

With a clear force signal it might be no problem to get a monitoring signal. But most times in practice it will not be the case because of the costs of the gun or measuring equipment. So we have to find robust filter algorithms in future.

7 Summary

The Aluminium Mode Classic (AMC) and AMF are methods to join aluminium materials. Typical materials used in automotive business are 5000 and 6000 series. It bridges the problem with the undefined oxide layers on any aluminium. Such layers caused problems with varying spot diameters. Stabilization of the weld process is very important.

Aluminium has for natural reasons oxide layers on the surface the AMC and AMF divide the weld process into sections:

- Conditioning phase
- Weld phase

During the conditioning phase with a lower current i.e. 10 kA the inverter monitors the resistance. When the resistance drops, it is assumed the aluminium is touched by the electrodes and the weld begins with me. i.e. 35 kA.

The AMF extends the joining process with a monitoring function for quality assurance.

There are some parameters to influence the weld, but important for the user is the maximum weld time is fixed. Typical weld times for aluminium are in the range of 100–120 ms plus conditioning phase.

However a method is only as good as the entire system consisting of gun, inverter and power supply. The examples in this paper showed the effect with too weak gun arms. However, in this case such gun has been used to show the effects.

New power demands come up with welding aluminium materials. More powerful inverters are needed.

References

1. Krevet A (2007) Untersuchungen zu Prozesseinflüssen beim Punktschweißkleben von Aluminiumwerkstoffen unter Verwendung einer widerstandsgesteuerten Prozessregelung. Die Verbindungsspezialisten, Tagungsband: Große Schweißtechnische Tagung, pp 638–639
2. Deutscher Verband für Schweißen und verwandte Verfahren e.V (1985) Widerstandspunkt- und –Rollnahtschweißen von Aluminium und Aluminiumlegierungen von 0.35 bis 3,5 mm Einzeldicke, Vorbereitung und Durchführung des Schweißers
3. Singh, S (1977) Beitrag zur Verbesserung und Sicherung des Tragverhaltens von Widerstandspunktschweißverbindungen an Aluminium- und Stahlwerkstoffen durch technologische Maßnahmen und durch Entwicklung einer Regelungseinrichtung, Dissertation RWTH Aachen (Germany)
4. NN Verbesserung der Prozesssicherheit des Punktschweißklebens von Aluminiumwerkstoffen und Ermittlung von Verbindungskennwerten für Konstruktion und Simulation, Aschlussbericht AiF 16.335N, GSI SLV, Duisburg (Germany), 2010

High Cycle Fatigue Investigations on High Strength Steels and Their GMA Welded Joints

Ádám Dobosy, János Lukács and Marcell Gáspár

Abstract High cycle fatigue tests were performed on two strength categories of high strength steels, on quenched and tempered (S690 and S960) and thermomechanical (S960) types, on base materials and their gas metal arc welded joints, and on different matching conditions. The planning and optimization of welding technologies based on investigations under cyclic loading conditions were built upon a large number of investigations and statistical evaluation of the test results. Statistical approach was already applied during the preparation of the investigations, which have been allowed the expansion of the scope of the results and the increasing of their reliability. The article demonstrates and evaluates the results comparing with each other and with literary data.

1 Introduction

In the automotive industry, as a dominant industry, one of the fundamental trends is the weight reduction and the environmental load reduction, which can be achieved with the application of high strength materials (steels [1, 2], aluminium alloys [3, 4] and composites [5, 6]). In the case of the various vehicles, and vehicle structures the dominant joining technology is the welding, with both the traditional and the modern processes, and with the fusion and the pressure welding processes. Failure statistics have long shown, that the third of the failures caused by fractures, where cracks initiated in welded structures, furthermore the four fifth of the failures occurred under cyclic loading condition [7].

Á. Dobosy (✉) · J. Lukács · M. Gáspár
Institute of Materials Science and Technology, University of Miskolc, Miskolc, Hungary
e-mail: metda@uni-miskolc.hu

J. Lukács
e-mail: janos.lukacs@uni-miskolc.hu

M. Gáspár
e-mail: gasparm@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_39

During the welding processes, the joining parts were affected by heat and force, which cause inhomogeneous welded joints. Because of these, zones with different microstructure and different properties can be developed, and additionally stress concentrator places can be formed. These changes can be caused faults, which define the properties of the welded joint. In the case of cyclic loading condition the failures have a more important role; however in the case of the vehicles the cyclic loading is very conventional.

All these together justify, that our investigations were focused on the high strength steels, from the joining methods on the welding technologies, and from the loading conditions on the cyclic loadings. Three types of high strength steels (S690QL, S960QL and S960TM), one fusion welding technology (shielded metal arc welding, ISO 4063: 135), and high cycle fatigue (HCF) loading condition were chosen for our investigations. In accordance with the welding challenges in our days, the matching problem of the base material and the welding consumable was examined as well. One of the main aims of our investigations is the determination of high cycle fatigue design or limit curves for the tested steels and their welded joints.

The wide range of our research work can be found in a book [8]; the details and the results of the low cycle fatigue (LCF) and the fatigue crack growth (FCG) investigations were summarized in [9, 10], and in [11], respectively.

2 High Cycle Fatigue Limit Curves

The wide variety of machines, equipment and structures are usually designed for long operation, often running for several decades under cyclic loading condition. In the case of the automobile structures, the number of cycles during the whole lifetime could be millions and tens of millions. We talk about high cycle fatigue when the loading is relatively low, and the number of cycles is relatively high, between 10^4 ($5 \cdot 10^4$) cycles and 10^8 (10^9) cycles. The cyclic loading of the components, structural elements and structures can be very different, from the simple mechanical stresses (tensile, tensile-compression, bending, torsion) to the complex stresses (for example tensile and bending), furthermore, cyclic thermal and/or environmental stresses can also occur.

The measurement results could be affected significantly by a number of factors. This clearly means that under the same testing conditions, more test specimens must be used, because of the relative high standard deviation of the test results. Furthermore, mathematical statistical methods must be applied to the evaluation of the results [12–14]. In the case, when test specimens from welded joints or whole welded joints were examined, the effects of the welding process increase the uncertainty further. The statistical approach is desirable in the preparation phase of the experiments, so the validity range of the results can be widened and the reliability can be increased as well.

In the area of the high cycle fatigue, one of the main research directions is the determination of the design or limit curves. The high cycle fatigue limit curves can be found in numerous prescriptions, and there are numerous known and applied methods for the derivation of these curves, based on data from different fatigue and non-fatigue examinations. A few, widely used regulations and design curves are as follows: Eurocode 3 [15]; BS 7608 [16]; BS 7910 [17]; AASHO direction [18]; design curves based on empirical correlations [19].

The design or limit curves, found in the different specifications, were determined based on experiments with different welded structures. The Basquin equation, which describes the life-time region of these curves, is the following:

$$N_t(\Delta\sigma)^m = a, \tag{1}$$

where m and a are constants, depending on the materials and the conditions.

3 Welding Experiments on High Strength Steels

3.1 Material Properties and Weldability

For our studies, high strength steels from 690 and 960 categories were chosen, welded with shielded metal arc welding (GMAW, ISO 4063: 135) process. In the 690 category quenched and tempered (Q + T) steels (S690QL and Weldox 700E), and in the 960 category quenched and tempered (S960QL) and thermo-mechanically (TM) rolled (S960TM) steels were chosen (EN ISO 10025-6). For the welding experiments different filler materials were selected. The chemical composition of the base and the filler materials and the mechanical properties can be seen in Tables 1 and 2, respectively.

These steels have a higher energy after the production, than in the equilibrium state, so the obtained microstructure can be irreversible altered during the welding process. The heat affected zone (HAZ) can be easily hardened, furthermore, in case of too large heat input the heat affected zone can be softened—because of the coarsen of the microstructure—compared to the base material, which can cause strength and hardness decreasing. It is a matter of course that both cases are intolerable. Because of these effects, additional microalloys are used on these steels, like aluminium, niobium, vanadium and titanium. Naturally, these features change also affects the fatigue properties; the different quality of the different zones in the welded joint (especially in the HAZ) significantly changes the fatigue resistance.

Besides all these, further undesirable phenomena can appear, e.g. different types of cracks. Primarily, to avoid the cold cracking the workpiece must be preheated before the welding, and it is necessary to limit the linear energy (E_v) during the welding [20].

The complex task of the weldability of the high strength steels summarizes the Graville diagram (Fig. 1) based on the carbon content and the carbon

Table 1 The chemical composition of the examined base and filler materials, wt%

Material designation	C	Si	Mn	Cr	Mo	Ni	S	P	Ti	V	Al
S690QL	0.14	0.30	0.96	0.60	0.19	–	0.002	0.009	0.02	0.005	0.05
Weldox 700E	0.14	0.30	1.13	0.30	0.167	–	0.001	0.007	0.009	0.01	0.03
S960QL[a]	0.16	0.23	1.25	0.2	–	0.04	0.001	0.008	0.004	0.04	0.06
S960TM	0.09	0.32	1.63	0.59	0.29	0.03	0.001	0.009	0.016	–	0.041
INEFIL NiMoCr[b]	0.08	0.50	1.60	0.30	0.25	1.50	0.007	0.007	–	0.09	–
Thyssen UNION 85[c]	0.07	0.68	0.61	0.29	0.61	1.73	0.010	0.006	0.08	0.01	0.01
Thyssen UNION X90	0.1	0.8	1.8	0.35	0.6	2.3	–	–	–	–	–
ESAB OK Tubrod 14.03[d]	0.08	0.51	1.61	0.02	0.55	2.27	–	–	–	0.01	–
Thyssen UNION X96	0.12	0.80	1.90	0.45	0.55	2.35	–	–	–	–	–

[a]Cu = 0.01, Nb = 0.016, B = 0.001, N = 0.003
[b]Cu = 0.12
[c]Cu = 0.06
[d]Cu = 0.02

Table 2 The mechanical properties of the examined base and filler materials

Material designation	$R_{p0.2}$, MPa	R_m, MPa	A_5, %	KV (−40 °C), J
S690QL	783	826	19	54
Weldox 700E	791	836	17	165
S960QL	1030	1076	16	56
S960TM	1051	1058	17	177
INEFIL NiMoCr	≥ 750	≥ 820	≥ 19	≥ 60
Thyssen UNION X85	≥ 790	≥ 880	≥ 16	≥ 53
Thyssen UNION X90	≥ 890	≥ 950	≥ 15	≥ 58
ESAB OK Tubrod 14.03	757	842	23	71
Thyssen UNION X96	≥ 930	≥ 980	≥ 14	≥ 40

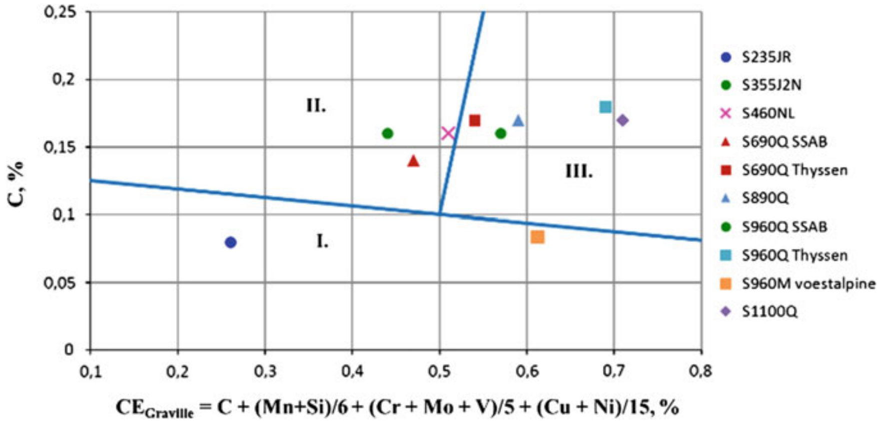


Fig. 1 Graville diagram [20]

equivalent (CE). It can be clearly seen, that the examined Q + T base materials are in the hardest weldable category (III.), which means that these steels require preheating and we must control the linear energy precisely; while the TM base materials locates in the easily weldable category (I.), so they do not require these sort of measurements.

In the case of these steels, one of the most important features of the successfulness of the welding is the heat input, which can be described with the linear energy. If the value of the linear energy is too low, the cooling rate of the welded joint may be too fast, and then cold cracks can occur. In the opposite case, strong coarse grain microstructure can be formed in the heat affected zone, which can cause the decreasing of the strength and the toughness. Therefore, we received a narrow welding lobe; inside this the quality of the joint may be suitable. Besides the previously mentioned characteristics, the material quality, the chemical composition and the applied thickness are important, too.

The $t_{8.5/5}$ cooling time was used for the common description of the welding conditions and parameters. Because of the previously mentioned reasons, the cooling time is a narrow range in the case of the Q + T steels, but it can be changed in a wide range in the case of the TM steels.

