

# FHWS SCIENCE JOURNAL

2014 (Jahrgang 2), Ausgabe 1

## Elektro- und Informationstechnik

Investigation of voltage transformers operating modes in networks with insulated neutral terminal at non-symmetric modes

**A. Popov, Y. Sayenko**

Ein Verfahren zur automatischen Abstimmung von Einseitenband-Empfängern

**L. Dastager, J. Weith**

A Correlation Based Method for Measuring and Monitoring the Impulse Response of Analog LTI Systems with low Realization Cost

**L. Dastager, J. Weith**

## Informatik und Wirtschaftsinformatik

Indoor Localization Using Step and Turn Detection Together with Floor Map Information

**L. Köping, F. Ebner, M. Grzegorzec, F. Deinzer**

Generic Open Source Gamification Framework

**S. Zink**

Modeling gamification rules with ontologies for achieving a generic platform architecture

**P. Mennig**

Automatische Bewertung von Spirogrammen durch Methoden der Mustererkennung

**T. Schulz**

## Maschinenbau und Kunststofftechnik

Untersuchungen zur Heizzeitreduzierung bei der Herstellung von dickwandigen Gummibauteilen

**V. Herrmann, F. Naß, S. Lemmich, F. Wachter**

Analysen von Polyoxymethylenen für Anwendungen in tribologischen Polymer-Polymer-Systemen der Fördertechnik – Teil 1

**R. Illek, M. Hering, F. Pohli, K. Faust**

Verschleißquantifizierungsverfahren für Polymer-Polymer-Gleitpaarungen

**R. Illek, F. Pohli, S. Kellner, K. Faust**

## Wirtschafts- und Sozialwissenschaften

Correlation between Airlines' use of foreign languages and their country of origin's acceptance of foreigners

**J. Dumetz, R. Wehner**





## INHALT

7 Vorwort

### Elektro- und Informationstechnik

- 8 **ELECTRICITY AND ELECTRICAL ENGINEERING**  
Investigation of voltage transformers operating modes in networks with insulated neutral terminal at non-symmetric modes  
Artem Popov, Yuriy Sayenko
- 16 **ELEKTRISCHE NACHRICHTENÜBERTRAGUNG**  
Ein Verfahren zur automatischen Abstimmung von Einseitenband-Empfängern  
Lida Dastager, Jürgen Weith
- 22 **SYSTEM CHARACTERISTICS, DIGITAL MEASUREMENT**  
A Correlation Based Method for Measuring and Monitoring the Impulse Response of Analog LTI Systems with low Realization Cost  
Lida Dastager, Jürgen Weith

### Informatik und Wirtschaftsinformatik

- 40 **INDOOR LOCALIZATION**  
Indoor Localization Using Step and Turn Detection Together with Floor Map Information  
Lukas Köping, Frank Ebner, Marcin Grzegorzec and Frank Deinzer
- 50 **IT-DRIVEN HUMAN RESOURCE MANAGEMENT**  
Generic Open Source Gamification Framework  
Stefan Zink
- 60 **ONTOLOGIES, PREDICATE LOGIC AND SEMANTIC NETWORKS**  
Modeling gamification rules with ontologies for achieving a generic platform architecture  
Patrick Mennig
- 72 **SOFTWARE MUSTERERKENNUNG**  
Automatische Bewertung von Spirogrammen durch Methoden der Mustererkennung  
Tobias Schulz

### Maschinenbau und Kunststofftechnik

- 84 ENERGY AND COST SAVING  
Untersuchungen zur Heizzeitreduzierung bei der Herstellung von dickwandigen Gummibauteilen  
Volker Herrmann, Felix Naß, Sebastian Lemmich, Florian Wachter
- 98 POLYMERTRIBOLOGIE UND -ANALYSE  
Analysen von Polyoxymethylenen für Anwendungen in tribologischen Polymer-Polymer-Systemen der Fördertechnik – Teil 1  
Rene Illek, Mareike Hering, Frank Pohli, Karsten Faust
- 110 Verschleißquantifizierungsverfahren für Polymer-Polymer-Gleitpaarungen  
Rene Illek, Frank Pohli, Stefan Kellner, Karsten Faust

### Wirtschafts- und Sozialwissenschaften

- 120 SOCIAL AND HUMAN SCIENCES  
Correlation between Airlines' use of foreign languages and their country of origin's acceptance of foreigners  
Jerome Dumetz, Rainer Wehner

DER HERAUSGEBER DANKT DEN GUTACHTERN DIESER AUSGABE

Prof. Dr. Bernhard Arndt	Hochschule Würzburg-Schweinfurt
Prof. Dr. Arndt Balzer	Hochschule Würzburg-Schweinfurt
Prof. Dr. Peter Braun	Hochschule Würzburg-Schweinfurt
Prof. Dr. Frank Deinzer	Hochschule Würzburg-Schweinfurt
Prof. Eberhard Grötsch	Hochschule Würzburg-Schweinfurt
Prof. Dr. Karsten Huffstadt	Hochschule Würzburg-Schweinfurt
Prof. Dr. Isabel John	Hochschule Würzburg-Schweinfurt
Prof. Dr. Roger Weinlein	Hochschule Darmstadt

## VORWORT

Nachdem die beiden ersten Ausgaben des neuen FHWS Science Journals jeweils als Sonderausgaben ein relativ enges Gebiet der Forschung an der FHWS behandelten – die erste Ausgabe 2013 des FHWS Sciences Journal enthielt Beiträge des Kongresses „Mobility moves Business“ in Würzburg, der sich mit Verbesserung mobiler Softwaresysteme befasste und daher überwiegend den Bereich Informatik und Wirtschaftsinformatik abdeckt und die zweite Ausgabe des Jahrgangs 2013 mit dem Themenbereich Energiewende, Energieeffizienz und Elektromobilität und damit überwiegend den Bereich Elektrotechnik und Maschinenbau behandelt – wird in der aktuell vorliegenden Ausgabe des FHWS Science Journals, Ausgabe 1, Jahrgang 2014, ein sehr viel breiteres Spektrum an wissenschaftlichen Gebieten der Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt abgedeckt.



*Vizepräsident  
Prof. Dr. Jürgen Hartmann*

Dieses breite Spektrum spiegelt – wenn auch nicht vollständig – die große Bandbreite der an der FHWS durchgeführten aktuellen Forschungsarbeiten wieder. Beginnend mit Verfahren und Methoden in der Elektrotechnik im Bereich Mobilfunk, Hochspannungstechnik und Signalverarbeitung über Quantifizierungsverfahren für Verschleißerscheinungen sowie Analysen bei Polymer-Systemen im Bereich Maschinenbau, hin zu den Untersuchungen zur energieeffizienten Herstellung von Gummibauteilen in der Kunststofftechnik wird die klassische ingenieurwissenschaftliche Expertise der FHWS deutlich. Dieser Teil wird eindrucksvoll ergänzt durch Beiträge zur Entwicklung von Plattformen und Software-Architekturen zur Gamifizierung von Anwendungen und Programmen, Indoor-Lokalisation ohne absolute Positionierungssysteme sowie automatisierte Bewertung von Spirogrammen aus dem Bereich Informatik und Wirtschaftsinformatik. Abgerundet wird die vorliegende Ausgabe durch einem Beitrag aus dem Bereich Wirtschaftswissenschaften zur Korrelation von Verwendung von Fremdsprachen und Akzeptanz ausländischer Passagiere bei internationalen Fluggesellschaften.

Auch wenn die vorliegende Ausgabe des FHWS Science Journals bereits einen Eindruck vermittelt, in welchen breiten Bereichen an der FHWS geforscht wird, deckt diese Ausgabe bei Weitem noch nicht das gesamte Forschungsspektrum der FHWS ab.

Ich wünsche Ihnen viel Freude und viele neue Informationen beim Durchlesen der vorliegenden Ausgabe und freue mich mit Ihnen bereits auf die nächsten Ausgaben des FHWS Science Journals mit neuen Forschungs-Highlights aus weiteren Bereichen der FHWS.

Prof. Dr. Jürgen Hartmann

Vizepräsident der Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt

## Investigation of voltage transformers operating modes in networks with insulated neutral terminal at non-symmetric modes

Artem Popov and Yuriy Sayenko  
Pryazovskyi State Technical University, Mariupol City, Ukraine

Bernhard Arndt  
University of Applied Sciences Würzburg-Schweinfurt

Ferroresonance and processes at one-phase grounding inside networks with insulated neutral terminal are investigated. It is shown that the mode with unidirectional arc, at which direct current passed in high voltage windings is the most difficult operating mode for the transformer. The influence of the secondary voltage transformer load upon the stability of ferroresonance oscillations has been also investigated.

Categories and Subject: Electricity and Electrical Engineering

Additional Key Words and Phrases: Ferroresonance, Magnetization Curve, One-phase Arc Grounding, Voltage Transformer, Zero-sequence Capacity

### 1 INTRODUCTION

Long-term exploitation of power networks with insulated neutral termination proved that the weakest point of such networks is the voltage transformer for insulation control. The data show that in the Ukraine 7-10% of all installed voltage transformers (VT) [1] get damaged every year. So, investigation of operating modes of VT seems to be of vital importance, as well as development of measures, aimed at raising their reliability.

Damage of VT is often accompanied with non-symmetric mode, particularly with one-phase arc grounding or ferroresonance. It results from substantial current flow in VT high-voltage winding, causing its overheating and thermal destruction. Both cases are to be discussed below.

### 2 MODELLING AND ANALYTICAL APPROACHES TO TRANSIENTS STUDIES

The method of imitational simulation, realized in MATLAB program was applied in this article for investigation of no-symmetric modes. The simplified equivalent circuit of the network under investigation is represented in Fig. 1, VT equivalent circuit, applied for simulation is

Author's address:

Yuriy Sayenko<sup>1</sup>, Artem Popov<sup>1</sup>, Bernhard Arndt<sup>2</sup>.

<sup>1</sup>Pryazovskyi State Technical University, 87500 Universytets'ka 7 str, Mariupol City, Ukraine

<sup>2</sup>University of Applied Sciences Würzburg-Schweinfurt

Email: yls@mail.ru, artempopov9@gmail.com, bernhard.arndt@fhws.de

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than FHWS must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission.

represented in Fig. 2, while Fig. 3 represents magnetization characteristics of VT of NTMI-6 type. Hysterisis-free VT model was used in our investigation.

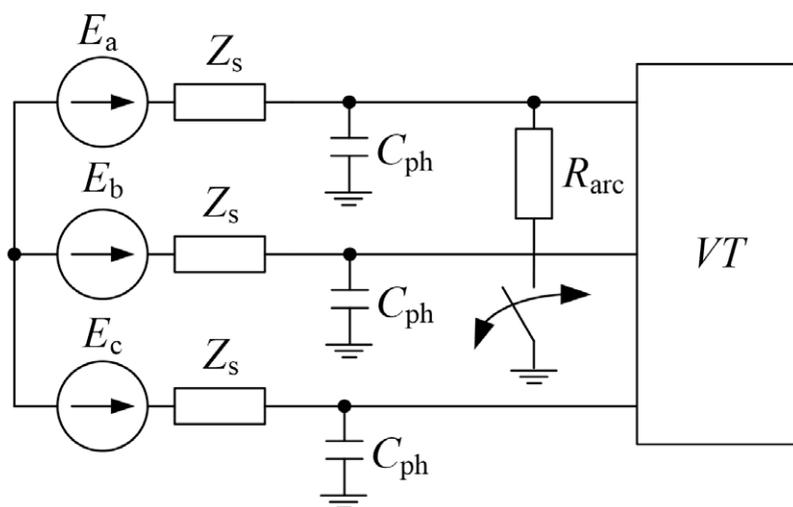


Fig. 1. Analytical model of substitution of the arc under investigation

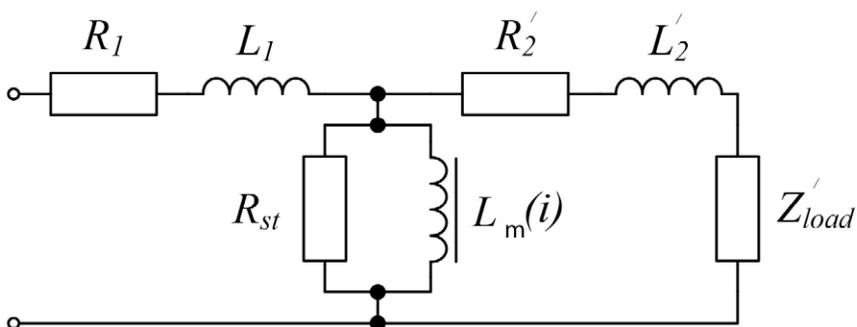


Fig. 2. The model of VT substitution

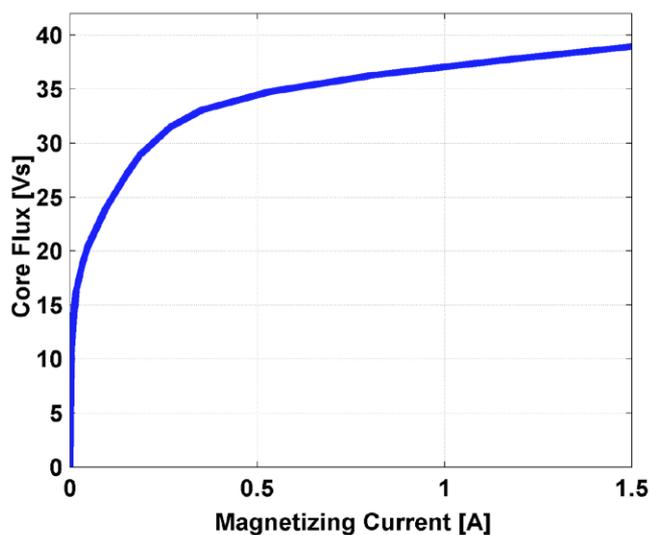


Fig. 3. Experimental the magnetization curve VT characteristics

The mode of one-phase metal shorting is not dangerous for VT. The voltage of not damaged phases is increased up to the value of linear voltage, the currents inside windings are but slightly raised, not exceeding the tolerable values. Much more serious consequences for VT happen to be in case of unstable shorting, i.e. at alternating arc. In this case overvoltages occur, the value of which could reach 2.6-3.2. p.u. [2], accompanied with VT core's saturation and the current inside high-voltage windings increases. Prolonged existence of this mode leads to thermal damage of high voltage winding. In Fig. 4 the curves, obtained by modelling of phase voltages and currents in the VT high voltage winding at one-phase arc shorting are summarized. The arc behavior is based on Petersen's theory, according to which the arc is started at the moment of the highest voltage in the damaged phase and is subsided at the moment the high-frequency component passes through zero. The current inside high voltage winding exceeds very much the long-term permissible value (0.3 A) and it inevitably leads to the damage of VT.

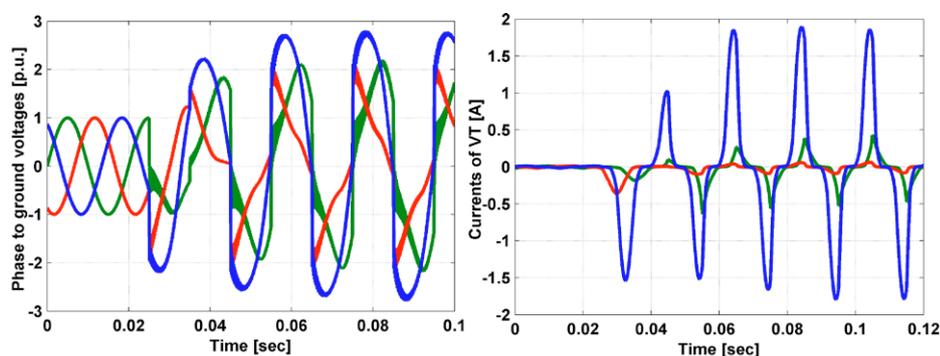


Fig. 4. Designed curves of phase voltages and currents in high-voltage VT windings at symmetrical arc shorting

Obviously, the value of currents in the winding becomes higher when repetition factor of overvoltages is raised. In the example discussed, the biggest value of repetition factor of overvoltages reached 3 p.u., this value very much exceeding the values, obtained by registering transitional processes in actual power networks. The main error at simulation is caused by overvoltage of arc-through, as at realization of Petersen's theory actual physical processes in insulation are not taken into account, that is why arc-through voltage could reach  $1.5 U_{pb}$  or higher, whilst the actual value of arc-through voltage fluctuates within  $0.6-1.0 U_{pb}$  range. The results of modeling showed that that if overvoltage reached 2.6 p.u. VT could operate for quite a long time without any thermal damage of high voltage winding.

The cases described above apply to symmetrical arc only, i.e. when the arc is struck at uniform voltage on positive and negative half-wave. Actual process of the arc flare can be non-symmetrical, resulting to different values of arc-through voltage on positive and negative half-waves. Unidirectional arc, which is ignited and subsides once throughout the period, is the ultimate example of non-symmetrical arc. Fig. 5 summarizes the appropriate designed curves of phase currents and voltages in high-voltage winding. Here, the currents are approximately equal and possess unidirectional character. It follows from oscillation diagrams of phase voltages that the area of positive and negative voltage half-waves differs. Hence, a constant component must be present in phase voltages.

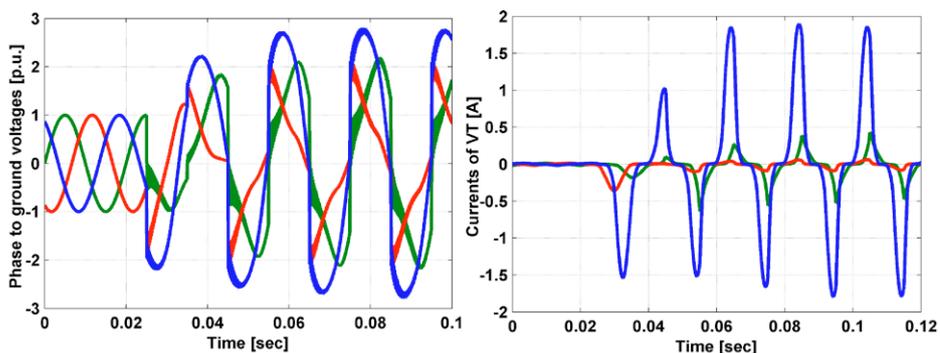


Fig. 5. Designed curves of phase voltages and currents in high-voltage VT windings at unidirectional arc

Fig. 6. summarizes amplitude-phase spectrum of *B*-phase voltage, from which it follows that the value of the constant component is equal to 650 V. That is why in this case magnetization of VT core with direct current happens, determined by the value of constant voltage and active resistance of high voltage winding, which for this VT is equal to 1200 Ohms.

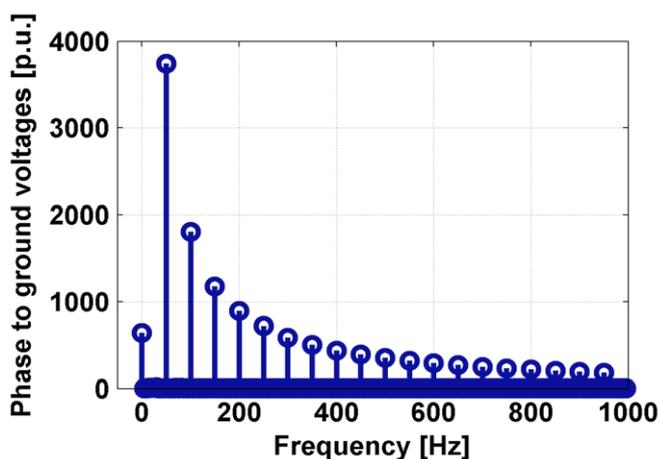


Fig. 6. Amplitude-spectrum of phase voltage at unidirectional arc

The mode of bidirectional non-symmetrical arc is less difficult but also dangerous. In this situation a constant component is also present in spectrum of phase voltages and VT is magnetized with direct current. However, the value of this constant component is smaller than in case of unidirectional arc. Investigations revealed that value of the constant component is determined by the relation of arc-through voltage on the positive ( $U_+$ ) and negative ( $U_-$ ) half-waves. The obtained dependence of the constant component  $U_{const}$  upon the relation of  $U_+/U_-$  is shown in Fig. 7. It should be noted here that experience of making oscillation diagrams, gained by the authors for many a year [2]-[4] showed that appearance of a non-symmetrical arc is an event, which is likely to happen, especially during the initial shorting stage.

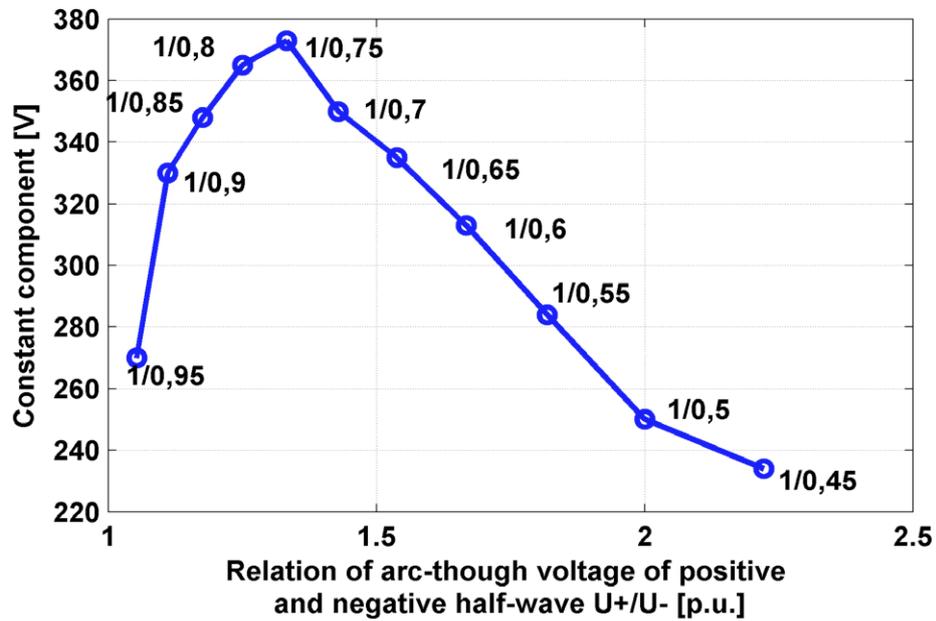


Fig. 7. The dependence of the DC component on the ratio of the breakdown voltage of the positive and negative half-wave

Ferroresonance mode is also dangerous for VT [5]-[7]. The biggest danger for VT is subharmonic ferro-resonance, which often occurs after one-phase shorting disappears. Low frequency component appears in phase voltages spectrum at subharmonic ferroresonance, which over-saturates VT rod, resulting in substantial currents passing through high voltage winding, leading to its thermal damage. An example of subharmonic ferroresonance is given in Fig. 8, where curves of phase voltages and currents in high voltage winding are shown. Phase grounding capacity ( $C_{ph}$ ) is equal to 0.6  $\mu$ F.

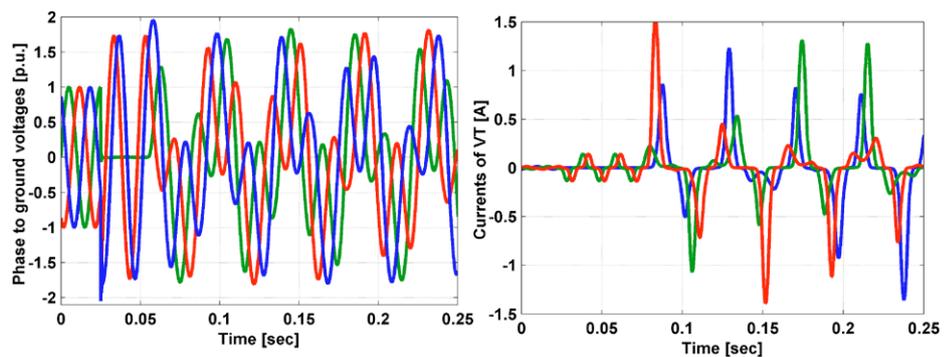


Fig. 8. Designed curves of phase voltages and currents in VT primary winding at subharmonic ferroresonance, which sprang up when grounding shorting was eliminated

Ferroresonance processes, developing at frequencies 50 Hz, or higher do not lead, as a rule, to thermal damage of VT. An example of ferroresonance, which originated when unloaded buses on VT were switched on, is given in Fig. 9, where phase voltages and currents in high VT voltage winding are shown. In comparison to subharmonic ferroresonance, ferroresonance processes at 50 Hz frequencies or higher are accompanied with bigger overvoltages, which could be dangerous of network insulation.

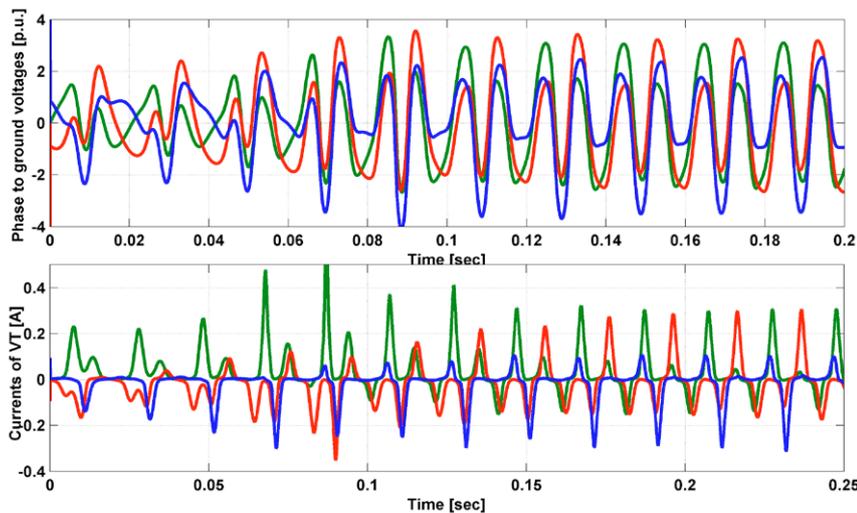


Fig. 9. Designed curves of phase voltages and currents in VT primary winding at ferroresonance, which sprang up when unloaded buses were switched on.  $C_{pb} = 0.012 \mu\text{F}$

When ferroresonance processes were modeled, the influence of the load of secondary VT winding upon the character of ferroresonance processes was investigated. The overall load of the secondary winding changed within 0-150 VA range at  $\cos \varphi = 0.8$  i.e. within the range, where VT operates in class 1 of precision. The results of investigation showed that increase of the load of the secondary winding diminished the probability of ferroresonance appearance, this being true for all types of ferroresonance. As an example, designed currents curves in the primary winding in case with unloaded VT are depicted in Fig. 10 (see Fig. 10 a) as well as in case with the load 150 VA (see Fig. 10 b). Phase grounding capacity is equal to  $0.6 \mu\text{F}$ .

It could be seen from curves, that, in case of loaded VT, ferroresonance fades away, i.e. the load damps ferroresonance. The damping efficiency gets higher alongside with the raise of the load in the secondary circuit. The obtained data are in conformity with the data obtained by exploitation [5], according to which the damaging rate of VT is inversely proportional to the load of the secondary winding. The Table 1 summarizes the dependence of failure flow of VT upon the load of the secondary winding.

Load of the secondary winding VT, VA	Up to 25	25 – 50	50 – 75	75 – 100	Over 100
Parameter of failure flow , 1/year	0.208	0.1383	0.1097	0.114	0.0513

Table 1: Dependence of VT failure rate upon the load of the secondary winding

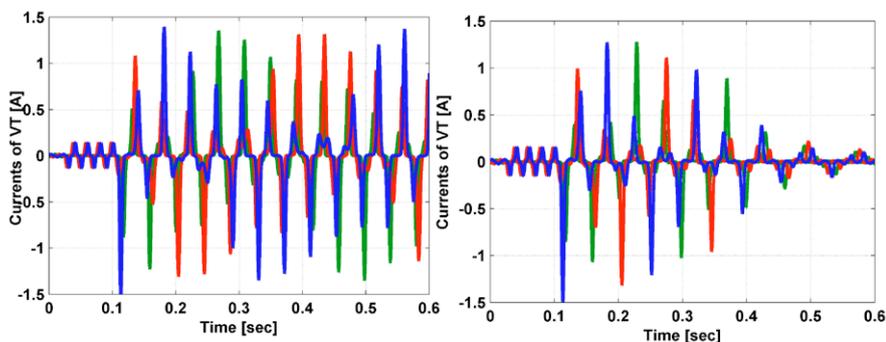


Fig. 10. Designed curves of currents in the high-voltage winding at unloaded VT (a) and when the load is 100 VA (b)

The calculations of ferroresonance, with unloaded buses on VT switched on, have also been made, VT's secondary load being 75 VA, phase grounding capacity being 0.015  $\mu\text{F}$ . The designed curves of phase voltages are summarized in Fig 11. Comparing them to the results, shown in Fig. 9 we may conclude, that amplitude of overvoltages has grown considerably smaller and the process of ferroresonance is fading.

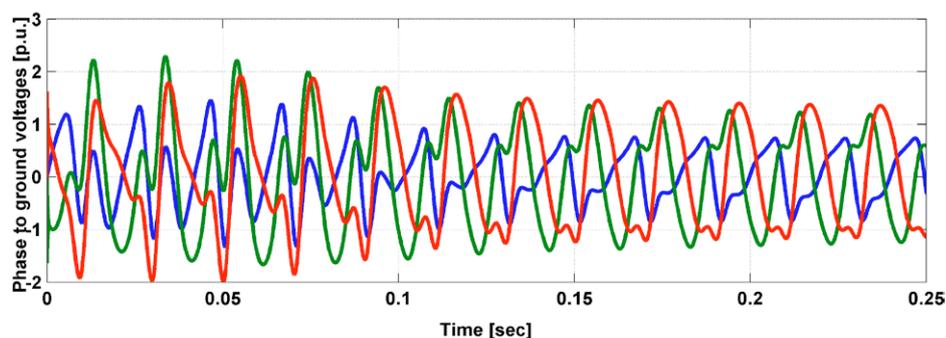


Fig. 11. Designed curves of phase to ground voltages, when VT with load on the secondary winding equal to 75VA upon no-load buses was switched on.  
 $C_{pb}=0.012 \mu\text{F}$

### 3 CONCLUSION

1. Damaging of VT can be caused by prolonged existence of arc shortings. The most difficult case is the mode of unidirectional arc, where rod becomes magnetized by direct current. In order to prevent VT damaging, measures should be specified for limitation of arc overvoltages, or limitation of time of intermittent arc existence.
2. The probability of ferroresonance appearance is inversely proportional to the load of VT secondary winding. So, we recommend to increase the load of the secondary winding up to maximal permissible value, according to the specified precision class, in order to diminish the risk of ferroresonance appearance.

## REFERENCES

- [1] ZHURAKHIVSKIY, A. V., YATSENKO, A. Y., AND MASLYAK, R. Y. 2009. Operating modes of voltage transformers inside power networks with insulated terminal. *Electroinform*, No. 1, pp. 8-11.
- [2] LIKHACHEV, F. A. 1971. Grounding shortings in networks with insulated neutral and compensation of capacity. Moscow: Power, p. 152.
- [3] KACHESOV, V. E., LARIONOV, V. N., AND OVSYANNIKOV, A. G. 2002. On the results of overvoltages monitoring at one-phase arc grounding shortings inside distribution cable networks. *Electric stations*, No 8, pp. 38-42.
- [4] SHIRKOVETS, A. I., SARIN, L. I., ILYNYKH, M. V. AND BUYANOV, E. V. 2008. Investigation of transitional processes at one-phase grounding shorting network of Thermal power station of iron and steel works with combined neutral grounding. *Power engineering and electrification*, No 8, pp. 39-47.
- [5] GANUS, A. I., GOVOROV, F. P. AND STARKOV, K. A. 2006. Influence of conditions of transitional processes in 6-10 kV electric power networks upon the character of damaging of voltage transformers. *Electrica*, No. 6, pp. 28-31.
- [6] JACOBSON, D. A. N. 2003. Examples of ferroresonance in high voltage power system. IEEE Power Engineering Society General Meeting Conf., pp. 1206-1212.
- [7] OSBORNE, M., JARMAN P., CHARALAMBOUS C. AND WANG Z. 2005. Study on Simulation of Ferroresonance Elimination in 10kV Power System. IEEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific Conf., pp. 1-7.

## Ein Verfahren zur automatischen Abstimmung von Einseitenband-Empfängern

Lida Dastager, Jürgen Weith  
Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt

Es wird ein Verfahren beschrieben, mit dessen Hilfe eine Korrektur der Demodulation von Einseitenband-Signalen (ESB-Signalen) vorgenommen werden kann. Eine solche Korrektur ist dann oft notwendig, wenn sich die Frequenzen, die zur Modulation und zur Demodulation verwendet worden sind, unterscheiden. Typischerweise tritt ein solcher Frequenzunterschied bei der Übertragung von Sprachsignalen auf, wenn kein festes Kanalaraster zugrunde liegt. Das Korrekturverfahren basiert auf einer Analyse stimmhafter Sprachabschnitte mit dem MUSIC-Verfahren.

Kategorie und Themenbeschreibung: Elektrische Nachrichtenübertragung

Zusätzliche Schlüsselwörter: Spektralanalyse mit Subspace-Verfahren, Super-Resolution-Analyse, MUSIC-Verfahren, Sprachqualität

### 1 EINFÜHRUNG

Zur Übertragung von Sprachsignalen über hochfrequente Kanäle wird in bestimmten Frequenzbereichen oft die Einseitenband-Modulation angewandt. Dabei wird – wie der Name sagt – lediglich ein Seitenband übertragen und das andere unterdrückt. Zusätzlich wird meist die Trägerschwingung ausgesiebt, um die Sendeleistung zu reduzieren. Man spricht dann von einem SSBSC-Signal (SSBSC = Single Side Band Subpressed Carrier).

Ein solches SSBSC-Signal besitzt aus übertragungstechnischer Sicht verschiedene Vorteile wie z. B.

- einen geringen hochfrequenten Bandbreitenbedarf (die hochfrequente Bandbreite ist hier gleich der niederfrequenten Signal-Bandbreite),
- eine hohe spektrale Effizienz,
- lineare Verzerrungen auf dem Übertragungsweg führen lediglich zu linearen Verzerrungen im demodulierten Signal (andere Modulationsverfahren erzeugen nichtlineare Verzerrungen, welche die Signalqualität gravierend verschlechtern können),
- ein SSBSC-Signal kann mit Hilfe eines digitalen Signalprozessors (DSP) in einfacher Weise erzeugt werden. Zu realisieren ist dabei primär ein System, das die Hilbert-Transformation berechnet.

Adresse des Autors:

Lida Dastager, Jürgen Weith, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Ignaz-Schön-Str. 11, 97421 Schweinfurt

Lida.Dastager@gmx.de, juergen.weith@fhws.de

Die Erlaubnis zur Kopie in digitaler Form oder Papierform eines Teils oder aller dieser Arbeit für persönlichen oder pädagogischen Gebrauch wird ohne Gebühr zur Verfügung gestellt. Voraussetzung ist, dass die Kopien nicht zum Profit oder kommerziellen Vorteil gemacht werden und diese Mitteilung auf der ersten Seite oder dem Einstiegsbild als vollständiges Zitat erscheint. Copyrights für Komponenten dieser Arbeit durch Andere als FHWS müssen beachtet werden. Die Wiederverwendung unter Namensnennung ist gestattet. Es andererseits zu kopieren, zu veröffentlichen, auf anderen Servern zu verteilen oder eine Komponente dieser Arbeit in anderen Werken zu verwenden, bedarf der vorherigen ausdrücklichen Erlaubnis.