3.2 Welding Experiments

The welding parameters and the welding conditions were determined based on the previously mentioned statements. For shielding gas, M21 designation with 18% CO_2 + 82% Ar composition was used, based on industrial experience. In the case of the filler material 1.2 mm diameter wire was applied in all cases. The cases of the base material-filler material pairing can be seen in Table 3. Further details of the matching problems can be read in the literature [21–23]. In case of matching condition, the evolved mechanical properties of the joints are equal or nearly the same as the base material. In the undermatching case the mechanical properties of the joints are lower, while in case of overmatching condition are higher than the base material features.

The welding equipment was a DAIHEN VARSTROJ WELBEE P500L power source. For the equal stress distribution X joint shape was used, with 80° opening angle and with 1.5 mm gap between the two plates. During the welding, the test pieces were rotated after each layer. The dimensions of the welded plates were 300 mm × 125 mm. The root layers were made by a qualified welder; while the other layers were made by automated welding car. The experimental assembly can be seen in Fig. 2.

The applied welding parameters were summarized in Table 4. The table shows the welding current (I), the voltage (U) and the welding speed (v_h) values, also the preheating (T_e) and the interpass (T_r) temperatures, with the linear energy (E_v) and the calculated critical cooling time ($t_{8.5/5}$) values. The parameters of the root and the filler layers were shown separately in each case.

In the case of S960TM material, wide ranges can be found due to the fact, that we made experiments with small and large linear energy. It is important to note, that both the interpass temperature and the linear energy are determining parameters in

Table 3 The base material-filler material pairing during our experiments

Base material	Filler material	Matching condition
S690QL	INEFIL NiMoCr	Matching (m)
Weldox 700E	Thyssen UNION X85	Matching (m)
Weldox 700E	Thyssen UNION X90	Overmatching (om)
S960QL	Thyssen UNION X96	Matching (m)
S960QL	ESAB OK Tubrod 14.03	Undermatching (um)
S960TM	Thyssen UNION X96	Matching (m)

Fig. 2 Assembly of the welding experiments

Table 4 The applied welding parameters

Base material	Layer	T_e/T_r, °C	I, A	U, V	v_h, cm/min	E_v, J/mm	$t_{8.5/5}$, s
S690QL	1–2	150	130–140	19.0–20.5	20	700–750	7–8
	3–8	180	280–300	27.5–28.5	40	1000–1100	9–11
S960QL	1–2	180	120–130	18.5–19.0	20	570–590	5–6
	3–8	150	280–305	28.5–31.0	45	900–1000	7.5–8.5
S960TM	1–2	60	130–140	19.0–20.5	20	700–750	7–8
	3–8	150	260–310	25.0–32.0	30–60	900–1400	7–15

the case of the critical cooling time, however, between those two values there are no close correlation.

4 High Cycle Fatigue Experiments and Their Results

4.1 Test Specimens

For the high cycle fatigue tests, flat specimens were used. The geometry of the specimens and the cut out orientation from the base materials and the welded joints

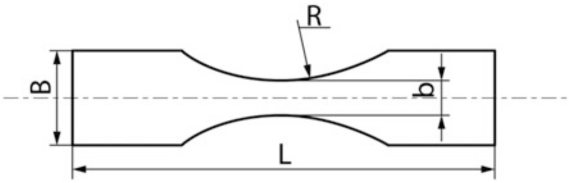


Fig. 3 Specimen geometry for the high cycle fatigue experiments

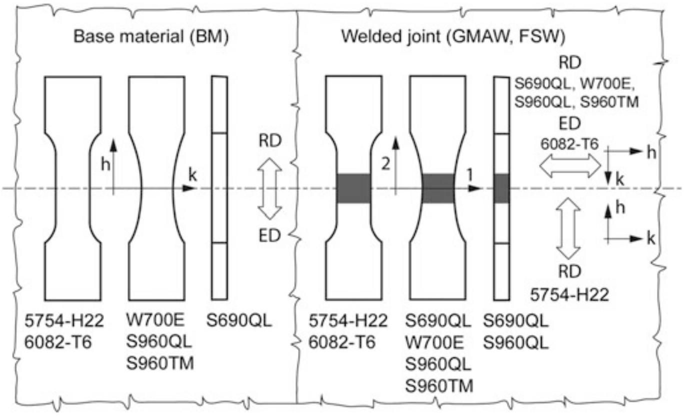


Fig. 4 The orientation of the test specimens and the welded joints for the experiments

can be seen in Figs. 3 and 4, respectively. (Figure 4 is applicable not only for this study, but also for our whole research work in this field.) In Fig. 3 the RD means the rolling direction of the sheets, h is the longitudinal and k is the transversal direction, furthermore, 1 and 2 mean the longitudinal and the transversal direction of the welded joints, respectively. The thickness of the sheets (v) and the welded joints (3) are perpendicular to the surface. Table 5 shows the nominal geometry of the test specimens, where the orientations can be identified with Fig. 4.

4.2 Testing Circumstances

High cycle fatigue experiments were performed according to [24], using MTS electro-hydraulic materials testing equipment (MTS 312 and MTS 810 (Fig. 5)), at room temperature and on laboratory environment. Constant load amplitude was applied during the experiments, with R = 0.1 stress ratio, f = 30 Hz loading frequency, and sinusoidal loading wave form. Test specimens cut from the welded joints were tested in as-welded condition.

Table 5 Main characteristics of the high cycle fatigue test specimens

Base material	Specimen axle/crack grow direction	t, mm	b, mm	B, mm	L, mm	R, mm
S690QL	BM h/v	6	8	24	175	92.5
	GMAWm k/3 W	6	8	18	135	60
	GMAWm k/1 W	6	7.5	18	135	60
Weldox 700E	BM h/k	6	7; 8	18	135	60
	GMAWm k/1 W	6	7; 8	18	135	60
	GMAWom k/1 W	6	8	18	150	60
S960QL	BM h/k	6	8; 5.5	18	150	60
	GMAWm k/3 W	6	7.75	18	150	60
	GMAWum k/1 W	6	7.5	18	150	60
S960TM	BM h/k	6	8	18	150	60
	GMAWm k/1 W	6	8	18	150	60

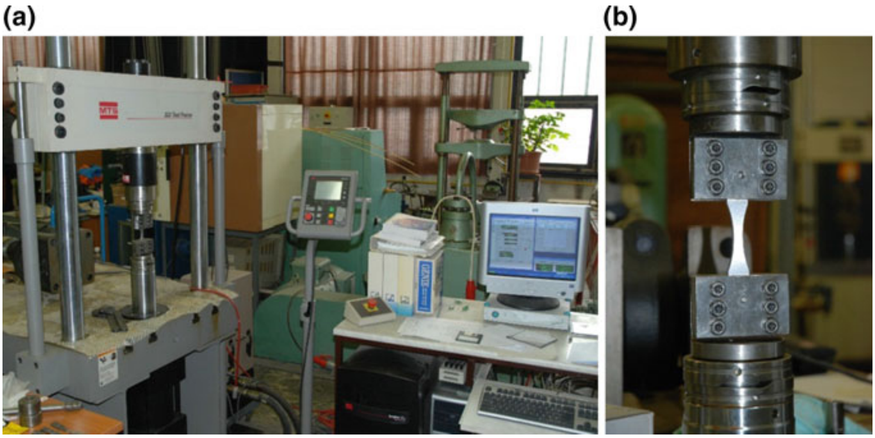


Fig. 5 High cycle fatigue tests: **a** MTS 810 universal testing equipment; **b** the test specimen and the grip

4.3 *The Results of the High Cycle Fatigue Experiments*

The results of the high cycle fatigue tests were shown in diagrams (Wöhler or S-N curves).

The results and the S-N curves of the 690 strength category base materials (amended with literature data [25–27]) can be seen in Fig. 6.

The results and the S-N curves of the 690 strength category gas metal arc welded joints (with literature data [26, 27]) were summarized in Fig. 7.

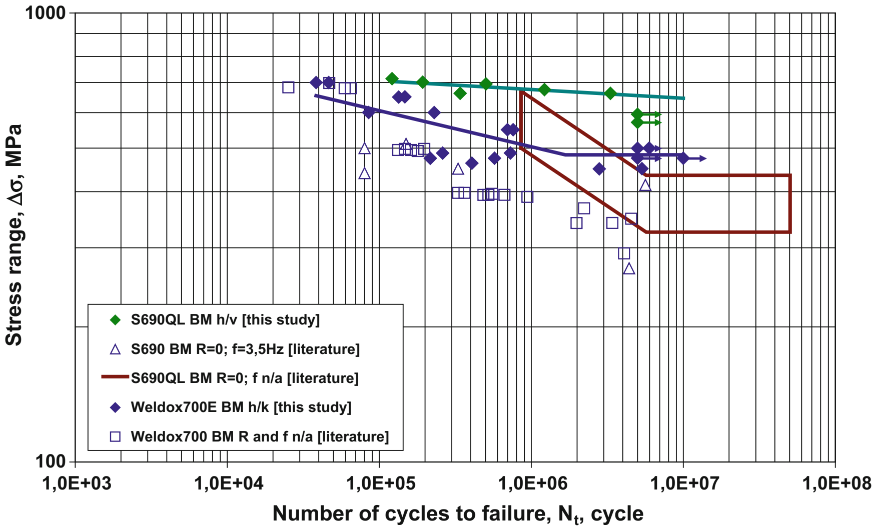


Fig. 6 The results of the 690 strength category base materials

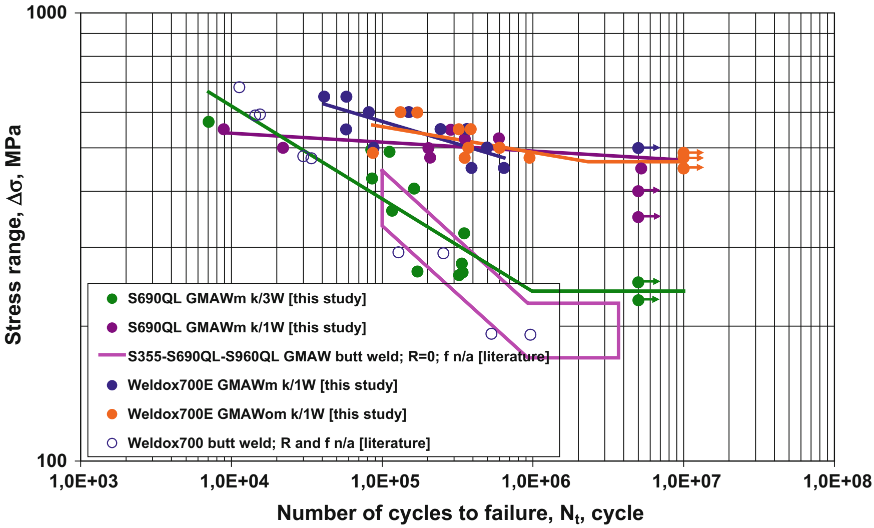


Fig. 7 The results of the 690 strength category gas metal arc welded joints

The results and the S-N curves of the 960 strength category base materials (integrated with literature data [27]) can be seen in Fig. 8.

Finally, the results and the S-N curves of the 960 strength category gas metal arc welded joints (integrated with literature data [27]) were demonstrated in Fig. 9. In

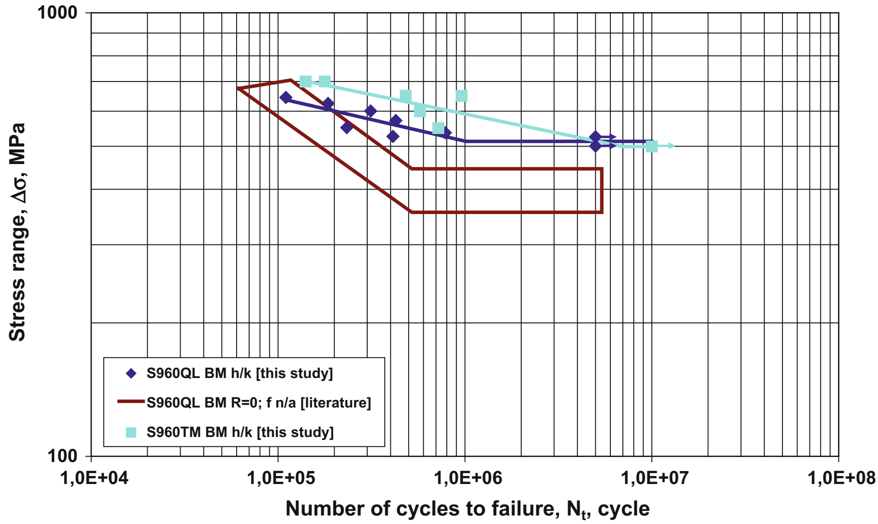


Fig. 8 The results of the 960 strength category base materials

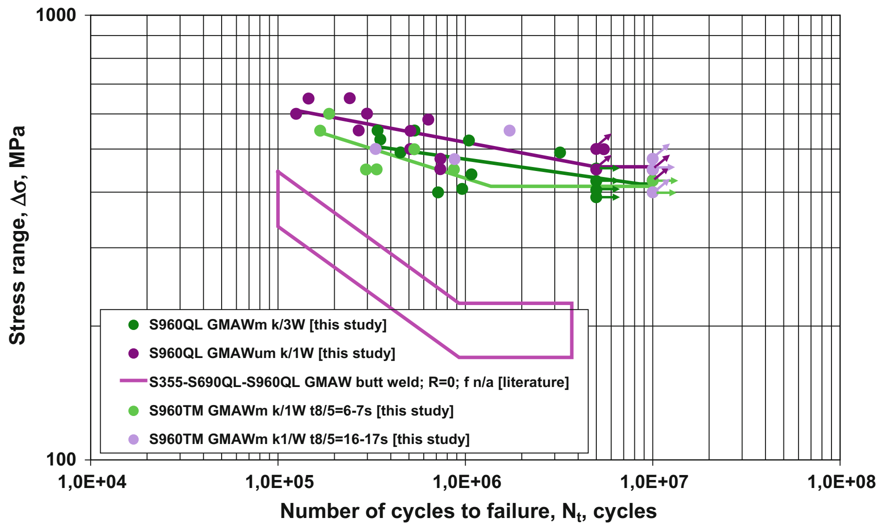


Fig. 9 The results of the 960 strength category gas metal arc welded joints

the case of the S960TM steels, two kinds of experiments were made; in the first case, the critical cooling time was $t_{8.5/5} = 6–7$ s, while in the second case it was $t_{8.5/5} = 16–17$ s.

In the cases, where the test specimens did not fracture to a given cycle value (usually $5 * 10^6$ or 10^7 cycles), the results (marks in the diagrams) were denoted with horizontal or a slope arrow.

5 Evaluation of the High Cycle Fatigue Experiments, the Fatigue Design or Limit Curves

The performed high cycle fatigue experiments were compared with literature data; accordingly we can evaluate the results more accurate. It is important to note, that during some literature examinations R = 0 stress ratio was used, while for our experiments R = 0.1 stress ratio was applied. This fact does not influence on our conclusions significantly. Furthermore, the used frequencies during the literature examinations are presumably smaller, than through our experiments; but the data do not indicate significant differences (for example [13]), so this fact also does not influence on our conclusions significantly. Finally, in the literature [27], the given S355, S690QL and S960QL data are too general, so these values can use only for general comparison.

The S690QL base material resistance against high cycle fatigue in thickness direction is more favorable, than the Weldox 700E base material resistance in transverse direction. The examined S690QL and Weldox 700E base materials were shown better results, than the same literature data.

The matching (m) joints from the S690QL steel in the 3 W orientation have lower resistance against the high cycle fatigue, than in the 1 W orientation; furthermore, this steel has lower resistance than the Weldox 700E steel, in the 1 W orientation. Both steel category have higher resistance against the high cycle fatigue in the 1 W orientation, than the overall results of the butt welded joints, in the case of the three steel category (S355, S690QL and S960QL [27]), in the 1 W orientation. This is fully consentaneous with the approach, because in the former case test specimens were used, and in the latter case the whole welded joints were examined. The high cycle fatigue resistance, in the case of test specimens from the S690QL steel matching (m) joints in 3 W orientation, is located in the range of the three steel category butt welded joints in the 1 W orientation. The resistance of the overmatched (om) joints against the high cycle fatigue in the case of the Weldox 700E steel is slightly higher, than the matched (m) joints resistance; but this statement requires further investigations.

The high cycle fatigue resistance of the examined S960QL steel is better, than the literature data in the same category. The high cycle fatigue resistance of these steels is higher in the case of undermatched (um) joints, than in the case of matching (m) joints. The high cycle fatigue resistance measured on specimens from matched (m) and undermatched (um) joints are considerable higher (and independent from the orientation), than the butt welded joints in the 1 W orientation, in the case of the three steel category. This is consentaneous with the approach because of the previously explained reason, but the differences are notable.

The high cycle fatigue resistance of the S960TM base material is more favourable, than the examined S960QL base material, furthermore, than the overall literature data. The results of matched (m) joints welded with the longer critical cooling time ($t_{8.5/5}$) are more advantageous, than welded with the shorter critical cooling time.

Statistical method [24] was used for the numerical evaluation of the high cycle fatigue experiments and for the determination of the fatigue design or limit curves. Using logarithmic form and rearranging the expression (1) we get the following equation:

$$\log(\Delta\sigma) = \frac{\log(a)}{m} - \frac{1}{m}\log(N_t). \tag{2}$$

This equation is a straight line on a double logarithmic scale, in our cases the life-time region of the design curve. The parameters of the equation and other characteristics of the curves can be found in Table 6. The N_k value is the number of cycles for the break point of the S-N curve, the $\Delta\sigma_D$ is the fatigue limit, and the $\Delta\sigma_{1E07}$ is the stress value belonging to $1 * 10^7$ cycles in the cases, when the horizontal (endurance limit) part of the curves cannot be determined.

As we can see in the Table 6, in four cases we cannot be determined the endurance limit part of the curves. In two cases (S690QL BM h/v and S690QL GMAWm k/1 W) the reason of this is the high resistance against the high cycle fatigue. In the third case (Weldox 700E GMAWm k/1 W), the available data were not enough for the determination of the second part of the S-N curve, while in the fourth case (GMAWm k/1 $W_{(16-17)}$), for the determination of the whole S-N curve. In these cases the $\Delta\sigma_{1E07}$ stress values also cannot be determined.

Table 6 The numerical results of the high cycle fatigue experiments

Base material	Specimen axle/crack grows direction	m	log(a)	N_k	$\Delta\sigma_D$	$\Delta\sigma_{1E07}$
		–	–	Cycle	MPa	MPa
S690QL	BM h/v	51.28	151.11	–	–	646
	GMAWm k/3 W	4.83	17.48	9.850E05	239	–
	GMAWm k/1 W	50.25	141.26	–	–	470
Weldox 700 E	BM h/k	12.45	39.65	1.677E06	483	–
	GMAWm k/1 W	9.96	32.47	–	–	–
	GMAWom k/1 W	17.45	52.93	2.320E06	466	–
S960QL	BM h/k	10.29	33.89	1.014E06	513	–
	GMAWm k/3 W	16.72	50.74	8.535E06	417	–
	GMAWum k/1 W	12.59	40.18	4.944E06	456	–
S960TM	BM h/k	11.51	37.90	6.995E06	500	
	GMAWm k/1 $W_{(6-7)}$	7.52	25.81	1.369E06	412	–
	GMAWm k/1 $W_{(16-17)}$	–	–	–	442	–

In accordance with these, the curves described by the given data in the Table 6, can be explained as fatigue design or limit curves.

6 Summary and Conclusions

Based on the accomplished investigations, their results, the elaboration of the results and the relevant literary data, furthermore, the performed comparisons the following conclusions can be drawn.

- The results of the executed investigations justified the necessity of the statistical approaches, especially referring to the directions of the welded joints and the determination of the number of the tested specimens.
- Applying the developed welding technologies eligible welded joints can be produced, where the appropriate quality contains the eligible resistance against high cycle fatigue, too.
- The resistance of the base materials against high cycle fatigue is more advantageous than the resistance of the welded joints. The quantity of the differences in some cases suggests the possibility and the necessity of the modification of the welding technology and/or the welding technological parameters.
- The matching problem has influence on the resistance of high strength steels against high cycle fatigue, depending on the strength category.
- The resistance of the investigated base materials and welded joints against high cycle fatigue is accordance with the resistance of the comparative materials can be found in the literature, moreover, in some cases is better at it.

References

1. Steel: a key partner in the European low-carbon economy of tomorrow. European steel technology platform (ESTEP), Brussels, March 2009, pp 1–16
2. European steel technology platform—Vision 2030. Report of the group of personalities. European Commission, Luxembourg, March 2004, pp 1–35. ISBN:92-894-5036-3
3. Miller WS et al (2000) Recent development in aluminium alloys for the automotive industry. Mater Sci Eng A 280:27–49
4. AluReport—AMAG customer and market information. 03. 2012, www.amag.at
5. Ghassemieh E (2011) Materials in automotive application. In: Chiaberge M (ed) New trends and developments in automotive industry InTech, pp 365–394. www.intechopen.com. (ISBN 978-953-307-999-8)
6. Plastics and polymer composites technology roadmap for automotive markets. American Chemistry Council, March 2014, pp. 1–59. www.americanchemistry.com
7. Metals Handbook, Volume 19. Fatigue and Fracture, ASM International, 1996
8. Balogh A, Dobosy Á, Frigyik G, Gáspár M, Kuzsella L, Lukács J, Meilinger Á, Nagy Gy, Pósalaky D, Prém L, Török I (2015) Hegeszthetőség és a hegesztett kötések tulajdonságai:

Kutatások járműipari acél és alumíniumötvözet anyagokon, (Szerk.) Balogh A., Lukács J., Török I., Miskolc (Hungary), p 324. (ISBN 978-963-358-081-3)

9. Dobosy Á, Nagy Gy (2015) Különböző folyáshatárú acélok és hegesztett kötéseinek kisciklusú fárasztása. XXII. OGÉT Nemzetközi Gépészeti Találkozó, Nagyszeben, 2014. 04. 24-27, pp 98–101, 2014 (ISSN 2068-1267)
10. Meilinger Á, Török I (2016) Lineáris dörzshegesztéssel készült kötések jellemzői kisciklusú fárasztó igénybevétel esetén, GÉP LXVII. évf. 1. szám, pp 63–71
11. Lukács J, Meilinger Á, Pósalaky D Fatigue curves for aluminium alloys and their welded joints used in automotive industry. Mater Sci Forum (In Press)
12. Lukács J, Nagy Gy, Harmati I, Koritárné FR, Kuzsella LnéKZs Szemelvények a mérnöki szerkezetek integritása témaköréből, (Szerk.) Lukács J, Miskolci Egyetem Miskolc (Hungary), pp 334, 2012 (ISBN 978-963-358-000-4)
13. Zsáry Á (1965) Méretezés kifáradásra a gépészetben. Műszaki Könyvkiadó, Budapest
14. Koncsik Zs, Lukács J (2013) Design curves for high-cycle fatigue loaded structural elements. Mater Sci Forum 729:135–144
15. MSZ-EN 1993-1-1:2009: EUROCODE 3: Acélszerkezetek tervezése. 1-1 rész: Általános és az épületekre vonatkozó szabályok
16. Stephens, RI, Fatemi A, Stephens RR, Fuchs HO (2001) Metal fatigue in engineering. Wiley (ISBN 0-471-51059-9)
17. BS 7910:2013 + A1:2015: Guide to methods for assessing the acceptability of flaws in metallic structures the British Standards Institution 2015. Published by BSI Standards Limited 2015 (ISBN 978 0 580 89564 7)
18. Barsom JM, Rolfe ST (1999) Fracture and fatigue control in structures: applications of fracture mechanics. ASTM Manual Series: MNL41. American Society for Testing and Materials, West Consthohocken, PA. (ISBN 0-8031-2082-6)
19. Lee YL, Pan J, Hathaway R, Barkey M (2005) Fatigue testing and analysis. Theory and Practice. Elsevier Butterworth-Heinemann. (ISBN-10 0-7506-7719-8)
20. Dobosy Á, Gáspár M (213) Welding of quenched and tempered high strength steels with heavy plate thickness. In: Proceedings 27th microCAD, international scientific conference, Miskolc (Hungary), Paper M7. (ISBN 978-963-358-018)
21. Gáspár M, Balogh A (2013) GMAW experiments for advanced (Q + T) high strength steels. Prod Process Syst 6(1):9–24
22. Balogh A, Gáspár M (2014) A matching kérdéskör: hozaganyagválasztás a konvencionális és korszerű nagyszilárdságú acélok hegesztéséhez. Hegesztéstechnika 25(3):75–80
23. Gáspár M, Balogh A (2014) Behaviour of mismatch welded joints when undermatching filler metal is used. Prod Process Syst 7(1):63–76
24. Nakazawa H, Kodama S (1987) Statistical S-N testing method with 14 specimens: JSME standard method for determination of S-N curves. In: Statistical research on fatigue and fracture. Current Japanese materials research. In: Tanaka T, Nishijima S, Ichikawa M (eds) Elsevier applied science and the society of materials science, Japan, vol. 2, pp 59–69 (ISBN 1-85166-092-5)
25. Pijpers RJM, Kolstein MH, Romeijn A, Bijlaard FSK (2007) Fatigue experiments on very high strength steel base material and transverse butt welds. Adv Steel Constr 5(1):14–32
26. Stemne D, Narström T, Hrnjez B (2010) Welding handbook. A guide to better welding of Hardox and Weldox, 1st edn. SSAB Oxelösund AB. (ISBN 978-91-978573-0-7)
27. Hamme U, Hauser J, Kern A, Schriever U (2000) Einsatz hochfester Baustähle im Mobilkranbau. Stahlbau 69(4):295–305

Toughness Examination of Physically Simulated S960QL HAZ by a Special Drilled Specimen

Marcell Gáspár, András Balogh and János Lukács

Abstract Based on the welding heat cycle models physical simulators are capable for the creation of critical heat-affected zones (HAZ). The simulated HAZ areas can be examined by various material testing methods (e. g. Charpy V-notch impact test) due to their increased homogeneous volume compared to their extension in real welding experiments. In our research work relevant technological variants ($t_{8.5/5}$ = 2.5…30 s) for gas metal arc welding technology were applied during the HAZ simulation of S960QL steel (EN 10025-6), and the effect of cooling time on the coarse-grained HAZ was analysed. In thermo-mechanical simulators the achievable cooling rate is always the function of specimen geometry and the presence of external cooling. Therefore a special drilled specimen with external cooling was developed for performing a shorter ($t_{8.5/5}$ = 2.5 s) cooling than 5 s, which cannot be realized on the conventional Gleeble specimen. Heat cycles were determined according to the Rykalin 3D model. The properties of the selected coarse-grained (CGHAZ) zone were investigated by scanning electron microscope, hardness test and Charpy V-notch impact test.