Der SSBSC-Modulation im Sender, bei der die Trägerfrequenz  $f_M$  eingesetzt wird, schließt sich nach der Übertragung über den Kanal im Empfänger die SSBSC-Demodulation mit der Frequenz  $f_D$  an. Für eine fehlerfreie Übertragung müssen dabei Sender und Empfänger die gleiche Frequenz verwenden, es muss also  $f_D = f_M$  sein.

Die Einhaltung der Bedingung  $f_D = f_M$  setzt voraus, dass sich Sender und Empfänger über die zu verwendende Frequenz verständigt haben. Das ist aber aus praktischen Gründen oft nicht möglich. Zum Auffinden der Frequenz  $f_D$  wird daher die demodulierende Frequenz solange verändert, bis das zurückgewonnene Sprachsignal „möglichst natürlich“ klingt.

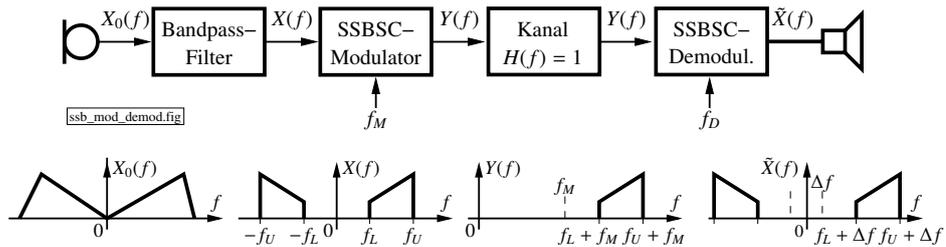


Abbildung 1: Skizze einer SSBSC-Übertragung zusammen mit den auftretenden Spektren der Teilsignale.

Ein solches Vorgehen bringt natürlich Nachteile mit sich. Zunächst kann eine Abstimmung „nach Klang“ erfahrungsgemäß recht zeitraubend sein, weil immer wieder kleine Frequenzänderungen vorgenommen werden müssen, um den natürlichsten Klangeindruck aufzufinden. Darüber hinaus ist eine derartige Frequenzwahl subjektiv gefärbt, unterschiedliche Personen werden i. a. auch unterschiedliche Demodulationsfrequenzen bevorzugen.

Es sind deshalb Verfahren interessant, welche aus dem demodulierten niederfrequenten Signal die Größe  $\Delta f = f_M - f_D$  zu bestimmen vermögen. Ist  $\Delta f$  bekannt, dann kann im Niederfrequenzbereich eine Korrektur des Sprachsignals erfolgen.

Ein Vorschlag zur Offset-Korrektur findet sich z. B. in [5]. Das nachfolgend beschriebene Verfahren hat (im Gegensatz zum Verfahren in der genannten Literatur) zum Ziel, eine Implementierung auf einem DSP, z. B. ausgehend von einem SIMULINK/MATLAB-Modell zu ermöglichen.

## 2 ESB-MODULATION UND DEMODULATION

Zu übertragen sei ein reelles niederfrequentes Signal  $x(t)$ , das durch Bandpass-Filterung aus einem Signal  $x_0(t)$  entstanden ist. Für das Spektrum  $X(f)$  gelte  $|X(f)| \equiv 0$  für  $|f| > f_U$ . Aus  $x(t)$  entsteht das Einseitenband-Signal  $y(t)$  gemäß

$$y(t) = \text{Re}\{[x(t) \pm j\hat{x}(t)] e^{j2\pi f_M t}\} \quad (1)$$

$$= x(t) \cos(2\pi f_M t) \mp \hat{x}(t) \sin(2\pi f_M t) \quad (2)$$

$\hat{x}(t)$  bedeutet die Hilbert-Transformierte der Zeitfunktion  $x(t)$ . Wird das positive Vorzeichen in Glg. (1) verwendet, so liegt ein USB-Signal (USB = Upper Side Band) vor und das obere Seitenband wird übertragen, das untere Seitenband dagegen unterdrückt. Im anderen Fall wird ein LSB-Signal (LSB=Lower Side Band) generiert. Die Gleichungen beschreiben die spektrale Verschiebung des in eckigen Klammern stehenden analytischen Signals um die Modulationsfrequenz  $f_M$ .

Die Demodulation erfolgt durch spektrale Rückverschiebung. Entsprechend erhält man das demodulierte Signal als

$$\tilde{x}(t) = 2 [y(t) \cdot \cos(2\pi f_D t)] * h_{TP}(t). \quad (3)$$

$h_{TP}(t)$  bezeichnet dabei die Impulsantwort eines Tiefpasses mit der Übertragungsfunktion

$$H_{TP}(f) = \begin{cases} 1 & \text{für } |f| \leq f_U + f_M - f_D \\ 0 & \text{für } |f| \geq f_L + f_M + f_D. \end{cases} \quad (4)$$

Durch dieses dem Multiplizierer (technisch: Mischer) nachzuschaltende Tiefpass-Filter werden solche Frequenzanteile unterdrückt, welche im Bereich  $f_L + f_M + f_D \leq f \leq f_U + f_M + f_D$  liegen<sup>1</sup>. Aus den Gleichungen (1) und (3) ergibt sich schließlich für das demodulierte niederfrequente Signal

$$\tilde{x}(t) = \text{Re}\{[x(t) + j\hat{x}(t)] e^{j2\pi\Delta f t}\} \quad (5)$$

$$= x(t) \cos(2\pi\Delta f t) - \hat{x}(t) \sin(2\pi\Delta f t). \quad (6)$$

Ein Diagramm des Modulations- und Demodulationsvorgangs zusammen mit den auftretenden Spektren ist in Abb. 1 gezeichnet. Für den Kanal wurde die Übertragungsfunktion  $H(f) = 1$  („Durchverbindung“) angenommen.

### 3 KORREKTUR DES FREQUENZ-OFFSETS

Das mit unterschiedlichen Frequenzen  $f_M$  bzw.  $f_D$  modulierte bzw. demodulierte SSBSC-Signal soll nun korrigiert werden. Eine derartige Korrektur wird dadurch erreicht, dass das zum Empfangssignal gehörende analytische Signal  $\hat{x}(t) + j\hat{\hat{x}}(t)$  bestimmt und anschließend um  $f = -\Delta f$  spektral verschoben wird. Abschließend  $\tilde{x}(t) + j\hat{\hat{x}}(t)$  ist noch der Realteil zu bilden. Es gilt dann für die Vorschrift der Frequenzoffset-Korrektur:

$$x(t) = \text{Re}\{[\tilde{x}(t) + j\hat{\hat{x}}(t)] e^{-j2\pi\Delta f t}\}. \quad (7)$$

Damit liegt wieder das im Sender durch Bandpass-Filterung gebildete, dem SSBSC-Modulator zugeführte Signal  $x(t)$  vor.

Auf die Bestimmung Wertes  $\Delta f$  wird nachfolgend näher eingegangen.

### 4 SCHÄTZUNG DES FREQUENZOFFSETS

Bei den meisten Übertragungssystemen, die SSBSC-Modulation einsetzen, sind die Frequenzen  $f_L$  und  $f_U$  nicht oder nicht genau genug bekannt. So werden bei der Sprachübertragung oft  $f_L = 330$  Hz und  $f_H = 3,4$  kHz verwendet. Ob diese Frequenzen aber bei einem aktuell vorliegenden Signal tatsächlich gültig sind, ist nicht bekannt. Eine Bestimmung des Frequenz-Offsets aus diesen Größen scheidet deshalb meist aus.

Eine Bestimmung von  $\Delta f$  ist jedoch auf anderem Weg möglich. Dazu wird ein Ausschnitt (Rahmen, Frame) aus dem Sprachsignal  $x_o(t)$  betrachtet, der stimmhaften Charakter hat. Ein solcher stimmhafter Abschnitt wird durch Stimmbandschwingungen erzeugt [1]. Es gilt in diesem Fall

<sup>1</sup>Auf den Tiefpass kann verzichtet werden, wenn alternativ das zum Signal  $y(t)$  analytische Signal gebildet und dieses spektral verschoben wird. Hier liegt der Aufwand im System zur Bildung der Hilbert-Transformierten.

<sup>2</sup>Unterstellt wird ein ideales Bandpass-Filter, das im Übertragungsband weder Amplitudenfehler noch Phasenänderungen bewirkt.

$$x_0(t) = \sum_{\mu=1}^n a_{\mu} e^{j\mu 2\pi f_p t}, \quad a_{\mu} \in \mathbb{C}, \quad (8)$$

wobei  $f_p$  die sog. Sprachgrundfrequenz („Pitch-Frequenz“) darstellt. Nach Bandpass-Filterung<sup>2</sup> ergibt sich

$$x(t) = \sum_{\mu=n_1}^{n_2} a_{\mu} e^{j\mu 2\pi f_p t}, \quad a_{\mu} \in \mathbb{C} \quad (9)$$

und am Ausgang des Demodulators erscheint schließlich das Signal

$$\tilde{x}(t) = \sum_{\mu=n_1}^{n_2} a_{\mu} e^{j2\pi(\mu f_p + \Delta f)t}, \quad a_{\mu} \in \mathbb{C}. \quad (10)$$

Eine Analyse der in  $\tilde{x}(t)$  vorkommenden Frequenzen  $f_{\alpha}$  liefert

$$f_{\alpha} = (\alpha - 1 + n_1)f_p + \Delta f, \quad \alpha = 1(1)n_2 - n_1 + 1. \quad (11)$$

Damit lässt sich zunächst aus den Frequenzen  $f_{\alpha}$  die Grundfrequenz  $f_p$  bestimmen (z. B. durch Differenzbildung benachbarter Frequenzen und Ausschluss von Mehrfachen der Grundfrequenz). Dann gilt für den gesuchten Frequenz-Offset

$$\Delta f = f_1 - \lambda f_p, \quad \lambda \in \mathbb{N} \quad (12)$$

mit einer ganzzahligen, positiven und noch unbekanntem Zahl  $\lambda$ . Es besteht also zunächst eine Mehrdeutigkeit bei der Bestimmung von  $\Delta f$ .

Diese Mehrdeutigkeit lässt sich auflösen, wenn weitere  $q - 1$  Bestimmungen der Frequenzkomponenten vorgenommen werden, wobei sich aber die Grundfrequenzen voneinander unterscheiden müssen. Man hat jetzt

$$\Delta f = f_1^{(v)} - \lambda^{(v)} f_p^{(v)}, \quad v = 1(1)q, \quad \lambda \in \mathbb{N} \quad (13)$$

( $v$  bedeutet die Nummer der Messung). Der gesuchte Offset ergibt sich daraus als

$$\Delta f = \arg \min_f \sum_{v=1}^q \left| \min_{\lambda^{(v)} \in \mathbb{N}} \{f - f_1^{(v)} + \lambda^{(v)} f_p^{(v)}\} \right|. \quad (14)$$

Damit ist der gesuchte Frequenz-Offset bekannt und es kann eine entsprechende Korrektur erfolgen.

## 5 BESTIMMUNG DER FREQUENZEN DER SPEKTRALKOMPONENTEN

Zur Bestimmung von  $\Delta f$  gemäß Glg. (14) werden die Werte  $f_1$  und  $f_p$  benötigt.  $f_1$  kann als Differenz zweier unmittelbar benachbarter Frequenzen aus Glg. (11) gewonnen werden:

$$f_p^{(v)} = f_{\alpha_0+1}^{(v)} - f_{\alpha_0}^{(v)}. \quad (15)$$

Bei gestörten Signalen kann auch an eine Mittelung benachbarter Frequenzen gedacht werden ( $\alpha_0 = 1, 2, \dots$ ).

Da sich die Frequenz  $f_p$  von Messung zu Messung (über mehrere Rahmen hinweg) zum Teil nur wenig ändert, ist eine möglichst genaue Bestimmung der Frequenzen  $f$  der stimmhaften Sprachabschnitte notwendig. Verschiedene Versuche führten zum Ergebnis, dass hierzu die Anwendung des MUSIC-Verfahrens<sup>3</sup> am besten geeignet ist.

<sup>3</sup>MUSIC = Multiple Signal Classification

Beim MUSIC-Verfahren handelt es sich um eine parametrische Methode zur Bestimmung der Frequenzen eines sich aus harmonischen Komponenten zusammensetzenden Signals  $y(t)$ . Details dazu finden sich in [2]. Verwendet werden dazu die Abtastwerte  $y[k] = y(kT_A)$ . Die Anzahl der Teilkomponenten muss dabei bekannt sein. Eine Beschreibung des Verfahrens findet sich z. B. in [3]. Das beim MUSIC-Verfahren unterstellte Signalmodell hat die Form

$$y[k] = \sum_{\mu=1}^N b_{\mu} e^{j\Omega_{\mu}k} + n[k], \quad (16)$$

wobei die  $b_{\mu}$  komplexe Koeffizienten und  $n[k]$  eine Störung mit konstanter Leistungsdichte, gaußförmiger Verteilungsdichtefunktion und der Varianz  $\sigma^2$  darstellen. Aus der Folge  $y[k]$  der Abtastwerte wird die Kovarianzmatrix  $\mathbf{R} = E\{yy^H\}$  mit  $\mathbf{y} = [y[0] y[1] \dots y[M-1]]^T$  gewonnen. Die hermitesche Matrix  $\mathbf{R}$  besitzt  $M$  Eigenwerte, die positiv sind und von denen  $M - N$  den Wert  $\sigma^2$  aufweisen.  $N$  Eigenwerte sind größer als  $\sigma^2$ . Man bestimmt nun die  $M - N$  zu den Eigenwerten  $\sigma^2$  gehörenden Eigenvektoren  $\mathbf{u}_{N+1}, \mathbf{u}_{N+2}, \dots, \mathbf{u}^M$  und fasst diese Spaltenvektoren in der Matrix  $\mathbf{U}$  zusammen. Mit dem Vektor  $\mathbf{a}(\Omega) = [1 \ e^{j\Omega} \ e^{j2\Omega} \ \dots \ e^{j(M-1)\Omega}]^T$  gilt dann die Aussage: *Die gesuchten  $N$  Frequenzen sind Lösungen der Gleichung*

$$\mathbf{a}^H(\Omega)\mathbf{U}\mathbf{U}^H\mathbf{a}(\Omega) = 0. \quad (17)$$

In der Praxis wird man die  $N$  kleinsten lokalen Minima dieser Funktion suchen, indem man die Frequenz  $\Omega$  in kleinen Schritten von Null an stetig vergrößert. Diejenigen Frequenzen, bei denen  $\mathbf{a}^H(\Omega)\mathbf{U}\mathbf{U}^H\mathbf{a}(\Omega)$  jeweils den Wert Null erreicht, sind dann die gesuchten Werte  $\Omega_{\mu}$ .

Die Bestimmung der Frequenzen  $\Omega_{\mu}$  kann manchmal einfacher durch Anwendung des sogenannten Root-MUSIC-Verfahrens erfolgen. Dazu definiert man zunächst die Vektoren  $\mathbf{z}^1 = [1 \ z \ z^2 \ \dots \ z^{M-1}]^T$  sowie  $\mathbf{z}^2 = [1 \ z^{-1} \ z^{-2} \ \dots \ z^{-(M-1)}]^T$  und setzt anschließend beide in die Glg. (17) ein. Dann stellt  $\mathbf{z}^{M-1} \cdot \mathbf{z}_1^T \mathbf{U}\mathbf{U}^H \mathbf{z}_2$  ein Polynom  $P(z)$  vom Grad  $2M-2$  dar. Von dessen  $2M-2$  Nullstellen  $z_{0,\lambda}$  liegen (im Idealfall)  $2N$  Nullstellen auf dem Einheitskreis (für sie gilt  $z_{0,\lambda} = e^{j\Omega_{\mu}}$ ). Diese Nullstellen – und damit die interessierenden Frequenzen – müssen nun durch ein geeignetes Suchprogramm gefunden werden.

## 6 ANMERKUNG ZUR PRAKTISCHEN AUSLEGUNG

Das hier beschriebene Korrektur-Verfahren wurde in Form einer Simulation mit Hilfe der Software OCTAVE untersucht. Dazu wurden zunächst SSBSC-Sprachsignale aus dem Kurzwellenbereich demoduliert. Es folgte eine Abtastung mit der Rate  $f_A = 8$  kHz bei einer Wortbreite von 16 Bit.

Das Korrekturverfahren verwendet eine Rahmenlänge von 62,5 ms, was 500 Abtastwerten entspricht. Zur Bestimmung der Kovarianzmatrix  $\mathbf{R}$  in Form einer Näherung  $\hat{\mathbf{R}}$  wurden aus den 500 Abtastwerten 349 Vektoren mit jeweils 160 Elementen gebildet, die um jeweils einen Abtastwert gegeneinander verschoben waren. Durch „Subspace Embedding“ wurde die Kovarianz-Näherung mit  $160 \times 160$  Elementen gebildet.

Zur Unterscheidung von stimmhaften und (für die Korrektur ungeeigneten) stimmlosen Segmenten wurde ein eigenes Verfahren angewandt, über das an dieser Stelle nicht berichtet werden soll. Zusammenfassend ist festzustellen, dass das beschriebene Verfahren eine Frequenzkorrektur liefert, die dem subjektivem Empfinden nach eine natürliche und angenehme Klangfarbe liefert. Dabei gelingt eine Korrektur auch bei stärker verrauschten Sprachsignalen. Über objektive Messergebnisse kann momentan noch nicht berichtet werden. Sie sind Gegenstand gegenwärtiger Untersuchungen.

## LITERATUR

- [1] W. Hess P. Vary, U. Heute. Digitale Sprachverarbeitung. Teubner-Verlag, 1998.
- [2] Harry L. Van Trees. Detection, Estimation, and Modulation Theory, Optimum Array Processing (Part IV). Wiley-Interscience, 2002.
- [3] Sophocles J. Orfanidis. Optimum Signal Processing. McGraw-Hill International Edition, 1990.
- [4] A. Jansen P. Clark, S. H. Mallidi and H. Hermansky. Frequency offset correction in speech without detecting pitch. IEEE ICASSP, May 2013.
- [5] H. J. Kolb T. Gölzow, U. Heute. ssb carrier mismatch detection from speech characteristics: Extensions beyond the range of uniqueness. EUSIPCO Vol 1, 2002.
- [6] H. Voelcker. Demodulation of single side-band signals via envelope detection. IEEE Trans. Communication Technology, Feb. 1966.

## A Correlation Based Method for Measuring and Monitoring the Impulse Response of Analog LTI Systems with low Realization Cost

Lida Dastager and Jürgen Weith  
University of Applied Sciences Würzburg-Schweinfurt,  
Dept. of Electrical Engineering

This article deals with the problem of measuring the impulse response of a linear time invariant (LTI) analog network. At the beginning three known measurement principles are discussed and compared with respect to their advantages and disadvantages. It is shown that good results can only be attained at the cost of measurement expense. Subsequently a new correlation based method is presented where low complexity and high quality results have not to be a contradiction. The measurement circuit basically consists of two m-sequence generators and an analog low pass filter. The characteristics of the proposed measurement circuit are examined and explained by two examples.

Categories and Subject Descriptors: System characteristics, digital measurement

Additional Key Words and Phrases: Impulse response, measurement, correlation, shift register, m-sequences

### 1 PRELIMINARY REMARK

The response  $y(t)_t$  of a linear time invariant system when an input signal  $x(t)$  is applied is given by Duhamel's well known convolution integral [1]

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h_0(t - \tau)d\tau.$$

$h_0(t)$  is the impulse response, representing the system characteristics in time domain.

It may be noted here that the impulse response is of great importance for both solving theoretical as well as practical problems.

In the field of engineering  $h_0(t)$  must often be determined by measurement. Which method to take? Generally speaking we can choose a time domain or a frequency domain based approach. We will discuss this point later on. The focus of this article will be on the presentation of a new low complexity time domain and correlation based measurement method.

Before proceeding it is to be expressed that the new method is suitable for linear systems with time invariance characteristic (LTI).

Author's address: Lida Dastager and Jürgen Weith, University of Applied Sciences Würzburg-Schweinfurt, Dept. of Electrical Engineering, Ignaz-Schön-Str. 11, 97421 Schweinfurt, Germany  
Email: lida.dastager@gmx.de, juergen.weith@fhws.de

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than FHWS must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission.

## 2 WAYS TO MEASURE THE IMPULSE RESPONSE OF AN LTI SYSTEM

### 2.1 Impulse excitation method

When an LTI-system is excited by a small rectangular impulse of duration  $T_0$  and amplitude  $A$ , the reaction  $\tilde{h}_0(t)$  approximates (except of a constant factor  $c$ ) the impulse response, i. e.

$$\tilde{h}_0(t) \approx c \cdot h_0(t) \quad (2)$$

where

$$c = AT_0. \quad (3)$$

When  $T_0$  decreases the approximation quality increases. This is true because the rectangular impulse gets more similar to the Dirac impulse (and vice versa). As a consequence the amplitude decreases as well. This can be compensated by increasing  $A$ , however the increase must be stopped before internal system overflow or overload occurs, otherwise non-linearities would be introduced. The maximum allowed width of the input rectangular pulse depends on how fast the impulse response changes. Generally speaking in most cases  $\tilde{h}_0(t)$  is more or less buried in the system noise.

As a consequence the impulse method is suitable only in some special cases (e. g. in low noise systems) because of its low noise immunity.

### 2.2 The spectral method

The convolution in time domain corresponds to a multiplication in frequency domain, i. e. the system characteristic can be expressed in terms of the quotient of the complex spectrum of output signal and input signal. There are some limitations when choosing the excitation signal. This detail is not be discussed here.

$$h_0(t) = \text{IFT} \left\{ \frac{\text{FT}\{y(t)\}}{\text{FT}\{x(t)\}} \right\} \quad (4)$$

(FT means Fourier-transformation and IFT the inverse Fourier-transformation).

The noise like phenomena are limited in this case.

The disadvantage of this method is the relatively large amount of technical build up needed to realize equ. (4): An AD-converter, a clock oscillator and a signal processor or additional hardware for realizing the Fourier transform.

### 2.3 The correlation method

The measurement set up is depicted in fig. 1. It consists of a signal source with excitation signal  $x(t)$ , a delay line (delay  $\tau$ ), the DUT (Device Under Test) with impulse response  $h_0(t)$ , a multiplier and an averaging filter [3].

Presuming the DUT to be causal we have (the symbol  $*$  designates convolution):

$$y(t) = x(t) * h_0(t) \quad (5)$$

$$= \int_0^{\infty} h_0(\sigma) x(t - \sigma) d\sigma \quad (6)$$

and

$$z(t) = \int_0^{\infty} h_0(\sigma)x(t-\sigma)x(t-\tau)d\sigma. \quad (7)$$

Averaging  $z(t)$  results in

$$\bar{z}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \int_0^{\infty} h_0(\sigma)x(t-\sigma)x(t-\tau)d\sigma dt \quad (8)$$

$$= \int_0^{\infty} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T h_0(\sigma)x(t-\sigma)x(t-\tau)dt d\sigma \quad (9)$$

$$= \int_0^{\infty} h_0(\sigma) \underbrace{\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t-\sigma)x(t-\tau)dt}_{\varphi_{xx}(\tau-\sigma)} d\sigma \quad (10)$$

$$= \int_0^{\infty} h_0(\sigma)\varphi_{xx}(\tau-\sigma)d\sigma. \quad (11)$$

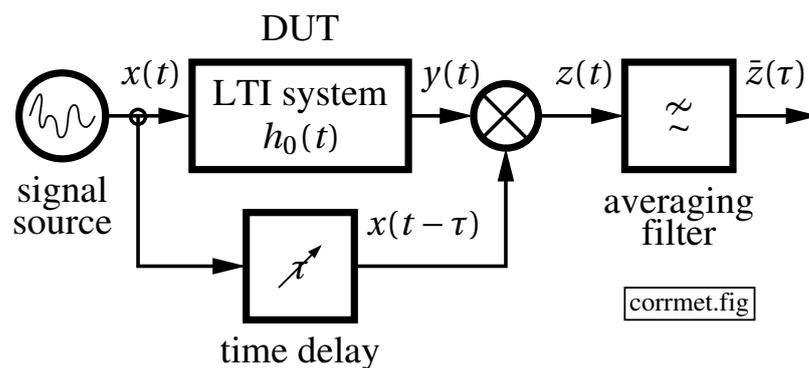


Figure 1: Setup for the determination of the impulse response based on the correlation method.

If the autocorrelation function  $\varphi_{xx}(\tau)$  is a small impulse with area  $F_\varphi$  centered around  $\tau = 0$  and if  $\varphi_{xx}(\tau) \equiv 0$  for  $|\tau| \geq T_\varphi$  (the impulse is small, if  $h_0(t) \approx h_0(t - T_\varphi) \forall t > T_\varphi$ ), we get

$$\bar{z}(\tau) \approx h_0(\tau) \int_{-\infty}^{\tau} \varphi_{xx}(\sigma)d\sigma \quad (12)$$

$$\approx h_0(\tau)F_\varphi. \quad (13)$$

We conclude that the average value  $\bar{z}(\tau)$  equals (apart from the constant  $F_\varphi$ ) the impulse response.

By choosing an appropriate integration time for the averaging filter the noise contained in the output  $\bar{z}(\tau)$  can be kept arbitrarily low. Of course a large integration time results in a long measurement cycle.

As can be seen using correlation for impulse response measurement we need

- an analog multiplier,
- a time delaying network with an *adjustable* delay time  $\tau$ , i. e. impulse response  $\delta_0(t - \tau)$ .

In practice the implementation of both systems can be done only at a great expense. E. g. the time delaying system can be realized with good quality only based on digital components: An analog to digital converter, an addressable memory and a digital to analog converter.

This makes clear that the correlation based method as well as the spectral method both suffer from a relatively high realization complexity.

### 3 A MODIFIED CORRELATION BASED MEASUREMENT METHOD

Correlation based measurements are attractive because of their excellent noise immunity. To overcome the relatively high complexity of this method in the original form shown in fig. 1 some modifications are now proposed to reduce the hardware effort. We start with a modification of fig. 1, leading to the circuit depicted in fig. 2.

#### 3.1 Substitution of the multiplier

The measurement set up in fig. 2 differs from that shown in fig. 1 by *two* sources with different signals  $x_1(t)$  and  $x_2(t)$ , while the former one has only one signal. We now have

$$\bar{z}(\tau) = \int_0^{\infty} h_0(\sigma) \varphi_{x_2 x_1}(\tau - \sigma) d\sigma \quad (14)$$

where  $\varphi_{x_2 x_1}(\tau)$  denotes the cross correlation function of  $x_1(t)$  and  $x_2(t)$  respectively. When choosing  $x_2(t)$  as a *binary* function with values  $x_2 \in \{a_2, b_2\}$ , the multiplier degenerates into a switch and two coefficient elements. Fig. 3 shows the circuit. It is easy and simple to realize.

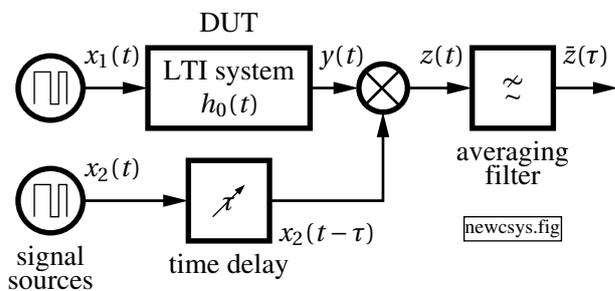


Figure 2: Measurement of the impulse response using the modified correlation method.

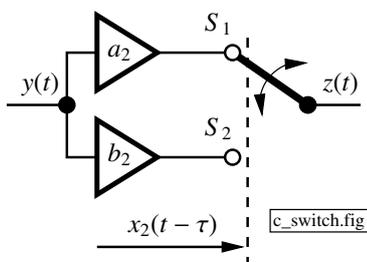


Figure 3: Substitution of the multiplier. The switch position is  $S_1$  for  $x_2(t - \tau) = a_2$  or  $S_2$  if  $x_2(t - \tau) = b_2$ .

There are two cases which are of special interest from a practical point of view:

1.  $a_2 = 1$  and  $b_2 = 0$ .  
The multiplier is substituted by an on-off-switch (SPST<sup>1</sup>). No need for coefficient elements here.
2.  $a_2 = 1$  and  $b_2 = -1$ .  
In this case a switch is used in addition to an amplifier with amplification  $V = -1$ .

Both cases are easy to put into practice, an important point for a realization with a low expenditure.

### 3.2 Considering the choice of $x_1(t)$ and $x_2(t)$

Regarding the excitation signals  $x_1(t)$  and  $x_2(t)$  from fig. 2 we have the following requirements:

1. As already mentioned  $x_2(t)$  should be a binary function to simplify the realization of the multiplier.
2. To achieve  $\bar{z}(\tau)$  approximates  $h_0(\tau)$  sufficiently well, the cross correlation function  $\varphi_{x_1x_2}(\tau)$  must resemble the Dirac impulse  $\delta_0(\tau)$  as close as possible.
3. It would be attractive if the impulse response could be measured continuously and periodically (i. e. not only once). In this case slow trends caused by ageing of the system elements or by temperature influence would be easy to recognize. To realize such a periodization of the impulse response  $\varphi_{x_1x_2}(\tau)$  must be periodic as well.

A suitable choice is to select  $x_1(t)$  and  $x_2(t)$  as m-sequences. A m-sequence can be realized very easily by using a shift register with an appropriate signal feedback<sup>2</sup> [4]. See fig. 4 for details.

The values 0/1 of the sequence  $x(t)$  are mapped to  $a_1/b_1$  to generate  $x_1(t)$  resp.  $a_2/b_2$  to generate  $x_2(t)$  i. e. for  $v \in \{1, 2\}$ :

$$x_v(t) = (b_v - a_v)x(t) + a_v. \quad (15)$$

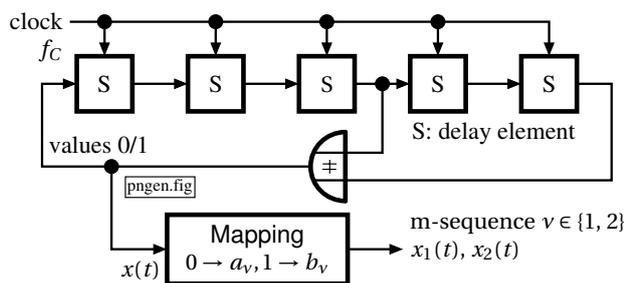


Figure 4: Example for a m-sequence generator.  $N = 5$  delay elements  $S$  are used. The EX-OR gate serves as a modulo-2 adder.

<sup>1</sup> Single pole single throw.

<sup>2</sup> Certain taps of the shift register chain are fed to a modulo-2 adder whose output is connected with the shift register's input.

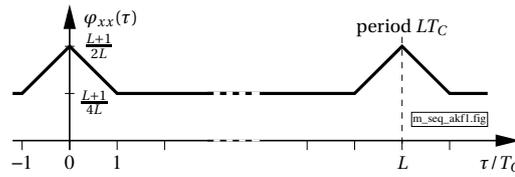


Figure 5: Autocorr. func.  $\varphi_{xx}(\tau)$  of the  $m$ -sequence  $x(t)$ .

For the following considerations the characteristics of the  $m$ -sequence are of interest:

1.  $x(t)$ ,  $x_1(t)$  and  $x_2(t)$  are periodic sequences with period  $L = T_C$ , where  $L = 2^N$  and  $T_C = 1/f_C$ .
2.  $x(t)$  has the autocorr. funct. depicted in fig. 5.
3. For the cross correlation  $\varphi_{xx}(\tau)$  we get

$$\begin{aligned} \varphi_{x_2x_1}(\tau) = & (b_2 - a_2)(b_1 - a_1)\varphi_{xx}(\tau) + \\ & + \bar{x}(\tau)a_1(b_2 - a_2) + \\ & + \bar{x}(\tau)a_2(b_1 - a_1) + \\ & + a_1a_2. \end{aligned} \quad (16)$$

The basic shape of the function  $\varphi_{x_1x_2}(\tau)$  is sketched in fig. 6.

4. The functions can be represented by a Fourier series ( $v \in \{1, 2\}$ ):

$$x_v(t) = \sum_{\mu=-\infty}^{\infty} c_{\mu}^{(v)} e^{j2\pi\mu\frac{f_C}{T}t} \quad (17)$$

where

$$|c_{\mu}^{(v)}| = \begin{cases} \frac{1}{2}\alpha_v + \frac{1}{2L}\beta_v & \text{for } \mu = 0 \\ \frac{\sqrt{L+1}}{2L}\beta_v \left| \frac{\sin\pi\frac{\mu}{L}}{\pi\frac{\mu}{L}} \right| & \text{for } \mu \neq 0 \end{cases} \quad (18)$$

with substitution  $\alpha_v = b_v + a_v$ ,  $\beta_v = b_v - a_v$ . It is  $c_{(-\mu)}^{(v)} = c_{\mu}^{(v)*}$  because the sequences are real valued. In table 1 the Fourier coefficients for both special cases are listed.

The desired impulse like behaviour of the cross correlation function results if for given values  $a_2$  and  $b_2$  the values  $a_1$  and  $b_1$  are selected as follows:

$$b_1 = -a_1 \left( 1 - \frac{4a_2}{(L+1)(a_2 + b_2)} \right). \quad (19)$$

Equation (19) results from the fact that the “offset” in fig. 6 is set to zero. For  $a_2 = -1$  and  $b_2 = 1$  the values  $a_1 = 0$  and  $b_1$  arbitrary, for  $a_2 = 0$  and  $b_2 = 1$  results  $b_1 = -a_1$ . In fig. 7 an example is given.

### 3.3 Elimination of the time delaying element

Up to now it is an open question how to realize the time delaying element in fig. 2. Below it is shown that this element is not required for the measurement of the impulse response. It can in fact be omitted using a little “trick”

If the two  $m$ -sequence generators are to be operated at slightly different clock frequencies the two sequences will – figuratively speaking – “slide” past each other. We achieve the same

effect as if the delay would increase (or decrease) permanently. Fig. 8 shows the realizing circuitry.

Before proceeding we must point to a general change concerning our measurement principle when using the new “sliding” method to substitute the time delay.

In the section 2.3 a fixed delay  $\tau$  is adjusted, then the measurement process takes place and then  $\tau$  was changed for the next measurement. Hence the measurement is done on a “step by step” base.  $\tau$  remains constant for a certain time step. In contrast to that the measurement process using the “sliding” sequences is based on a continuously changed delay. As a consequence we do not get  $\tilde{z}(\tau)$  at the low pass output exactly. We get  $w(t)$  instead, i. e. a signal differing more or less from the expected impulse response  $h_0(t)$ .