1 Introduction

Quenched and tempered (Q+T) high strength steels, applied in mobile structures (e.g. trucks, mobile cranes) belong to the highest category of structural steels. Their spreading application in vehicle industry is motivated by their outstanding strength

M. Gáspár (✉) · A. Balogh · J. Lukács
University of Miskolc, Miskolc, Hungary
e-mail: gasparm@uni-miskolc.hu

A. Balogh
e-mail: metweld@uni-miskolc.hu

J. Lukács
e-mail: janos.lukacs@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_40

and toughness characteristics. Due to their outstanding mechanical properties significant weight reduction can be achieved by their application.

These steels have a non-equilibrium tempered martensitic (in lower strength categories tempered martensitic + bainitic) microstructure due to the water cooling applied in the quenching cycle and to the high temperature tempering [1]. In order to realize the quenching condition in the whole cross section, alloying components (Cr, Mo) are added to the steel, which move the CCT curves to the right. Microalloying elements (Nb, V and Ti) are also used in order to ensure a fine grain microstructure and to prevent grain coarsening. The examined S960QL [2] steel has a guaranteed 27 J Charpy-V energy at −40 °C thanks to the fine grained microstructure (approx. 10 μm grain size) and low impurity level (minimized S and P content). The excellent strength of Q + T high strength steels is originated primarily from the fine tempered lath-like martensitic microstructure (Fig. 1) and the dislocation density. The effect of the dissolved carbon is generally lower [3].

The chemical composition and the calculated carbon equivalents of the investigated base material according to the material certification is summarized in Table 1.

The required and the measured mechanical properties of the examined S960QL heat according to the material certification is summarized in Table 2.

The strength and toughness characteristics of the base material can significantly decrease during the welding heat input due to the irreversible microstructural

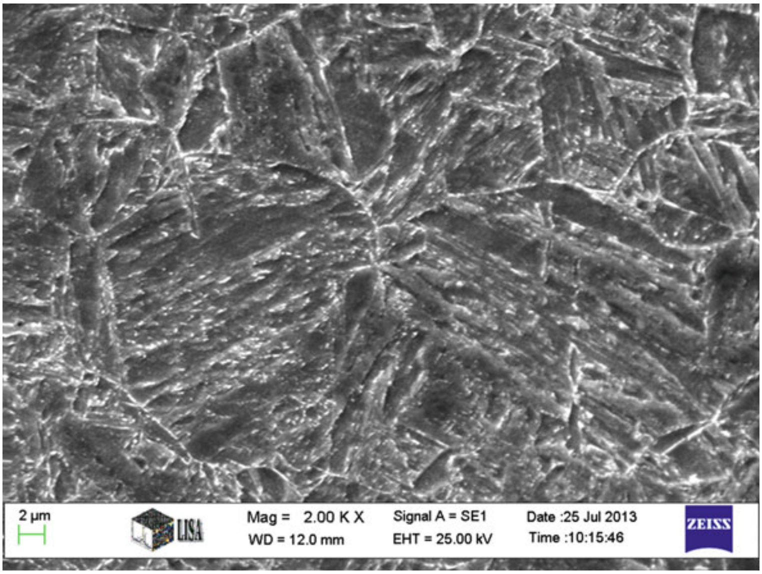


Fig. 1 Tempered martensitic microstructure of the base material, Nital 2%

Table 1 Chemical composition and carbon equivalents of S960QL in mass percent

S960QL	C	Si	Mn	P	S	Cr	Ni	CEV
	0.17	0.23	1.23	0.011	0.001	0.20	0.04	0.54
Mo	V	Ti	Cu	Al	Nb	B	N	CET
0.588	0.041	0.004	0.01	0.061	0.017	0.001	0.002	0.36

Table 2 Mechanical characteristics of the investigated S960QL

$R_{p0.2}$	R_m	$R_{p0.2}/R_m$	A_5	KV (at −40 °C)
MPa	MPa	–	%	J
≥ 960	980–1150	–	≥ 10	≥ 27
1014	1053	0.96	14	75

changes in the heat-affected zone. Besides softening hardness peaks can also locally form in the HAZ. Next to the fusion line the material is heated much above A_{c3} temperature therefore homogeneous austenite forms. When a coarse-grained zone (CGHAZ) forms the peak temperature is above 1100 °C where the grains start to exponentially expand in the function of the presence of different microalloying elements [4]. The decreased toughness of this zone has two reasons in quenched and tempered high strength steels. On the one hand the grain size can approximately be more than 10 times higher than the base material (>100 µm). Another reason derived from the alloying elements, resulting martensite with rough lath-like structure at shorter cooling times. Long cooling times may result coarse, upper-bainitic microstructure where M-A constituents can occasionally form, resulting in extremely low impact energy values [5]. Besides the weld metal CGHAZ has the highest risk of cold cracking since the hydrogen can diffuse from the fusion line to the brittle, coarse-grained microstructure resulting cold cracks with the residual tensile stress of the welded joint. Microalloying elements (in the present case Nb and V) can form small, disperse precipitates at the grain boundaries restraining excessive grain growth [4].

In real welding experiments the critical parts (e.g. CGHAZ) of the heat-affected zone is limitedly identified and investigated due to their narrow extension. Nowadays, physical simulation opened a wide range of examination possibilities for the precise analysis of heat-affected zone [6]. In the presented experimental work HAZ tests have been performed in a new generation of simulators, called Gleeble 3500, installed at the Institute of Materials Science and Technology of the University of Miskolc (Fig. 2). By the application of HAZ test the desired part of the heat-affected zone can be precisely and homogeneously created in a volume sufficient for the further material tests, e.g. Charpy V-notch impact test.

Fig. 2 Gleeble 3500 thermo-mechanical, physical simulator installed at the University of Miskolc

2 Experimental Plan

In our research work the aim was to analyse the effect of arc welding parameters on the HAZ properties of S960QL. Although more welding heat cycle models are available in the QuickSim software developed for the simulator, in the present case the GSL programs were manually written, using the time and temperature points of the total heating and cooling phase calculated by the Rykalin 3D model [7]. This model describes the temperature field generated by a moving point-like heat source on the surface of a semi infinity body. In this case 3D thermal conductivity is dominant while surface heat transfer (convection) is negligible. There were more reasons for the selection of 3D model. Firstly, this equation is independent from plate thickness, therefore the number of variables can be decreased during the analysis. Secondly the investigated S960QL is often used in medium and heavy plate thickness (>15 mm) where the 3D model may give more precise result.

Thermophysical properties were determined for the whole relevant temperature range (between 20 and 1400 °C) by the application of JMatPro software, and then their average values were used in the Rykalin equation (λ = 37.8 W/(m°C), c_p = 690.2 J/(kg°C), ρ = 7614.7 kg/m^3). JMatPro software was applied at the University of Oulu during the internship of one of the authors. Preheating and interpass temperatures (T_0) were set to 150 °C according to our previous welding experiments [8].

Our aim was to simulate the CGHAZ of the investigated S960QL for a $t_{8.5/5}$ cooling range (cooling time from 850 to 500 °C), typical for the generally applied gas metal arc welding (GMAW) process: 2.5…30 s. Short cooling times (2.5…5 s) can be relevant at root welding or high speed robotic welding. Longer cooling times

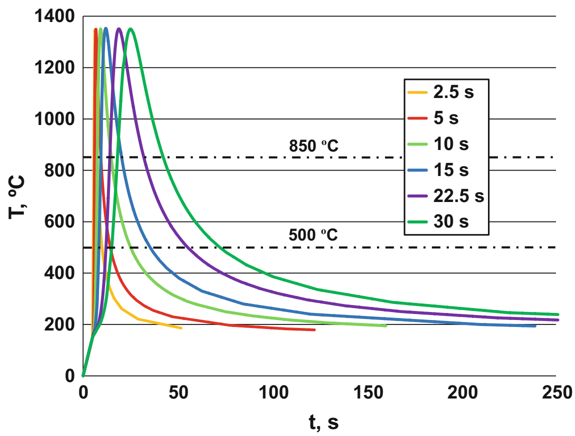


Fig. 3 CGHAZ heat cycles between $t_{8.5/5}$ = 2.5 and 30 s

(>20 s) may also happen in the industrial practice when weaving technique is applied by the welder. Considering the selection of peak temperatures our motivation was to generate the most critical part of CGHAZ, having the lowest toughness. Hence, the peak temperature was selected for 1350 °C in order to produce the occurring largest grains. This value is safely lower than the nil-strength temperature of the investigated steel (NST = 1408 °C [9]).

The designed CGHAZ cycles for the technological variants between 2.5 and 30 s are illustrated in Fig. 3.

Before running the simulation in Gleeble the optimal specimen geometry should be also selected. The cooling rate is generally determined by the copper grips and the water-cooled jaws surrounding the specimen. The specimen geometry and the application of external cooling can also significantly affect the cooling. At HAZ test when Charpy V-notch impact tests are planned, the conventional experimental set-up demands 10 × 10 × 70 mm square bar specimen. By the application of this specimen geometry the achievable shortest cooling time is approximately $t_{8.\ 5/5} \approx 5$ s (around 70 °C), whilst in the case of the GMAW shorter cooling time may happen. Therefore, there was a demand for a new specimen geometry capable of determining the impact energy of CGHAZ at short cooling times.

3 Development of a Special Drilled Specimen

3.1 Preliminary Tests

Since there is a technical possibility in Gleeble 3500 for the application of external cooling, the development of a drilled specimen geometry was planned which can be connected to the cooling systems through the holes on its both ends. It is important to note that cylindrical drilled specimen with a reduced cross section at the middle is generally applied for the determination of CCT diagrams where the extremely high cooling rate is often needed. The cooling set of this cylindrical specimen is planned to be applied to the developed new specimen geometry.

The first step was to analyse the effect of the holes at the end of the standard Charpy V-notch 10 × 10 × 55 mm specimen, which will be manufactured from the planned new Gleeble specimen. It is obvious that the diameter and the depth of the holes may affect the fracture during the impact test. Therefore, standard Charpy specimens were manufactured from the investigated base material, and holes were drilled with several depths (*l*) and diameter (*d*) sizes. Three specimens were prepared from each geometry. Then the tests were performed at −15 °C by a PSD 300/150 type impact testing equipment. The results of the absorbed energy (*CVE*) measured on the modified Charpy specimen (Fig. 5) are summarized in Table 3.

The fracture surface of the specimens was also analysed. In case of the deepest hole (l = 25 mm) the crack initiated at the radius of the V-notch, but then propagated partially towards the holes. When l = 22.5 mm was applied the hole did not have an effect to the crack propagation and hence to the impact energy. By considering that notable difference was not observed between the tested 6 and 7 mm hole diameter, the latter physical simulation specimens have d = 7 mm in order to realize a more intensive cooling.

3.2 Geometry

After performing the preliminary tests the specimen geometry was finalized (Fig. 4), and the necessary specimens were manufactured from the S960QL plate.

Table 3 The effect of the hole geometry on CVE at −15 °C

d, mm	l, mm	CVE_1, J	CVE_2, J	CVE_3, J	CVE_{mean}, J	D, J
6	0	133	160	133	142	15.6
	17.5	145	81	172	133	46.7
	20	176	187	201	188	12.5
	22.5	128	173	165	155	24.0
	25	146	117	94	119	26.1
7	22.5	136	148	149	144	7.2

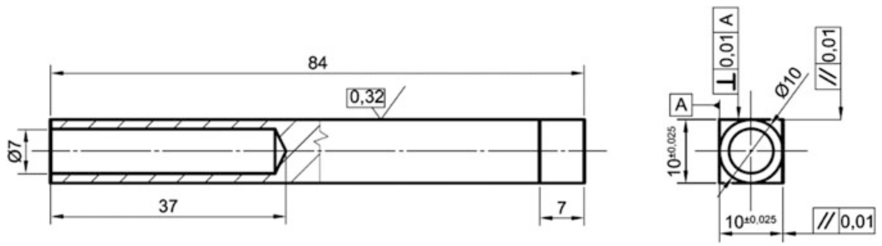


Fig. 4 The developed Gleeble specimen for the analysis of HAZ toughness at short cooling times

After the successful Gleeble tests the specimen geometry has to be modified (Fig. 5) according to the requirements of Charpy V-notch impact test.

3.3 Test Results

Generally, one thermocouple is enough for the process control, but in present case, two K-type (NiCr–Ni) thermocouples were welded to the opposite sides of the specimens at the middle in order to control whether the heat input symmetrical enough.

Then the specimens were placed into the grips and the cooling device was joined at the two sides (Fig. 6) before fixing it into the jaws.

Free span between the grips was set to 7.5 mm in order to maximize the cooling rate without wagering the success of the simulation (Gleeble Application Note allows even 5 mm, although the optimal is around 10 mm). Compressed air was applied to the cooling, although water can also be used.

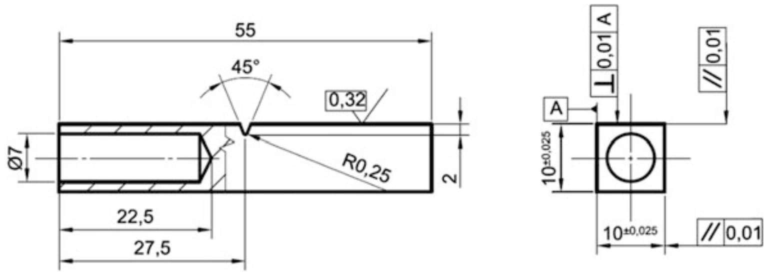


Fig. 5 Modified Charpy V-notch specimen

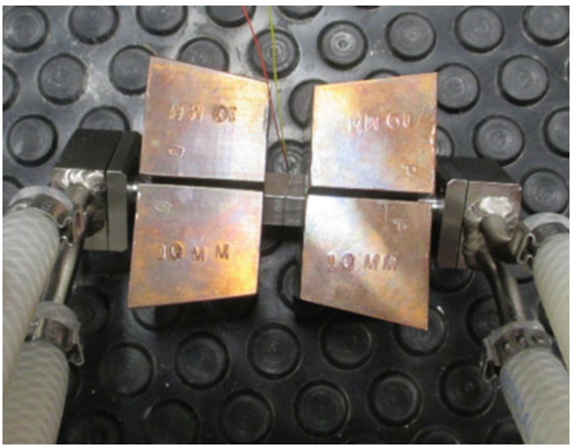


Fig. 6 Drilled square bar specimen joined to the cooling device in the copper grips

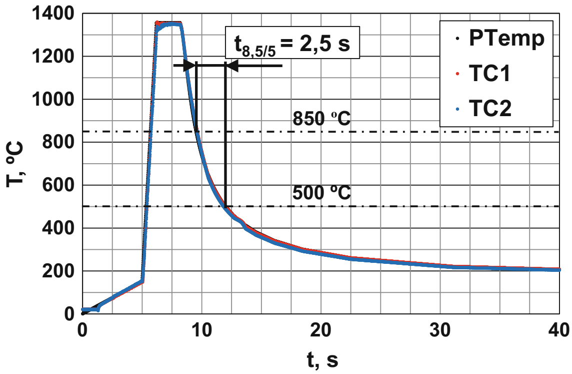


Fig. 7 Programmed (PTemp) and measured (TC1 és TC2) heat cycles

After the first simulations inhomogeneity was noticed in the grain size during the microscopic tests. The grains closer to the surface had larger size compared to the inner part, so the maximum temperature was not equal at the middle and the surface. After further experiments the conclusion was to apply a 2 s holding time at the peak temperature, which can equalize the temperature between the middle and the surface. Therefore the heat cycle of the $t_{8.5/5}$ = 2.5 s (Fig. 3) had to be modified, and only the cooling part was described by the Rykalin 3D. Heating rate was set to 1000 °C/s, followed by a 2 s holding time. The programmed and the measured thermal cycles by the two thermocouples are illustrated in Fig. 7.

In Fig. 7 the red (*TC1*) and blue (*TC2*) curves, indicating the measured heat cycles, nearly overlapped the programmed cycle (black curve). This verifies that by the application of this special drilled specimen with external compressed air cooling $t_{8.5/5}$ = 2.5 s can be achieved, which is a double cooling rate compared to the conventional specimen set-up at HAZ test.

4 Material Tests

4.1 Microscopic Tests

After the physical successful simulations with the two specimen geometries (10 × 10 × 70 mm for 5…30 s, and special geometry presented in Fig. 4 for 2.5 s) the samples were perpendicularly cut at the thermocouples and the surfaces were prepared for the microstructural analysis. The properties of CGHAZ and the effect of welding parameters were examined with a ZEISS EVO MA10 scanning electron microscope. Samples were coated with a thin gold layer in order to develop picture quality. The microstructures of CGHAZ at 2.5, 5 and 30 s are presented between Figs. 8, 9 and 10.

All microscopic tests verified that the desired CGHAZ was successfully simulated in all cases. Regarding CGHAZ rough lath-like martensitic microstructure

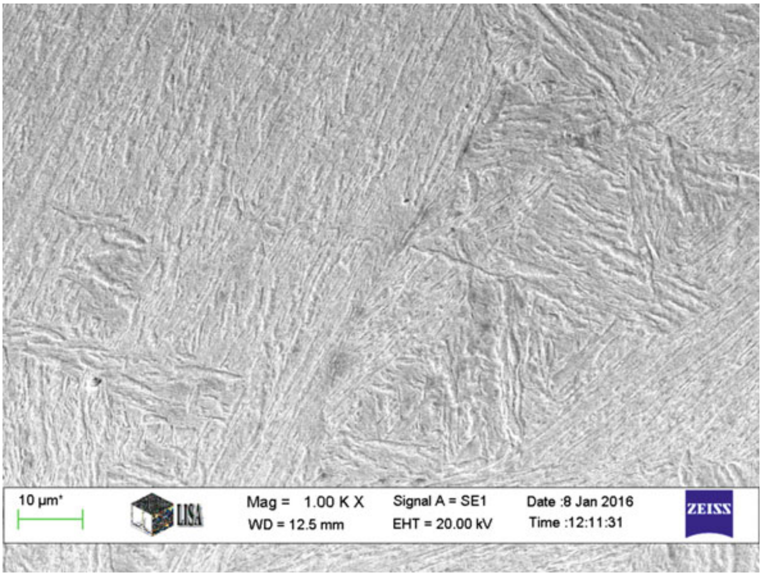


Fig. 8 CGHAZ microstructure at $t_{8.5/5}$ = 2.5 s, Nital 2%

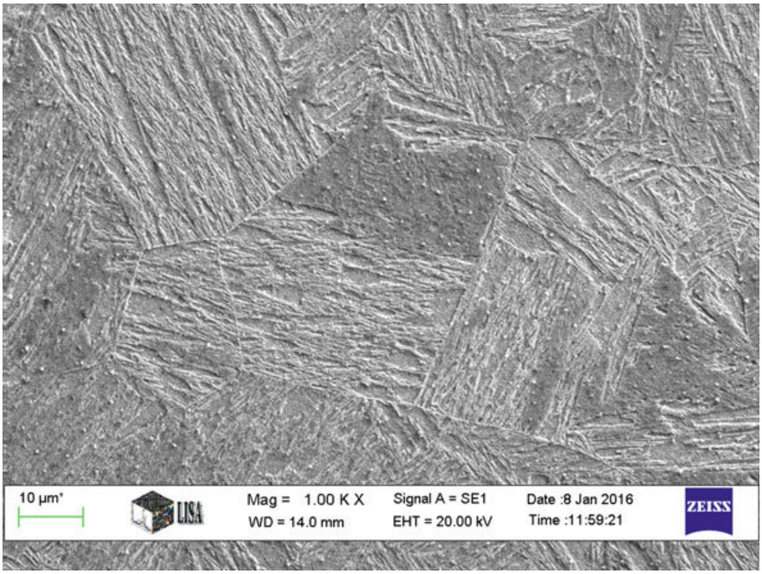


Fig. 9 CGHAZ microstructure at $t_{8.5/5}$ = 5 s, Nital 2%

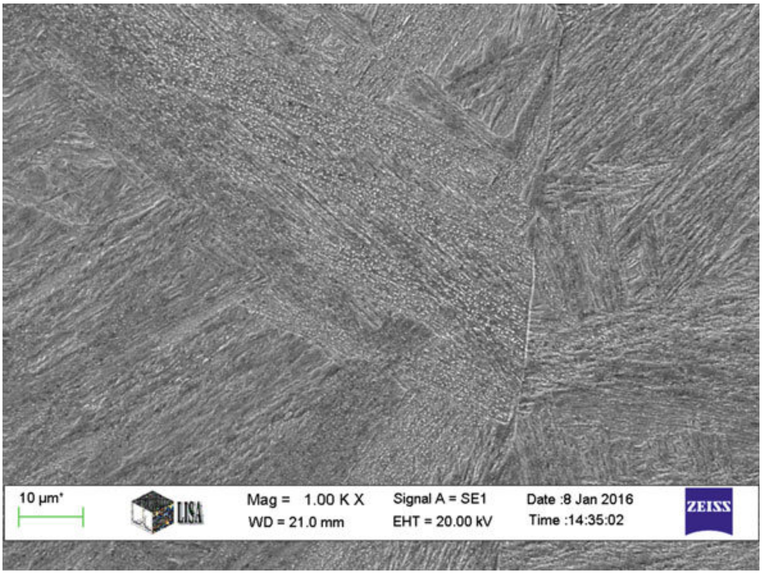


Fig. 10 CGHAZ microstructure at $t_{8.5/5}$ = 30 s, Nital 2%

with large (>100 µm) prior austenite grain size was noticed between the investigated whole cooling time range: $t_{8.5/5}$ = 2.5…30 s. Occasionally the measured grain size was even 200 µm. The grain size increased with the longer cooling times. It is important to note that in physical simulations the grain size of CGHAZ is generally larger than in the real welded joints [6]. The type of microstructure did not change significantly in the function of the cooling time, although the characteristics of the self-tempering of martensite were noticed marked by small carbide precipitates at $t_{8.5/5}$ = 30 s, where the conditions of carbon diffusion are more favourable. The results correlated with the JMatPro calculations.

4.2 Hardness Tests

A Reichter UH250 universal macro hardness tester was applied for the examination. Five measurements were done on the cut surface. The average macro hardness values of the simulated CGHAZ are summarized in Fig. 11.

An evaluation was performed according to the EN 15614-1 [10] standard, which permits HV_{max} = 450 HV10 for the non-heat treated welded joints (including HAZ) of quenched and tempered high strength steels belonging to the 3rd group of CR ISO 15608. In Fig. 11 340 HV10 indicates the average hardness of the base material. In the investigated cooling time range CGHAZ fulfilled the requirement of the governing standard, and critical hardening or softening there was not noticed.

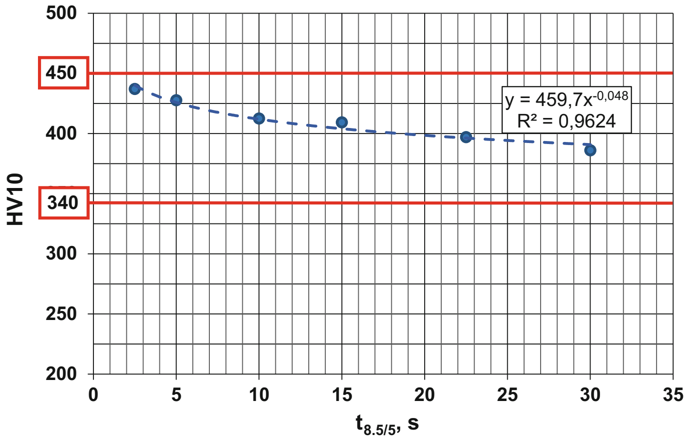


Fig. 11 Hardness of CGHAZ in the function of $t_{8.5/5}$ cooling time, HV10

The maximum hardness (437 HV10) was measured at $t_{8.5/5}$ = 2.5 s in CGHAZ. The effect of cooling time can be clearly identified with the hardness values of the CGHAZ specimens. The average hardness reduced as the cooling time was longer. Although the requirement was fulfilled in CGHAZ it can be important to mention that in quenched and tempered high strength steel the fine-grained zone (FGHAZ) can have similarly high hardness, which may exceed the CGHAZ hardness [11].

4.3 *Charpy V-Notch Impact Tests*

Minimum five specimens from each heat cycle were used for the impact tests performed in PSD 300/150 equipment. According to EN 10025-6 the required minimum impact energy is 27 J at −40 °C for S960QL. According to the material certificate the investigated steel plate had 75 J average impact energy, although 167 J was measured during our impact test on the base material. The individual and the mean *CVE* values are illustrated in Fig. 12.