#### 3.4 The output signal $w(t)$

The m-sequence time functions  $x_1(t)$  and  $x_2(t)$  are periodic with periods  $T_1 = L/f_C$  and  $T_2 = L/((1-\delta)f_C)$ . For the Fourier series of  $x_1(t)$  and  $x_2(t)$  we get using  $f_{p1} = 1/T_1 = f_C/L$  and  $f_{p2} = 1/T_2 = (1-\delta)f_C/L$

$$x_1(t) = \sum_{\mu=-\infty}^{\infty} c_{\mu}^{(1)} e^{j2\pi\mu f_{p1} t} \quad (24)$$

$$x_2(t) = \sum_{\mu=-\infty}^{\infty} c_{\mu}^{(2)} e^{j2\pi\mu f_{p2} t}. \quad (25)$$

Assuming steady state at the output of the LTI-system excited by  $x_1(t)$  the following time function  $y(t)$  results:

$$y(t) = \sum_{\mu=-\infty}^{\infty} H(j2\pi\mu f_{p1}) c_{\mu}^{(1)} e^{j2\pi\mu f_{p1} t}. \quad (26)$$

We now form

$$z(t) = y(t) \cdot x_2(t) \quad (27)$$

$$= \left( \sum_{\mu=-\infty}^{\infty} H(j2\pi\mu f_{p1}) c_{\mu}^{(1)} e^{j2\pi\mu f_{p1} t} \right) \cdot \left( \sum_{\nu=-\infty}^{\infty} c_{\nu}^{(2)} e^{j2\pi\nu f_{p2} t} \right). \quad (28)$$

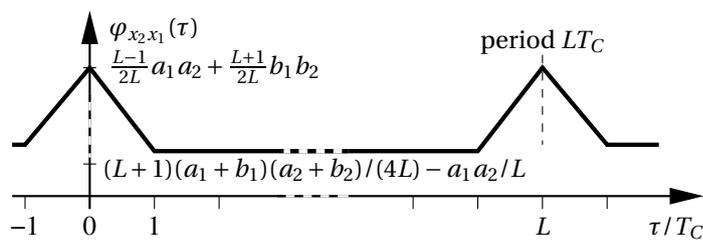


Figure 6: Cross correlation function  $\varphi_{x_2x_1}(\tau)$ .

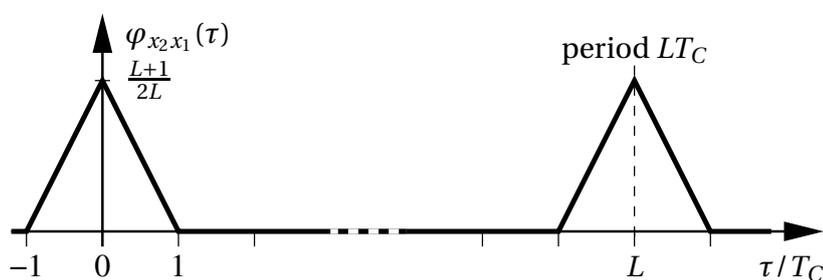


Figure 7: Cross correlation function  $\varphi_{x_2x_1}(\tau)$  for the special case  $a_2 = 0, b_2 = 1, a_1 = -1$  and  $b_1 = 1$ .

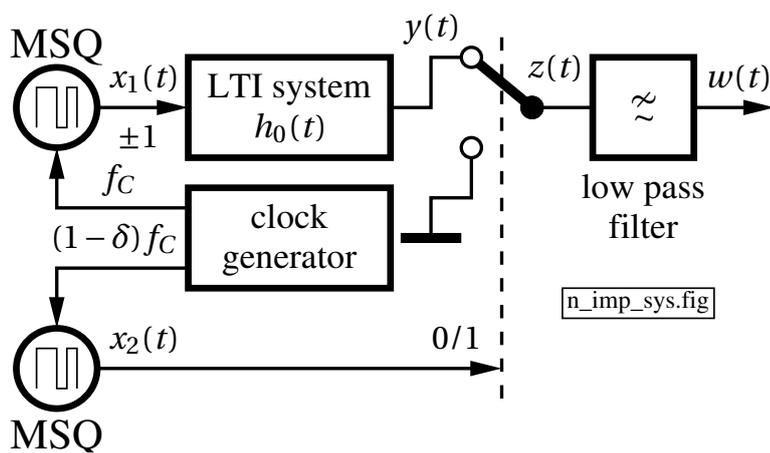


Figure 8: The proposed method for measuring the impulse response with  $a_1 = -1, b_1 = 1, a_2 = 0$  and  $b_2 = 1$ . Used are two m-sequence generators MSQ (sequence length  $L$ ) with different clocking, a clock oscillator, a SPST switch and a low pass filter.

---

**Case 1:**  $a_1 = -1, b_1 = 1, a_2 = 0, b_2 = 1$

---

$$|c_\mu^{(1)}| = \begin{cases} \frac{1}{L} & \text{for } \mu = 0 \\ \frac{\sqrt{L+1}}{L} \left| \frac{\sin \pi \frac{\mu}{L}}{\pi \frac{\mu}{L}} \right| & \text{for } \mu \neq 0 \end{cases} \quad (20)$$

and

$$|c_\mu^{(2)}| = \begin{cases} \frac{1}{2} \left(1 + \frac{1}{L}\right) & \text{for } \mu = 0 \\ \frac{1}{2} \frac{\sqrt{L+1}}{L} \left| \frac{\sin \pi \frac{\mu}{L}}{\pi \frac{\mu}{L}} \right| & \text{for } \mu \neq 0. \end{cases} \quad (21)$$


---

**Case 2:**  $a_1 = 0, b_1 = 1, a_2 = -1, b_2 = 1$

---

$$|c_\mu^{(1)}| = \begin{cases} \frac{1}{2} \left(1 + \frac{1}{L}\right) & \text{for } \mu = 0 \\ \frac{1}{2} \frac{\sqrt{L+1}}{L} \left| \frac{\sin \pi \frac{\mu}{L}}{\pi \frac{\mu}{L}} \right| & \text{for } \mu \neq 0. \end{cases} \quad (22)$$

and

$$|c_\mu^{(2)}| = \begin{cases} \frac{1}{L} & \text{for } \mu = 0 \\ \frac{\sqrt{L+1}}{L} \left| \frac{\sin \pi \frac{\mu}{L}}{\pi \frac{\mu}{L}} \right| & \text{for } \mu \neq 0 \end{cases} \quad (23)$$


---

*Table 1: Fourier coefficients of the functions  $x_1(t)$  and  $x_2(t)$  for the special cases under consideration.*

By rearranging the summation terms follows

$$z(t) = \sum_{\mu=-\infty}^{\infty} H(j2\pi\mu f_{P1}) c_\mu^{(1)} \cdot \sum_{\nu=-\infty}^{\infty} c_\nu^{(2)} e^{j2\pi(\mu+\nu)f_{P1}t} e^{-j2\pi\nu\delta f_{P1}t}. \quad (29)$$

Because we expect low pass behaviour of the DUT, using  $f_H = \lambda f_{P1} = \lambda f_C / L, \lambda \in \mathbb{N}$

$$z(t) = \sum_{\mu=-\lambda}^{\lambda} H(j2\pi\mu f_{P1}) c_\mu^{(1)} \cdot \sum_{\nu=-\infty}^{\infty} c_\nu^{(2)} e^{j2\pi(\mu+\nu)f_{P1}t} e^{-j2\pi\nu\delta f_{P1}t}. \quad (30)$$

If the low pass filter used for averaging has a cut-off frequency  $f_g$  with

$$\lambda\delta f_{P1} \leq f_g \leq (1 - \lambda\delta) f_{P1}, \quad (31)$$

then only such summation terms of eqn. (30) appear at the output, for which  $\mu + \nu = 0$  holds. From this the time signal  $w(t)$  results:

$$w(t) = \sum_{\mu=-\lambda}^{\lambda} c_{\mu}^{(1)} c_{-\mu}^{(2)} H(j2\pi\mu f_{p1}) e^{j2\pi\mu\delta f_{p1} t}. \quad (32)$$

If we put eqn. (20) and eqn. (21) resp. Glg. (22) and (23) into eqn. (32) it follows

$$w(t) = \frac{L+1}{2L^2} \sum_{\mu=-\lambda}^{\lambda} \left( \frac{\sin \pi \frac{\mu}{L}}{\pi \frac{\mu}{L}} \right)^2. \quad (33)$$

$$\cdot H(j2\pi\mu f_{p1}) e^{j2\pi\mu\delta f_{p1} t}. \quad (34)$$

When the period length  $L$  of the shift register sequence was chosen to hold

$$\left( \frac{\sin \pi \frac{\lambda}{L}}{\pi \frac{\lambda}{L}} \right)^2 \approx 1 \quad (35)$$

the calculation leads to

$$w(t) = \frac{L+1}{2L^2} \sum_{\mu=-\lambda}^{\lambda} H(j2\pi\mu f_{p1}) e^{j2\pi\mu\delta f_{p1} t}. \quad (36)$$

For eqn. (36) we can write

$$\begin{aligned} w(t) &= \frac{L+1}{2L^2} \sum_{\mu=-\infty}^{\infty} \left( \int_{-\infty}^{\infty} h_0(\tau) e^{-j2\pi\mu f_{p1} \tau} d\tau \right) \cdot e^{j2\pi\mu\delta f_{p1} t} \\ &= \frac{L+1}{2L^2} \int_{-\infty}^{\infty} h_0(\tau) \sum_{\mu=-\infty}^{\infty} e^{-j2\pi\mu f_{p1}(\tau - \delta t)} d\tau \\ &= \frac{L+1}{2L^2 f_{p1}} \int_{-\infty}^{\infty} h_0(\tau) \sum_{\mu=-\infty}^{\infty} \delta_0(\tau - \delta t - \frac{\mu}{f_{p1}}) d\tau \\ &= \frac{L+1}{2L^2 f_{p1}} \sum_{\mu=-\infty}^{\infty} h_0(\delta t - \frac{\mu}{f_{p1}}) \\ &= \frac{L+1}{2L f_C} \sum_{\mu=-\infty}^{\infty} h_0(\delta t - \frac{\mu L}{f_C}). \end{aligned} \quad (37)$$

*Summarizing result: At the low pass output the periodized, stretched and scaled (by an amplitude factor) impulse response of the DUT appears. The amplitude factor has the value  $(L+1)/(2L f_C)$ , the stretching factor is  $\Gamma = 1/\delta$  and the period time is  $\Gamma L/f_C = \Gamma L T_C$ .*

### 3.5 Dimensioning specification

Prior to dimensioning the following variables must be known:

- the cut off frequency  $f_H$  of the device under test,
- the length (duration)  $T_p$  of the impulse response,
- the desired spreading factor  $\Gamma$ .

The determination is as follows:

1. First of all it must be fixed, "how well" eqn. (35) is to be satisfied. For this the quotient  $l_0 = L/\lambda$ , must be chosen suitable.
2. With a known value for  $l_0$  and known cut off frequency  $f_H$  of the DUT we get  $f_C$  from

$$f_C = l_0 f_H \quad (38)$$

3. To avoid a temporal overlap of the impulse response in equ. (37),

$$\frac{L}{f_C} \geq T_h \quad (39)$$

must be satisfied. Because for the period length  $L$  there is the condition  $L = 2^N - 1$ ,  $N \in \mathbb{N}$ , results

$$N = \lceil \lg(fc T_h + 1) \rceil \quad (40)$$

( $\lceil x \rceil$ : Largest integer not exceeding  $x$ ).

4. The sequence length  $L$  now results from

$$L = 2^N - 1. \quad (41)$$

5. Now  $\lambda$  can be fixed:

$$\lambda = \left\lfloor \frac{L}{l_0} \right\rfloor \quad (42)$$

( $\lfloor x \rfloor$ : Smallest integer not falling below  $x$ ).

6. Finally we need the cut off frequency  $f_g$  of the averaging low pass. It can be calculated as

$$\frac{\lambda}{\Gamma L} f_C \leq f_g \leq \left( \frac{1}{L} - \frac{\lambda}{\Gamma L} \right) f_C. \quad (43)$$

## 4 EXAMPLES

### 4.1 Impulse response of an RC circuit

For testing purposes the investigation of an RC circuit (fig. 9) is of principal interest. The RC circuit under consideration has the transfer function  $H(p) = Y(p)/X_1(p)$  where

$$H(p) = \frac{10^3 s^{-1}}{p + 10^3 s^{-1}} \quad (44)$$

and its impulse response  $h_0(t)$  is

$$h_0(t) = \delta_{-1}(t) 10^3 s^{-1} e^{-t/(10^{-3}s)}. \quad (45)$$

The frequency response  $|H(j2\pi f)|$  as well as the impulse  $h_0(t)$  have no sharp limitation. Therefore it is demanded arbitrarily

$$H(j2\pi f_H) = \frac{1}{40} \quad (46)$$

and

$$h_0(T_h) = \frac{1}{40} h_0(+0). \quad (47)$$

From this we get

$$T_b = 5.3\text{ms} \quad f_H = 6.36\text{kHz}. \quad (48)$$

For  $\Gamma$  we again use arbitrarily  $\Gamma = 10^3$  (extension of the impulse response from 6.36ms to 6.36sec).

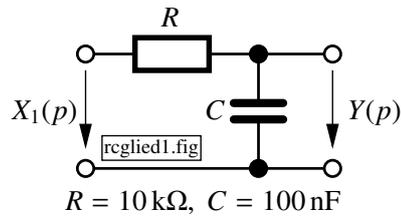


Figure 9: RC circuit, whose impulse response is to be determined by the proposed method.

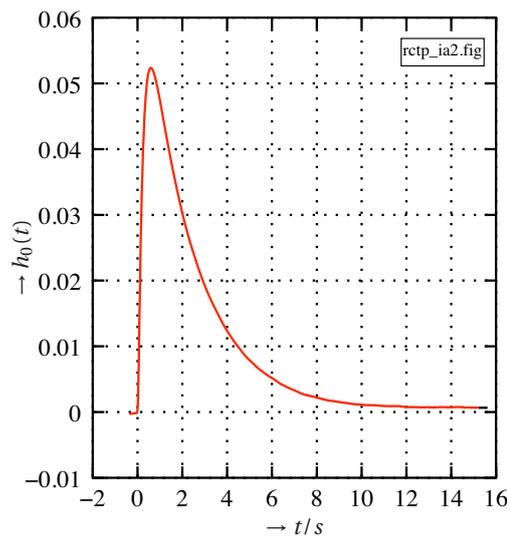


Figure 10: Measured impulse response of the RC circuit.

Further it will be requested that the function  $\sin(\pi\lambda/L)/(\pi\lambda/L)$  equ. (35) should have the value 0.9, and from this

$$l_0 = \frac{L}{\lambda} = 4 \quad (49)$$

Now applies

$$f_C = 25,44\text{kHz}. \quad (50)$$

For this case it follows

$$N = \lceil \text{ld}(135,83) \rceil = 8 \quad (51)$$

as well as

$$L = 255. \quad (52)$$

and for  $\lambda$ :

$$\lambda = \left\lfloor \frac{L}{l_0} \right\rfloor = \left\lfloor \frac{255}{4} \right\rfloor = 63. \quad (53)$$

If  $f = 0 \dots f_{gp}$  is the passband and  $f = f_{gs} \dots \infty$  the stop band of the averaging low pass, we get

$$f_{gp} = 6,28 \text{ Hz} \quad \text{and} \quad f_{gs} = 93,48 \text{ Hz}. \quad (54)$$

In this case we can use a very simple filter.

In fig. 10 the measured impulse response of the RC circuit is drawn. It is in good agreement with the expected response. Deviations are observed at time  $t = 0$ , where there should occur a step. But generally speaking a step can not occur because of the low pass characteristic of the averaging filter. A low pass has a rise time greater than zero and thus no step is possible.

#### 4.2 Impulse response of a Cauer (elliptic) filter

The next device under test was a Cauer low pass filter (also known as an elliptic low pass filter) with pass band frequency  $f_D = 4.5 \text{ kHz}$ , where the pass band attenuation is  $a_D = 0.8 \text{ dB}$ . The stop band starts at  $f_S = 5 \text{ kHz}$ , the stop band attenuation is  $a_S = 55 \text{ dB}$ .

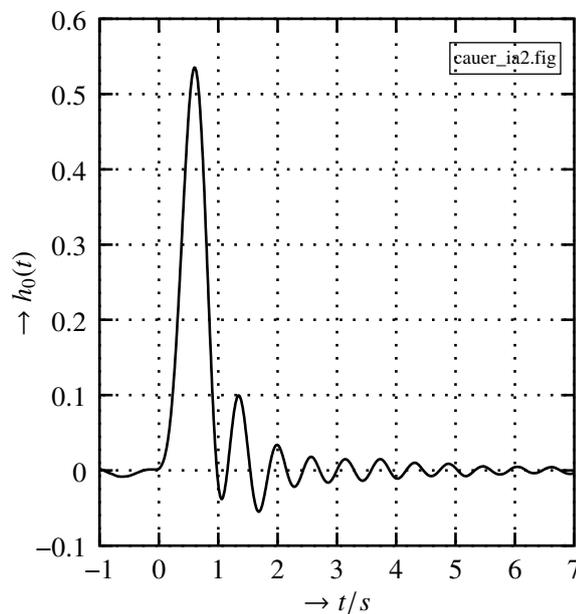


Figure 11: Measured impulse response of an elliptic filter.

The measured impulse response for this type of filter is depicted in fig. 11. It can be seen that the period duration of 8 s was chosen a bit too short, because left of the time  $t = 0$  the transients of the preceding period is still present.

## 5 CHARACTERISTICS OF THE PROPOSED MEASURING METHOD

### 5.1 Preliminary remark

How exactly can the measurement circuit determine the impulse response? To answer the question it is assumed that the device under test shows a noise like interfering signal  $n(t)$ , whose spectral density  $N_0(f)$  will be

$$N_0(f) = \begin{cases} D & \text{for } |f| \leq f_H \\ 0 & \text{else.} \end{cases} \quad (55)$$

Fig. 12 shows the model of the DUT used below.

Now the signal  $y'_0(t)$  is fed to the multiplier (switch). Caused by the interfering noise being now present at the low pass output appears

$$w'(t) = w(t) + n'(t), \quad (56)$$

where  $w(t)$  means the periodized impulse response which results in the absence of the noise signal.  $n'(t)$ , denotes the interference generated by the low pass filter when excited by  $n(t)$ .

### 5.2 Definition of a figure of merit

The quality of the measurement is to be described by a figure of merit  $Q$  defined by

$$Q = \frac{\int_0^{\Gamma T_h} w^2(t) dt}{\Gamma T_h E\{(n'(t))^2\}}. \quad (57)$$

$Q$  means the quotient of useful energy and noise like energy of the output signal during time  $\Gamma T_b$  of the spread impulse response.

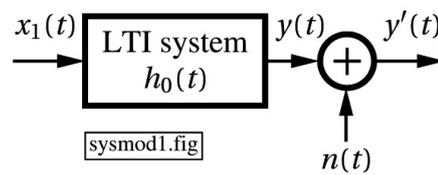


Figure 12: System model with noisy output signal.

### 5.3 The case $a_2 = 0$ and $b_2 = 1$

The numerator of  $Q$  equals

$$\int_0^{\Gamma T_h} w^2(t) dt = \Gamma \frac{L+1}{2Lf_C} \int_0^{T_h} h_0^2(t) dt. \quad (58)$$

For the expectation value in the denominator of equ. (57) we get<sup>3</sup>

$$E\{(n'(t))^2\} = D \left[ |c_0^{(2)}|^2 + |c_\lambda^{(2)}|^2 + 2 \sum_{\mu=1}^{\lambda-1} |c_\mu^{(2)}|^2 \right] \frac{2f_H}{\Gamma} \quad (59)$$

and with equ. (21)

$$\approx D \frac{L+1}{L} \left[ \frac{1}{2} + \frac{\lambda}{L} \right] \frac{f_H}{\Gamma}. \quad (60)$$

Now  $Q$  is fixed:

$$Q \approx \frac{\frac{\Gamma}{f_C} \int_0^{T_h} h_0^2(t) dt}{2D \left[ \frac{1}{2} + \frac{\lambda}{L} \right] f_H T_h}. \quad (61)$$

The gain of the procedure can be judged if  $Q$  is compared with the quality figure  $Q_r$  of the impulse method explained in 2.1. It is

$$Q_r = \frac{\frac{1}{f_C} \int_0^{T_h} h_0^2(t) dt}{2D f_H T_h}. \quad (62)$$

For the gain factor  $q = Q/Q_r$  we have

$$q = \Gamma \frac{2}{1 + 2\frac{\lambda}{L}}, \quad \frac{\lambda}{L} = \frac{f_H}{f_C}. \quad (63)$$

*Summarizing result: The gain factor  $q$  is proportional to the stretching factor  $\Gamma$ . It depends on the quotient of the cut off frequency  $f_H$  and the clock frequency  $f_C$  of the  $m$ -sequence generator.*

<sup>3</sup> In the interest of a simple and clear final result in the following the  $\sin(x)/x$  dependency of the coefficients  $c_\mu^{(2)}$  will be ignored.

#### 5.4 The case $a_2 = -1$ and $b_2 = 1$

For this case the numerator of the figure of merit is in conformity with the numerator of eqn. (61). The denominator follows from eqn. (59) together with eqn. (23). We get

$$Q \approx \frac{\frac{\Gamma}{f_c} \int_0^{T_h} h_0^2(t) dt}{4D \frac{\lambda}{L} f_H T_h}. \quad (64)$$

If  $Q$  normalized as above by  $Q_r$ , eqn. (62), we get for  $q$

$$q = \Gamma \frac{L}{2\lambda}.$$

For such values  $\lambda/L$  actually used in measuring tasks (e. g.  $\lambda/L = 1/4$ ),  $q$  approximately equals the figure of merit of (Glg. 63).

*In principle using the proposed method the figure of merit can be increased arbitrarily, but only on the expense of a growing measuring time.*

## 6 SUMMARY

A correlation based method for measuring the impulse response of a linear and time invariant analog system is described.

It is characterized by a low technical expense: The circuitry consists only of two m-sequence generators, a clock oscillator, an electronic switch (possibly with a inverting amplifier) and a simple RC circuit (low pass).

The measurement results in a stretched and periodized form of the actual impulse response of the DUT. The acquisition and further processing of this signal is simple and easy.

It was shown that the figure of merit is governed by the stretching factor. It can be made arbitrarily large, but this is true at the expense of measuring time.

## REFERENCES

- [1] R. Unbehauen: Systemtheorie. Oldenbourg-Verlag MünchenWien 1997
- [2] H.W. Schüßler: Netzwerke, Signale und Systeme. Springer-Verlag Berlin 1985
- [3] E. Philippow: Taschenbuch der Elektrotechnik, Band 3. VEB Verlag Technik Berlin 1969
- [4] S.W. Golomb: Shift Register Sequences. Aegean Park Press Laguna Hills California 1982
- [5] E.Mommertz, S. Müller: Measuring impulse responses with digitally pre-emphasized pseudorandom noise derived from maximum-length sequences Applied Acoustics Elsevier Volume 44, Number 3, 1995, pp.
- [6] Stan, Guy-Bart; Embrechts, Jean-Jacques; Archambeau, Dominique: Comparison of Different Impulse Response Measurement Techniques JAES Volume 50 Issue 4 pp. 249-262; April 2002



## Indoor Localization Using Step and Turn Detection Together with Floor Map Information

Lukas Köping, Frank Ebner, Marcin Grzegorzek, Frank Deinzer  
University of Applied Sciences Würzburg-Schweinfurt,  
Pattern Recognition Group, University of Siegen

In this work we present a method to estimate an indoor position with the help of smartphone sensors and without any knowledge of absolute positioning systems like Wi-Fi signals. Our system uses particle filtering to solve the recursive state estimation problem of finding the position of a pedestrian. We show how to integrate the information of the previous state into the weight update step and how the observation data can help within the state transition model. High positional accuracy can be achieved by only knowing that the pedestrian makes a foot step or changes her direction together with floor map information.

Categories and Subject Descriptors: Indoor Localization

Additional Key Words and Phrases: Particle Filter, Step Detection, Turn Detection

### 1 INTRODUCTION

Indoor localization considers the problem of finding the position of a person, robot or, more general, of a target object inside buildings. While Global Navigation Systems (GNSS) like GPS solved the localization problem in outdoor areas almost completely, this technique cannot be used indoor. This is because walls, windows, roofs and other obstacles attenuate the GNSS signals too heavily. However, a successful realization of an indoor localization system would pave the way for dozens of new applications. Imagine the possibility of guiding firemen through smoky buildings leading them directly to people in emergency situations. In commercial scenarios one can think of systems that guide customers through big shopping malls or museums to name a few.

A distinction must be made between indoor localization systems for pedestrians and indoor localization systems for robots. In the latter case the intention of movement is known to the system, e.g., if the robot wants to go left or right in the next time step. In pedestrian systems this is typically an unknown factor. Additionally, robots can be equipped with dozens of sensors, allowing to receive different forms of information of the environment like laser scans or video sequences and processing this information directly on an embedded high-performance computer. On the other side, pedestrian indoor localization systems can only rely on sensor data collected by devices like smartphones, since it is unlikely that people will agree to wear

Author's address: Lukas Köping, Frank Ebner; Fakultät für Informatik und Wirtschaftsinformatik, Hochschule für Angewandte Wissenschaften, Sanderheinrichsleitenweg 20, Würzburg, Deutschland

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than FHWS must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission.

additional devices. This assumption might be softened in case of emergency teams for whom it is common to wear additional equipment, nevertheless, it would be asked too much from firemen to wear heavy laser scanners with them. Furthermore, the whole processing must be doable on a small device with the additional restriction of low energy consumption.

In this work we will introduce an approach that utilizes only sensor data of a smartphone. The process of indoor localization will be formulated as a state estimation problem in a dynamic, non-linear system. Since state estimations rarely can be calculated in an analytically closed form, we will use particle filtering as an approximation technique. For this, we will use floor maps as a priori information.

The paper is organized as follows. Section 2 gives a brief overview over indoor localization systems. Section 3 summarizes recursive state estimation and particle filtering. In addition we give a theoretical formulation for the integration of observation data into the state transition model and for the integration of old state information into the weight update step. Section 4 matches the recursive state estimation with indoor localization. In section 5 we show some experimental results. Finally, section 6 summarizes our work and illustrates future work.

## 2 RELATED WORK

Today's indoor positioning systems mostly make use of absolute positioning with Wi-Fi signals. In an on-line phase the received signal strengths (RSS) are measured at different locations and listed in a radio map. During the localization process the signal strengths at the current position of the target object are measured and compared against the radio map entries. For this, [Bagosi and Baruch 2011] evaluated different methods of matching the current fingerprint against the signal strength map. Among them, deterministic approaches like the Centroid, K nearest neighbour and weighted K nearest neighbour were compared against a probabilistic Gaussian method, showing the latter one outperforming all the other methods and working with an average error rate of 1 meter.

On the other side, dead reckoning systems [Woodman and Harle 2008; Link et al. 2011; Wang et al. 2012; Zhang et al. 2013] estimate the current position with respect to a starting or reference point. Inertial navigation for example relies on the information of gyroscopes, accelerometers and magnetometers to infer a position by using Newton's laws of motion. However, even small errors in measurement will sum up during integration at every time step and finally lead to heavily defective positional estimations [Köping et al. 2012]. It was shown that the integration of floor map information together with methods like map-matching can improve this approach [Davidson 2010]. Nevertheless, the positional estimation of dead reckoning systems must be corrected by external signals [Woodman and Harle 2008]. A popular approach in this context is to combine Wi-Fi localization and readings from inertial sensors. A major drawback of this method lies in the development of a signal strength map - a process that has to be repeated periodically to incorporate environmental changes.

In the last few years the use of smartphones continuously rose because they provide most of the necessary sensors like accelerometers, gyroscopes, magnetometers, Wi-Fi modules, Bluetooth and cameras. Additionally, smartphones constitute a practical way for people to navigate through buildings, since they do not have to wear additional, unhandy devices like laptops or laser scanners with them. [Link et al. 2011] uses step detection and path matching to infer a

position. A step is detected by utilizing specific patterns in accelerometer data. The detected step and its heading are matched against an expected route, which is a path, suggested by the system. It describes the shortest distance between the user's current position and its desired destination. For this, they propose two different path matching algorithms that utilize dynamic programming techniques and incorporate detected steps, the heading and map information. [Zhang et al. 2013] uses a modified Kalman filter to achieve accurate orientation. During localization also step detection as well as step length detection is performed. In contrast to [Link et al. 2011], where step length is derived from a person's height, [Zhang et al. 2013]'s system estimates step length with the help of accelerometer data. In [Wang et al. 2012] the dead-reckoning system is stabilized by so called internal landmarks. These are special locations on a map, where characteristic signals can be measured. For example elevators and stairs reveal unique signal curves in accelerometer data. Given that such a characteristic signal is measured, it is possible to reset the defective position estimation and use this location as new starting point for dead-reckoning position.

### 3 RECURSIVE STATE ESTIMATION

Recursive state estimation is a technique to infer a hidden state  $\mathbf{q}_t$  at a given time  $t$  with the help of observations  $\mathbf{o}_t$ . Typically this technique is often found in the field of robotics, where different states like, e.g., the robot's position must be estimated given some sensor measurements [Thrun et al. 2006]. Thereby,  $\mathbf{q}_t$  contains all relevant information to describe the present system and the set of all observations up to time  $t$  can be written as  $\langle \mathbf{o} \rangle_t = \{\mathbf{o}_{t'} \mid 0 \leq t' \leq t\}$ . Often enough the goal is to estimate a most probable state  $\hat{\mathbf{q}}_t$  at time  $t$ :

$$\hat{\mathbf{q}}_t = \arg \max_{\mathbf{q}_t} p(\mathbf{q}_t | \langle \mathbf{o} \rangle_t) \quad (1)$$

Given Bayes rules,  $p(\mathbf{q}_t | \langle \mathbf{o} \rangle_t)$  can be written as

$$p(\mathbf{q}_t | \langle \mathbf{o} \rangle_t) = \frac{p(\langle \mathbf{o} \rangle_t | \mathbf{q}_t) p(\mathbf{q}_t)}{p(\langle \mathbf{o} \rangle_t)} \quad (2)$$

It describes the probability density of a state at time  $t$  given a series of observations. A convenient way to rewrite (2) is given by [Deinzer 2005]:

$$p(\mathbf{q}_t | \langle \mathbf{o} \rangle_t) = p(\mathbf{o}_t | \mathbf{q}_t) \int p(\mathbf{q}_t | \mathbf{q}_{t-1}) p(\mathbf{q}_{t-1} | \langle \mathbf{o} \rangle_{t-1}) d\mathbf{q}_{t-1} \quad (3)$$

Three essential parts can be identified in (3). The likelihood of an observation while in a given state is denoted as  $p(\mathbf{o}_t | \mathbf{q}_t)$ , the state transition probability as  $p(\mathbf{q}_t | \mathbf{q}_{t-1})$  and  $p(\mathbf{q}_{t-1} | \langle \mathbf{o} \rangle_{t-1})$  contains the information of all previous time steps.

In our approach we deviate from the standard form given by (3) and integrate observation data into the state transition model: The information of the pedestrian making a foot step or changing its direction will be used to estimate a new state:

$$p(\mathbf{q}_t | \mathbf{q}_{t-1}, \mathbf{o}_{t-1}) \quad (4)$$

In addition to that, state information of the previous time step  $\mathbf{q}_{t-1}$  is added to  $p(\mathbf{o}_t|\mathbf{q}_t)$ . For this, the position of the last time step is used to evaluate the current position. In section 4 this approach is described in more detail. Formally, the integration of the last time step's information leads to:

$$p(\mathbf{o}_t|\mathbf{q}_t, \mathbf{q}_{t-1}) \quad (5)$$

Subsequently we will show how to integrate (4) and (5) into (3). Our state at time  $t$  does not only depend on the current observation but also on the previous state, so that we are looking for  $p(\mathbf{q}_t|\langle \mathbf{o} \rangle_t, \mathbf{q}_{t-1})$ :

$$p(\mathbf{q}_t|\langle \mathbf{o} \rangle_t, \mathbf{q}_{t-1}) = \frac{p(\mathbf{q}_t, \langle \mathbf{o} \rangle_t, \mathbf{q}_{t-1})}{p(\mathbf{q}_{t-1}, \langle \mathbf{o} \rangle_t)} \quad (6)$$

$$= \frac{p(\mathbf{q}_t, \mathbf{o}_t, \langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1})}{p(\mathbf{q}_{t-1}, \langle \mathbf{o} \rangle_{t-1}, \mathbf{o}_t)} \quad (7)$$

$$= \frac{p(\langle \mathbf{o} \rangle_{t-1})p(\mathbf{q}_{t-1}|\langle \mathbf{o} \rangle_{t-1})p(\mathbf{q}_t|\langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1})p(\mathbf{o}_t|\mathbf{q}_t, \langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1})}{p(\langle \mathbf{o} \rangle_{t-1})p(\mathbf{q}_{t-1}|\langle \mathbf{o} \rangle_{t-1})p(\mathbf{o}_t|\mathbf{q}_{t-1}, \langle \mathbf{o} \rangle_{t-1})} \quad (8)$$

$$= \frac{p(\mathbf{q}_t|\langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1})p(\mathbf{o}_t|\mathbf{q}_t, \langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1})}{p(\mathbf{o}_t|\langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1})} \quad (9)$$

$$= \frac{1}{k} p(\mathbf{q}_t|\langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1})p(\mathbf{o}_t|\mathbf{q}_t, \langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1}) \quad (10)$$

$$= \frac{1}{k} p(\mathbf{q}_t|\langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1})p(\mathbf{o}_t|\mathbf{q}_t, \mathbf{q}_{t-1}) \quad (11)$$

$$= \frac{1}{k} p(\mathbf{o}_t|\mathbf{q}_t, \mathbf{q}_{t-1}) \int p(\mathbf{q}_t|\mathbf{q}_{t-1}, \langle \mathbf{o} \rangle_{t-1})p(\mathbf{q}_{t-1}|\langle \mathbf{o} \rangle_{t-1})d\mathbf{q}_{t-1} \quad (12)$$

$$= \frac{1}{k} p(\mathbf{o}_t|\mathbf{q}_t, \mathbf{q}_{t-1}) \int p(\mathbf{q}_t|\mathbf{q}_{t-1}, \mathbf{o}_{t-1})p(\mathbf{q}_{t-1}|\langle \mathbf{o} \rangle_{t-1})d\mathbf{q}_{t-1} \quad (13)$$

For (6) the definition of the conditional probability is used, while (8) is the result of applying the Multiplication Theorem of Probabilities. (8) can then be reduced to (9), whereby the denominator is a constant and denoted as  $k$  in (10). Using Markov's assumption,  $p(\mathbf{o}_t|\mathbf{q}_t, \langle \mathbf{o} \rangle_{t-1}, \mathbf{q}_{t-1})$  is rewritten as  $p(\mathbf{o}_t|\mathbf{q}_t, \mathbf{q}_{t-1})$ . This is true since the probability of a current observation is independent of previous observations. Applying the Total Law of Probability for  $p(\mathbf{q}_t|\mathbf{q}_{t-1}, \langle \mathbf{o} \rangle_{t-1})$  finally results in  $p(\mathbf{q}_t|\mathbf{q}_{t-1}, \mathbf{o}_{t-1})$ .