It can be concluded that all of the investigated heat-affected zones in the examined $t_{8.5/5}$ cooling time interval had significantly lower toughness than the base material. In almost all cases except $t_{8.5/5}$ = 30 s there were individual values that did not reach the required minimum level. Above $t_{8.5/5}$ = 15 s the impact energy slightly increased by the cooling time when martensitic microstructure was noticed that may be originated from the self-tempering effect presented at the microstructural analysis. Although the average absorbed energy was slightly above the required 27 J at almost every cooling time (except $t_{8.5/5}$ = 15 s), the specimens fractured brittle, which means that 27 J requirement cannot fully guarantee the ductile behaviour of the heat-affected zone.

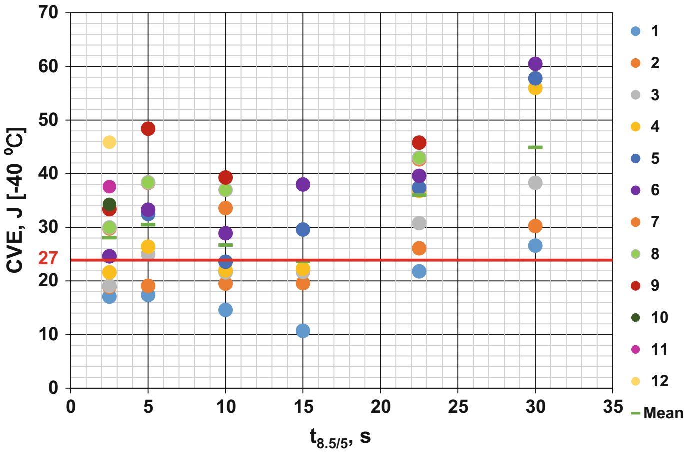


Fig. 12 Charpy-V energy of CGHAZ in the function of the cooling time

5 Conclusions

The CGHAZ of S960QL was successfully performed by the Gleeble 3500 physical simulator between $t_{8.5/5}$ = 2.5…30 s cooling time.

For the physical simulation and the toughness analysis of $t_{8.5/5}$ = 2.5 s a drilled specimen geometry was developed and tested.

Comparing to mild steels (e.g. S355J2+N) quenched and tempered high strength steels are more sensitive to the welding heat input. In the CGHAZ of S960QL relatively high (>400 HV10) hardness values were measured between $t_{8.5/5}$ = 2.5… 22.5 s. However, the measured macro hardness values did not exceed the 450 HV10 requirement of the governing standard.

In the examined $t_{8.5/5}$ cooling time interval CGHAZ had significantly lower toughness (the mean was slightly above 27 J at −40 °C) than the base material.

Acknowledgements The research work presented in this paper based on the results achieved within the TÁMOP-4.2.1.B-10/2/KONV-2010-0001 and TÁMOP-4.2.2.A-11/1/KONV-2012-0029 projects in the framework of the New Széchenyi Plan. The realization of these projects was supported by the European Union, and co-financed by the European Social Fund.

References

1. Weglowsky M (2012/2) Modern toughened steels—their properties and advantages, Biuletyn Instytutu Spawalnictwa, pp 25–36
2. EN 10025-6: Hot rolled products of structural steels. Technical delivery conditions for flat products of high yield strength structural steels in the quenched and tempered condition

3. Porter DA, Easterling KE (1996) Phase transformations in metals and alloys, 2nd edn. Chapman and Hall, 2-6 Boundary Row, London SE1 8HN, UK. (ISBN 0-412-45030-5)
4. Bhadesia HKDH, Honeycombe RWK (2006) Steels microstructure and properties, 3rd edn. Elsevier Linacre House, Jordan Hill, Oxford OX2 8DP, UK
5. Nevasmaa P (1996) Evaluation of HAZ toughness properties in modern low carbon low impurity 420, 550 and 700 MPa yield strength thermomechanically processed steels with emphasis on local brittle zones. Lisensiaatintyö, University of Oulu, p 176
6. Adony Y (2006) Heat-affected zone characterization by physical simulations. Welding J 42–47
7. Rykalin NN (1953) Teplovüe processzi pri szvarke, Vüpuszk 2 Izdatelsztvo Akademii Nauk SzSzSzR, Moszkva, p 56
8. Gáspár M, Balogh A (2013) GMAW experiments for advanced (Q+T) high strength steels. Prod Process Syst 6(1). University of Miskolc, Department of materials processing technologies, pp 9–24
9. Kuzsella L, Lukács J, Szűcs K (2013) Nil-strength temperature and hot tensile tests on S960QL high-strength low-alloy steel. Prod Process Syst 6(1):67–78
10. EN ISO 15614-1: Specification and qualification of welding procedures for metallic materials—welding procedure test—Part 1: arc and gas welding of steels
11. Dunne DP, Pang W (2015) Heat affected zone hardness of welded low carbon, quenched and tempered steels. IIW IX-2523-15

Innovation Methods for Residual Stress Determination for the Automotive Industry

Máté Sepsi, Dávid Cseh, Ádám Filep, Márton Benke and Valéria Mertinger

Abstract Determination of the residual stress state in a loaded automotive component is highly important because of its strong effect on the lifetime of the element. Nowadays, the residual stress characterization of products became an everyday requirement in the automotive industry, and the quality control is impossible to imagine without it. Forasmuch as every producing process (casting, heat treating, different kinds of metal deformation processes and surface compressing methods, etc.) influences the residual stress state, therefore it can be very complex and various within the materials. If we conscious in the effect of these processes it is possible to reach such a state in the material which can enhance its lifetime and quality, and with an optimized process, the costs could be reduced. Several methods exist for measuring the residual stress and each of them has its own advantage and disadvantage. With this paper our purpose is to introduce nowadays' available stress measuring methods such as X-ray diffraction, magnetic Barkhausen noise and the hole drilling methods and a few more alternatives. Some useful results from the practice are also presented.

M. Sepsi (✉) · D. Cseh · Á. Filep · M. Benke · V. Mertinger
Institute of Physical Metallurgy, Metalforming and Nanotechnology,
University of Miskolc, Miskolc, Hungary
e-mail: mate.sepsi@gmail.com

D. Cseh
e-mail: cseh88david@gmail.com

Á. Filep
e-mail: adam9032@gmail.com

M. Benke
e-mail: fembenke@uni-miskolc.hu

V. Mertinger
e-mail: femvali@uni-miskolc.hu

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4_41

1 Introduction

Residual stress is introduced into solid state materials during almost all shape modifying processing types including casting, plastic deformation, machining, welding, etc. For most cases, the presence of residual stress is not designed, but is a consequence of the required manufacturing step [1, 2]. In the automotive industry, however, residual stress is introduced within a machine component intentionally to enhance the component's properties, such as resistance against fatigue [3, 4]. The most common techniques to induce compressive stresses are shot peening and ball burnishing. In such cases, the stress state (type and magnitude) of the component is a requirement and regulated similarly to mechanical properties or geometry. Because of that, providing accurate data about residual stress is of great interest. There are many methods that can qualitatively describe the residual stress state of a component [5, 6]. One of them is the Magnetic Barkhausen Noise measurement, which could be the quickest and the easiest measurement from the side of the operator. It is based on a ferromagnetic phenomenon which induced by a changing magnetic field in a ferromagnetic material. Forasmuch the characteristic of the Barkhausen noise is influenced by the microstructure, the hardness and the stress state of the magnetized material, therefor it is suitable to use it in the industry for quality control, where the measurement time is very important [7]. Nevertheless, the description of the residual stress state quantitatively is not easy with this method [1]. Another one is the Hole-drilling method, where the definition of the residual stress based on the macroscopic deformation of a solid material upon changing the stress balance through material removal. The largest advantage of this kind of measurement is that not just the ferromagnetic and the polycrystalline materials are analysable with it. But, as it is mentioned, material should be removed to get result, so stress state on the surface is not definable. However, diffraction based methods are the only ones that provide accurate quantitative information about the stress state from the surface to the deep. While neutron diffraction requires a neutron source, X-ray diffraction based residual stress measurement can be performed in laboratory conditions, or, even as on-site measurements. Using non-destructive X-ray diffraction, no sample has to be cut from the component for the examination. Using this technique, not only remains the component usable after the examination, but the true stress state is not distorted with sample cutting. Thus, the true stress state is provided after such examination. The residual stress measurement by X-ray, as all of the other diffraction measurements is based on Bragg's law [8]:

$$n\lambda = 2d \sin\theta \tag{1}$$

This equation gives the diffraction angle (θ) where the diffracted X-ray photons with a given wavelength (λ) meet in the same phase after scattered on the lattice with a given plane distance (d), (n is an integer.) An intensity peak appears as a function of the diffraction angle in the recorded interference function where Bragg's law is fulfilled. The diffraction angle of a stressed material is shifted relative to the relaxed value. The lattice parameter change is the source of this phenomenon which is generated by the residual stress. Assuming linear elastic distortions, the related normal stresses can be calculated by Eq. (2) which is called $\sin^2\psi$ method [8]:

$$\sigma = \frac{d_\psi - d_0}{d_0} * \frac{E}{(1+\nu)\sin^2\psi} \tag{2}$$

where d_0 and d_ψ are the lattice plane distances of the strained (stressed) material in the normal and in the tilted position defined by the angle ψ, respectively. E (Young modulus) and ν (Poisson number) are the elastic parameters of the material. In practice, several tilting positions (± () in one measuring point are usually applied and after the interferential functions are recorded, d), normal stress and its scatter, shear stress and its scatter, the average 2 (values and the average full with half maximum (FWHM) can be calculated. Figure 1 provides the process of an XRD residual stress data in one measuring point. In this case 5 tilt positions (3/3) were applied and the interferential functions were recorded at each (Ψ = −45, −30, 0, +30, +45°) tilt positions. The sloop of the parabola axis on the plot of calculating d as the function of $\sin^2\Psi$ gives the value of normal stress data. The direction of the measured stress is given by the direction of the applied tilting was used during the measurement. The FWHM and the Bragg angle are calculated as an average value

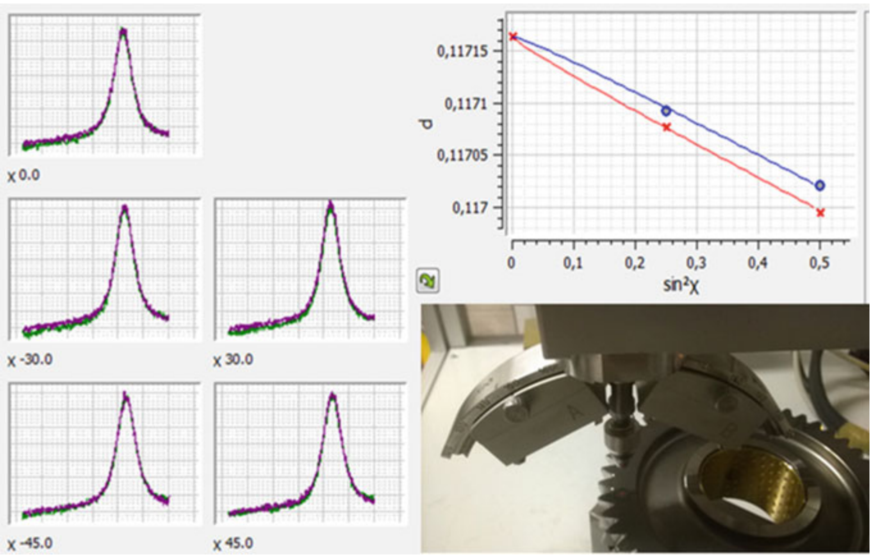


Fig. 1 The process of an XRD residual stress data. The recorded interferential functions at different (Ψ = −45, −30, 0, +30, +45°) tilt positions. The plot of calculating d as the function of $\sin^2\Psi$. The gears under the measuring process at the A location. The calculated results are: σ: −448.7 ± 9.6 MPa, τ: −16.0 ± 2.1 MPa, FWHM: 4.57° ± 0.07°, 2θ:155.77°

of the data of interferential functions which were recorded at different tilting positions.

The X-ray diffraction technique is a reliable method for stress measurements. However, the residual stress state of a component can be measured accurately only if the true stress state is not affected by sample cuts. This requirement demands a special equipment and measurement method [2, 8]. Non-destructive X-ray diffraction is generally accepted as the ultimate tool for residual stress measurements and the method meets the requirements of the regarding Machinery and Low Voltage directives and applies the following standards: SFS-EN ISO 12100-1, 12100-2, 13857, 60204-1. Even the method is standardized, the effect of measurement parameters on the obtained result is still debated. The objective of the present manuscript is to examine the effect of measurement parameters, namely tilting range, number of tilting positions and spot size of the measured stress values of a shot-peened gear wheel. Our results also give evidence about the reproducibility and sensitivity of the diffraction method.

2 Experimental

The residual stress was measured non-destructively by a Stresstech Xstress 3000 G3R type centerless X-ray diffractometer, owned by the Institute of Physical Metallurgy, Metalforming and Nanotechnology at the University of Miskolc. This method provides residual stress values from ~10 μm depth of the surface layer (depending on material type, applied generator power, Bragg-angle). A Cr radiation was used to measure the {211} reflections of the samples. For stress calculations, the following material constants were used: Poisson's ratio $\nu = 0.3$, Young's modulus $E = 211000$ MPa. A carburized and shoot peened gear wheel was chosen as test component. All of the measurements were performed at the same two locations (D and A location) and the same direction of the gear wheel. The measured parameters (the investigated variables) such as maximum tilting angle, the number of tilting and spot size (the diameter of the beam on the examined surface) were systematically changed during the test. The exposure time was chosen based on the requirement of proper signal/background ratio, so it was changing, but it is not the investigated variable in this paper. The spot size was 1, 3, 5, mm in diameter. The maximum tilting angles were ±20° or ±40°. The 3/3, 5/5, 7/7 and 11/11 number of tilts were chosen which refers that one stress data were calculated based on 3, or 5, or 7, or 11 tiltings in both plus and minus directions. Three parallel measurements were performed in each case of the measuring conditions. Thus, reproducibility was also tested.

3 Results and Discussion

Figure 1 provides one of the measured data at location A as an example and Table 1 summarizes all of the calculated data, such as normal stress (σ), shear stress (τ), FWHM and Bragg angle (2θ). The normal and shear stresses refer to the stress state in the measured point and direction, which was perpendicular to the axis. The FWHM value is larger if the micro stress caused by dislocations or interstitials is also larger. The Bragg angle is associated with the difference of chemical composition, especially the carbon solution of the carburized gear surface zone.

Figures 2, 3, 4, 5, 6, 7, 8, 9 and 10 show the results of Location A and D which were the equivalent parts of the gear. The error bars of each data are also presented. The significant different data of the two locations are originated from the treatment of the gear which can also be important information. The sufficient number of tilts has been always a question for the users. Too high number increases the measuring time while the too low causes wide error bars. Figures 2 and 3 show that it is a false theory. The stress results of 11 tilts number shows only smaller error bar, but the difference is much smaller than the difference of the result from the equivalent parts (location A and D) of the gear. The FWHM data show no difference.

The optimal spot size is also a regular question for the operators. The larger collimator (larger spot size) gives higher reflected intensity so the measuring time is decreased which is convenient situation. Also, a feeling of the users that using larger collimator gives better statistic for the material results that the commonly used spot size 3–4 mm in diameter. As we see from Figs. 4 and 5 the stress data deviation is within the error bars, even if 6 times larger spot size was used, so the thumb rule works, but the FWHM data are strongly increased as the spot size is increasing. Which does not originate from the microstructure, but is an artificial effect. So the FWHM data can be compared in the case of same spot size measurements only.

The result in Figs. 6, 7, 8 and 9 has a serious message. The ±45° maximum tilting angle is usually used, but in some cases the special geometry limit does not yield such a big tilting. Measuring in the bottom land of a gear is a typical example. In that case the operators cut the gears or apply different measuring mode (psi or omega) or decrease the maximum tilting angle. As we have already mentioned in the introduction, the sample cutting can modify the stress state so the measured data do not represent the real initial situation. The question is always the effect which generated the stress state and the volume fraction of the sample where this state is in equilibrium. Applying the omega mode can be a good solution, but some of the available equipment are not able to fulfil this. Figure 6 shows that applying ±20° tilting angle instead ±45° results an extremely large increase in the error bars of the measured points and also the large scatter of the parallel measurements and the spot size variation does not help at all (Fig. 7). In front of this, it is obvious that the FWHM data unaffected by the tilting angles (Fig. 8) only the spot size effect is detectable (Fig. 9).

Table 1 Calculated data

Location	Var.	No.	Tilt max, ″	No. of tilt	Spot size, mm	σ, Mpa	(±)	τ, MPa	(±)	FWHM, °	(±)	2θ, °
Location D	a	1	45	3	3	−444.20	11.60	−10.30	2.60	4.52	0.09	155.76
	a	2	45	3	3	−448.70	9.60	−16.00	2.10	4.57	0.07	155.77
	a	3	45	3	3	−435.90	14.50	−11.60	3.20	4.53	0.10	155.80
	b	4	45	5	3	−437.90	7.90	−7.10	1.60	4.53	0.06	155.79
	b	5	45	5	3	−443.90	4.80	−7.30	1.00	4.53	0.07	155.76
	b	6	45	5	3	−447.40	8.40	−8.50	1.70	4.51	0.06	155.76
	c	7	45	7	3	−443.60	5.90	−7.30	1.10	4.50	0.06	155.80
	c	8	45	7	3	−432.50	5.20	−7.60	1.00	4.50	0.07	155.76
	c	9	45	7	3	−452.10	7.50	−7.30	1.40	4.52	0.07	155.77
	d	10	45	11	3	−450.10	6.00	−9.40	1.10	4.51	0.06	155.77
	d	11	45	11	3	−448.50	4.70	−8.20	0.90	4.50	0.06	155.77
	d	12	45	11	3	−441.70	5.70	−9.40	1.10	4.51	0.08	155.77
	e	13	20	5	5	−459.50	30.50	−1.70	2.40	4.43	0.04	155.76
	e	14	20	5	5	−445.60	28.80	−5.90	2.30	4.46	0.04	155.76
	f	15	45	5	3	−437.90	7.90	−7.10	1.60	4.53	0.06	155.79
	f	16	45	5	3	−443.90	4.80	−7.30	1.00	4.53	0.07	155.76
	f	17	45	5	3	−447.40	8.40	−8.50	1.70	4.51	0.06	155.76
	g	18	20	5	3	−459.50	30.50	−1.70	2.40	4.43	0.04	155.76
	g	19	20	5	3	−445.60	28.80	−5.90	2.30	4.46	0.04	155.76
	h	20	45	5	3	−438.4	11.30	−9.10	2.20	4.53	0.08	155.77
	h	21	20	5	3	−438.3	35.80	−7.70	2.80	4.44	0.03	155.76
	h	22	20	5	3	−445.3	27.70	−5.40	2.20	4.48	0.03	155.76
	i	23	20	5	1	−429.4	26.90	5.30	2.10	4.36	0.03	155.83
	i	24	20	5	1	−408	27.60	2.70	2.20	4.35	0.03	155.84
	i	25	20	5	1	−464.5	28.90	3.40	2.30	4.36	0.04	155.81

(continued)

Table 1 (continued)

Location	Var.	No.	Tilt max, ″	No. of tilt	Spot size, mm	σ, Mpa	(±)	τ, MPa	(±)	FWHM, °	(±)	2θ, °
Location A		1	20	5	5	−498.4	16.80	11.50	1.30	5.06	0.03	155.71
		2	20	5	5	−448.6	11.10	−10.10	0.90	5.07	0.02	155.71
		3	20	5	5	−504.7	15.60	−10.00	1.20	5.08	0.03	155.72
		4	20	5	4	−481.9	28.30	−8.40	2.20	4.91	0.05	155.71
		5	20	5	4	−446.9	22.60	−10.60	1.80	4.91	0.03	155.71
		6	20	5	4	−503.1	16.20	−8.40	1.30	4.91	0.02	155.72
		7	20	5	3	−467.6	16.00	−6.80	1.30	4.73	0.03	155.73
		8	20	5	3	−492.6	28.20	−7.60	2.20	4.73	0.03	155.73
		9	20	5	3	−463	20.80	−6.60	1.60	4.73	0.04	155.73
		10	20	5	1	−485.2	20.70	−7.80	1.60	4.65	0.06	155.75
		11	20	5	1	−498.1	41.00	−6.70	3.20	4.65	0.06	155.75
		12	20	5	1	−462.4	24.20	−8.60	1.90	4.68	0.05	155.75
		13	20	5	1	−498.8	24.40	−5.60	1.90	4.58	0.04	155.74
		14	20	5	1	−461.9	20.50	−5.70	1.60	4.58	0.05	155.74
		15	20	5	1	−493.1	13.00	−6.00	1.00	4.60	0.06	155.75
		16	45	5	5	−497.6	8.20	−10.90	1.90	5.15	0.10	155.71
		17	45	5	5	−501.7	9.90	−12.50	2.00	5.14	0.09	155.70
		18	45	5	5	−500.3	6.60	−11.40	1.30	5.13	0.09	155.71
		19	45	5	4	−491.5	8.30	−9.70	1.60	4.95	0.09	155.69
		20	45	5	4	−498	9.00	−12.70	1.80	4.94	0.07	155.73
		21	45	5	4	−488.9	7.80	−10.90	1.50	4.95	0.08	155.70
		22	45	5	3	−490.8	5.70	−8.50	1.10	4.74	0.05	155.73
		23	45	5	3	−488	5.40	−9.10	1.10	4.74	0.05	155.73
		24	45	5	3	−479.2	6.20	−7.80	1.20	4.74	0.06	155.72

(continued)

Table 1 (continued)

Location	Var.	No.	Tilt max, ″	No. of tilt	Spot size, mm	σ, Mpa	(±)	τ, MPa	(±)	FWHM, °	(±)	2θ, °
		25	45	5	1	−469.9	11.70	−5.10	2.30	4.63	0.05	155.74
		26	45	5	1	−474.6	11.70	−4.50	2.30	4.62	0.05	155.75
		27	45	5	1	−470.3	11.30	−3.50	2.20	4.60	0.05	155.75
		28	45	5	1	−494.8	8.30	−5.30	1.60	4.54	0.06	155.74
		29	45	5	1	−485.7	7.20	−4.30	1.40	4.56	0.07	155.72
		30	45	5	1	−485	6.00	−6.10	1.20	4.54	0.05	155.74

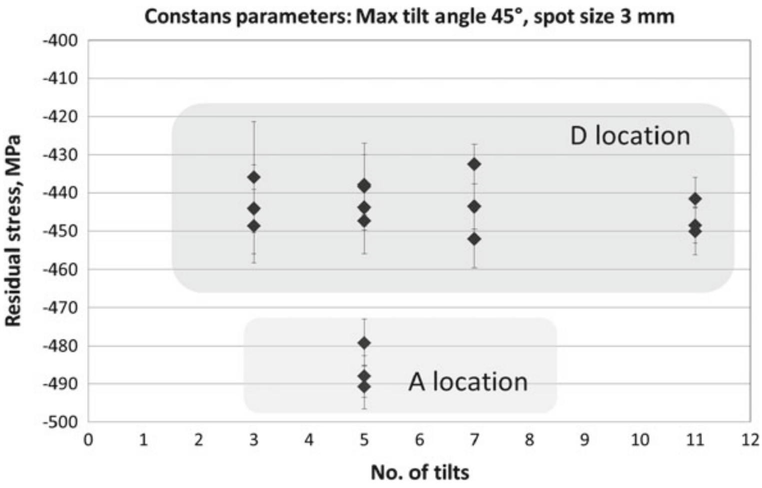


Fig. 2 Residual stress data in function of number of tilts

Fig. 3 FWHM data in function of number of tilts

The difference of the two locations (A and D location) can be explained based on the results of Fig. 10. All the residual stress and FWHM data at A and D locations are plotted. The spot size, the max tilting angle and the number of tilt are variables. Three parallel measurements were performed with a given set of variables. The Bragg angles have a small scatter in both locations due to the measuring parameters, but the data of the two locations are markedly different. The A location represents smaller Bragg angles and larger FWHM values. It can happen, if the soluted carbon concentration is higher at the A location. The intensity difference of shot peening

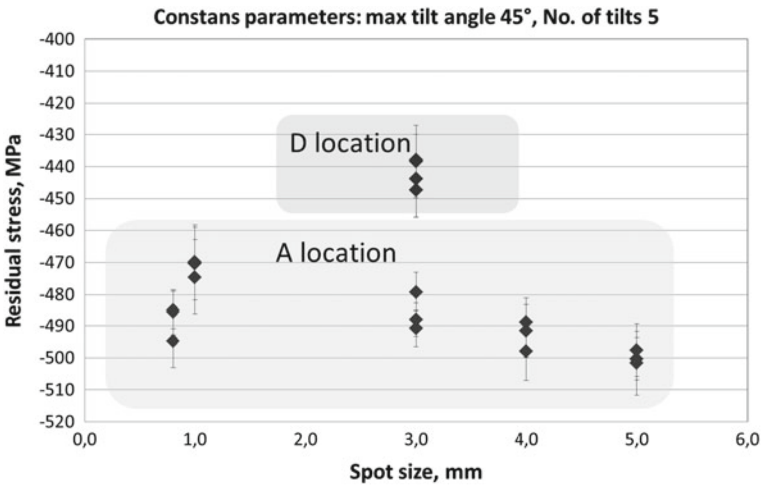


Fig. 4 Residual stress data in function of spot size

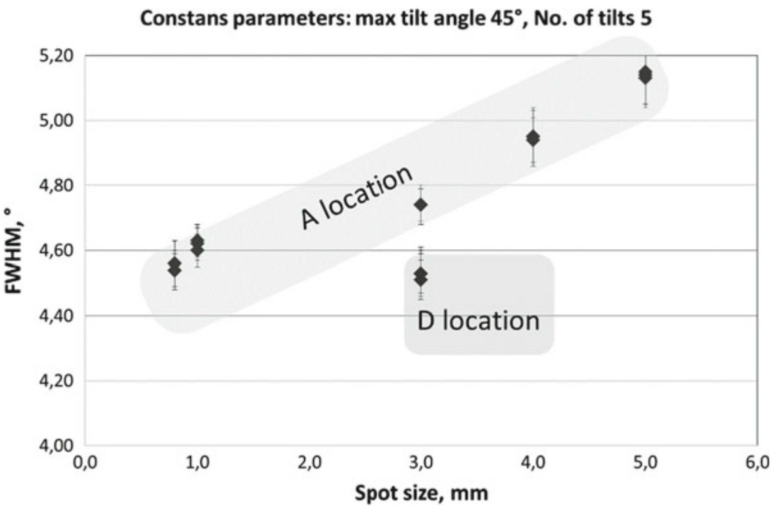


Fig. 5 FWHM data in function of spot size

can cause a shift of FWHM value, but in that case the Bragg angle is unaffected. Because reporting stress results together the other parameters (FWHM, θ) gives a full description with much additional information is always highly recommended in case of any component investigation.