Since an analytical solution for densities of the form of (13) exists only in rare cases, a common approach to approximate such densities is particle filtering [Arulampalam et al. 2002; Gordon et al. 1993; Isard and Blake 1998; Doucet and Johansen 2011]. A particle  $\pi_t^i, i \in 1, \dots, N$  represents a potential state at time  $t$ , whereby  $N$  different particles are used. Every particle is defined as tuple

$$\pi_t^i = (\mathbf{q}_t^i, w_t^i). \quad (14)$$

$\mathbf{q}_t^i$  is the represented state of particle  $\pi_t^i$ ,  $w_t^i$  is the weight of the particle telling about its importance. Additionally, weights are normalized so that

$$\sum_{i=1}^N w_t^i = 1. \quad (15)$$

The posterior density is then equivalent to

$$p(\mathbf{q}_t|\langle \mathbf{o} \rangle_t) \approx \sum_{i=1}^N w_t^i \delta(\mathbf{q}_t - \mathbf{q}_t^i), \quad (16)$$

where  $\delta(\cdot)$  is Dirac's delta function and  $N \rightarrow \infty$  holds true. Practically, in the particle filter at every time step new particles are drawn according to

$$\pi_t^i \propto p(\mathbf{q}_t | \mathbf{q}_{t-1}, \mathbf{o}_{t-1}), \quad (17)$$

which is called the state transition step. In the evaluation step every particle is then weighted with

$$w_t^i = p(\mathbf{o}_t | \mathbf{q}_t, \mathbf{q}_{t-1}). \quad (18)$$

Finally, in a resampling step a new particle generation for the next time step  $t+1$  is generated by drawing  $N$  new particles according to

$$\pi_t^i \propto w_t^i \quad (19)$$

#### 4 RECURSIVE STATE ESTIMATION IN INDOOR LOCALIZATION

Our approach uses recursive state estimation with particle filtering as realization to infer the position of a pedestrian at a given time  $t$  with the help of different sensor data collected by a smart phone. If particle filtering is used, one must be aware of the state's dimension, since it has a vast impact on the number of samples to be simulated. For our approach it is enough to incorporate the x- and y-coordinate of the searched position as information, so that

$$\pi_t^i \propto w_t^i \quad (20)$$

The incoming sensor data at time  $t$  is treated as observation  $\mathbf{o}_t$ , thereby  $\mathbf{o}_t$  consists of

$$\mathbf{o}_t = \begin{pmatrix} \text{step}_t \\ \text{turn}_t \\ \alpha_t \end{pmatrix}, \quad (21)$$

where  $\text{step}_t$  and  $\text{turn}_t$  are both boolean variables and  $\text{step}_t = \text{true}$ , if it was detected that the target object made a footstep in the last time step and  $\text{turn}_t = \text{true}$ , if the target object made a left or right turn in the last time step. If a turn was detected the direction of the change can be determined by the peak value in the turn detection process (see chapter 4.1).  $\alpha_t$  denotes the heading information. The heading information is initialized at the beginning of the position tracking at a known starting position. We assume to know the initial direction of the pedestrian's body regarding the floor map. This direction usually can be estimated due to the context, e.g., if the pedestrian walks upstairs, her body direction is known at the end of the stairway and can be used as heading information. Another possibility initializing the heading information is to use the orientation data at a given time step, which today is easily available on every smartphone.

##### 4.1 Step and Turn Detection

**Step detection:** We use a step detection process to estimate if our target object made a move in the last time interval, whereat the method we use is described in detail in [Link et al. 2011]. Basically, the linear acceleration towards the ground is measured. In this signal a step is detected by picking a data point, which exceeds a certain threshold. In addition this data point must be greater than one of its following data points by a certain amount, whereby the following data point must lie within a given time interval. After a certain dead time, in which no step can be recognized, the step detection process starts from the beginning (see Fig.1).

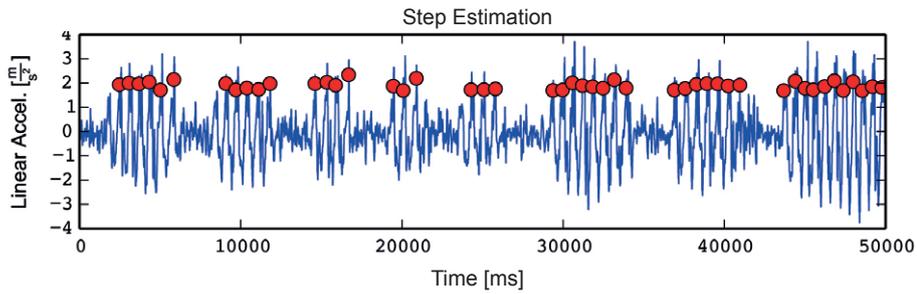


Fig. 1. Step detection process: The blue curve represents the linear acceleration in ground direction. Red dots represent detected steps.

**Turn detection:** In addition to the step detection process, turn detection is integrated into our system. Based on the gyroscope data of the smartphone it is possible to recognize if the target object made a left or right turn or is still walking straight. Although the smartphone orientation data offers similar information, this data cannot be reliably used for pedestrian localization systems. This is because the geomagnetic sensors suffer heavily from small influences outgoing from the building’s characteristic or parts of the building, e.g., metal doors or metal railing, not to mention computers or monitors. The gyroscope on the other side does not suffer from such influences. This is why a left or right turn can easily be detected in the data signal of the smartphone’s yaw angle. For this, we can use a similar approach like in the step detection process:

We denote  $g_t$  as a data point of the gyroscope at the measurement time  $t$ , which tells how fast the phone is turning around a given axis and is given in  $\frac{\text{rad}}{\text{sec}}$ . We are then looking for a measurement  $g_t$ , which is bigger by a threshold  $\Delta g$  than one of the measurements in the time interval  $(t; t + \Delta t]$ . If a measurement  $g_t$  was found, a new turn can only be detected after a time  $t_{\text{noTurn}}$ . To distinguish between left and right turns, we can use the sign of  $g_t$ . If  $g_t$  is positive a left turn is detected, otherwise a right turn is detected. We have achieved good results using  $\Delta g = 1.0 \frac{\text{rad}}{\text{sec}}$ ,  $\Delta t = 300\text{ms}$  and  $t_{\text{noTurn}} = 2500\text{ms}$ . Furthermore, we assume detected turns to have an angle of  $90^\circ$ . While this might not be realistic in most cases, such a rough estimation is enough as input for the recursive state estimation. If a turn was detected, the global heading  $\alpha_i$  in  $\mathbf{o}_i$  is updated accordingly by adding or subtracting  $90^\circ$ .

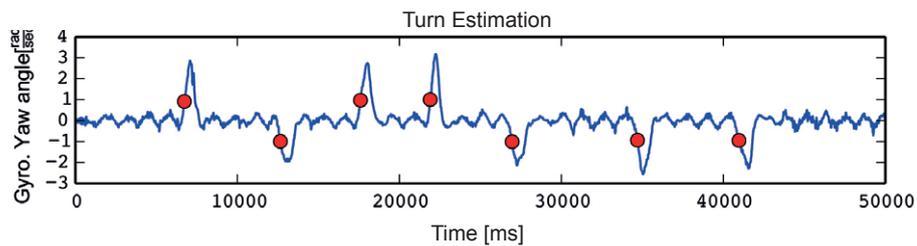


Fig. 2. The estimation of turns is shown using the gyroscope signal. Positive peaks are left turns, while negative peaks are right turns.

#### 4.2 State Transition and Weight Update

**State transition:** The state transition simulates a possible movement of a pedestrian during two time steps. A crucial information that is integrated into this model is the knowledge taken from floor maps. Utilizing this information, the movement estimation can be restricted due

to walls, doors or other obstacles. In addition to that, we assume pedestrians to move with a typical walking speed, which results in an average step length. The step detection process is integrated into the state transition model such that we differentiate between two cases when drawing a new particle:

$$\pi_t^i \propto p(\mathbf{q}_t | \mathbf{q}_{t-1}, \mathbf{o}_{t-1}) = \begin{cases} N_{\text{step}}(d_t | \mu_{\text{step}}, \sigma_{\text{step}}) & \text{if } \text{step}_t = \text{true} \\ N_{\text{noStep}}(d_t | \mu_{\text{noStep}}, \sigma_{\text{noStep}}) & \text{otherwise} \end{cases} \quad (22)$$

where  $\mu_{\text{step}}$  denotes the average step length of a pedestrian and is assumed to be 70 cm,  $\sigma_{\text{step}}$  is the standard deviation and assumed to be 50cm,  $\mu_{\text{noStep}}$  to be the mean value when no step was detected and  $\sigma_{\text{noStep}}$  to be the standard deviation in this case, and assumed to be 0 cm and 10 cm respectively.  $d_t$  is the walking distance between two time steps.

One must be aware that drawing from both distributions also takes floor map information into account. This means that particles cannot cross walls and, therefore, all distance measurements from one point to another also must incorporate this restriction [Köping et al. 2012]. In the phase of the state prediction we do not rely on any heading information, which leads the particles to spread in every possible direction.

**Weight update:** Information of direction is integrated in the weight update step. At every time step  $t$  a particle is evaluated as follows: Based upon the position  $x_{t-1}$  and  $y_{t-1}$  of the previous state  $q_{t-1}$ , a position  $x_{t,\text{exp}}^i, y_{t,\text{exp}}^i$  for each particle  $i$  is computed. This is the position for the new state at time  $t$ , if it followed a deterministic model. Essentially, this expected position is calculated with

$$x_{t,\text{exp}}^i = x_{t-1}^i + l' \cdot \cos(\alpha_t + \Delta\alpha_t) \quad (23)$$

$$y_{t,\text{exp}}^i = y_{t-1}^i + l' \cdot \sin(\alpha_t + \Delta\alpha_t) \quad (24)$$

where  $l'$  is the step length. Its value depends on whether a footstep was detected or not and, therefore, its value is either 0 cm or 70 cm.  $\alpha_t$  is the global heading information contained in  $\mathbf{o}_t$  and  $\Delta\alpha_t$  is the change of heading. The value of  $\Delta\alpha_t$  is either  $\pm 90^\circ$ , if a left or right turn was detected in the last time step, or  $0^\circ$ , if no turn was detected. Finally, we can evaluate each state using a bivariate Gaussian

$$N(\mathbf{q}_t | (x_{t,\text{exp}}^i, y_{t,\text{exp}}^i)^T, \Sigma) \quad (25)$$

where  $\Sigma$  represents the covariance matrix and its values are estimated empirically with  $\mu_x = x_{t,\text{exp}}^i, \mu_y = y_{t,\text{exp}}^i, \sigma_x = \sigma_y = 2.0$  and correlation of 0.0.

## 5 EXPERIMENTS

We tested our method with two different walking paths (see Fig. 3), the first one with a total distance of 23m and a walking time of about 36 seconds, the second with a total distance of 42m and a walking time of about 55 seconds. The sequences contain also periods, where the target object stands still to prove that our method is also capable of dealing with no movement. We walked both sequences 10 times each, to compensate for measurement variance of the sensors. All sensor data was collected with a Samsung Galaxy SII smartphone while the latter was laying flat on the hand during walking. After recording the data the particle filter computation was done on a Core i-5 2520M CPU@2.50GHz and 4.00GB RAM. All particles

were updated every 150ms with a total number of 10,000 particles. Since the calculation for one whole run only took few seconds, we expect that an implementation of our system on an up-to-date smartphone can calculate a position estimation in real time.

For estimation of the total error we evaluated the distance between the real position, where the walk ended, and the estimated position at the end of the simulation run. On average our system produces an error of about 80cm for walk 1 and 92cm for walk 2. The worst position estimation had an error of about 1.4m for walk 1 and 1.9m for walk 2, the best estimation had an error of 35cm and 30cm for walk 1 and walk 2, respectively. The ground truth data was generated by walking from one turning point to the next turning point in a straight line.

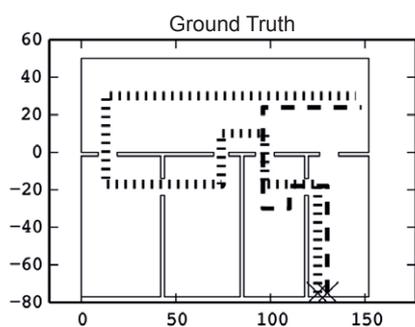


Fig. 3. Ground truth data for walk1 and walk2

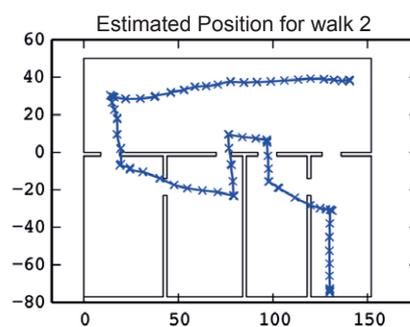


Fig. 4. Exemplary walking path for walk2

One source of error is the duration of the walk. The errors of wrongly estimated steps and step length sum up after some time. This was especially evident during a walk in a circle in the corridor, where we could not rely on map information too much. Nevertheless it should be mentioned that the main reason for the positional error is not necessarily the duration of a walk, but the missing information of walls and obstacles while in the long corridor. This is because the information of walls and doors often is enough to overcome the detection of too many or too little foot steps or a natural variation of step length. In the lower right of Figure 4 this principle can be shown. Clearly, either not enough steps are detected or the step length was estimated incorrectly. Without map information, after the first left turn, the estimated route would be a horizontal line to the left. Instead the position gets corrected to a position a few meters northwards.

Figure 4 shows an exemplary positioning estimation. It can be seen that the positional estimation corresponds with the real walking sequence, but the visualization of the estimated path crosses a wall. At the beginning some steps are detected. After a detected left turn and some newly detected steps particles try to cross the wall, which is not possible because of the map information. Nevertheless, because of the distribution of the particles over the map in the transition model some particles have the possibility to go far enough and to walk through the door. The new estimated position correctly is left of the wall, but simply connecting the estimated positions leads to a visualization of the estimated position, which suggests that the pedestrian walks through the wall. Such problems may be overcome if an appropriate smoothing algorithm will be applied [Doucet and Johansen 2011].

One major drawback of our system is the possibility that a left or right turn is not or wrongly detected. This leads to massive errors in the system because the state evaluation assigns a low weight to particles, which do not follow the expected route. In consequence these particles will most likely disappear after resampling.

## 6 CONCLUSION AND FUTURE WORK

In this paper we presented an indoor navigation system based upon the detection of steps and turns as well as on floor map information. To improve the weight update step of the particle filter we showed that it is possible to weaken the Markov assumption and explicitly integrate the information of a previous step. We showed how observation data can be used within the state transition model. To improve our system for the future, we plan to extend our current turn detection. At this time we are not capable to give a good estimation of the angle the pedestrian turned, e.g., by now we cannot detect if someone made a 180° turn. Given such information would significantly improve the current system. In addition to that we will also integrate a step length detection and enrich our system with new sensors like Wi-Fi and cameras. The additional sensors will stabilize the system in situations where a step and turn detection is only hard to achieve like in escalators and elevators.

## REFERENCES

- BAGOSI, T., BARUCH, Z. 2011. Indoor Localization by WiFi. IEEE International Conference on Intelligent Computer Communication and Processing (ICCP) (August 2011) 449-452
- WOODMAN, O., HARLE, R. 2008. Pedestrian Localisation for Indoor Environments. Proceedings of the 10th International Conference on Ubiquitous computing (September 2008) 114-123
- LINK, J.A.B., SMITH, P., WEHRLE, K. 2011. FootPath: Accurate map-based indoor navigation using smartphones. International Conference on Indoor Positioning and Indoor Navigation (IPIN) (September 2011) 1-8
- WANG, H., SEN, S., ELGOHARY, A., FARID, M., YOUSSEF, M., CHOUDHURY, R.R. 2012. No need to war-drive: unsupervised indoor localization. Proceedings of the 10th international conference on Mobile systems, applications, and services (2012) 197-210
- ZHANG, R., BANNOURA, A., HÖFLINGER, F. 2013. Indoor localization using a smart phone. IEEE Sensors Applications Symposium (SAS) (February 2013) 38-42
- KÖPING, L., MÜHSAM, T., OFENBERG, C., CZECH, B., BERNARD, M., SCHMER, J., DEINZER, F. 2012. Indoor Navigation Using Particle Filter and Sensor Fusion. Annual of Navigation 19(2) (December 2012) 31-40
- DAVIDSON, P., COLLIN, J., TAKALA, J. 2010. Application of particle filters for indoor positioning using floor plans. Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS) (October 2010) 1-4
- THRUN, S., BURGARD, W., FOX, D. 2006. Probabilistic Robotics. MIT Press
- DEINZER, F. 2005. Optimale Ansichtenauswahl in der aktiven Objekterkennung. Volume 20 of Studien zur Mustererkennung.
- ARULAMPALAM, M.S., MASKELL, S., GORDON, N., CLAPP, T. 2002. A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking. IEEE Transactions on Signal Processing 50(2) (February 2002) 174-188
- GORDON, N.J., SALMOND, D.J., SMITH, A.F.M. 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proceedings F Radar and Signal Processing 140(2) (April 1993) 107-113
- ISARD, M., BLAKE, A. 1998. Condensation - conditional density propagation for visual tracking. International Journal of Computer Vision 29(1) (August 1998) 5-28

## Generic Open Source Gamification Framework

Stefan Zink

Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt

Gamification wurde in den letzten Jahren zu einem immer größeren Hype-Thema, sodass immer mehr Unternehmen versuchen dies einzusetzen. Aufgrund von mangelnder Kenntnis und fehlerhaftem Design sind diese Umsetzungen zum größten Teil zum Scheitern verurteilt. Es existieren zudem keine allgemeinen technischen Lösungen, welche offen zugänglich und frei implementiert sind. Aufgrund dessen ist es nötig, eine allgemeingültige und freie Art der Umsetzung zu generieren, welche in dieser Arbeit umrissen wird.

Kategorie und Themenbeschreibung: IT-Driven Human Resource Management

Zusätzliche Schlüsselwörter: openGamification, Generic, Open Source, Gamification, Framework, Human Resource Management (HRM), Mechanics-Dynamics-Aesthetics (MDA), Pointification)

### 1 AUSGANGSSITUATION UND ZIELSETZUNG

Archäologische Funde beweisen, dass die Menschen schon seit über 5000 Jahren spielen. [Tyldesley 2007] Das Spielen ist demnach tief in uns verankert. Videospiele machen jedoch erst auf einfache Weise messbar, wie viel Zeit Menschen in Spiele investieren. So wurden alleine in World of Warcraft 6 Millionen Spielerjahre aufgewendet. Ein Spieler von WoW spielte 2010 durchschnittlich 22 Stunden pro Woche. Bis zum 21. Lebensjahr hat der typische Gamer bereits 10000 Stunden alleine in Online-Spiele investiert. [McGonigal 2010] Dies entspricht in etwa der Zeit, die ein Mensch bis zum Abitur in der Schule verbringt. [Zink 2014] Doch Zeit, die in spielerische Kontexte eingebracht wird, ist nicht immer wirtschaftlich verloren. Gelegentlich wurden Spielmechaniken selbst in der Antike genutzt um Probleme zu lösen. [McGonigal 2010]

Der Einsatz von Gamification ist demnach nichts Neues, kam jedoch erst in den letzten Jahren in die Schlagzeilen und damit in den Forschungsfokus (siehe Abbildung 1). In der heutigen Zeit werden wöchentlich drei Milliarden Stunden in Online-Spiele anstatt in produktive Tätigkeiten gesteckt. Das zeigt uns, welchen Status die Motivation durch Spielen einnimmt. [McGonigal 2010]

Adresse des Autors: Stefan Zink  
Würzburg, Deutschland  
[www.openGamification.org](http://www.openGamification.org)

Die Erlaubnis zur Kopie in digitaler Form oder Papierform eines Teils oder aller dieser Arbeit für persönlichen oder pädagogischen Gebrauch wird ohne Gebühr zur Verfügung gestellt. Voraussetzung ist, dass die Kopien nicht zum Profit oder kommerziellen Vorteil gemacht werden und diese Mitteilung auf der ersten Seite oder dem Einstiegsbild als vollständiges Zitat erscheint. Copyrights für Komponenten dieser Arbeit durch Andere als FHWS müssen beachtet werden. Die Wiederverwendung unter Namensnennung ist gestattet. Es andererseits zu kopieren, zu veröffentlichen, auf anderen Servern zu verteilen oder eine Komponente dieser Arbeit in anderen Werken zu verwenden, bedarf der vorherigen ausdrücklichen Erlaubnis.

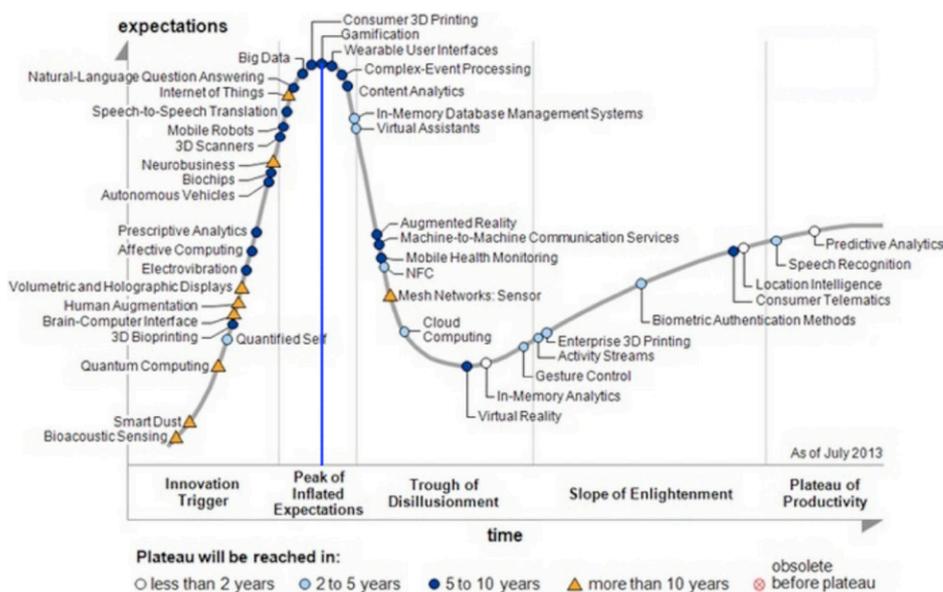


Abb. 1: Gartner Hype Cycle 2014, [o.V. 2013c]

Gamification wird zwar bereits umgesetzt, ist jedoch häufig lediglich abstrakt beschrieben, auf einzelne UseCases zugeschnitten, schlecht umgesetzt (vgl. 3.4), proprietär oder abhängig von einer Plattform oder Ressource. Um dies zukünftig zu verhindern soll deshalb ein Framework geschaffen und verwendet werden, welches durch dessen generische Eigenschaft diese Probleme in einem Open Source-Projekt lösen kann.

In dieser Arbeit wird zunächst eine Einführung zum Thema Gamification gegeben. Dabei wird eine normative Handlungsempfehlung für die Erstellung eines Gamification Frameworks anhand der vorgestellten Prinzipien erstellt. Darüber hinaus wird das openGamification Framework beschrieben, welches auf diesen Prinzipien aufbaut.

## 2 GRUNDLAGEN

Das Oxford Dictionary definiert den Begriff **Generic**: “(It is a) characteristic of or relating to a class or group of things; not specific” [Stevenson 2010] Dieses Adjektiv kann wohl am besten mit „im allgemeingültigen Sinne“ [Scholze-Stubenrecht et al. 2011] umschrieben werden. Unter der generischen Eigenschaft werden in dieser Arbeit mehrere Eigenschaften subsumiert, welche für die allgemeingültige Betrachtungsweise der Thematik notwendig sind. Die Generik des Projektes hängt aus technischer Sicht im großen Maße von der Menge der unterstützten Plattformen ab. Hierzu muss die Umsetzung möglichst plattformunabhängig sein. Da nicht alle Programmiersprachen von jeder Plattform unterstützt werden, hängt die Anzahl der unterstützten Plattformen maßgeblich von der Wahl der Programmiersprache ab, in der das programmierte Framework entwickelt wird. Um allgemein umsetzbar zu sein, darf der Einsatz nicht von einer bestimmten Ressource, wie einer aktiven Datenverbindung, abhängig sein. Des Weiteren muss auch Gamification in einer allgemeingültigen Form zur Umsetzung beschrieben werden. Dazu müssen die Grundlagen sowie die Umsetzung so formuliert sein, dass diese sich nicht auf einen speziellen Bereich beschränken. Unter einer generischen Veröffentlichung des Quellcodes wird eine freie und übliche Form verstanden um eine maximale Beteiligungsmöglichkeit von Dritten nach Abgabe dieser Arbeit zu ermöglichen.

Wissen, das nicht öffentlich ist, ist verloren. Das Oxford Dictionary erklärt **Open Source** Software als: “Denoting software for which the original source code is made freely available and

may be redistributed and modified.” [Stevenson 2010] Das bedeutet, dass der Quellcode der Software frei zugänglich offengelegt wird und dieser weiterverbreitet und modifiziert werden darf. Die deutsche Definition ergänzt dies noch um die freie Nutzung derselben. [Scholze-Stubenrecht et al. 2011] Open Source definiert sich im Wesentlichen durch folgende zehn Punkte: [Perens 2008]

1. Freie Weiterverbreitung
2. Frei modifizierbarer Quellcode
3. Freie Bildung von Zweigen unter gleicher Lizenz
4. Integritätswahrung des Sourcecodes des Urhebers
5. Keine Einschränkung von Personen oder Gruppen
6. Keine Einschränkung der Einsatzfelder
7. Freie Lizenzverbreitung
8. Keine Produktspezifizierung der Lizenz
9. Keine lizenzbedingte Beeinträchtigung
10. Beispiellizenzen: GNU GPL, MPL, BSD, Apache

Ein **Framework** ist, per Definition des Oxford Dictionary of English, “a basic structure underlying a system, concept, or text”. [Stevenson 2010] Das bedeutet frei übersetzt eine Basisstruktur, die das ihr zugrundeliegende System, Konzept oder einen Text stützt, wie beispielsweise das Konzept des Nutzerverständnisses. Im Kontext dieser Arbeit wird als Framework die reine technische Umsetzung der Spielelemente mit deren Implementierung und Design im Sinne des Softwareengineering verstanden. Bei der Referenzierung des Frameworks dieser Arbeit wird demnach das im Rahmen dieser Arbeit entstandene Projekt openGamification gemeint.

### 3 GAMIFICATION

#### 3.1 Definition und Bedeutung

Gamification bezeichnet das Nutzen von Spiel- und Belohnungsmechaniken auf normalerweise nicht-spielerische Tätigkeiten, [Deterding 2011] um Nutzer zu einem bestimmten Verhalten zu bringen oder ihr Engagement zu fördern. [Duggan K. and Shoup K. 2013] Das Wort Game bezeichnet dabei ein Spiel nach festen Regeln und ist somit kein ‚Herumspielen‘. Dazu werden, im Gegensatz zu Serious Games, lediglich einzelne Elemente eines Spieles genutzt. [Deterding 2011]

Gamification zielt darauf ab, das Engagement zu erhöhen und die oben aufgelisteten Faktoren messbar zu beeinflussen. Dass dies möglich ist, zeigt ein Beispiel, welches seit 1987 ein großer Erfolg ist: Das McDonalds Monopoly. Die Spieler haben dabei nicht nur den Anreiz auf einen schnellen Gewinn durch kostenlose Artikel, sondern ebenso den Langzeiteffekt durch das Sammeln ganzer Sets, den „Monopolies“. Dieses härter zu erreichende Ziel verspricht noch größere Preise und führt Nutzer nach eigenen Angaben dazu weiter Geld für Essen auszugeben obwohl sie nicht mehr hungrig sind. [Zichermann and Linder 2013] Abbildung 2 zeigt die Auswirkungen von Gamification, nach einer Untersuchung von Badgeville. Dabei werden verschiedene Aspekte einer Seite darauf untersucht, wie diese durch den Einsatz von Gamification gesteigert werden konnten. [Duggan K. and Shoup K. 2013]

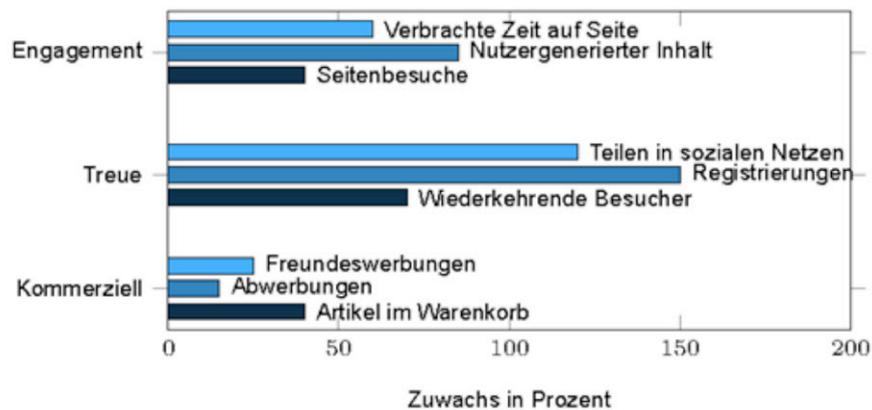


Abb. 2: Nutzerressourcenzuwachs durch Gamification

### 3.2. Motivation

Um diese Steigerung der Nutzerinteraktion zu erreichen muss der Nutzer motiviert werden. Motivation ist keine absolute Größe, die vorhanden ist oder nicht, sondern muss weiter nach ihrem Grad der Selbstbestimmung kategorisiert werden. Dies ermöglicht eine Qualifizierung der einzelnen Merkmale, wie sie im Gamification benötigt wird. Wir unterscheiden daher zwischen Amotivation, extrinsischer und intrinsischer Motivation. [Benzig and Beyerhaus 2012]

Nach der Selbstbestimmungstheorie der Motivation liegt die **Amotivation** bei Abwesenheit der Motivation vor. [Benzig and Beyerhaus 2012]

**Extrinsisch** motivieren lässt sich dadurch, dass die Handlung als Nebeneffekt einen vom Nutzer gewünschten Effekt hat. Die Belohnung kann dadurch von der Handlung separiert werden. Eine spontane Leistung wird daher bei einer rein extrinsisch motivierten Handlung nicht zu erwarten sein, da zum in Gang setzen dieser zuvor eine Belohnung versprochen werden muss. Eine Verstärkung des Verhaltens kann erfolgen, indem etwas Positives hinzugefügt wird (Belohnung), oder eine negative Eigenschaft aufgehoben wird. [Benzig and Beyerhaus 2012]

**Intrinsisch** motiviertes Verhalten kann im Gegensatz dazu spontan auftreten oder verstärkt werden, da hier die Tätigkeit um der Sache willen ausgeführt wird. Dies liegt daran, dass die Ausübung der Tätigkeit als angenehm empfunden wird. [Benzig and Beyerhaus 2012] Dazu gehören Neugier, Spontaneität, emotionaler Anreiz und Erfolgserwartung und dienen damit Zwecken der Identität, der Selbstverwirklichung, des Status, Fortschritts und des Erreichens einer Errungenschaft. [Benzig and Beyerhaus 2012; Duggan and Shoup 2013]

### 3.3 Betrachtungsweisen der Spielentwicklung

Das MDA-Framework ist eine generische Betrachtung auf die Umsetzung eines Spieles. Es lässt sich jedoch auch auf spielnahe Bereiche, wie Gamification, ausweiten. Die Abkürzung MDA steht für Mechanics, Dynamics und Aesthetics. Der Name stellt damit die drei Bestandteile des Frameworks dar. Diese werden vom Spieledesigner erstellt und vom Spieler genutzt. Dabei werden die drei Grundbausteine eines Spiels - Die Regeln, das System und Spaß - umgemünzt auf die zuvor genannten Elemente. Die Elemente sind dabei keine Schichten des Frameworks, sondern lediglich eine andere Betrachtungsweise desselben Spiels. Diese beschreiben die Betrachtungsweise des Spielers mittels der Aesthetics, über die Dynamics als hybride Sichtweise bis hin zu den Mechanics, welche vom Entwickler umgesetzt werden können. [Hunike et al. 2004]

**Aesthetics** beschäftigt sich mit der emotionalen Auffassung durch den Spieler, während dieser das Spiel spielt. Dies verkörpert den Spaß, muss jedoch etwas aufgespalten werden, um damit

weiter arbeiten zu können. Dazu verwendet man in der Regel eines oder mehrere dieser Elemente: Sie stellen die Sichtweise des Nutzers auf den spielerischen Kontext dar. [Hunike et al. 2004]

1. Sensation: Spiel zum Verwöhnen der Sinne.
2. Fantasy: Spiel mit surrealen Elementen.
3. Narrative: Geschichte verkörperndes Spiel.
4. Challenge: Spiel, bei dem Hindernisse überwunden werden müssen.
5. Fellowship: Spiel unter sozialen Aspekten.
6. Discovery: Spiel auf unerforschtem Land.
7. Expression: Spiel zur Selbstfindung.
8. Submission: Spiel zum Zeitvertreib.

Dynamics tragen die Beschreibung des Laufzeitverhaltens eines Spieles. Bei diesen wird geregelt, welcher Input welchen Output erzeugen kann. In den Dynamics trifft sich die technische Seite der Umsetzung durch den Designer mit der spielerischen Seite des Konsumenten. Mit den Spieldynamiken werden die ästhetischen Komponenten beschrieben. Eine Challenge kommt beispielsweise durch Zeitdruck oder einen gegnerischen Spieler zustande. Abbildung 3 zeigt eine Auflistung einiger Game Mechanics und Aesthetics der auf Gamification spezialisierten Firma Bunchball. Die grünen Punkte markieren hierbei die Primärabsicht der jeweiligen Mechanik. Die blauen Punkte weitere Bereiche, die durch dieses Element beeinflusst werden. [Hunike et al. 2004]

Game Mechanics	Human Desires					
	Reward	Status	Achievement	Self Expression	Competition	Altruism
Points	●	●	●		●	●
Levels		●	●		●	
Challenges	●	●	●	●	●	●
Virtual Goods	●	●	●	●	●	
Leaderboards		●	●		●	●
Giftng & Charity		●	●		●	●

Abb. 3: Interaction der Game Mechanics und Aesthetics [Bun10, S. 9]

Mechanics sind die Komponenten eines Spiels, welche dieses in Form von Algorithmen und Daten beschreiben und somit dessen genaues Verhalten bestimmen. Die Erstellung der Spielmechaniken ist der Punkt, von dem aus der Spieldesigner an die Erstellung des Spiels herangeht. In den Mechaniken wird alles technisch umgesetzt, was in den anderen Bereichen beschrieben wird. Hierzu gehören Countdowns, Punkte, Spieler, etc. Durch Anpassung der Spielmechaniken kann die Spieldynamik darüber hinaus einen Feinschliff erhalten. Selbst kleine Änderungen in diesen Regel gebenden Spielelementen haben große Auswirkungen auf das Gameplay. [Hunike et al. 2004] Die häufigsten spieltypischen Elemente sind: [Heidt 2012]

- Punktesysteme und virtuelle Güter
- Level
- Ranglisten
- Rätsel und Abzeichen

### 3.4 Kritikpunkte

Gamification ist, wie jede Technologie, mit Risiken behaftet. Aufgrund dessen werden zwei der Hauptkritikpunkte in diesem Abschnitt aufgelistet. Eine allgemeine Kritik an den typischen Spielelementen, wie sie in Gamification verwendet werden ist, dass sie Abhängigkeiten der Spieler fördern. Dies wird auch als **Exploitationware** bezeichnet. Obwohl eine Verhaltenssucht nicht Ziel des Gamificationprozesses ist, wird eine gewisse Abhängigkeit in Form des Engagement-Zyklus vom Spieldesign bezweckt. Dem Einsatz von Spielelementen wird daher unterstellt, Spieler in eine virtuelle Skinner-Box zu versetzen und somit auszunutzen. [Heidt 2013] Hierzu wird die Parallele zu den Spielelementen gezogen, wenn diese Belohnungen ausgeben, welche lediglich als vollständige Zusammenstellung interessant sind und somit den Nutzer zur ständigen Rückkehr zwingen. [Heidt 2013]

Die ausschließliche exzessive Nutzung von simpelsten Spielmechaniken wie Punkten wird pejorativ **Pointsification** genannt. [Stampfl 2012] Hierbei wird Gamification als das bloße Hinzufügen von Spielelementen an Anwendungen verstanden, ohne es geplant in einen Prozess zu integrieren. [Heidt 2013] Pointsification ist ein Mangel an Design und Implementierung einer Anwendung à la „heute machen wir mal Gamification“. Die simplen Spielmechaniken wirken dabei primär extrinsisch motivierend und fördern damit kein nachhaltiges Engagement. Um dies zu erreichen muss auch die intrinsische Motivation genutzt werden. [Heidt 2013] In der Realität kommen gamifizierte Anwendungen selten über die Vergabe von Punkten und Badges hinaus, [Stampfl 2012] sodass Statistiken wie die von Gartner, dass 80% fehlschlagen werden, [o.V. 2013d] durchaus plausibel erscheinen. Die Vergabe von Punkten und anderen Spielelementen bleibt ohne die Gestaltung ein rein extrinsischer Motivator. [Stampfl 2012] Pointsification gilt es daher durch ein gutes Design und gründliche Recherche zu vermeiden.

## 4 UMSETZUNG

### 4.1 Vergleichbare Arbeiten

Es existieren mehrere Plattformen und Frameworks, welche sich mit Gamification auseinandersetzen. Von diesen wurden in der nachstehenden Tabelle einige aufgeführt und auf die in Sektion 2 aufgeführten Prinzipien untersucht. Dabei lässt sich feststellen, dass keines der aufgeführten Systeme die gestellten Anforderungen erfüllt. Deshalb wurde das openGamification Framework geschaffen, welches im nachfolgenden Abschnitt grob erläutert wird.

	UserInfuser [o.V. 2012]	Google Play Services [o.V. 2014]	Badgeville [o.V. 2013a; o.V. 2013b]
Generic	Nein	?	Ja
Open Source	Ja	Nein	Nein
Gamification	Ja	Ja	Ja
Framework	Nein	Ja	Ja

### 4.2 Framework

Die Nutzung der verschiedenen Spielelemente wird im Rahmen von openGamification immer nach einem vergleichbaren Grundschema abgehandelt. Deshalb gibt es einen vereinfachten

Anwendungsfall, der die Grundzüge dessen beschreibt, was bei den einzelnen Spielelementen im Detail ausgeführt wird. Der Kern hiervon ist, dass wie beim MDA-Framework im Abschnitt 3.3 der Designer bzw. der Publisher ein gamifiziertes System mit integrierten Spielelementen generiert und bereitstellt, welches vom Nutzer, dem Gamer, konsumiert wird. Das Spielelement liegt dabei eingebettet in der Applikation des Publishers in Form des openGamification Frameworks. (siehe Abbildung 4) Die Nutzung eines eingesetzten Spielelementes wird im Rahmen von openGamification immer nach einem standardisierten Schema abgehandelt. Nach diesem greift ein Nutzer über die Applikation des Publishers auf den gamifizierten Inhalt zu. Hierbei steht die Ausgabe des GameElements als JSON oder HTML5 zur Verfügung um dieses weiter zu verarbeiten oder direkt grafisch auszugeben. Verwendete Spielelemente können darüber hinaus auf weiteren Spielelementen aufbauen oder durch die Nutzung eines Spielelementes entstehen. Die hierbei umgesetzten Spielmechaniken werden im Abschnitt 4.2 aufgeführt.

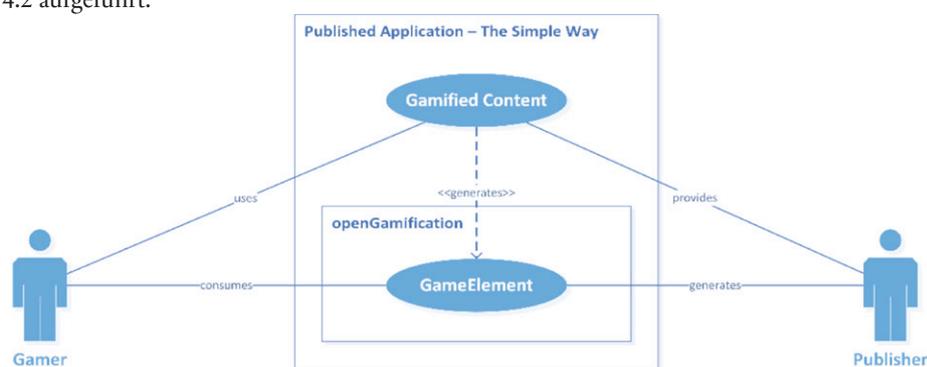


Abb. 4: Abstrakter Use Case

#### 4.3 Gamification

Es existieren nach aktuellem Stand fünf zentrale Spielmechaniken, welche ein Publisher seinen Nutzern zur Verfügung stellen kann. Diese Spielelemente werden nach einem einheitlichen Schema umgesetzt und beziehen sich auf die gemeinsame Wurzelklasse GameElement, welche den kleinsten gemeinsamen Nenner aller Spielelemente darstellt. Durch diese Systematik wird das Verhältnis von Spielelementen zu den Nutzern geregelt. Dies umfasst eine Referenz auf einen Publisher sowie eine Menge von Gamern, welche einem Objekt des Spielelements zugeordnet werden können. Kindklassen von GameElement sind nach aktuellem Stand die häufigsten Spielmechaniken des Gamedesigns [Heidt 2013] nach dem Beispiel zum MDA Framework in 3.3. Abbildung 3 macht ersichtlich, welche menschlichen Bedürfnisse durch die folgenden Spielmechaniken erfüllt werden können.

- Points zum Sammeln oder als virtuelle Währung
- Leaderboards
- Quests/Challenges
- Levels
- Achievements

Spielmechaniken sind nicht zwangsläufig eigenständige Entitäten. So werden Punkte typischerweise in allen Instanzen anderer Mechaniken wiederverwendet oder Achievements als Belohnung auf einen erfolgreich absolvierten Quest ausgegeben. Durch diesen geschachtelten Aufbau werden weniger komplexe Spielmechaniken automatisch bei diffizileren generiert.

#### 4.4 Generic

Das Framework openGamification kann auf zwei Arten integriert werden und ist dadurch vielseitig einsetzbar. Es kann konventionell in eine Applikation integriert werden. Bei dieser

Methode sollte das Framework als gekapseltes Modul innerhalb der Applikation verwendet werden. Die Daten sind dabei lokal verfügbar und können unabhängig von einer Netzwerkverbindung verwendet werden. Die Daten sind jedoch auf dem Gerät gespeichert und somit vom verwendeten Gerät abhängig. Hierzu muss eine Programmiersprache verwendet werden, welche offline auf jedem üblichen Anwendungssystem funktioniert. Eine weitere Möglichkeit ist der serverseitige Einsatz von openGamification. Dazu wird es in die Plattform integriert und über diese angesteuert. Die Daten sind dadurch serverseitig verfügbar, unabhängig vom verwendeten Gerät. Die Nutzung beschränkt sich jedoch auf Anwendungen mit vorhandener Datenverbindung zum Server. Um dies umzusetzen, muss die eingesetzte Programmiersprache so gewählt werden, dass diese serverseitig genutzt werden kann. Als Zielsprache wurde daher JavaScript gewählt, da diese als einzige auf allen untersuchten Plattformen zur Anwendungs- oder Plattformentwicklung zur Verfügung stand. [Zink 2014]

#### 4.5 Open Source

Unter Sektion 2 wurden zehn Aspekte von Open Source aufgelistet, welche durch die Distribution umgesetzt werden müssen. Hierzu wurden im Rahmen des openGamification Projektes Plattformen zur Distribution, sowie ein Open Source Lizenzmodell ausgewählt.

Eine Studie untersuchte 2011 Open Source Entwickler auf ihre präferierte Plattform. Nach der Auswertung dieser lässt sich sagen, dass die bevorzugte Plattform, je nach Betriebssystem, Codeplex und Github sind. [o.V. 2011] Um möglichst viele Entwickler abzudecken ohne zu viele Plattformen verwalten zu müssen, wurde die Distribution auf beiden genannten Plattformen gestartet. Dabei wurde diese zur freien Distribution unter der Apache 2.0 Lizenz veröffentlicht, welche die Anforderungen der Open Source Initiative aus Abschnitt 2 erfüllt.

## 5 FAZIT UND AUSBLICK

Zukünftig wird sich jede Arbeit quasi von alleine erledigen, da wir diese nicht mehr als solche wahrnehmen. Diese Worte stellen eine utopische Sichtweise dar, welche das theoretische Potential von Gamification widerspiegelt. Prognosen sagen jedoch, dass Gamification, zumindest in den nächsten Jahren, nicht ein derartiges Wunderwerk vollbringen wird. [o.V. 2013c] In einem Interview mit Gartner Ende 2013 bestätigte das Institut seine Ansicht, dass Gamification in 2014 in die Phase der Desillusionierung gelangen wird. Das bedeutet, dass Organisationen ihren Fokus auf den erfolgreichen und praktischen Einsatz legen werden. Gartner prognostiziert dabei, dass Gamification auch 2020 noch eine Rolle spielen wird und spekuliert auch auf eine längere Zeit. [o.V. 2013d] Der übermäßig produktive Einsatz wird jedoch voraussichtlich erst in fünf bis zehn Jahren sein. [o.V. 2013c]

Das Ziel, ein technisches Framework auszuarbeiten, welches zum Einsatz von Gamification genutzt werden kann und dabei die Prinzipien von Open Source erfüllt sowie einen generischen Charakter hat, wurde erreicht. Um dies zu erfüllen wurde das Framework beschrieben, organisiert, implementiert und für die weitere Verwendung aufbereitet. Damit wird versucht, einen zuvor genannten, produktiveren Einsatz von Gamification zu ermöglichen, indem Standards der Implementierung geschaffen werden. Diese sollen sich durch die freie Anpassungsmöglichkeit einer Community in den Folgejahren elaborieren und hierzu die generische Einsetzbarkeit auch bei neuen Technologien weiter ausbauen. Um dies zu ermöglichen müssen jedoch zunächst Beiträge einfließen, die den vollen Umfang von Gamification in dessen theoretischer wie praktischer Natur abdecken. Der frühe Entwicklungsstand von Gamification verkompliziert eine genaue Abschätzung von zukünftigen Problematiken und Einsatzmöglichkeiten. [o.V. 2013c]

## REFERENZEN

BENZIG M. AND BEYERHAUS C. 2012. Bedeutung von Gamification für eine nachhaltige Konsumentenverhaltensänderung, S. 8-9

DETERDING S ET AL. 2011. Gamification: Toward a Definition, S. 2

DUGGAN K. AND SHOUP K. 2013. Business Gamification For Dummies

HEIDT M. 2012. Gamification - Grundlage der Implementierung spieltypischer Elemente im Wirtschaftskontext, S. 15-64

HUNIKE R. ET AL. 2004. MDA: A formal approach to game design and game research

MCGONIGAL, J. 2010. Gaming Can Make A Better World, TED Conferences

O.V. 2012. Userinfuser API Documentation, Google.com, URL [https://code.google.com/p/userinfuser/wiki/API\\_Documentation](https://code.google.com/p/userinfuser/wiki/API_Documentation), Aufgerufen am 20. März 2014.

O.V. 2011. Userinfuser API Documentation, Google.com, URL <http://blogs.msdn.com/b/codeplex/archive/2011/07/11/survey-results-open-source-developer-preferences-june-2011.aspx>, Aufgerufen am 23. März 2014.

O.V. 2013A. Development Tools & Visualisations, Badgeville.com, URL <http://badgeville.com/products/developer-tools-apis>, Aufgerufen am 27. März 2014.

O.V. 2013B. Game Mechanics. Badgeville.com, URL <http://badgeville.com/products/gamification>, Aufgerufen am 27. März 2014.

O.V. 2013C. Gartner's 2013 Hype Cycle for Emerging Technologies Maps Out Evolving Relationship Between Humans and Machines. Gartner.com, URL <http://www.gartner.com/newsroom/id/2575515>, Aufgerufen am 15. Mai 2014.

O.V. 2013D. Gartner Sticks to its Failing Gamification Prediction, Cmswire.com, URL <http://www.cmswire.com/cms/social-business/gartner-sticks-to-its-failing-gamification-prediction-023464.php>, Aufgerufen am 15. Mai 2014.

O.V. 2014. Google Play Game Services, Google.com, URL <https://developers.google.com/games/services>, März 2014, Aufgerufen am 10. April 2014

PERENS B. 2008. The Open Source Definition, S. 6

SCHOLZE-STUBENRECHT W. ET AL. 2011. Duden Deutsches Universalwörterbuch, Bd. 7, Dudenverlag Mannheim/Zürich

STEVENSON A. 2010. Oxford Dictionary of English, Bd. 3, Oxford University Press

STAMPFL N. 2012. Die verspielte Gesellschaft - Gamification oder Leben im Zeitalter des Computerspiels, S. 113-114

TYLDESLEY J. 2007. Egyptian Games and Sports (Shire Egyptology), S. 15

ZINK S. 2014. Generic Open Source Gamification Framework, Bachelorthesis, Hochschule für Angewandte Wissenschaften Würzburg-Schweinfurt

## Modeling gamification rules with ontologies for achieving a generic platform architecture

Patrick Mennig

Fakultät Informatik und Wirtschaftsinformatik

Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt

Gamification is an emerging trend for both, consumer and enterprise applications. The latter feature additional complexity due to their often highly distributed and interlaced nature. Herzig et. al. addressed this problem and proposed a generic platform for enterprise gamification. Based on their work, the problem of formulating gamification rules is taken up in this paper. Ontologies and semantic technologies are proposed as a possible solution. An ontology for gamification rules is created. Its applicability to solve the problem is validated with a thought experiment.

Categories and Subject Descriptors: Ontologies, Predicate Logic and Semantic Networks

Additional Key Words and Phrases: Gamification, Gamification Rules

### 1 GAMIFICATION IN ENTERPRISES

Playing is a popular human activity. The term ‘Play’ can be defined in various ways. An interesting aspect is the gratuitousness and the feeling of joy and fun [Kark 2011]. Not alone based on the roots in human history, applying game aspects and mechanics to – in the first place – atypical situations is an emerging trend. The rise began around 2010, mostly coined under the term gamification. Other variants are for instance “productivity games”, “applied gaming” or “playful design” [Deterding et al. 2011]. In its 2011 hype cycle, the technology advisory company Gartner introduced gamification as a technology trigger [Petty and Goasduff 2011]. Despite gamification being a new trend in HCI, one will find several practical implementations. Among the popular examples is ‘Foursquare’, a platform to interact with friends and strangers based on check-ins to various places [Anon. 2014]. One aspect of which is said to have led to its success (50 million users worldwide by May 2014) is the immersive use of gamification in its application. One may earn points, badges and other rewards for using the application and testing the boundaries of what is possible, finding secret rewards [Zichermann and Cunningham 2011]. Not only consumer applications benefit from gamification, but enterprises seek the integration of these concepts to increase employee engagement and productivity. The enterprise segment in the gamification market is projected to exceed the consumer segment by the end of 2013 [Peterson 2012]. Newer data could not be accessed due to high prices.

Author’s address:

Patrick Mennig, University of Applied Science Würzburg-Schweinfurt, Germany.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than FHWS must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission.

Gamification is often referred to as the “use of video game elements in non-gaming systems to improve user experience (UX) and user engagement” or the “use of game design elements in non-gaming contexts” [Deterding et al. 2011]. Another definition delineates it as the “process of game-thinking and game mechanics to engage users and solve problems” [Zichermann and Cunningham 2011]. The common denominator in all is the term ‘game’ or variants of it. Caillois distinguishes between two types of play activities: “paidia” and “ludus”, the former being more expressive and free form (referring to ‘play’) whilst the latter is bound to rules and a competitive rivalry (referring to ‘game’)[Caillois 1961]. Mapping this to the available definitions of gamification shows that ‘game’ other than ‘play’ was chosen correctly, as any gamification implementation has to include certain rules.

Seeking to define a generic platform architecture for enterprise gamification, Herzig et. al. introduce a dualism between business and game rules. They demand an event engine, capable of describing work artifacts in a way, that any relevant action throws events  $E = (u, x, p)$  with  $u$  being the user doing the action and  $x$  the step in a given process  $p$ . The events are routed to the business rule engine, which stores game rules  $R = (r_1, \dots, r_n)$  and processes received events  $E$  by choosing which rule  $r_i \in R$  should be applied. In their work, each rule is defined as  $P \Rightarrow C$ , with  $P$  being the premise and  $C$  the consequence to be executed (e.g. rewarding a certain amounts of points to the triggering user  $u$ ). They further demand the rules to be “creatable and changeable for business users without technical background” allowing dynamical changes at runtime. Herzig and Schill propose a domain specific language, based on a taxonomy or ontology of game rules and mechanics [Herzig et al. 2012].

An ontology itself is a formalism that allows finding knowledge in data, construe it and enables a computer system to act based on semantic information. Taking the unavailability of such generic gamification platform and the coherence of ontologies being able to act based on generic semantics, leads to hypothesis H1 by abduction.

*H1.* A generic gamification architecture can be implemented with the usage of semantic technologies and ontologies.

In this paper, one valid ontology for concepts of ‘game’ relevant for enterprise gamification is crafted. An approach to define game rules as instances of ontology concepts is shown. It takes into account the propositional logic of the ontology itself to express gamification rules. The usage of an ontology as a base also helps to reduce overhead for creation of partial similar rules. To allow evaluation, requirements towards such an implementation, extracted from literature analysis, are shown and lead to a thought experiment, inductively validating the approach. Due to the restricted time available, the study abstains from a full deductive approach and limits itself to the inductive thought experiment.

## 2 CREATION OF AN ONTOLOGY FOR GAMIFICATION

Ontologies seek to define unambiguous representations of relations, symbols, real world objects, concepts (or classes) and their characteristics. The goal of those is to allow processing of information and data to understand its meaning and inherent semantics to generate knowledge. Allowing computer systems to understand data rather than parsing it following strict schemas, might finally lead to intelligent behavior. It is still required to add metadata to information, describing the context and semantics of given information [Dengel and Bernardi 2012]. Formalizing knowledge is called knowledge representation. Instructions for conclusions are

called inference rules. Facts describe the domain and represent the inherent knowledge. See listing 1 for three example rules. They are written as ‘subject-predicate-object’ sentences with, for instance ‘player’ being the subject, ‘actor’ the object and ‘is an’ the predicate. All rules follow this formalized syntax, in other words a formal specification.

The relations (predicates) in (A) and (C) are for humans obviously transitive, leading to the conclusion (E). That is the derivation of yet unknown from established knowledge and itself a fact. With that said, inference is fundamental and formal reasoning. The formality required to use it in computer systems is given by explicitly expressing the transitive nature of such relations. The rule (F) defines this and allows inference mechanisms in programs [Dengel et al. 2012].

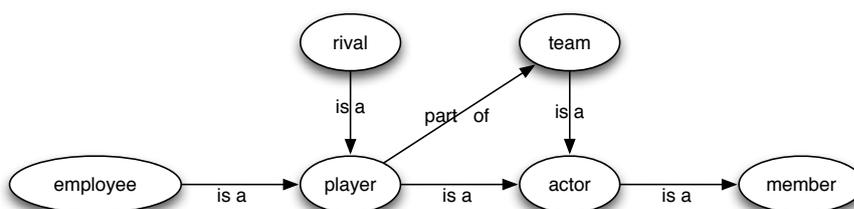


Fig. 1. Example of a game ontology as directed graph

*Listing 1: Example facts*

// facts given a-priori

- (A) Player is-a actor.
- (B) Team is-a actor.
- (C) Jon is-a player.
- (D) Jane is-a administrator.

// conclusion based on transitive relations (A) and (C)

(E) Jon is an actor.

// rule necessary for formal reasoning

(F) if x is -a y and y is -a z then x is -a z

Representations contain assertions about a given domain. In order to do this, facts, in form of objects and relations, and rules are defined. Inference mechanisms allow to conclude further facts. Mathematical logic is used to abstract mechanical knowledge gaining from human thinking [Dengel et al. 2012]. Consider the facts in listing 1 for an example: Let Jon be a sales agent, creating a new lead for his company. With just enough metadata a computer system could be able to conclude that he is participating in gamification while the actions of Jane might not be relevant for point calculation.

To give computer systems access to knowledge, it has to be projected from reality to a formalized semantic representation. The selection of concepts and properties modeled is defined by the goal of the system itself. Hence the resulting semantic model is less complex than the original domain through inherent abstraction. Still the structure and semantic of the resulting representation is to the greatest extent similar to the observed domain section [Dengel et al. 2012]. An ontology is the explicit specification of such domain concepts. The following findings F1 and F2 summarizes above remarks.

*F1.* Knowledge representation as an ontology allows to formalize domain concepts in a way that it is understandable for humans and computers.

F2. The formalized representation allows to conclude knowledge and facts from the given specification and (meta-)data.

The techniques to visualize and note ontologies vary. An ontology can be displayed as a semantic net, which itself is a graph  $G = (V, E, l, L, s, t)$ .  $V \neq \emptyset$  is the set of objects  $O = (o_1, \dots, o_n)$  and concepts  $C = (c_1, \dots, c_n)$ ,  $E \neq \emptyset$  is the set of relations between objects  $O$  and/or concepts  $C$ . The function  $l : V \cup E \rightarrow L$  assigns objects and/or concepts a name out of  $L \neq \emptyset$ . Further  $s : E \rightarrow V$  and  $t : E \rightarrow V$  are functions referencing the start and end vertex  $V$  of an edge  $E$ . Considering an edge  $e$ ,  $t(e)$  is the successor of  $s(e)$ . A directed path  $P$  within the graph  $G$  is the sequence  $e_1, \dots, e_n$  so  $s(e_i = n_i)$  and  $t(e_i = n_{i+1})$  for  $i \in \{1, \dots, n - 1\}$  [Stuckenschmidt 2009]. See figure 1 for an example of an ontology represented as a graph. This technique is human-readable and allows a quick reference and changes. The threshold for domain users to accept and work with graphs in a visual way is low [Rostanin and Weber 2012]. It is expected that domain experts and business users with basic understanding of formalisms used are able to create and maintain such ontologies, with regards to their domain understanding. This leads to the following finding F3.

F3. Knowledge representation as an ontology allows non-technical users to create and maintain computer-understandable specifications.

The same ontology can be expressed in various description languages like OWL, RDFSchema, F-Logic or SHIQ. Some of them are more restrictive than others, but the common point is the possibility to formalize ontologies and make them queryable [Baader et al. 2004]. They require a specific amount of knowledge about the technique itself which might not correlate with business users level of skills to stay maintainable. The tools used to express and visualize concepts and relations might affect this threshold positively. As the tool support depends on the individual decisions made in enterprises, this aspect is not further discussed here.

A straight way to translate semantic nets into logic formulas is based on the direct correspondence between graphs and relations, which itself can be expressed as a predicate. Each edge (i.e. an instance of a relationship) is understood as an atomic formula with two terms. The relationship ‘is-a’ would be transformed into the formula  $\text{is-a}(x, y)$  which results in *true* or *false*, depending on  $x$  and  $y$ . The concatenation of all atomic formulas in a given graph is done via the conjunction operator  $\wedge$ . The full example in figure 1 would result in the following formula (1).

$$\begin{aligned} & \text{is-a}(\text{employee}, \text{player}) \wedge \text{is-a}(\text{rival}, \text{player}) \\ & \wedge \text{is-a}(\text{player}, \text{actor}) \wedge \text{is-a}(\text{team}, \text{actor}) \\ & \wedge \text{is-a}(\text{actor}, \text{member}) \wedge \text{part-of}(\text{player}, \text{team}) \end{aligned} \quad (1)$$

Transforming this concrete formula into an universally valid deduction rule is done by replacing concrete instances with variables and adding operators for ‘exists’  $\exists$  and ‘for all’  $\forall$  increases the value of those. (2) and (3) are rules taken from the example in figure 1. (2) describes the transitive rule that each employee is a player and each player is an actor, hence each employee is an actor. The second rule describes the part-of relationship between a player and a team which again leads to transitive behavior.

$$\forall x, y, z \text{ is-a}(x, y) \wedge \text{is-a}(y, z) \rightarrow \text{is-a}(x, z) \quad (2)$$

$$\forall x, y, z \text{ part-of}(x, y) \wedge \text{is-a}(y, z) \rightarrow \text{is-a}(x, z) \quad (3)$$

A shorter form to show transitive inclusion in classes is the implication by using concept names as unary predicates on concrete instances expressed by variables. The path  $P = (e_1, e_2)$  with  $t(e_2)$  being transitively dependent of  $s(e_1)$  can be found in figure 1 as following:  $\forall x \text{ Employee}(x) \rightarrow \text{Player}(x) \wedge \forall x \text{ Player}(x) \rightarrow \text{Actor}(x) \wedge \forall x \text{ Team}(x) \rightarrow \text{Actor}(x)$ . Characteristic, i. e. necessitative properties of concepts are displayed as  $\forall x C(x) \rightarrow \phi(x)$  which reads as ‘for all instances of a concept C applies the property phi’. Sufficient properties are expressed as  $\forall x \phi(x) \rightarrow C(x)$ , reading as ‘for all instances with the property phi applies that they belong to the concept C’. Necessitative and sufficient properties are expressed as  $\forall x C(x) \leftrightarrow \phi(x)$ . Cardinality can be expressed by combining the exists operator with binary comparison, for example:  $\forall x \text{ Team}(x) \rightarrow \exists \geq a y \text{ part-of}(x, y)$ ,  $a = 2$  describes that a team consists of at least two members for whom applies the part-of relation [Stuckenschmidt 2009].

Using the possibilities of predicate logic to further formalize ontologies, it is possible to create powerful formulas. For a given instance  $i$  of a concept  $C$ , one can decide whether transitive relationships or properties apply. By given properties, one can decide whether an instance  $i$  is part of a concept  $C$  or not. Hence, by formulating an ontology and parsing it into description logic it is possible to process it with computer programs. This leads to finding F4 and F5.

F4. Ontologies visualized as semantic nets (graphs) can be parsed into formulas.

F5. These formulas can be used to decide for membership of classes and check necessitative/sufficient properties.

The findings F1-F5 lead to the suggested approach to implement a gamification rule engine, by creating a basic ontology for game and gamification concepts. To assert this approach, two further hypotheses are constructed:

H1.1. The ontology will be understandable for both, business users and information systems as well. Any adjustments that need to be made to an existing rule system will be done in a graphical environment, thus reducing the inhibition threshold of users entrusted with maintaining gamification rules.

H1.2. Transitive rule definitions can be avoided by inheritance mechanisms of ontologies, e.g. ‘salesman is a actor’ and ‘Jane Doe is a salesman’) ‘Jane Doe is a actor’. Hence all rules for actors apply to Jane.

Stuckenschmidt proposes a process for creating an ontology. He differentiates between the conceptual level of an ontology and its formal level. The first targets the structure and terminology, while the latter focuses on the formal depiction in a valid description language. No coherent procedure model has been established yet, but his proposed process is an iterative approach of several steps that allow cyclic refinement [Stuckenschmidt 2009].

At first, the to be modeled section of the domain has to be defined in order to restrict the boundaries. Taking a closer look at what shall be modeled shows ‘game’, ‘gamification’ and ‘enterprise’ as the root domains. It is unnecessary to model all aspects of these, keeping in mind that certain specific aspects could be added once the need arises. Approaching what is not to be modeled helps reducing domain complexity. Bista et. al. propose a gamification model  $G$  as follows:  $G = \langle M, A, C, P, B, R_p, R_b \rangle$  while  $M$  is a member of a given community,  $A$  is the set of possible actions one can perform in a context  $C$ . He earns points  $P$  and badges  $B$  by rules  $R_p$  and  $R_b$  [Bista et al. 2012]. The rules  $R_p$  and  $R_b$  will be implicitly and explicitly contained in the ontology itself and serve – as well as all other components of the seven tuple – as the constraints of domain concepts. The ontology will focus on members, their possible actions in respective contexts and the possibility to earn badges. Other concepts like

challenges or quests should not be considered in order to reduce complexity [Zichermann and Cunningham 2011].

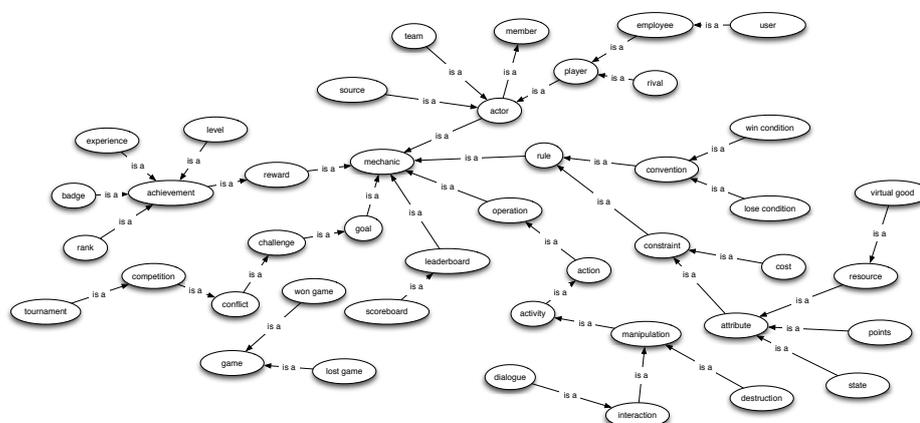


Fig. 2: Class hierarchy of game ontology.

Following the proposed process requires setting up competency questions. These are exemplary but still constructivistically for the resulting ontology. Defining beforehand what knowledge is desired to be contained. To create the competency questions domain experts are interviewed. This is done in form of an online questionnaire published on [www.boardgamegeek.com](http://www.boardgamegeek.com), a community for board game enthusiasts. To focus on participants with high expertise in the field of games and game rules, their experience is recorded and only those with at least three years of regular game consumption are taken into consideration. One objective of the expert interviews is to gather competency questions, describing beforehand what the ontology should be able to answer afterwards. Additional questions are found through literature research. The following list provides an overview of identified ones.

- Q1. What is/are the goal(s) of the game?
- Q2. Is a given X a player?
- Q3. What are the consequences of action Y from player X.
- Q4. Is player X part of a team?
- Q5. Which rank has player X on a leaderboard Z?

Ensuing from these interviews and the study of literature, relevant terms of the domain are identified. Besides publications covering gamification ([Herzig et al. 2012], [Bista et al. 2012], [Zichermann and Cunningham 2011]), games ([Caillois 1961]) and game design ([Schell 2008]) rulebooks of board games are taken as research sources. The major findings are listed together with information about the sources and the frequency, i.e. how often they are mentioned across different sources. These are then sorted into concepts and relationships by figuring out which represent sets of objects and which not. This methodology lacks formalism, however it is best practice for setting up ontologies [Stuckenschmidt 2009].

After setting up the list of relevant terms and sorting out concept candidates, the class hierarchy is built. By asking 'all X are Y', subsumption between concepts has to be noticed. This is constructivistically motivated and the result cannot be understood as 'the' class hierarchy rather 'one possible'. This should not be considered as a drawback since this accounts for the flexibility in using ontologies to model concepts and relationships. Figure 2 shows the identified hierarchies of concepts. It does not yet include any other relationship than 'is a'. Hence it only accounts for the understanding of subsumption between concepts. The complete hierarchy

is shown in figure 2 and its full discussion would leave the scope of this paper. An interesting aspect that has to be mentioned is that all concepts except for 'game' are subconcepts of 'mechanic'. The targeted ontology shall reflect games against the background of gamification in enterprise contexts. 'Mechanics' as the parts and pieces of any gamification implementation split into 'rewards', 'goals', 'operations', 'rules', 'leaderboards' and 'actors'. Reconsidering the defined boundaries  $G = \langle M, A, C, P, B, R_p, R_B \rangle$  [Bista et al. 2012] shows high conformity.

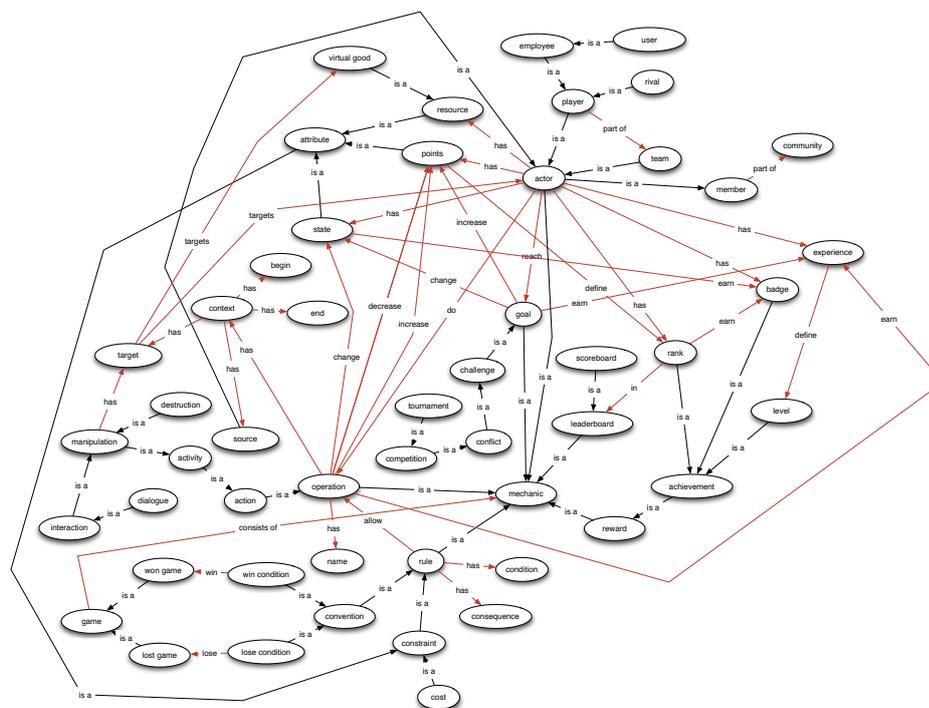


Fig. 3: Full created gamification ontology.

Identified terms that did not match well into concepts are considered candidates for relationships. Their set is not limited to the list of terms as characteristic attributes are considered. With respect to the class hierarchy, this targets common properties of all members of a concept as well as attributes unique to members of a concept or unique when compared to other concepts. From the point of view of philosophy, these properties might be intrinsic or extrinsic. The first are constructivistically for a given object, losing such attributes means losing the very own character. The latter are those properties that are given to an object, but might change over time, without affecting the objects character [Stuckenschmidt 2009]. Both types are contained in the final ontology but used equitable. The closeness to IT-systems allows to consider extrinsic attributes being 'nullable' while intrinsic properties force a value other than 'null'. Relationships are attached as high as possible in the concept hierarchy to account for inheritance and minimize repetition.