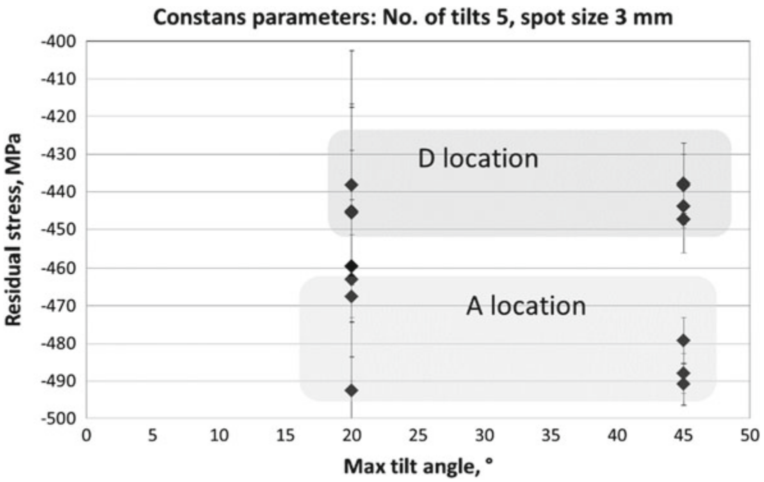


Fig. 6 Residual stress data in function of max tilt angle

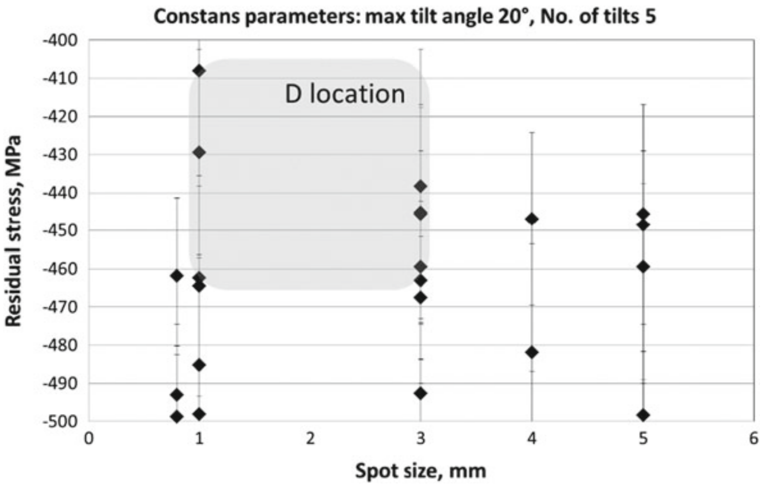


Fig. 7 Residual stress data in function of spot size

The reproducibility of the measurement has been always the question in any case of measurements. Figure 11 shows the reproducibility of stress measurement at the D location in 9 different sets of measurements (a–i). Three parallel measurements were carried out with a given set of variables. Now it is evident that the reproducibility is very good, the scatter of the parallel measured data is smaller than the error bar value.

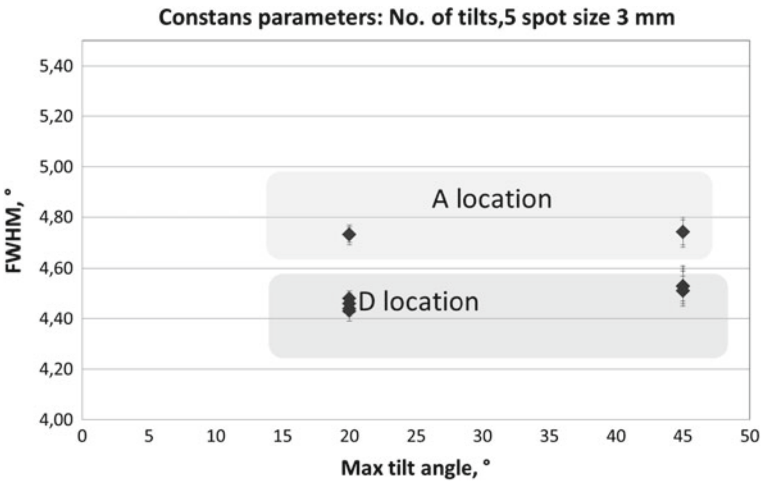


Fig. 8 FWHM data in function of max tilt angle

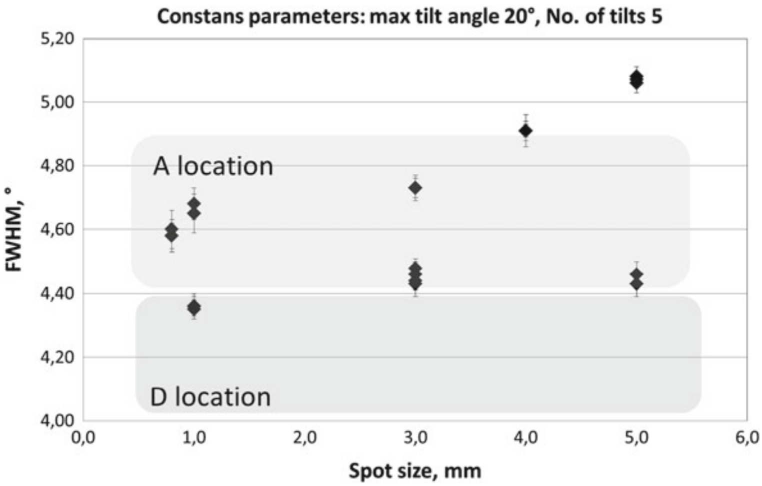


Fig. 9 FWHM data in function of spot size

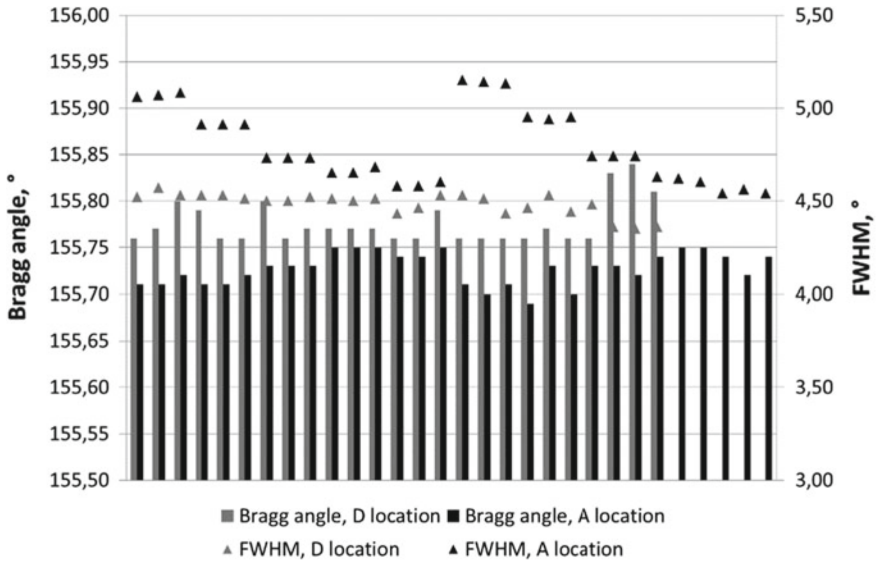


Fig. 10 All the Bragg angle and FWHM data at A and D locations, the spot size, the max tilting angle and the number of tilt are variables. Three parallel measurements were done with a given set of variables

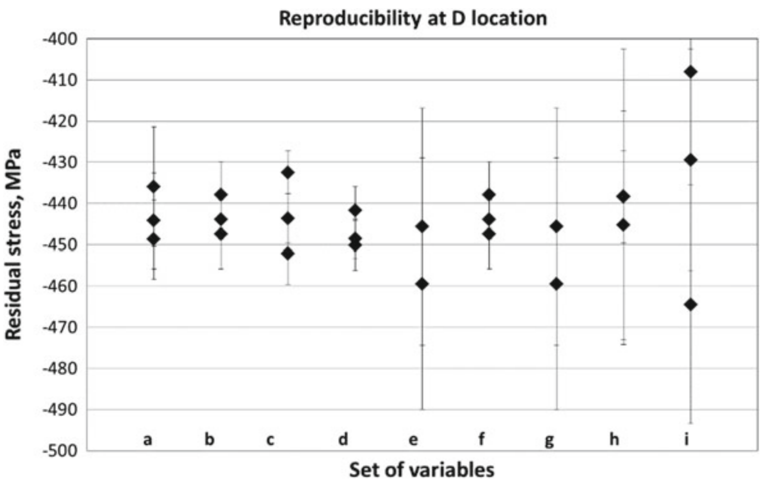


Fig. 11 Reproducibility of stress measurement at the D location in 9 different sets of variables (*a*–*i*). Three parallel measurements were done with a given set of variables

4 Summary

The residual stress was measured non-destructively with a Stresstech Xstress 3000 G3R type centerless X-ray diffractometer on a carburized and shoot peened gear wheel using the modified $\sin^2\psi$ method. The data were recorded at two equivalent parts of the gear in the perpendicular direction to the axis. The normal stress (σ), the shear stress (τ), FWHM and Bragg angle (2θ) were calculated at every measuring time. To affect the applied measured parameters such as the spot size, the number of tilting, and the tilting range were showed on based on the highest number of measured data. The result can be summarized as follows:

- The two locations of the gear contain different stress and other results due to the different solute carbon concentration and it is evident on based of simultaneous investigation of normal stress, FWHM, and θ.
- The increase of number of tilting has slightly effect on stress data. The stress results of 11 tilts number shows only smaller error bar, but the difference is much smaller than the difference of the result from the equivalent parts (location A and D) of the gear. The FWHM data are unaffected by the number of tilting.
- The spot size effect on the stress data deviation is within the error bars, if even 6 times larger spot size was used, but the FWHM data are strongly increased as the spot size is increasing. So the FWHM data can be comparable in case of same spot size measurements only.
- The tilting range has the strongest effect on the stress data. In case of $\pm 20°$ tilting range the error bar can be 10 times larger than in case of $\pm 45°$ which is the usually used value. So, the sample cutting may have a smaller modification effect on the original stress state and gives better results than the decrease of the maximum tilting angle.
- The reproducibility was good enough in any case of a set of variables.

Acknowledgements The described article was carried out as part of the EFOP-3.6.1-16-00011 “Younger and Renewing University—Innovative Knowledge City—institutional development of the University of Miskolc aiming at intelligent specialisation” project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund. The research work of Mate Sepsi was supported through the New National Excellence Program of the Ministry of Human Capacities.

References

1. Handbook of residual stress and deformation of steel. ASM Int 2008, USA pp 347–358
2. Macherauch E, Hauk V (1987) Residual Stress Sci Technol 2:697
3. Soady KA, Mellor BG, Shackleton J, Morris A, Reed PAS (2011) The effect of shot peening on notched low cycle fatigue. Mater Sci Eng, A 528(29):8579–8588
4. Cseh D, Mertinger V, Lukács J (2013) Residual stress evolution during fatigue test of a shoot peened steel sample. Mater Sci Forum 752:95–104

5. Withers PJ, Bhadeshia H (2001) Residual stress Part1—measurement techniques. Mater Sci Technol 17:355–365
6. Schajer GS (2013) Practical residual stress measurement methods. Wiley, pp 140–161
7. In: Proceeding of 1st international conference on barkhausen noise and micromagnetic testing, Sept 1–2, 1998, Hannove–Germany, Organized by Stresstech Group and IFW
8. Krawitz AD (2001) Introduction to diffraction in materials science and engineering. Wiley, pp 119–143, 278–318

Author Index

K. Jármai and B. Bolló (eds.), *Vehicle and Automotive Engineering*,
Lecture Notes in Mechanical Engineering, DOI 10.1007/978-3-319-51189-4