Once the relationships between concepts are taken into consideration, the ontology's complexity increases. The terms identified as relationships but yet unassigned between concepts are added. Defining beforehand which are applicable for certain concepts limits the possibilities. Figure 3 shows the ontology for gamification. One may notice certain aspects left out, like a property 'name' for 'player'. This is due to the fact that the ontology describes concepts, not instances and the 'name' of a 'player' X as concept is irrelevant. Concrete implementations might add certain aspects typical for IT systems. Checking back with the competency questions

proves as an assertion of the created result. Considering Q1, the answer could be ‘the goals are mechanics of the game’ which is correct but unrewarding. Adding concrete instances of goals as well as quantification operators will result in a true assertion of (4). Q2 can be proved answerable the same way. Checking the class hierarchy and the existence of intrinsic properties allows to decide whether a given  $v$  is a player or not. The identical approach leads to the conclusion that the ontology created, can answer the competency questions Q1-Q5 and proves usefulness for the desired goals.

$$\begin{aligned}
 \forall w, x, y, z \text{ Goal}(w) \rightarrow \text{Mechanic}(w) \\
 \wedge \exists \geq^1 x \text{ Goal}(w) \text{ change}(w, \text{State}(x)) \\
 \wedge \exists \geq^1 y \text{ Goal}(w) \text{ earn}(w, \text{Experience}(y)) \\
 \wedge \exists^1 z \text{ Goal}(w) \text{ increase}(w, \text{Points}(z)) \quad (4)
 \end{aligned}$$

The ontology is still to be translated into predicate logic for formal reasoning. To take this into account, quantification operators as in (4) have to be added. Ontology engineering environments like Protégé assist the user to a great extent [Mizoguchi 2004]. Custom toolsets integrated in the domain software business users are familiar with, are imaginable. The targeted induction, i.e. the thought experiment following does not require to have these operators present for all relationships. Hence it is refrained from adding them to the whole ontology.

Summarizing previous section, ontologies have been examined for aspects of usability and convenience to fulfill requirements. F1-F5 led to the hypotheses H1.1 and H1.2. In order to approve or refuse them, an ontology for gamification has been created. The following section will investigate gamification implementation requirements and construct one concrete case for a thought experiment. This will indicate towards validation of H1.1 and H1.2.

### 3 GAMIFICATION THOUGHT EXPERIMENT

The posited approach of using gamification rules based on semantics in enterprise software and ontologies shall be evaluated for one specific case with a thought experiment. Referring to the outset, the proposed architecture of Herzig et. al. is taken into consideration. They demand an event system with a message broker and a separate gamification rule engine reacting once receiving events. The goal is lose coupling of components and instant feedback [Herzig et al. 2012]. Active broadcasting of events is mainly based on the demand for decoupling single system elements and shall be replaced by adding semantic information to system logs, database tables and existing events. The requirements towards gamification highly vary, depending on the source used, from points, redeemable points, social points, badges, leaderboards, player interaction, level, progress, challenge, reward, instant feedback, story [Scheiner et al. 2012] [Zichermann and Cunningham 2011] [Hamari et al. 2014] to virtual goods, teams, time pressure, gift and marketplaces [Herzig et al. 2012]. Due to the limited boundaries of the ontology fixed to  $G = \langle M, A, C, P, B, R_p, R_b \rangle$  [Bista et al. 2012] the following case is constructed:

Observed is a company  $Co$  that has  $n$  employees  $E_{Co} = \{E_1, \dots, E_n\}$  working with an enterprise software system  $S_{Co}$ . They perform actions  $A = \langle S_{Co}, E_x, C \rangle$  within  $S_{Co}$ . The context  $C$  describes the source, the target and temporal aspects of  $A$  and is added together with semantic information to a specific logfile  $L_{S_{Co}}$  that is accessible by  $S_{Co}$  and the gamification rule engine  $G$ . It has a set of rules  $R_G = \{R_b, R_p\}$  defined as instances of the concept ‘rule’, a leaderboard (instance of concept ‘leaderboard’)  $Lb_C$  and a list of  $m$  badges (instance of concept ‘badge’)  $B_{E_{Co}} = \{B_1, \dots, B_m\}$  each employee has earned.

Assume Jane Doe is the employee  $E_1$  at Co. She administrates  $G$  and creates a new instance  $R_{p_1}$  of 'rule' covering  $A_1$ .  $E_1$  utilizes a graphical environment that shows the existing ontology's concepts.  $E_1$  is able to add the new instance via drag-and-drop or a wizard. Once  $R_{p_1}$  is created Jon Doe,  $E_2$  starts working in  $S_{Co}$ . He is working in sales and has a new potential customer, hence he generates a lead ( $A_1$ ) in  $S_{Co}$  which outputs a new entry containing  $C$  in  $L_{S_{Co}}$ .  $G$  detects the change in  $L_{S_{Co}}$  and parses the entry. Due to semantic information,  $G$  finds the following knowledge:  $E_2$  is an employee, hence he is an actor, hence he has an amount of points  $P_{E_2} = 0$ ;  $R_{p_1}$  allows  $A_1$ ;  $A_1$  is an operation which increases points;  $C$  determines that  $E_2$  has ended  $A_1$  in less than 5 minutes. Based on this knowledge  $G$  decides to increase  $P_{E_2} += 10$ . The change in  $P_{E_2}$  defines  $E_2$  in  $Lb_{G_1}$  and allows him to advance one rank, overtaking another colleague.

Above thought experiment inductively shows for one concrete case the applicability of semantics and ontologies for gamification rule engines. Therefore H1.1 cannot be rejected – at least in this case it has to be considered as valid. For a general acceptance, a correct deduction has to be done. This paper also shows that H1.2 can be accepted in this case. Inheritance mechanisms used in the semantic net ('is-a') allow automatic parsing of transitive relationships. As H1.2 was mainly accepted due to findings in literature, it is suggested to be accepted generally until falsified. This case also allows to accept H1, as within the defined boundaries, the suggested approach was able to fulfill the requirements. To further support H1, one should do an implementation and perform field experiments.

#### 4 PROSPECTS ON A GENERIC GAMIFICATION RULE PLATFORM

As discussed in the beginning, gamification is an emerging trend, increasing the need for software support. Reducing coupling between enterprise systems points towards a generic gamification platform approach. This paper investigated one possibility towards a generic gamification rule platform by using ontologies and semantics to loosely couple the involved systems. The hypotheses crafted in section 1 and 2 could be accepted for now, based on an induction with a thought experiment. It is necessary to further investigate ontologies as an approach to generic platforms, in order to falsify or back up the proposed hypotheses.

## REFERENCES

- Anon. 2014. About foursquare.
- Baader, F., Horrocks, I., and Sattler, U. 2004. Description logics. In Handbook on Ontologies. International Handbooks on Information Systems. Staab, S. and Studer, R., 3–28.
- Bista, S., Nepal, S., Colineau, N., and Paris, C. 2012. Using gamification in an online community. In Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2012 8th International Conference on. 611–618.
- Caillois, R. 1961. Man, Play, and Games. Sociology/Sport. University of Illinois Press.
- Dengel, A. and Bernardi, A. 2012. Einleitung. In Semantische Technologien. Grundlagen - Konzepte - Anwendungen, A. Dengel, Ed. Spektrum, Heidelberg, 3–19.
- Dengel, A., Bernardi, A., and van Elst L. 2012. Wissensrepräsentation. In Semantische Technologien. Grundlagen - Konzepte - Anwendungen, A. Dengel, Ed. Spektrum, Heidelberg, 21–72.
- Deterding, S., Dixon, D., Khaled, R., and Nacke, L. 2011. From game design elements to gamefulness: Defining “gamification“. In Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. MindTrek ’11. ACM, New York, NY, USA, 9–15.
- Deterding, S., Sicart, M., Nacke, L., O’Hara, K., and Dixon, D. 2011. Gamification. using game-design elements in non-gaming contexts. In CHI’11 Extended Abstracts on Human Factors in Computing Systems. ACM, 2425–2428.
- Hamari, J., Koivisto, J., and Sarsa, H. 2014. Does gamification work? a literature review of empirical studies on gamification. In Proceedings of the 47th Hawaii International Conference on System Sciences.
- Herzig, P., Ameling, M., and Schill, A. 2012. A generic platform for enterprise gamification. In Software Architecture (WICSA) and European Conference on Software Architecture (ECSA), 2012 Joint Working IEEE/IFIP Conference on. 219–223.
- Kark, R. 2011. Games managers play: Play as a form of leadership development. Academy of Management Learning Education 10, 3, 507–527.
- Mizoguchi, R. 2004. Ontology engineering environments. In Handbook on Ontologies - International Handbooks on Information Systems, S. Staab and R. Studer, Eds. Springer, 275–296.
- Peterson, S. 2012. Gamification market to reach 2.8 billion in 2016.
- Pettey, C. and Goasduff, L. 2011. Gartner’s 2011 hype cycle special report evaluates the maturity of 1,900 technologies.

Rostanin, O. and Weber, G. 2012. Semantische netze, thesauri und topic maps. In Semantische Technologien. Grundlagen - Konzepte - Anwendungen, A. Dengel, Ed. Spektrum, Heidelberg, 73–107.

Scheiner, C. W., Witt, M., Voigt, K.-I., and Robra-Bissantz, S. 2012. Einsatz von spielmechaniken in ideenwettbewerben: Einsatzmotive, wirkungen und herausforderungen. In Multikonferenz Wirtschaftsinformatik 2012: Tagungsband der MKWI 2012. Braunschweig: Institut für Wirtschaftsinformatik, 2012, 781–792.

Schell, J. 2008. The Art of Game Design. A Book of Lenses. Morgan Kaufman.

Stuckenschmidt, H. 2009. Ontologien. Konzepte, Technologien und Anwendungen, 1 ed. Springer, Dodrecht, Heidelberg, London, New York.

Zichermann, G. and Cunningham, C. 2011. Gamification by Design. Implementing Game Mechanics in Web and Mobile Apps. O'Reilly, Sebastopol.



## Automatische Bewertung von Spirogrammen durch Methoden der Mustererkennung

Tobias Schulz

Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt

Der technische Aufwand zur Durchführung von Spirometrie-Messungen ist im Vergleich zu anderen Methoden der Lungenfunktionsdiagnostik relativ gering. Um eine ausreichende Qualität der Spirogramme sicherzustellen, ist jedoch während der Spirometrie eine Anleitung des Patienten durch geschultes Personal notwendig. Eine automatische Qualitätsbewertung kann bei der Beurteilung der Spirogramme unterstützen. So kann es dem Patienten ermöglicht werden, selbstständig, mit Kontrolle durch eine automatische Qualitätsbewertung Spirometrie-Messungen durchzuführen. In dieser Arbeit soll geklärt werden, ob mit Methoden, wie sie in der Mustererkennung angewendet werden, die Qualität von Spirogrammen zuverlässig beurteilt werden kann. Die Beurteilung des Spirogramms findet dabei auf Grundlage von klassifizierten Trainingsdaten statt.

Kategorie und Themenbeschreibung: Software: Mustererkennung

Zusätzliche Schlüsselwörter: Lungenfunktionsdiagnostik, Spirometrie

### 1 EINLEITUNG

Die Spirometrie stellt eine zentrale Untersuchungsmethode zur frühzeitigen Erkennung von Lungenfunktionsstörungen dar. Gerade für differenzialdiagnostische Untersuchungen stellen spirometrische Messungen, verglichen mit anderen Methoden der Lungenfunktionsprüfung, eine einfach durchzuführende Alternative dar [Bösch und Criée 2009]. Der technologische Fortschritt ermöglicht stetige Verbesserungen der Spirometer hinsichtlich der Benutzerfreundlichkeit sowie der Kompaktheit des Messgeräts. So kann der Patient zuhause oder unterwegs mit mobilen Geräten, ohne Anleitung eines Arztes oder Technikers selbst Spirometrie-Messungen durchführen. Die in digitaler Form aufgezeichneten Spirogramme können später vom Arzt ausgewertet werden.

Spirometrie-Messungen erfordern eine aktive Mitarbeit des Patienten. Der Patient muss unter anderem eine Ausatmung mit maximaler Anstrengung leisten, damit auf Basis des Spirogramms eine aussagekräftige Diagnose erstellt werden kann. Führt der Patient die Messung ohne Anleitung durch geschultes Personal durch, kann die Qualität der Messung nicht

Adresse des Autors:

Tobias Schulz, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Deutschland, Fakultät Informatik und Wirtschaftsinformatik, [www.fhws.de](http://www.fhws.de).

Die Erlaubnis zur Kopie in digitaler Form oder Papierform eines Teils oder aller dieser Arbeit für persönlichen oder pädagogischen Gebrauch wird ohne Gebühr zur Verfügung gestellt. Voraussetzung ist, dass die Kopien nicht zum Profit oder kommerziellen Vorteil gemacht werden und diese Mitteilung auf der ersten Seite oder dem Einstiegsbild als vollständiges Zitat erscheint. Copyrights für Komponenten dieser Arbeit durch Andere als FHWS müssen beachtet werden. Die Wiederverwendung unter Namensnennung ist gestattet. Es andererseits zu kopieren, zu veröffentlichen, auf anderen Servern zu verteilen oder eine Komponente dieser Arbeit in anderen Werken zu verwenden, bedarf der vorherigen ausdrücklichen Erlaubnis.

sichergestellt werden. Abhilfe schafft eine automatische, computergestützte Bewertung des Spirogramms, die Qualitätsmerkmale automatisch überprüft [Pfüller 2007]. Somit können typische Fehler bei der Durchführung der Spirometrie frühzeitig erkannt werden.

Bisherige Algorithmen zur Qualitätskontrolle von Spirogrammen basieren auf der Anwendung von domänenspezifischem Wissen [Pfüller 2007]. Dazu werden Vitalparameter, wie etwa Lungenvolumen oder Maximalluftstrom, aus den Spirogrammen abgeleitet und auf Plausibilität überprüft. In dieser Arbeit soll untersucht werden, ob eine qualitative Bewertung von Spirogrammen mithilfe von Methoden der Mustererkennung möglich ist. Die Bewertung soll dabei alleine auf einem Vergleich des Spirogramms mit vorgegebenen Trainingsdaten stattfinden.

## 2 LUNGENFUNKTIONSPRÜFUNG MITTELS SPIROMETRIE

Als Hilfsmittel zur Diagnose von Lungenerkrankungen stehen dem Arzt verschiedene Methoden zur Verfügung. Der Aufwand für eine Untersuchung mittels Spirometrie ist vergleichsweise gering. Andere Methoden der Lungenfunktionsdiagnostik erfordern den Einsatz von technisch komplexeren Messeinrichtungen oder einen höheren Zeitaufwand (z. B. Bodyplethysmographie oder Bronchospasmodysetestung) [Bösch und Criée 2009]. Im Folgenden werden digitale Spirometrie-Systeme, der Ablauf der Spirometrie-Messung und die Interpretation der daraus gewonnenen Daten, der Spirogramme, beschrieben.

### 2.1 Digitale Spirometrie-Systeme

Digitale Spirometrie-Systeme messen die Strömungsgeschwindigkeit des Atemluftstroms mit einer konstanten Abtastfrequenz. So wird der Verlauf des Atemluftstroms über eine bestimmte Zeit quantisiert, als Folge digitaler Werte zeitlich äquidistant aufgezeichnet.

Die Abtastfrequenz muss nach dem Nyquist-Shannon-Abtasttheorem mindestens doppelt so hoch sein wie die maximale Frequenz des zu messenden Signals [Smith 2003]. Eine typische Abtastrate von Spirometern, wie auch bei dem in dieser Arbeit verwendeten Gerät, ist 100 Hz.

Zur Umwandlung des Betrags und der Richtung des Luftstroms in digitale Werte haben sich die drei folgenden Messprinzipien durchgesetzt.

*Turbine.* Der zu messende Luftstrom versetzt das Flügelrad einer Turbine in Bewegung. Aus der Drehzahl des Flügelrades wird die Luftstromstärke berechnet.

*Pneumotachograph.* Der Luftstrom wird durch eine Verengung mit definiertem Durchmesser geleitet. Vor und hinter der Verengung entsteht proportional zur Stärke des Luftstroms ein Druckunterschied. Drucksensoren messen diese Druckdifferenz. So kann auf Luftstromstärke geschlossen werden.

*Ultraschall.* Das Ultraschall-Wandler-Messprinzip beruht auf dem Mitnahmeeffekt. Der zu messende Luftstrom beschleunigt oder verringert die Geschwindigkeit von zwei gegenläufigen Ultraschall Impulsen. Aus der Laufzeitdifferenz der beiden Impulse wird auf die Luftstromstärke geschlossen [Fiedler 1992].

Zur Erzeugung von Trainings- und Testdaten wird in dieser Arbeit ein PT (Pneumotachograph)-System verwendet. Systeme nach dem Turbinen-Messprinzip geben die Dynamik des

Luftstroms typischerweise nicht exakt wieder. Die Form der Luftstromkurve wird dadurch verfälscht. Da die Spirogramme auf Basis ihrer Form klassifiziert werden sollen, stellt das PT-System somit die optimale Wahl dar. Durch den Einsatz eines Ultraschall-Systems ergäben sich für diese Arbeit keine weiteren Vorteile.

## 2.2 Durchführung der Spirometrie

Spirometrie-Messungen sollten unter Anleitung einer geschulten Person vorgenommen werden. Dabei wird der Patient am Spirometer angewiesen, die folgenden Schritte durchzuführen.

- (1) Eine schnelle, maximale Einatmung bis die Lungen vollständig gefüllt sind
- (2) Ein sehr schnelles Ausatmen mit maximaler Kraft
- (3) Weiteres Ausatmen bis die Lungen maximal entleert sind

Der Mund des Patienten muss das Mundstück des Spirometers während diesem Manöver stets fest umschließen. Das Ausatmen muss mit maximaler Anstrengung ausgeführt werden. Die anleitende Person motiviert den Patienten während des Manövers, das Ausatmen mit der nötigen Kraft durchzuführen [Miller et al. 2005].

## 2.3 Qualitätsbewertung der Spirogramme

Abbildung 1 zeigt beispielhaft drei verschiedene Spirogramme als mögliche Resultate von Spirometrie-Messungen. Die vertikale Achse des Diagramms in Abbildung 1 zeigt die Stärke des Luftstroms an. Zu beachten ist, dass auf die horizontale Achse nicht die Zeit, sondern das Atemvolumen aufgetragen ist. Relevante Abschnitte des Spirogramms lassen sich so besser beurteilen.

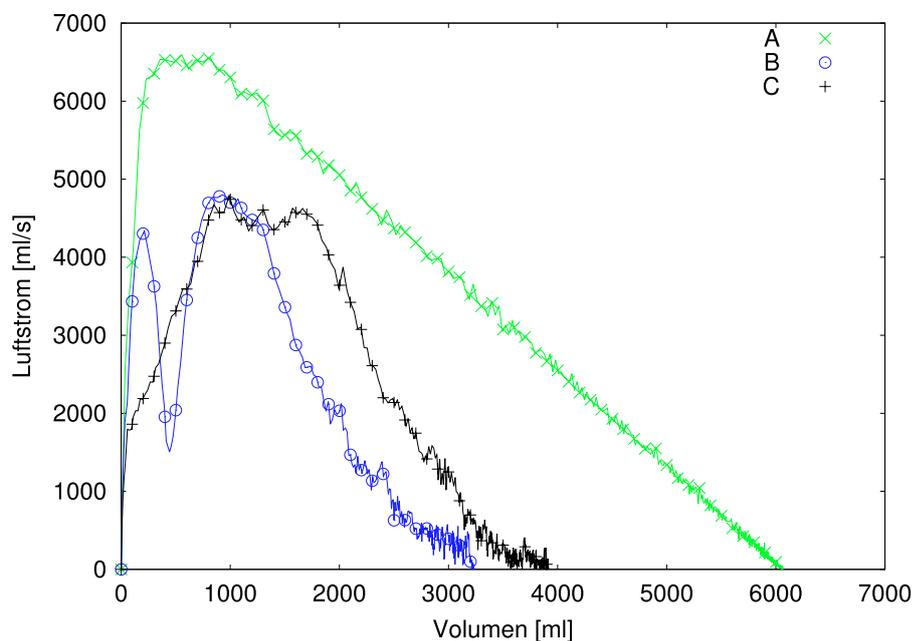


Abb. 1: Drei mögliche Resultate einer Spirometrie-Messung: optimaler Verlauf (A), mit Husten- Artefakt (B) und flacher Anstieg wegen zu geringer initialer Anstrengung (C).

Das Spirogramm A in Abbildung 1 zeigt einen idealen Verlauf. Zu Beginn ist deutlich ein steiler Anstieg bis zum maximalen Luftstrom zu erkennen, gefolgt von einem gleichmäßigen Abklingen. Spirogramm B zeigt ebenfalls zu Beginn einen steilen Anstieg, wird jedoch vor Erreichen des maximalen Luftstroms durch ein Artefakt unterbrochen. Solche Artefakte

können beispielsweise durch ein Husten des Patienten entstehen. Spirogramm C weist einen zu flachen Anstieg auf - ein deutlicher Hinweis auf mangelnde Anstrengung des Patienten.

Nicht immer ist eine mangelhafte Qualität so eindeutig zu erkennen, wie bei den Spirogrammen B und C in Abbildung 1. Abweichungen von der idealen Luftstromkurve A können auch Hinweise auf eine Atemwegs- oder Lungenfunktionsstörung sein. Eine gute Qualität des Spirogramms ist daher die Bedingung für eine korrekte Diagnose [Pellegrino et al. 2005].

Diese Arbeit beschränkt sich auf die zwei Arten von Qualitätsmängeln, wie sie in den Spirogrammen B und C in Abbildung 1 dargestellt sind. Die Aufgabe des zu implementierenden Algorithmus ist die korrekte Zuordnung von Spirogrammen in eine der folgenden Klassen:

- A Akzeptable Spirogramme
- B Spirogramme mit Artefakten
- C Hinweise auf mangelnde Mitarbeit

### 3 AUSWAHL DER TRAININGS- UND TESTDATEN

Die Entscheidung, welcher Klasse ein Spirogramm zugeordnet wird, bestimmt in dem zu implementierenden Mustererkennungs-System der Klassifikator. Dem Klassifikator müssen die Bereiche der Klassen, die er zuordnen soll, bekannt sein. Die Bereiche werden durch klassifizierte Trainingsdaten definiert.

Mit den Testdaten kann die Qualität bzw. die Zuverlässigkeit des Klassifikators, die Klassifikationsrate, überprüft werden. Dazu wird der Anteil der zuvor manuell klassifizierten Testdaten gezählt, die der Klassifikator der korrekten Klasse zuordnet. Test- und Trainingsdaten müssen disjunkt sein [Niemann 1983].

Als Trainingsdaten werden 24 Standard-Testkurven (ATS Spirometry Waveforms (ATS24)) verwendet [o.V. 2014]. Diese Test-Luftstromkurven werden üblicherweise dazu verwendet, die korrekte Funktionsweise von Spirometern zu verifizieren. Bei der Spirometer-Verifikation wird vor allem die korrekte Messung des vorgegebenen Luftstroms und -volumens überprüft [Miller et al. 2005]. Die 24 Standard-Testkurven repräsentieren typische Ergebnisse von Spirometrie-Messungen, darunter auch Spirogramme mit minderer Qualität oder Hinweisen auf Atemwegserkrankungen. Die Testkurven werden mithilfe einer computergesteuerten Pumpe erzeugt und mit einem PFSpirometer aufgezeichnet. Die so digitalisierten Luftstromkurven werden danach auf einen PC übertragen. Tabelle 1 zeigt die Zuordnung der ATS24-Kurven zu den Klassen aus Abschnitt 2.3, die für diese Arbeit getroffen wurde. Die Zuordnung wurde von einer sachkundigen Person vorgenommen.

Klasse	ATS24-Kurven
A	1,2,3,4,15,17,18,19,21,23,24
B	5,7,8,9,11,20,22
C	6,8,10,12,13,14,16,22

*Tab. 1: Die einzelnen ATS24-Kurven wurden von einer sachkundigen Person den Klassen aus Abschnitt 2.3 zugeordnet.*

Eine ATS24-Kurve kann der Klasse B und C zugeordnet sein, falls sie Merkmale beider Klassen besitzt. Einige ATS24-Kurven weisen Verläufe auf, die typisch bei Atemwegserkrankungen

sind. Falls jedoch keine Mängel bezüglich der Qualität des Spirogramms erkennbar sind, werden die ATS24-Kurven mit Hinweisen auf Atemwegserkrankungen der Klasse A zugeordnet.

Für die Testdaten werden von verschiedenen Personen jeweils mehrere Spirogramme aufgenommen. Die Probanden werden bei den Messungen aufgefordert, die Spirometrie so durchzuführen, dass Spirogramme entsprechend den Klassen aus Abbildung 1 entstehen. Insgesamt werden 24 Spirogramme von sechs Probanden in die Testdaten aufgenommen. Die Zuordnung der Testdaten zu den Klassen wurde ebenfalls von einer sachkundigen Person vorgenommen.

#### 4 IMPLEMENTIERUNG EINES PROTOTYPS

Abbildung 2 zeigt eine typische Verarbeitungskette, wie sie zur Klassifikation von Mustern eingesetzt wird [Niemann 1983].

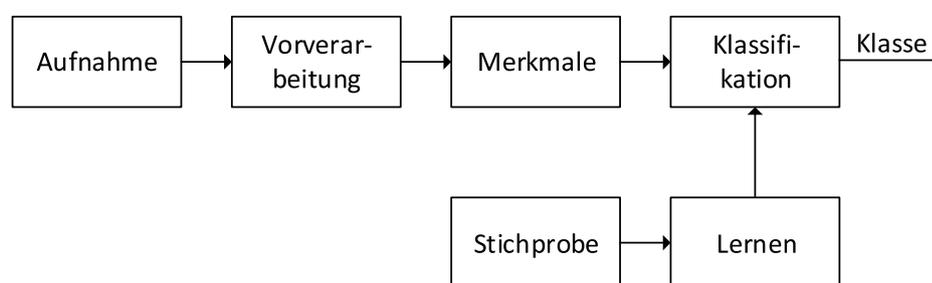


Abb. 2: Das System zur Klassifikation der Spirogramme wird in mehrere Module einer Verarbeitungskette aufgeteilt [Niemann 1983].

Das Modul *Aufnahme* in Abbildung 2 bezieht sich auf die Messung der Trainings- und Testdaten, wie in Abschnitt 3 beschrieben. Die weiteren Module werden in den folgenden Abschnitten beschrieben.

##### 4.1 Vorverarbeitung

Bei der Vorverarbeitung werden die Rohdaten so bearbeitet, dass eine Merkmalsgewinnung und Klassifikation erleichtert wird. Ein typisches Verfahren ist z. B. das Entfernen von Störsignalen durch digitale Filter. Eine Filterung der Spirogramm Daten wird bereits durch das verwendete Spirometer-System vorgenommen. Die Vorverarbeitung beschränkt sich deshalb auf die Anpassung der Darstellungsform und die Normierung.

Die Daten des Spirogramms werden als Luftstrom Werte im 10 ms Zeitraster gespeichert. Wertet man diese Daten direkt aus, erhält man eine Fluss-Zeit-Darstellung wie im oberen Diagramm in Abbildung 3. Das mittlere Diagramm in Abbildung 3 zeigt dieselben Daten in Fluss-Volumen-Darstellung. Relevante Bereiche, wie der langsame Anstieg des Luftstroms, sind in der Fluss-Volumen-Darstellung leichter zu erkennen. Das gilt einerseits für die Beurteilung durch Menschen, andererseits auch für die automatische Klassifikation durch die Mustererkennung. Die relevanten Abschnitte nehmen einen größeren Teil des gesamten Musters ein.

Da in dieser Arbeit die Spirogramme alleine aufgrund ihrer Form klassifiziert werden sollen, werden die Informationen der absoluten Luftstrom- und Volumenwerte entfernt. Dies wird durch eine Normierung beider Absolutwerte erreicht. Das untere Diagramm in Abbildung 3

zeigt die normierten Daten des darüber liegenden Spirogramms. Der Luftstrom wird auf den Wert eins, das Luftvolumen auf den Wert 1000 normiert. Die Größe der Spirogramm-Daten beträgt nach der Vorverarbeitung 1000 Datenpunkte. Eine konstante Länge über alle Muster erleichtert die weitere Verarbeitung im Musterkennungs-System.

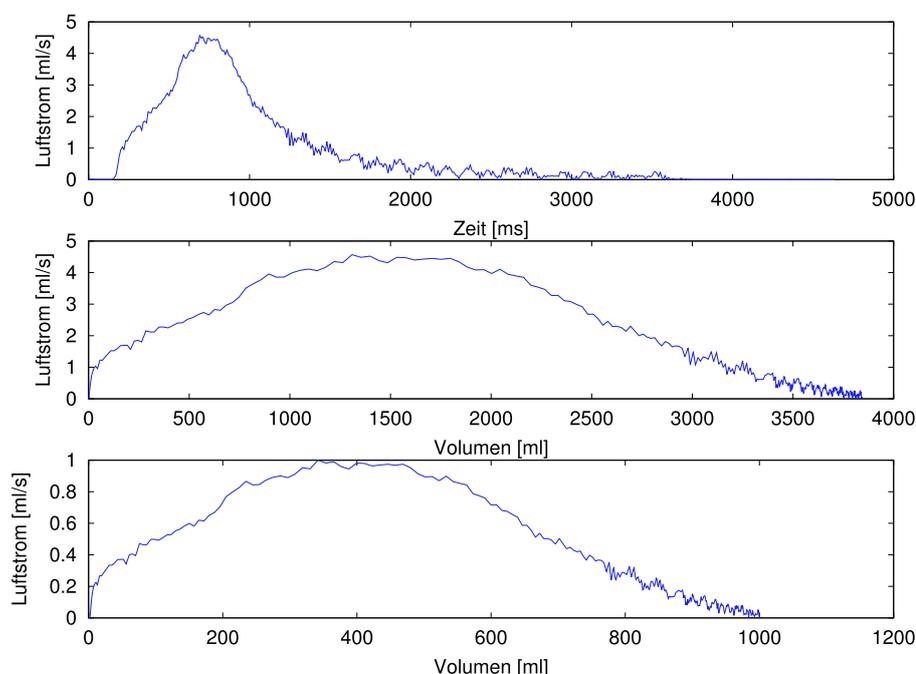


Abb. 3: Schritte der Vorverarbeitung: Fluss-Zeit-Darstellung (oben), Fluss-Volumen-Darstellung (mittig), Fluss-Volumen-Darstellung normiert (unten).

Normiert man die Spirogramme auf diese Weise, wird ihre ursprüngliche Form verändert. Wesentliche Bereiche der Luftstrom Kurve, wie etwa der zu langsame Anstieg in Abbildung 3, bleiben in der normierten Darstellung erhalten. Der große Vorteil dieses Vorgehens ist, dass Attribute des Patienten, wie etwa Alter oder Geschlecht, nicht berücksichtigt werden müssen. Andere Verfahren zur Auswertung von Spirogrammen verwenden Referenzwerte aus statistischen Erhebungen, um zu überprüfen, ob die erreichten Werte in einem erwartetem Bereich liegen [Pellegrino et al. 2005].

#### 4.2 Merkmale

Die Klassifikation beruht auf der Annahme, dass die Ähnlichkeit zweier Muster aus einer Klasse größer ist, als die Ähnlichkeit von zwei Mustern aus verschiedenen Klassen. Ein quantitatives Maß für die Ähnlichkeit von zwei Mustern kann berechnet werden, indem man Merkmale der Muster als Vektoren betrachtet und den Abstand zwischen beiden Vektoren ermittelt. Elemente des Merkmal-Vektors  $\mathbf{c}$  sind dabei numerische Werte, die man aus den Daten des Musters  $f$  gewinnt. Der Abstand zwischen zwei Vektoren kann mit der  $p$ -Norm aus Formel 1 und  $p \in \mathbb{R}^{>0}$  berechnet werden.

$$\|\mathbf{c}_1 - \mathbf{c}_2\|_p = \|\mathbf{c}\|_p := \left( \sum_{i=1}^n |c_i|^p \right)^{1/p} \quad (1)$$

Das Koordinatensystem in Abbildung 4 stellt beispielhaft die Daten von Mustern aus drei beliebigen Klassen dar. Die 30 Kreuz-, Kreis- und Sternsymbole markieren Werte der Muster als

Endpunkte von zweidimensionalen Ortsvektoren  $f_1$  bis  $f_{30}$ . Muster, die zu einer Klasse gehören, treten gehäuft an einer Stelle auf. Eine Klassifikation ist in diesem Fall leicht möglich, die Muster sind gut separiert. Als Merkmale können direkt die Daten der Muster übernommen werden ( $c_n = f_n$ ). Abbildung 4 stellt somit zugleich den Merkmalsraum dar.

Aus den Spirogrammen müssen daher Merkmale gewählt werden, die eine möglichst gute Trennung zwischen den Klassen ermöglichen. Ein einfaches Beispiel für Elemente eines zweidimensionalen Merkmals  $c_{QV}$  eines Spirogramms wäre etwa der absolute Luftstrom  $Q$  und das absolute Luftvolumen  $V$  der Luftstromkurve (Formel 2). Eine Klassifizierung nach den gewählten Klassen aus Abschnitt 3 ist mit diesem Merkmal jedoch nicht möglich. Aus den Werten des absoluten Luftstroms und Luftvolumens alleine kann nicht auf die Qualität des Spirogramms geschlossen werden. Außerdem werden diese Informationen durch die Normierung entfernt (vgl. Abschnitt 4.1).

$$c_{QV} = \begin{pmatrix} Q \\ V \end{pmatrix} \quad (2)$$

Eine einfache Übernahme aller 1000 Spirogramm-Datenpunkte in einen Merkmal-Vektor  $c$  ist nicht praktikabel. Es ist wünschenswert, dass die gewählten Merkmale eine möglichst niedrige Dimension besitzen. Zum einen wird die zu verarbeitende Datenmenge bei niedrigen Dimensionen geringer. Zum anderen ist eine Abstandsbestimmung nach Formel 1 im hochdimensionalen Raum problematisch. Abhängig von der verwendeten  $p$ -Norm verringern sich die Abstände der Merkmale ab einer bestimmten Anzahl an Dimensionen. Die Leistung des Klassifikators kann sich deshalb durch eine Wahl einer höheren Anzahl an Dimensionen verringern [Aggarwal et al. 2001].

Will man bei vorhandenen Merkmalen die Anzahl der Dimensionen verringern, bietet sich ein Vorgehen wie bei den Beispiel-Merkmalen in Abbildung 4 an. Um die Dimensionen von zwei auf eine zu reduzieren, werden die durch Vektoren dargestellten Merkmale mittels Skalarprodukt orthogonal auf den Vektor  $a$  projiziert [Teschl 2008]. Die Dreiecksymbole repräsentieren die projizierten zweidimensionalen Vektoren als eindimensionale Merkmale entlang des Vektors  $a$ .

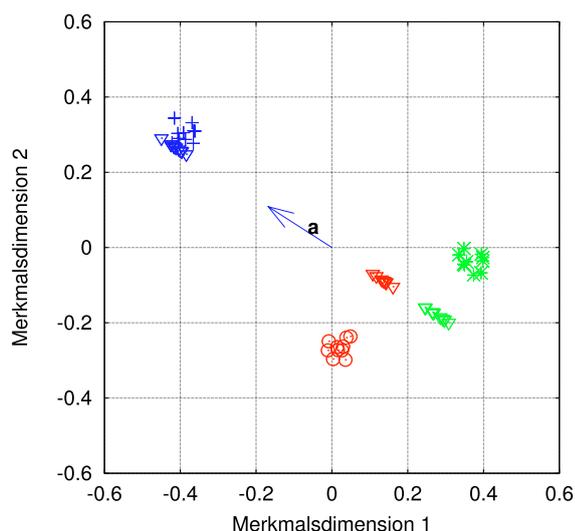


Abb. 4: Beispiel für einen Merkmalsraum und Reduktion der Merkmalsdimensionen: Kreuz- Kreis und Sternsymbole repräsentieren Merkmale im zweidimensionalen Merkmalsraum. Dreiecksymbole repräsentieren die in den eindimensionalen Raum projizierten Merkmale.

Die eindimensionalen Merkmale in Abbildung 4 sind immer noch gut separiert. Merkmale verschiedener Klassen überlagern sich nicht. Vektor  $a$  muss so gewählt werden, dass Merkmale verschiedener Klassen gut getrennt sind. Eine analytische Methode zur Ermittlung des Vektors  $a$  ist die Hauptachsentransformation (engl. PCA (Principal Component Analysis)). Mittels Eigenwert- oder Singulärwertzerlegung erhält man eine Transformationsmatrix  $A$  bzw. den Vektor  $a$ . Der mittlere Abstand bzw. die Varianz ist durch die Projektion mit  $a$  maximal. Die Kompaktheit der Merkmale innerhalb der Klassen bleibt dennoch erhalten [Shlens 2003].

Formel 3 zeigt vereinfacht, wie die Hauptachsentransformation im Prototyp benutzt wird. Der erste Eingangsparameter ist für das Beispiel aus Abbildung 4 eine  $2 \times 30$  Matrix. Der zweite Eingangsparameter  $m$  gibt die gewünschte Anzahl der Dimensionen an, auf die die Merkmale reduziert werden. Das Ergebnis ist eine  $l \times m$  Transformationsmatrix  $A$  mit  $l$  als Länge der Muster-Vektoren  $f$ . Für  $m = 1$  ist das Ergebnis ein Spaltenvektor  $a$  mit  $l$  Elementen. Mit dem Skalarprodukt aus Formel 4 werden die Muster  $f$  in einen  $m$ -dimensionalen Merkmalsraum transformiert.

$$PcaMatrix([f_1 \dots f_n], m) = A \quad (3)$$

$$(A, f) = c \quad (4)$$

So können die kompletten Datenpunkte eines Spirogramms als Muster-Vektor  $f$  interpretiert und mittels der Transformationsmatrix  $A$  auf Merkmale mit einer beliebigen Anzahl an Dimensionen  $m$  reduziert werden.  $A$  wird dabei mithilfe der Trainingsdaten, den ATS24-Kurven, berechnet. Je nach Zuverlässigkeit des Klassifikators kann die verwendete Anzahl an Dimensionen angepasst werden.

#### 4.3 Stichprobe und Lernen

Ziel des Lernens ist es, den Klassifikator so zu parametrieren, dass er Merkmal-Vektoren von Mustern den richtigen Klassen zuordnet. Für die Merkmal-Vektoren in Abbildung 4 lassen sich leicht Klassengrenzen definieren, etwa durch drei Dichtefunktionen der Normalverteilung mit Mittelwerten im Zentrum der jeweiligen Merkmal-Wolken. Die Varianzen lassen sich ebenfalls für jede Klasse aus den Daten der jeweiligen Merkmale schätzen. Durch die Position eines zu klassifizierenden Merkmals im Merkmalsraum lässt sich so die Wahrscheinlichkeit der jeweiligen Klassenzugehörigkeit berechnen. Die zu lernenden Parameter wären in diesem Fall die Mittelwerte und Varianzen für jede Klasse. Beim überwachten Lernen können diese Parameter anhand einer Stichprobe von klassifizierten Mustern bzw. den Trainingsdaten geschätzt werden [Niemann 1983].

Abbildung 5 zeigt die Merkmal-Vektoren der Trainingsdaten nach der Hauptachsentransformation in einem dreidimensionalen Merkmalsraum. Die Klassengrenzen lassen sich nicht durch einfache Dichtefunktionen von Wahrscheinlichkeitsverteilungen definieren. Die Klassengrenzen sind ineinander verzahnt. Die ATS24-Kurven, welche als Trainingsdaten dienen, weisen pro Kurve oft Eigenschaften von mehreren der in Tabelle 1 aufgeführten Klassen auf.

Für die Klassifikation der Spirogramme wird ein nichtparametrisches Klassifikationsverfahren, die Nearest-Neighbor-Klassifikation gewählt. Somit müssen keine Parameter durch Lernen ermittelt werden, die Lernphase entfällt. Ein weiterer Vorteil dieses Klassifikators ist außerdem, dass komplexe Klassengrenzen wie in Abbildung 5 nicht durch eine parametrische Funktion beschrieben werden müssen.

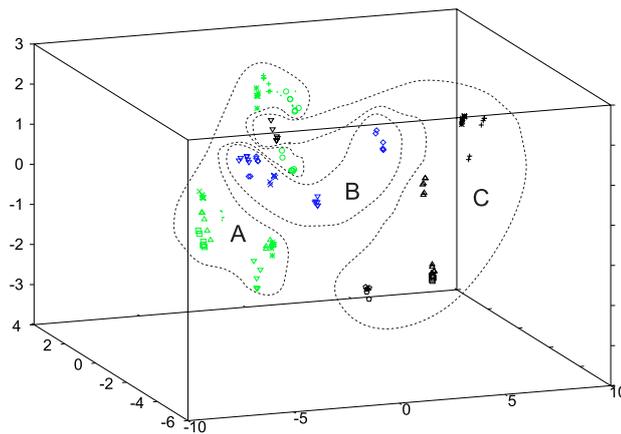


Abb. 5: Dreidimensionale Merkmale der ATS24-Kurven nach der Hauptachsentransformation: Die Merkmale der drei Klassen A, B und C können nur durch komplexe Klassengrenzen voneinander separiert werden. Merkmale, die mehreren Klassen zugeordnet sind, werden hier nicht abgebildet. Für die Trainingsdaten wurde jede der ATS24-Kurven mehrfach aufgezeichnet. Deshalb sind mehr als 24 Datenpunkte abgebildet.

#### 4.4 Klassifikation

Der Nearest-Neighbor-Klassifikator (NN) geht davon aus, dass ein Merkmal von Merkmalen derselben Klasse umgeben ist. Um die Klasse eines Merkmals zu bestimmen, wird der Abstand des Merkmals zu allen bekannten und klassifizierten Merkmalen der Trainingsdaten berechnet. Das zu klassifizierende Merkmal wird derselben Klasse zugeordnet, welche das Trainingsdaten-Merkmal mit dem geringsten Abstand besitzt. Es können auch mehrere nächste Nachbarn gesucht werden und eine Mehrheitsentscheidung getroffen werden. Bei der geringen Anzahl an Trainingsdaten, die für diese Arbeit zur Verfügung stehen, ist eine Mehrheitsentscheidung jedoch nicht sinnvoll.

Der NN-Klassifikator erlaubt beliebig komplexe Klassengrenzen. Allerdings müssen die Merkmale aller Trainingsdaten gespeichert werden. Bei jeder Klassifikation muss zudem der nächste Nachbar in den gesamten Trainingsdaten durch Abstandsbestimmung ermittelt werden. Die Laufzeitkomplexität der NN-Suche  $O(n \cdot m)$  steigt mit der Anzahl der Trainingsdaten  $n$  und deren Anzahl an Dimensionen  $m$ . Geeignete Datenstrukturen ermöglichen bei größeren Mengen an Trainingsdaten eine effiziente Suche [Shibata et al. 2003].

## 5 ERGEBNISSE

Für die Bewertung des Prototyps wird der Anteil vom Klassifikator richtig zugeordneten Merkmal-Vektoren  $c$  aus den Testdaten zur korrekten Klasse betrachtet ( $c_{\kappa} \rightarrow \kappa$ ). Test- und Trainingsdaten sind zuvor von einer sachkundigen Person den Klassen A, B und C zugeordnet worden ( $\kappa \in \{A, B, C\}$ , vgl. Abschnitt 3). Die Klassifikationsrate gibt den Anteil der vom Prototypen korrekt zugeordneten Mustern an. Außerdem wird die Zuverlässigkeit des Klassifikators in Abhängigkeit von der Anzahl der Merkmaldimension und der verwendeten p-Norm gemessen (vgl. Formel 1). Abbildung 6 zeigt die Ergebnisse der Klassifikation der Testdaten durch den Prototyp.

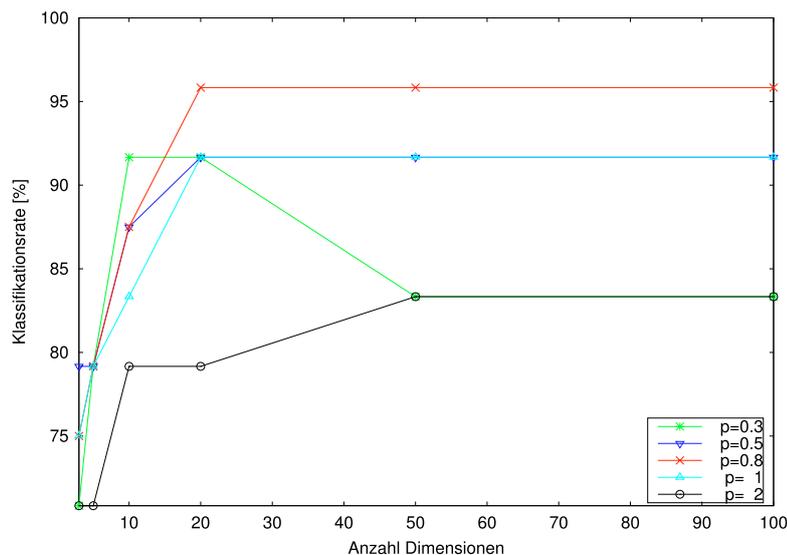


Abb. 6: Klassifikationsrate des Prototyps: Die Klassifikationsrate steigt bei zunehmender Anzahl an Dimensionen, abhängig von der verwendeten  $p$ -Norm.

Erwartungsgemäß erhöht sich der Anteil der richtig zugeordneten Spirogramme mit der Anzahl der Merkmalsdimensionen. Bei der Verwendung der  $p$ -Norm mit  $p = 0,3$  entsteht ein Optimum von 92% bereits bei 10 Merkmalsdimensionen. Die höchste Klassifikationsrate von 96% wird für  $p = 0,8$  ab 20 Merkmalsdimensionen erreicht, es wird lediglich ein Spirogramm der Testdaten falsch zugeordnet.

## 6 FAZIT

Die Ergebnisse aus Abbildung 6 zeigen, dass sich bereits durch das Prototypen-System hohe Trefferquoten erzielen lassen. Der Klassifikator ordnet bis zu 23 von 24 Testdaten korrekt zu. Somit kann ein solches System Unterstützung bieten, falls Spirogramme von dem Patienten als Laien selbstständig aufgenommen werden. Die Wahrscheinlichkeit, dass Spirometrie-Messungen mit zu wenig initialer Anstrengung oder Artefakten erkannt werden, liegt bei 96%. Das Spirometrie-System kann den Patienten daraufhin auffordern, die Messung zu überprüfen bzw. zu wiederholen. Ein weiterer Einsatz wäre bei der automatischen Vorauswahl bei einer größeren Menge von Spirogrammen denkbar. Wie etwa bei der Durchführung medizinischer Studien. Weiterhin hat sich bei der Arbeit mit den Testdaten gezeigt, dass ebenfalls eine Klassifikation nach einer Atemwegsobstruktion möglich ist. Testdaten mit leichten Hinweisen einer solchen Einengung der Atemwege konnten den entsprechenden Trainingsdaten zugeordnet werden. Für eine valide Auswertung der Spirogramme im Prototyp bezüglich einer Obstruktion lagen jedoch nicht ausreichend Testdaten mit diesem Krankheitsbild vor.

Abbildung 5 in Abschnitt 4.3 zeigt, dass die gewählten Merkmale der Trainingsdaten nicht optimal separiert sind. Durch die komplexen Klassengrenzen ist die Auswahl an Klassifikatoren deutlich eingeschränkt. In den verwendeten Trainingsdaten, den ATS24-Testkurven, treten in vielen Spirogrammen Merkmale verschiedener der in dieser Arbeit verwendeter Klassen auf. Außerdem entspricht die Aufzeichnungsqualität der ATS24-Testkurven nicht dem aktuellen Stand der Technik [Reinstädler 2012]. Eine bessere Klassentrennung lässt sich evtl. durch eine gezielte Wahl bzw. Generierung von Trainingsdaten erreichen. Somit könnte ein Klassifikator mit geringerer Laufzeit- und Speicherkomplexität als der NN-Klassifikator verwendet werden. Dies würde eine effiziente Implementierung auf Embedded-Systemen mobiler Spirometer ermöglichen.

## LITERATUR

Charu C. Aggarwal, Alexander Hinneburg, und Daniel A. Keim. 2001. On the Surprising Behavior of Distance Metrics in High Dimensional Space. In Database Theory — ICDT 2001, Jan Bussche and Victor Vianu (Bearbeiter). Lecture Notes in Computer Science, Band 1973. Springer Berlin Heidelberg, Berlin und Heidelberg, 420–434. DOI:[http://dx.doi.org/10.1007/3-540-44503-X\\_27](http://dx.doi.org/10.1007/3-540-44503-X_27)

o.V. 2014. American Thoracic Society Spirometry Waveforms. (2014). Zuletzt aufgerufen am 03.05.2014 von <http://www.thoracic.org/education/pulmonary-function-testing/spirometry-waveforms/index.php>

Dennis Bösch und Carl-Peter Criée. 2009. Lungenfunktionsprüfung: Durchführung - Interpretation - Befundung (2. Auflage). Springer Medizin, Heidelberg.

Otto Fiedler. 1992. Strömungs- und Durchflussmesstechnik. Oldenburg, München and Wien.

M. R. Miller, J. Hankinson, V. Brusasco, F. Burgos, R. Casaburi, A. Coates, R. Crapo, P. Enright, van der Grinten, C P M, P. Gustafsson, R. Jensen, D. C. Johnson, N. MacIntyre, R. McKay, D. Navajas, O. F. Pedersen, R. Pellegrino, G. Viegi, und J. Wanger. 2005. Standardisation of spirometry. *The European respiratory journal* 26, 2 (2005), 319–338. DOI:<http://dx.doi.org/10.1183/09031936.05.00034805>

Heinrich Niemann. 1983. Klassifikation von Mustern. Springer, Berlin.

R. Pellegrino, G. Viegi, V. Brusasco, R. O. Crapo, F. Burgos, R. Casaburi, A. Coates, van der Grinten, C P M, P. Gustafsson, J. Hankinson, R. Jensen, D. C. Johnson, N. MacIntyre, R. McKay, M. R. Miller, D. Navajas, O. F. Pedersen, und J. Wanger. 2005. Interpretative strategies for lung function tests. *The European respiratory journal* 26, 5 (2005), 948–968. DOI:<http://dx.doi.org/10.1183/09031936.05.00035205>

Mark Pfüller. 2007. Validierung eines Algorithmus zur automatischen Qualitätskontrolle spirometrischer Fluss-Volumen-Kurven und Messgrößen. (2007). Zuletzt aufgerufen am 25.03.2014 von [http://edoc.ub.uni-muenchen.de/7778/1/Pfueller\\_Mark.pdf](http://edoc.ub.uni-muenchen.de/7778/1/Pfueller_Mark.pdf)

Jürgen Reinstädler. 2012. How Physiological Are The ATS Flow/Volume-Loop Waveforms? In A66. MODELING, MECHANICS AND GAS EXCHANGE. Band 185. A2087–A2087. DOI:[http://dx.doi.org/10.1164/ajrccm-conference.2012.185.1\\_MeetingAbstracts.A2087](http://dx.doi.org/10.1164/ajrccm-conference.2012.185.1_MeetingAbstracts.A2087)

T. Shibata, T. Kato, und T. Wada. 2003. K-d decision tree: an accelerated and memory efficient nearest neighbor classifier. In Third IEEE International Conference on Data Mining. 641–644. DOI:<http://dx.doi.org/10.1109/ICDM.2003.1250997>

Jonathon Shlens. 2003. A Tutorial On Principal Component Analysis. (2003). Zuletzt aufgerufen am 03.05.2014 von [http://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition\\_jp.pdf](http://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf) Steven W. Smith. 2003. Digital signal processing: A practical guide for engineers and scientists. Newnes, Amsterdam.

Gerald Teschl. 2008. Mathematik für Informatiker (3. Auflage). eXamen.press, Band 1. Springer, Berlin und Heidelberg und New York.



## Untersuchungen zur Heizzeitreduzierung bei der Herstellung von dickwandigen Gummibauteilen

Volker Herrmann, Felix Naß, Sebastian Lemmich, Florian Wachter  
Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt

Die Untersuchung beschäftigt sich mit einer neuen Spritzgießtechnologie zur Herstellung von Bauteilen aus Elastomeren. Mit dieser Technologie sind kürzere Zykluszeiten bei gleichbleibenden Bauteileigenschaften möglich. Die Höhe der Heizzeitreduzierung hängt von der Art der Bestimmung derselben ab und wird in der vorliegenden Arbeit auf unterschiedliche Weise quantitativ bestimmt. Je nach Verfahrensparameter und Bestimmungsmethode treten Heizzeitreduzierungen von maximal etwa 40 % auf. Weiterhin werden Untersuchungen an den hergestellten Bauteilen durchgeführt, welche die Interpretation des Ausmaßes und der Qualität der Vernetzungsdichte über dem Probenquerschnitt erlauben. Es können deutliche Heizzeiteinsparungen bei mindestens gleichbleibenden und teilweise homogeneren Vernetzungsdichten über dem Querschnitt eines dickwandigen Elastomerbauteils attestiert werden.

Kategorie und Themenbeschreibung: Energy and Cost Saving

Zusätzliche Schlüsselwörter: Zykluszeitreduzierung, Schererwärmung, Reaktionskinetik, NMR Bildgebung, Vernetzungsdichte, Reversion

### 1 PROBLEMATIK DER HERSTELLUNG DICKWANDIGER GUMMIBAUTEILE

Aus dem täglichen Umfeld und vielen technischen Anwendungsfällen sind Kunststoffe und Elastomere nicht mehr wegzudenken. Aufgrund ihrer geringen Dichte und Funktionalität und damit verbunden geringem Gewicht sind Polymerwerkstoffe beispielsweise in der Luft- und Raumfahrt, im Automobilbereich sowie in der Transport- und Verpackungsindustrie unentbehrlich geworden. Während bei der Bauteilherstellung von thermoplastischen Kunststoffen teilweise sehr kurze Zykluszeiten erreicht werden können, ist bei der Herstellung von Bauteilen aus Elastomeren die Vulkanisationsphase zu berücksichtigen, die einen Großteil der Zykluszeit im Spritzgießprozess ausmacht. Im Gegensatz zu Thermoplasten finden in dieser Phase der Vernetzung chemische Prozesse statt, die je nach Rezeptur ein gewisses Temperatur-/Zeitfenster benötigen. Insbesondere bei dickwandigen Bauteilen aus Elastomeren führt dies aufgrund der niedrigen Wärme-, bzw. Temperaturleitfähigkeit von Kautschukmischungen [Baldt et al. 1978] häufig zu langen Zykluszeiten. Darüber hinaus sind daraus resultierend nicht selten inhomogene Vernetzungsverläufe über dem Probenquerschnitt die Folge.

Volker Herrmann, Felix Naß, Sebastian Lemmich, Florian Wachter  
Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Deutschland, Fakultät Kunststofftechnik und Vermessung, Studiengang Kunststoff- und Elastomertechnik  
<http://fkv.fhws.de>  
Zusammenfassung von drei Veröffentlichungen, die im Fachmagazin „GAK“ (Gupta-Verlag, Ratingen) gedruckt wurden (Ausgaben Januar, Februar und März 2013).

Die Erlaubnis zur Kopie in digitaler Form oder Papierform eines Teils oder aller dieser Arbeit für persönlichen oder pädagogischen Gebrauch wird ohne Gebühr zur Verfügung gestellt. Voraussetzung ist, dass die Kopien nicht zum Profit oder kommerziellen Vorteil gemacht werden und diese Mitteilung auf der ersten Seite oder dem Einstiegsbild als vollständiges Zitat erscheint. Copyrights für Komponenten dieser Arbeit durch Andere als FHWS müssen beachtet werden. Die Wiederverwendung unter Namensnennung ist gestattet. Es andererseits zu kopieren, zu veröffentlichen, auf anderen Servern zu verteilen oder eine Komponente dieser Arbeit in anderen Werken zu verwenden, bedarf der vorherigen ausdrücklichen Erlaubnis.

Die vorliegende Arbeit beschäftigt sich mit einer neuen Methode der Spritzgießtechnik zur Herstellung von Elastomerbauteilen, die das Potential hat, diese Nachteile zu kompensieren.

## 2 ANSATZ DER NUTZUNG VON SCHERERWÄRMUNG

Im Verarbeitungstechnikum des Studiengangs Kunststoff- und Elastomertechnik der FHWS steht seit November 2008 eine Gummiinjektionspresse der Firma LWB Steinl für Lehre und Forschung zur Verfügung, Abbildung 1. Die Besonderheit dieser Anlage ist die sogenannte EFE-Spritztechnologie, die es ermöglicht, während der Einspritzphase der Kautschukmischung dieser ein einstellbares Maß an Schererwärmung zuzuführen. Dies geschieht durch gezielte Verengung der Drosselstelle während der Einspritzphase mittels Schnecke von der EF-Einheit in die E-Einheit. Durch die in der Regel relativ hohen Viskositäten von Kautschukmischungen sind somit beträchtliche Temperaturerhöhungen möglich. Die Folge daraus sind höhere Formeintrittstemperaturen, somit ist ein geringerer Wärmefluss von der heißen Werkzeugwand in das Formteil für die Vulkanisationsphase notwendig. Daraus resultieren schließlich kürzere Vulkanisations-/Zykluszeiten sowie eine gleichmäßigere Temperatur über dem Querschnitt eines dickwandigen Bauteils während der Vulkanisationsphase. Darüber hinaus stellt sich in diesem Zusammenhang die Frage, ob ein geringerer Wärmefluss von der heißen Werkzeugwand in die kühlere Kautschukmischung einen kleineren Gradienten der Vernetzungsdichte über dem Bauteilquerschnitt bewirkt. Aufgrund der breiten Akzeptanz dieses Systems in der Kautschukindustrie wurde die Fa. LWB Steinl im Jahre 2009 mit dem „DKG Product Award“ der Deutschen Kautschuk-Gesellschaft ausgezeichnet.

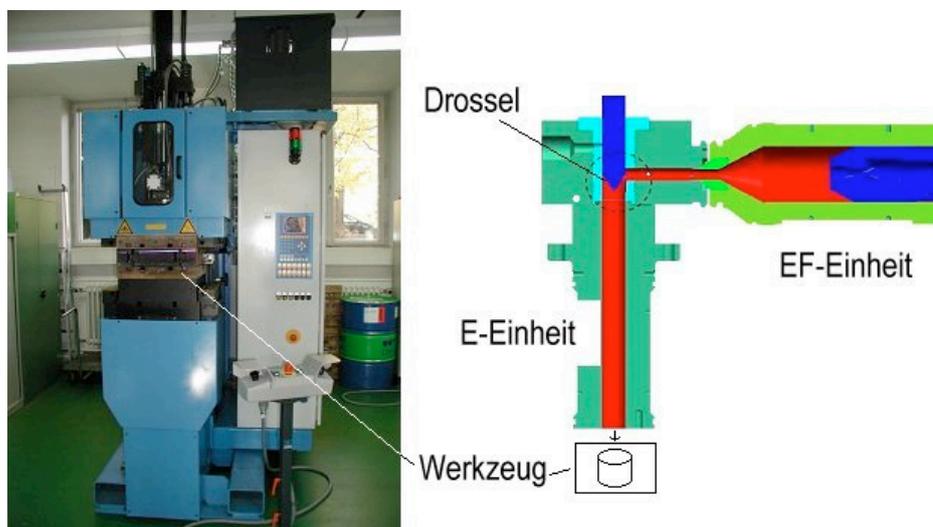


Abbildung 1: Gummiinjektionspresse der Fa. LWB Steinl (Vertikalmaschine mit C-Rahmen, 50 t, 115 ccm) im Verarbeitungstechnikum des Studiengangs Kunststoff- und Elastomertechnik (links) mit EFE-Spritzaggregat (rechts) [Gogolin 2008]

## 3 ZIELSETZUNG UND EXPERIMENTELLES VORGEHEN

In der vorliegenden Untersuchung sollen die vorangegangenen Aspekte der Heizzeitreduzierung und des homogenen Profils der Vernetzungsdichte quantitativ und qualitativ bewertet werden. Hierzu wird folgendes dickwandiges Bauteil ausgewählt: ein Zylinder mit einer Höhe von 40 mm und einem Durchmesser von 55 mm (Abbildung 2). Es wird eine typische

Rezeptur verwendet, die beispielsweise in der Schwingungstechnik von Automobilen zum Einsatz kommen könnte (Tab. 1). In der Kavität sind Thermoelemente eingesetzt, um den Temperaturverlauf in der Mitte des Zylinders über der Vulkanisationszeit zu messen. Dies ist wichtig für eine anschließende reaktionskinetische Nachberechnung des Bauteils und zur Bestimmung des Vernetzungsgrades am kritischsten Punkt des Bauteils zum Zeitpunkt der Entformung. Zur Bestimmung der reaktionskinetischen Parameter werden mit einem rotorlosen Torsionsschubvulkameter SIS V 50 der Fa. Scarabaeus Vulkanisationsisothermen zwischen 120 und 190 °C ermittelt.



Abbildung 2: Für die Untersuchung ausgewähltes dickwandiges Bauteil (Zylinder mit  $h = 40 \text{ mm}$ ,  $\varnothing = 55 \text{ mm}$ ); die Sicht erfolgt hier auf die Unterseite des Zylinders mit den aufgrund der Thermoelemente entstandenden Löchern.

Naturkautschuk	100
ZnO aktiv	5
Stearinsäure	2,5
Ruß N660	25
Kaolin	15
Sunpar 2280	5
PEG4000	1,5
Vulkanox 4020 (6PPD)	1,5
Vulkacit CZ (CBS)	1,5
Vulkacit Mercapto/C (MBT)	0,4
Schwefel	2

Tab. 1: Verwendete Rezeptur für den Elastomerzylinder

An ausgewählten gespritzten und vulkanisierten Zylindern werden u.a. folgende Prüfungen durchgeführt [NAß 2011]:

- Dichtemessungen
- Shore A Messungen an aufgeschnittenen Zylindern (Härtescan, siehe Abbildung 3)
- NMR Bildgebung an den ganzen Zylindern (zerstörungsfrei)

Darüber hinaus werden photographische Aufnahmen der aufgeschnittenen Zylinder erstellt, um möglicherweise vorhandene Blasen sichtbar zu machen.

Dichtemessungen:

Die Dichtemessungen erfolgen nach dem Archimedischen Prinzip (Auftriebsmethode nach DIN EN ISO 1183-1:2004-05) an den gespritzten Zylindern. Nachdem die Zylinder unmittelbar nach dem Entformen im Wasserbad abgekühlt werden, wird die Dichtemessung erst nach

einem weiteren Tag Lagerungszeit durchgeführt um den Einfluss von Restfeuchtigkeit zu eliminieren. Die Messung erfolgt dann durch Wägung des gesamten Zylinders in Luft und in Wasser [NAß 2011].

Shore A Messungen an aufgeschnittenen Zylindern (Härtescan):

Obwohl im Verarbeitungstechnikum mittlerweile eine pneumatische Stanzvorrichtung zum Aufschneiden der Zylinder vorhanden ist, wurden die Zylinder, deren Härteprofile über dem Querschnitt in dieser Abhandlung dargestellt sind noch mittels Wasserstrahlschneiden (Fa. Steinbrenner GmbH, Insingens) durch die Mitte der Probe gemäß Abbildung 3 geschnitten. Die lokalen Messungen der Shore A Werte erfolgen in einem Abstand von 15 mm zur oberen Kante des Zylinders. Die Messpunkte liegen in einem Abstand von jeweils 2 mm zueinander. Die Shore A Werte werden nach 15 s abgelesen. Für die exakte Positionierung des Prüfkörpers steht eine Probenhalterung aus Epoxidharz zur Verfügung, die im Rahmen der Diplomarbeit von Naß angefertigt wurde [NAß 2011].

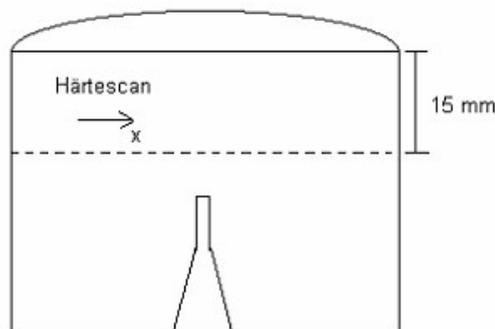


Abbildung 3: Schematische Darstellung der durch Wasserstrahlschneiden erzeugten Schnittfläche der Zylinder und Lage der Scanlinie für die Shore A Messungen.

NMR-Bildgebung:

Die  $^1\text{H}$ -NMR-Messungen erfolgen an einem Bruker DSX 200 NMR Spektrometer mit  $\mu$ -Imaging Ausstattung des ITMC an der RWTH Aachen. Die Pulssequenz ist eine für die Bildgebung optimierte und an die CPMG-Pulsfolge angelehnte Sequenz (Paravision MSME) mit den folgenden Echozeiten: 3,713 ms, 6,344 ms, 10,057 ms und 13,770 ms. Die Anzahl der Pixel in einem Bild beträgt  $128 \times 128$  Pixel in der Fläche bei einer Schichtdicke von 5 mm. Der beobachtete Bereich beträgt  $64 \times 64$  mm. Wiederholzeit: 800 ms bei 4 Wiederholungen zur Verbesserung des Signal/Rausch-Verhältnisses.

Die Bildgebungen erfolgen als Differenzbild der Signalintensität der  $^1\text{H}$ -Spindichte zwischen den Echozeiten 3,713 und 6,344 ms.

## 4 ERGEBNISSE UND DISKUSSION

### 4.1 Heizzeitreduzierungen

#### Bestimmung der Heizzeitreduzierungen

Im Mittelpunkt der Untersuchung steht zunächst die Quantifizierung der maximal möglichen Heizzeitreduzierung für das gegebene Bauteil durch die Drosselwirkung des EFE-Spritzaggregats. Die ermittelte Heizzeitreduzierung unter Einsatz der Drossel wird bezogen auf den Einsatz des EFE-Aggregats ohne Drossel.

Die Heizzeitreduzierung HR ergibt sich somit gemäß Gl.(1):

$$HR = \frac{t_{H,o} - t_{H,m}}{t_{H,o}} \cdot 100 \% \quad (1)$$

wobei gilt:

$t_{H,o}$  = notwendige Mindestheizzeit in der Form ohne Einsatz der Drossel

$t_{H,m}$  = notwendige Mindestheizzeit in der Form mit Einsatz der Drossel

Während Gl.(1) eine leicht nachvollziehbare Gleichung darstellt, stellt sich jedoch die Frage nach der praxisnahen und akzeptablen Ermittlung der „notwendigen Mindestheizzeit in der Form“. Hierfür werden nachfolgend zwei Möglichkeiten verfolgt. Sind die reaktionskinetischen Parameter nach DIN 53529-2 bekannt, so lässt sich für eine gegebene Temperatur-/Zeit-Kurve der Vernetzungsumsatz berechnen. Aufgrund des Wärmestroms von der heißeren Werkzeugwand in die je nach Verfahrensparameter um 50 – 100 °C kühlere Kautschukmischung ergibt sich ein inhomogenes Temperaturprofil, wie in Abbildung 4 mittels Thermographiekamera an einem aufgeschnittenen Zylinder dargestellt [Wachter 2012]. Im Inneren des Bauteils sind die niedrigsten Temperaturen, bzw. langsamsten Temperaturanstiege zu verzeichnen. Somit ist die Bauteilmitte als die kritischste Stelle für reaktionskinetische Berechnungen heranzuziehen. Die randnahen Bereiche sind hinsichtlich der Bestimmung der notwendigen Mindestheizzeit unkritisch, da hier der Vernetzungsprozess früher startet und diese daher nicht die Zykluszeit bestimmenden Bereiche sind.

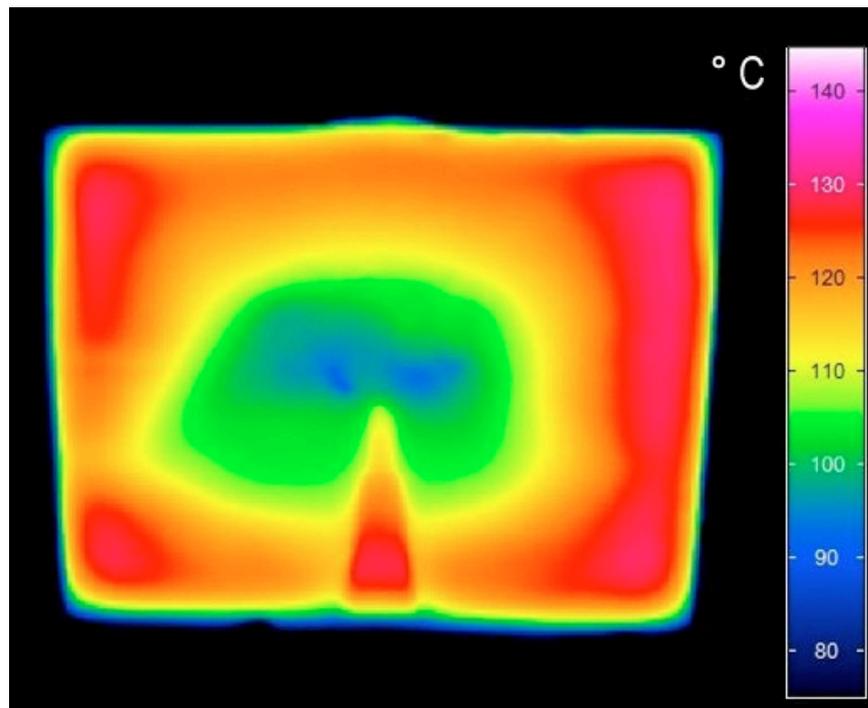


Abbildung 4: Thermographische Aufnahme an der Schnittfläche eines durch die Mitte geschnittenen Zylinders nach einer Verweilzeit von 580 s in der Werkzeugkavität bei einer Werkzeugtemperatur von 160 °C ohne Einsatz der Drossel; die Zeitdauer vom Schnitt des Zylinders mit einer pneumatischen Stanze bis zur Aufnahme des Bildes beträgt etwa 10 s.

Durch den Einsatz der Drossel erhöht sich nun das gesamte Temperaturniveau für den Vernetzungsprozess und dieser startet entsprechend früher. Interpretiert man die Inkubationsphase

als die Zeit, die bis zum Beginn des Vernetzungsprozesses notwendig ist, kann man diese zunächst als die notwendige Mindestheizzeit in Gl.(1) einsetzen. Zylinder, welche zur Inkubationszeit entformt werden, weisen jedoch noch Blasen im Bauteilinneren auf. Diese Blasen entstehen, wenn das chemisch gebildete Netzwerk noch nicht die Festigkeit aufweist, die notwendig ist, um dem inneren Druck standzuhalten. Erfahrungsgemäß sind in der industriellen Praxis Vernetzungsumsätze von etwa 20 – 30 % notwendig, um Blasenfreiheit zu garantieren. Abbildung 5 zeigt aufgeschnittene Zylinder, die mit unterschiedlichen Vulkanisationszeiten hergestellt wurden. Ab einem gewissen Zeitpunkt sind keine Blasen mehr im Bauteilinneren zu erkennen. Dieser Zeitpunkt wird in der englischen Fachliteratur als „Blow Point“ oder „Porosity Point“ bezeichnet. Mittels Dichtemessungen an den Probekörpern ist das Auftreten von Blasen im Inneren indirekt zu erkennen: so steigt die Dichte der Zylinder mit der Vulkanisationszeit durch die Abnahme der Blasen (Abbildung 6). Dieser Dichteanstieg ist nicht mit der Zunahme der Vernetzungsdichte zu erklären, bzw. zu verwechseln. Will man die notwendige Mindestheizzeit über die Blow-Point Methode bestimmen, so ist dies zwar die praxisnähere Methode, erfordert aber einen größeren experimentellen Aufwand. Um nun die Heizzeitreduzierungen, welche mittels EFE-Spritzaggregat möglich sind, über Gl.(1) zu berechnen, ist das Kriterium zur Wahl der notwendigen Mindestheizzeit in der Form ausschlaggebend. Bezieht man sich beispielsweise in Abbildung 6 auf eine notwendige Dichte von  $1,1075 \text{ g/cm}^3$ , so ist dies ohne Einsatz der Drossel bei einer Heizzeit von 8 min erreicht, bei der Drosselstellung 85 % bereits nach 4 min, was gemäß Gl.(1) einer Heizzeitreduzierung von 50 % entspricht. Dieser Darstellung ist schon die Größenordnung der möglichen Heizzeitreduzierung zu entnehmen und entspricht den von Anwendern des EFE-Systems berichteten Werten [RADOSAI 2008, N.N. 2005]. In der Praxis gängige Methoden zur Bestimmung der Mindestheizzeit sind reaktionskinetische Berechnungen und/oder das Aufschneiden der Bauteile zur Ermittlung der Blasenfreiheit („Blow Point-Methode“). Nachfolgend sind die Heizzeitreduzierungen für unterschiedliche Verfahrenseinstellungen wie Einspritzgeschwindigkeit und Werkzeugtemperatur über diese beiden Methoden dargestellt. Auf die Bestimmung der Heizzeitreduzierung mittels Dichteanstieg wie in Abbildung 6 dargestellt, wird verzichtet, da aufgrund der durch die Thermolemente hinterlassenen Aussparungen in den Zylindern eingeschlossene Luftblasen bei der Dichtemessung nicht zu vermeiden sind. Diese führen zu großen Schwankungen (siehe Fehlerbalken in Abbildung 6), welche eine genaue Bestimmung der Heizzeitreduzierung verhindern.



*Abbildung 5: In der Mitte aufgeschnittene Zylinder mit unterschiedlichen Heizzeiten. Von links nach rechts nimmt die Heizzeit zu. Deutlich sind in den untervulkanisierten Bereichen Blasen zu erkennen. In der Mitte des Zylinders entsteht eine Aussparung durch das Thermolement, in welcher sich bei der Dichtemessung Luftblasen der gleichen Größenordnung der Blasen im Bauteil ansammeln.*

*Dies verhindert eine genaue Bestimmung der Heizzeitreduzierung über die Dichtemessung.*

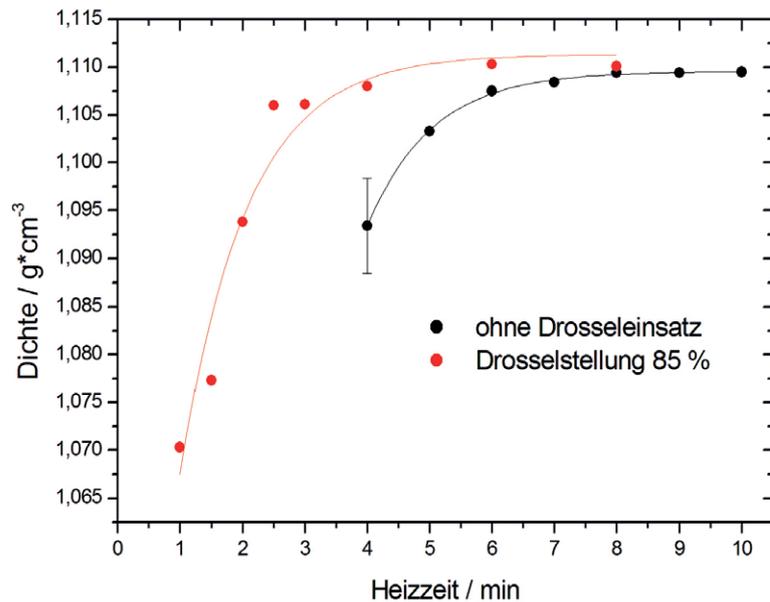


Abbildung 6: Dichte der gespritzten Zylinder für unterschiedliche Heizzeiten und Drosselstellungen (Einspritzgeschwindigkeit  $15 \text{ cm}^3/\text{s}$  und Werkzeugtemperatur  $190 \text{ }^\circ\text{C}$  [NAß 2011])

#### Ermittelte Heizzeitreduzierungen für unterschiedliche Verfahrenseinstellungen

Um für jede Verfahrenseinstellung die maximale Heizzeitreduzierung zu finden, wird für jede Einstellung zunächst die maximal mögliche Drosselstellung ermittelt. Durch die Druckbegrenzung der Anlage auf einen Spritzdruck von 1950 bar ist diese jedoch abhängig von der Einspritzgeschwindigkeit. Mit zunehmender Einspritzgeschwindigkeit der Schnecke nimmt die maximal mögliche Drosselstellung ab, siehe Tabelle 2 [Lemmich 2012].

Tabelle 2 zeigt die für die unterschiedlichen Verfahrensparameter (Einspritzgeschwindigkeit der Schnecke 20, 40 und  $60 \text{ cm}^3/\text{s}$  und Werkzeugtemperaturen 160, 170 und  $180 \text{ }^\circ\text{C}$ ) ermittelten Heizzeitreduzierungen mittels reaktionskinetischer Berechnung anhand der jeweiligen Inkubationszeit für den anisothermen Fall. Im Vergleich hierzu sind die Heizzeitreduzierungen über die experimentelle Ermittlung („Blow Point-Methode“) dargestellt. Deutlich sind niedrigere Heizzeitreduzierungen für den experimentell ermittelten Fall erkennbar. Dies ist zu erklären durch den Mindest-Vernetzungsumsatz, der zum Zeitpunkt der Entformung die Blasenfreiheit erst ermöglicht. Dadurch erhöhen sich die Heizzeiten insgesamt und die Heizzeitreduzierung gemäß Gl.(1) nimmt ab. In Abbildung 7 sind für die Verfahrenseinstellung  $180 \text{ }^\circ\text{C}$  und  $20 \text{ cm}^3/\text{s}$  der jeweils gemessene Temperaturverlauf in der Kavität unmittelbar nach dem Einspritzen sowie der gemäß Reaktionskinetik der Mischung zugehörige berechnete Vernetzungsumsatz dargestellt. Durch den Einsatz der Drossel wird eine Erhöhung der Formeintrittstemperatur von über 15 K erreicht. Für beide Fälle ist der Zeitpunkt der Blasenfreiheit eingetragen. Daraus ist erkennbar, dass man im vorliegenden Fall erst bei einer Entformung des Bauteils bei 24, bzw. 25 % Vernetzungsumsatz ein blasenfreies Produkt erhält.

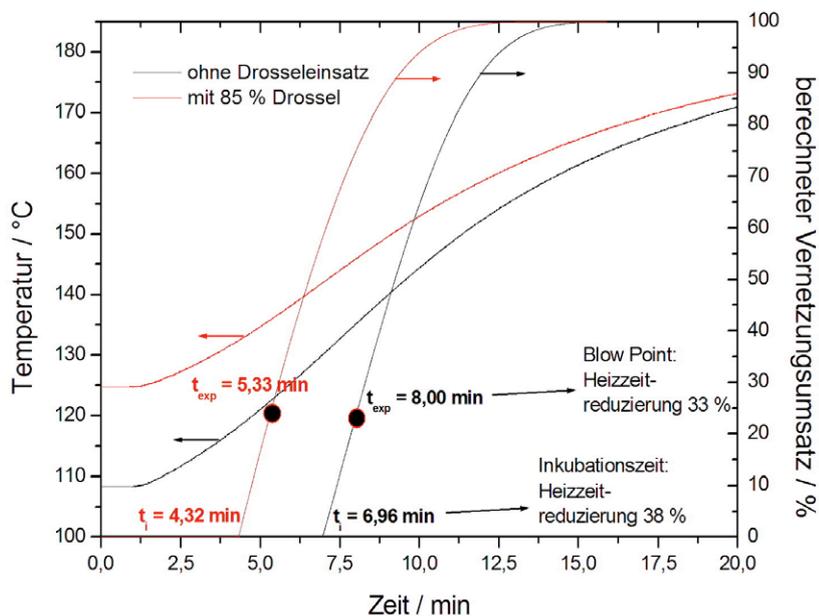


Abbildung 7: Temperaturverlauf in der Mitte des Zylinders unmittelbar nach dem Einspritzvorgang ohne und mit Drosseleinsatz sowie der zugehörige, aus der Reaktionskinetik resultierende berechnete Verlauf des Vernetzungsumsatzes. Die Zeiten  $t_{exp}$ , zu denen jeweils gerade keine Blasen mehr experimentell im Bauteil nachweisbar waren, sind markiert (Einspritzgeschwindigkeit  $20 \text{ cm}^3/\text{s}$  und Werkzeugtemperatur  $180 \text{ °C}$ ) [Lemmich 2012].

Bezieht man sich bei der Berechnung der Heizzeitreduzierung nun auf die jeweiligen Inkubationszeiten, so erhält man eine Reduzierung von 38 %. Über die „Blow Point-Methode“ ergibt sich eine kleinere Heizzeitreduzierung von 33 % aufgrund der längeren notwendigen Mindestheizzeit in der Form. Bei Betrachtung aller auf diese Weise ermittelten Heizzeitreduzierungen fällt nun jedoch auf, dass vereinzelt bei der „Blow Point-Methode“ notwendige Mindestheizzeiten ermittelt wurden, welche sogar unterhalb der rechnerisch bestimmten Inkubationszeit liegen. Die Gründe hierfür sind vielfältig und liegen u.a. auch an folgenden experimentellen Randbedingungen [Lemmich und Herrmann 2012]:

- Die Reaktionskinetik wurde an einem bestimmten Batch durchgeführt und ist u.U. nicht auf jedes Batch übertragbar (Batch-zu-Batch-Statistik). Aufgrund der Größe des Bauteils und der zahlreichen Versuche mussten im Rahmen der Untersuchung mit einem Laborinnenmischer mehrere Batches hergestellt werden. Abbildung 8 zeigt eine statistische Untersuchung der Vulkanisationsisothermen aller verwendeten Batches bei  $160 \text{ °C}$ . Das eingefügte Bild zeigt die Verteilung der Zeiten für einen Vernetzungsumsatz von 10 % zwischen 0,95 und ca. 1,18 min, also einer Schwankungsbreite von etwa 20 %.

- Bei den Spritzversuchen wird daher die Nummer des Batches berücksichtigt und die entsprechende Vulkameterkurve zugrunde gelegt. Die reaktionskinetischen Berechnungen erfolgen daher stets individuell für die entsprechenden Batches. Bei dem Probenstück zur Messung der Vulkanisationsisothermen und dem Spritzling ist jedoch zusätzlich die Schwankung innerhalb des Batches zu berücksichtigen.
- Die Temperaturkurve zur Berechnung der Inkubationszeit stammt von einem vorher definiert durchgeführten Spritzvorgang mit längerer Heizzeit zur Aufnahme des Temperaturverlaufs in der Form für die Messung einer Vulkanisationssimulation, jedoch nicht vom

Spritzvorgang bei der Betrachtung der Blasenfreiheit im Zylinder. Unterschiede in den Temperatur-/Zeitverläufen sind beispielsweise durch unterschiedliche Werkzeugöffnungszeiten nicht auszuschließen.

- Die gemessenen Temperaturen in der Form liegen zu Beginn des Vulkanisationsverlaufs in der Form noch unterhalb des Messbereichs zur Bestimmung der Reaktionskinetik (120 – 190 °C). Die Aktivierungsenergien, welche in diesem Bereich ermittelt wurden, werden auf den Temperaturbereich der Anwendung extrapoliert.
- Effekte durch die Lagerung der Batche können das Vulkanisationsverhalten beeinflussen. So erfolgen die Messungen der Vulkameterkurve und der Reaktionskinetik zu einem früheren Zeitpunkt (ca. 2 Wochen) als der Spritzvorgang und die Aufnahme der Temperaturkurve.
- Aufgrund der Temperierung des Werkzeugs ist mit einem zusätzlichen Temperatureintrag durch den eingebauten Thermofühler in die Kautschukmischung zu rechnen, wodurch sich die Lage der kritischen Stelle aus der Probenmitte heraus nach oben verschiebt. Einen Hinweis darauf liefert auch die Tatsache, dass die Blasen im Zylinder nicht direkt am Thermofühler auftreten, sondern oberhalb des Fühlers (siehe Abbildung 5, links).

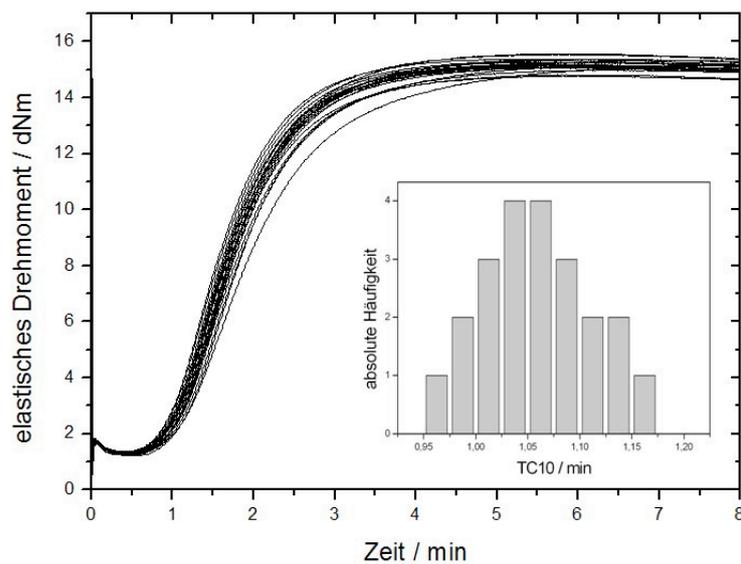


Abbildung 8: Statistik der Vulkanisationsisothermen bei 160 °C aller 22 verwendeten Batche; das eingefügte Bild zeigt die Verteilung der absoluten Häufigkeit der TC10-Zeit, d.h. dem Zeitpunkt der Vernetzung von 10 %.

Mit der Annahme, dass die in Abbildung 8 dargestellten Schwankungen im Vulkanisationsverhalten der unterschiedlichen Batche ausschließlich auf die Präexponentialkonstanten  $t_{i,0}$  der Inkubationszeit und nicht auf die Aktivierungsenergien zurückzuführen sind, lässt sich über die Referenzmethode (Methode der äquivalenten Heizzeit) der Einfluss auf die Inkubationszeit für den anisothermen Fall berechnen [Kleemann 1982]. So kann über die Inkubationszeiten  $t_{eq}$  für jedes Batch bei der Mess-, bzw. Referenztemperatur  $T_{ref} = 160$  °C mittels

$$t_{eq} = \int_0^{t_x} e^{\frac{E_{at}}{R} \left( \frac{1}{T_{ref}} - \frac{1}{T(t)} \right)} dt \quad (2)$$

die Inkubationszeit für den anisothermen Fall  $t_x$  für den jeweiligen Temperaturverlauf  $T(t)$  für jede verfahrenstechnische Einstellung berechnet werden. Hierbei stellt  $R$  die allgemeine Gaskonstante ( $8,313 \text{ Jmol}^{-1}\text{K}^{-1}$ ) dar. Im Falle einer vergleichbaren Aktivierungsenergie für die Inkubationsphase  $E_{at}$  und der Vernetzungsphase  $E_{ak}$  (d.h.  $E_{at} \approx E_{ak}$ ) ist Gl.(2) über die Inkubationsphase hinaus weiter verwendbar und es können auch für die gegebenen Heizzeiten die Vernetzungsumsätze berechnet werden. Im vorliegenden Fall gilt  $E_{at} = 62,20 \text{ kJ/mol}$  und  $E_{ak} = 65,63 \text{ kJ/mol}$ . Somit wird für jeden Zylinder anhand der entsprechenden Vernetzungsisothermen bei  $160 \text{ °C}$  für das verwendete Batch der sich in der Form ergebende theoretische Vernetzungsumsatz in der Mitte des Bauteils berechnet. Man erhält so über alle Verfahreneinstellungen hinweg für den jeweiligen Zeitpunkt rechnerisch Werte zwischen 7 und 25 % Vernetzungsumsatz für die Entformung eines blasenfreien Bauteils (siehe Tab. 2). Es stellt sich heraus, dass sich unter den hier gegebenen Umständen (d.h. bei der gegebenen Mischung und Bauteil, bzw. Werkzeug) die Kombination aus niedrigerer Einspritzgeschwindigkeit und hoher Drosselstellung günstiger auf die Heizzeitreduzierung auswirkt. Ebenso ist eine Tendenz zu höheren Heizzeitreduzierungen bei niedrigeren Werkzeugtemperaturen erkennbar. Bei niedrigeren Werkzeugtemperaturen findet ein geringerer Wärmefluss von der heißen Werkzeugwand in die Mischung statt als bei höheren Werkzeugtemperaturen, wodurch sich gerade in diesem Fall eine erhöhte Starttemperatur durch Schereintrag deutlicher günstig auf die Heizzeitreduzierungen auswirkt.

Verfahrensparameter:

Werkzeugtemperatur / °C	170	170	170	160	180
Einspritzgeschwindigkeit / $\text{cm}^3\text{s}^{-1}$	60	40	20	20	20
Maximal mögliche Drosselstellung / %	55	70	85	85	85

Reaktionskinetische Berechnung:

Inkubationszeit ohne Drosseleinsatz / min	8,1	8,1	8,5	8,8	7,0
Inkubationszeit mit Drosseleinsatz / min	5,4	4,9	4,6	5,0	4,3
daraus resultierende Heizzeitreduzierung / %	33	39	46	43	38

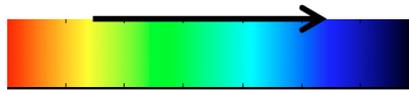
Experimentelle Ermittlung („Blow Point“):

„Blow Point“ ohne Drosseleinsatz / min	8,0	7,9	8,3	9,0	8,0
Vernetzungsumsatz bei der Entformung / %	9	8	7	8	24
„Blow Point“ mit Drosseleinsatz / min	6,2	5,1	5,3	5,3	5,3
Vernetzungsumsatz bei der Entformung / %	19	10	21	9	25
daraus resultierende Heizzeitreduzierung / %	22	35	36	41	33

*Tab. 2: Ermittelte Heizzeitreduzierungen für unterschiedliche Verfahreneinstellungen und unterschiedliche Bestimmungsmethoden. Für den Zeitpunkt der Entformung sind zudem die über Reaktionskinetik („Methode der äquivalenten Heizzeit“) rechnerisch ermittelten Vernetzungsumsätze in der Zylindermitte dargestellt.*

#### 4.2 NMR-Untersuchungen zur qualitativen Darstellung der lokalen Vernetzungsdichte

Die NMR-Bildgebung ist eine Methode, die für medizinische Anwendungen sowie in der Materialforschung ein breites Einsatzfeld hat. Im vorliegenden Fall eignet sich diese Methode hervorragend, um die Zylinder zerstörungsfrei zu vermessen. Darüber hinaus wurde bereits in vielen Untersuchungen die Sensitivität der NMR hinsichtlich der Vernetzungsdichte von Vulkanisaten nachgewiesen [Fedotov und Schneider 1989, Fülber et al. 1996, Blümich und Blümmler 1997]. In den Abbildungen 9 und 10 sind schließlich für jeweils vier Heizzeiten die Photographien der mittigen Schnittflächen sowie NMR-Aufnahmen der ungeschnittenen Zylinder für die Drosselstellung 0 und 85 % dargestellt. Zur Interpretation der NMR-Aufnahmen lässt sich sagen, dass gemäß der Farbreihenfolge rot → gelb → grün → hellblau → dunkelblau die Vernetzungsdichte zunimmt, was mittels einer Eichreihe bestätigt wurde (hier nicht dargestellt).



*Zunahme der Vernetzungsdichte*

In Abbildung 9, links sind für die Heizzeiten 4 und 6 min in dem rot eingezeichneten Bereich noch deutlich untervulkanisierte Zonen zu erkennen, was sich einerseits durch die Blasen erkennbar macht, andererseits an den gelben und roten Bereichen in den NMR-Bildern. Nach 8 min Heizzeit kann in der Schnittfläche lediglich noch eine kleine Blase erkannt werden. Das NMR-Bild hierfür zeigt einen ausvulkanisierten Zylinder, der allerdings im Vergleich zu den bläulichen Bereichen in der Probenmitte noch nicht seine höchste Vernetzungsdichte erreicht hat. Hingegen treten an den Rändern bereits wieder vereinzelte gelbe Stellen auf. Diese Stellen sind auf Reversion zurückzuführen, d.h. dies sind Bereiche in denen aufgrund der langen Temperatureinwirkung durch die 190 °C heiße Werkzeugwand die gebildeten Vernetzungsstrukturen wieder abgebaut werden. Nach 10 min Heizzeit wird dieses noch deutlicher, hier tritt gut erkennbar Reversion am Rand auf. Für diese Vulkanisationszeit sind jedoch keine Blasen im Bauteil mehr erkennbar. Im Vergleich mit Abbildung 6 (gleiche Verfahrensparameter!) wird deutlich, dass nach 8 min das Dichteplateau erreicht wird und je nach Sicherheitsbedürfnis unter Umständen diese Zeit als Heizzeit für diese Einstellung herangezogen werden kann. Betrachtet man nun Abbildung 9, rechts, so ist hier durch den Drosseleinsatz entsprechend bereits bei niedrigeren Heizzeiten die Blasenfreiheit zu erkennen. Beim Zylinder nach 4 min Heizzeit können drei Blasen in der Schnittfläche erkannt werden, das NMR-Bild zeigt noch einen untervulkanisierten Bereich. Nach 8 min ist der Zylinder frei von Blasen und auch in der Mitte ausvulkanisiert. Auch hier zeigt das NMR-Bild einsetzende Reversion am Rand, im vergleichbaren Ausmaß wie auch nach 8 min ohne Drosseleinsatz. In Abbildung 10 sind schließlich die NMR-Bilder für die Heizzeiten 4 und 8 min gegenübergestellt. Als schwarze Linie sind hier zusätzlich die Shore AHärten über dem Probenquerschnitt eingezeichnet. Deutlich ist in beiden Fällen nach 4 min noch eine Untervulkanisation erkennbar, wohingegen mit Einsatz der Drossel in der Zylindermitte bereits ein Shore A Wert von > 40 erreicht wird. Nach 8 min wird für den mit Drosseleinsatz gespritzten Zylinder eine ausvulkanisierte Probenmitte erhalten. Beim Spritzgießen des Zylinders ohne Drosseleinsatz ist die Probenmitte noch nicht ausvulkanisiert, der Probenrand jedoch erfährt bereits Reversion. Somit ist durch die NMR-Bilder gut erkennbar, dass bei der Herstellung dickwandiger Gummibauteile mit Einsatz der Drossel des EFE-Systems und damit verbunden höheren Formeintrittstemperaturen grundsätzlich die Möglichkeit gegeben ist, eine homogenere Vernetzungsdichte über dem Probenquerschnitt zu ermöglichen. Darüber hinaus ist durch den Einsatz kürzerer Heizzeiten generell die Reversion an der Bauteiloberfläche reduziert [NAß 2011]. Eine optimale Vulkanisationszeit für den Einsatz mit Drossel scheint im vorliegenden Fall schließlich nach 7 min gegeben, hier ist eine augenscheinlich homogene Vernetzungsdichte über dem Zylinderquerschnitt zu erkennen (Abbildung 11).

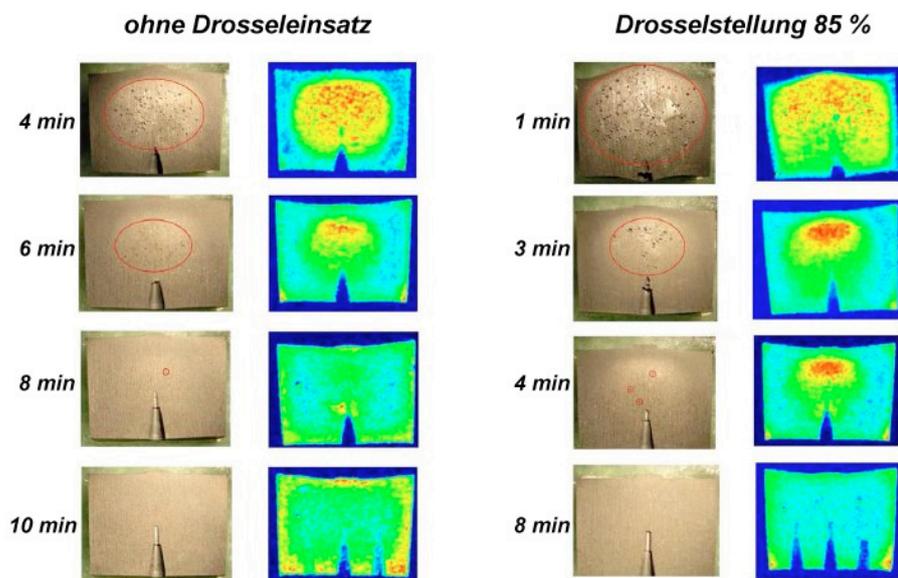


Abbildung 9: Vergleich der Schnittflächen und NMR-Bilder der Zylinder ohne und mit 85 % Drosselstellung (Einspritzgeschwindigkeit  $15 \text{ cm}^3/\text{s}$  und Werkzeugtemperatur  $190 \text{ }^\circ\text{C}$ , [NAß 2011])

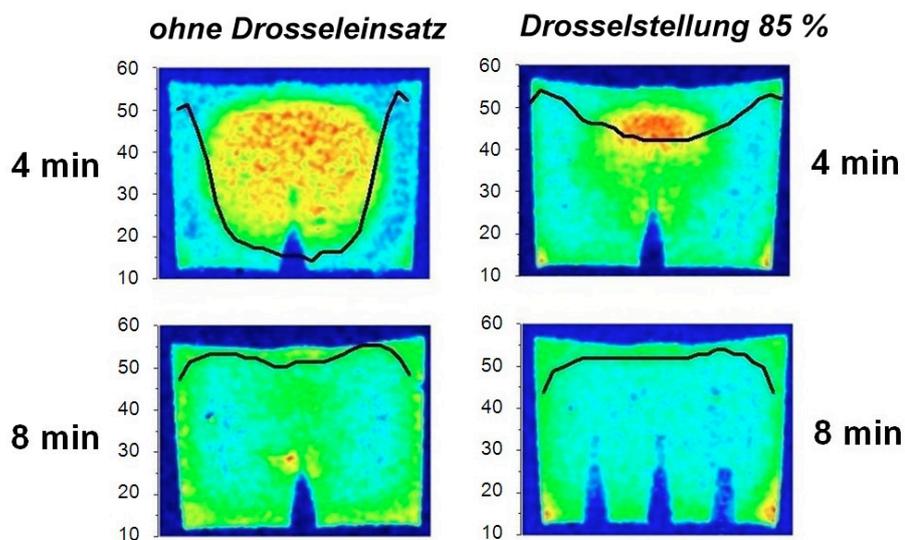


Abbildung 10: NMR-Bilder und Shore A-Verlauf über dem Probenquerschnitt für 4 min (oben), bzw. 8 min (unten) Heizzeit ohne (links), bzw. mit (rechts) 85 % Drosselstellung; am linken Bildrand ist jeweils die Skalierung für Shore A angebracht.

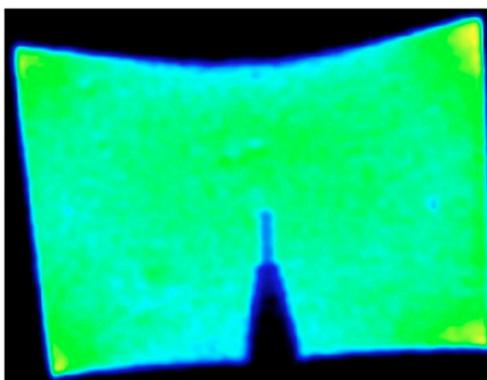


Abbildung 11: NMR-Bild des Zylinders nach 7 min Vulkanisationszeit in der Form bei einer Drosselstellung von 85 %, Einspritzgeschwindigkeit 15 cm<sup>3</sup>/s und Werkzeugtemperatur 190 °C.

## 5 FAZIT

Mittels Schereintrag, der durch das EFE-Spritzaggregat möglich ist, sind Temperaturerhöhungen von über 15 K realisierbar. Die dadurch erzielbaren Heizzeitreduzierungen hängen deutlich von der Bestimmungsmethode sowie der Verfahrenseinstellung (Drosselstellung, Werkzeugtemperatur, Einspritzgeschwindigkeit) ab. Realistische Werte liegen bis bei etwa maximal 40 % Heizzeitreduzierung. Hierbei ist jedoch der Aspekt der Heizzeit, bzw. Zykluszeit nicht getrennt von der Qualität des Bauteils zu betrachten. Für zu kurze Heizzeiten stellt sich bei dickwandigen Bauteilen in jedem Fall – unabhängig von der Drosselstellung – aufgrund der niedrigen Wärmeleitfähigkeit von Kautschukmischungen ein inhomogenes Profil der Vernetzungsdichte über dem Probenquerschnitt ein. Aufgrund der höheren Formeintrittstemperaturen ist durch den Einsatz der Drossel jedoch eine Möglichkeit gegeben, eine gleichmäßigere Vernetzungsdichte einzustellen. Für reversionsgefährdete Kautschukmischungen ist darüber hinaus durch die Verkürzung der Vulkanisationszeiten die Reversion in jedem Fall vermindert, was zudem ein homogeneres Vernetzungsdichteprofil begünstigt. Durch die Untersuchungen von Lemmich konnte gezeigt werden, dass für die hier behandelte Kombination aus Bauteil und Rezeptur bei Heizzeitreduzierungen zwischen 17 und 19 % in der Regel immer noch homogenere Vernetzungsdichten über dem Bauteilquerschnitt durch den Einsatz der Drossel erreicht werden können [Lemmich 2012]. Bis zu etwa 35 % Heizzeitreduzierung werden noch mittels NMR-Bildgebung vergleichbare Vernetzungsdichteprofile erhalten. Erst ab Heizzeitreduzierungen  $\geq 35$  % entstehen tendenziell bei Einsatz der Drossel wieder inhomogenere Vernetzungsdichteprofile, wobei die mechanischen Eigenschaften der Zylinder (z.B. Kompressionsmodul) jedoch noch nicht negativ beeinflusst werden [Lemmich 2012, Wachter 2012].

### Danksagung:

Die Autoren bedanken sich bei der Deutschen Kautschuk-Gesellschaft e.V. (DKG) für die finanzielle Unterstützung bei dem Projekt DKG-7/10-11 sowie dem Arbeitskreis um Prof. Dr. Dr. B. Blümich von der RWTH Aachen für die NMR-Aufnahmen. Dem Dr. Gupta-Verlag, Ratingen sei gedankt für die Genehmigung zur wiederholten Veröffentlichung der Ergebnisse. Diese wurden bereits in einer umfangreichen Trilogie von Januar – März 2013 im Fachmagazin „Gummi Fasern Kunststoffe“ veröffentlicht. Hier sind zusätzliche Informationen und weitere, über die vorliegende Veröffentlichung hinausgehende Ergebnisse zu finden.

## REFERENZEN

BALDT, V., KRAMER, H., KOOPMANN, R., „Temperaturleitzahl von Kautschukmischungen – Bedeutung, Meßmethode und Ergebnisse“, Bayer-Mitteilungen für die Gummi-Industrie Nr. 50 (1978)

GOGOLIN, J., „Rheologische Auslegung eines Rapid Cure Systems, Theorie und Praxis“, Tech Days 2008, 11.-12. September 2008, Landshut, Burg Trausnitz

NASS, F., „Möglichkeiten der Heizzeitreduzierung und Eigenschaftsbeeinflussung eines dickwandigen Bauteils beim Spritzgießen von Kautschukmischungen mittels EFE-Spritzaggregat“, Diplomarbeit, Fachhochschule Würzburg-Schweinfurt, 2011

WACHTER, F., „Untersuchungen zum Einfluss von Mischungsbestandteilen auf die Heizzeitreduzierung und Bauteileigenschaften beim Spritzgießen mittels EFE-Spritzaggregat“, Diplomarbeit, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, 2012

RADOSAI, P. „EFE-Spritzsystem – Eine Technologie setzt sich durch“, Tech Days 2008, Landshut, 11./12.9.2008

N.N., „Kürzere Heizzeiten beim Gummispritzgießen“, Kautschuk Gummi Kunststoffe 58 (2005) S.422

LEMMICH, S., „Untersuchungen zum Einfluss der Verfahrensparameter auf die Heizzeitreduzierung und Bauteileigenschaften beim Spritzgießen mittels EFE-Spritzaggregat“, Diplomarbeit, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, 2012

LEMMICH, S., HERRMANN, V., „Einfluss der Verfahrensparameter auf die Heizzeitreduzierung und Bauteileigenschaften beim Spritzgießen mittels EFE-Aggregat“, Posterbeitrag auf der DKT2012, 2.-5.7.2012, Nürnberg

KLEEMANN, W., „Mischungen für die Elastverarbeitung“, VEB Deutscher Verlag für Grundstoffindustrie, Leipzig (1982)

FEDOTOV, V.D., SCHNEIDER, H., „Structure and Dynamics of Bulk Properties by NMR-Methods“, aus: DIEHL, P. (ed.): NMR Basic Principles and Progress, 1989, Springer Verlag, Heidelberg

FÜLBER, C., UNSELD, K., HERRMANN, V., JAKOB, K.H., BLÜMICH, B., „In situ investigation of SBR vulcanization: A combined study of <sup>1</sup>H-NMR and vulcametry“, Colloid and Polymer Science 274 (1996) 191-196

BLÜMICH, B., BLÜMLER, P., „NMR Imaging of Elastomers: A Review“, Rubb. Chem. & Technol. 70 (1997) 468 – 518

## Analysen von Polyoxymethylenen für Anwendungen in tribologischen Polymer-Polymer-Systemen der Fördertechnik – Teil 1

Dipl.-Ing. (FH) Rene Illek, Dipl.-Ing. (FH) Mareike Hering, Frank Pohli,  
Prof.-Dr.-Ing. Karsten Faust  
Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt  
Fakultät Maschinenbau  
Professur Kunststofftechnik und Produktentwicklung, Fachbereich Tribologie

Aufgrund der großen Bedeutung von Polyoxymethylen (POM) für fördertechnische Anwendungen befassen sich die aktuellen Forschungsarbeiten u.a. mit der Charakterisierung und spezifischeren Analytik von POM für den Einsatz in tribologischen Polymer-Polymer-Systemen (PPS). Auf Basis von umfangreichen Produktrecherchen wurden für aktuelle Polymerentwicklungen verschiedene POM-Typen tribologisch, mechanisch und morphologisch untersucht und gegenübergestellt. In der vorliegenden Veröffentlichung soll aus diesen zahlreichen Untersuchungen eine exemplarische Darstellung zum Zusammenhang einiger für Polyoxymethylen typischen Eigenschaften dargestellt werden. Des Weiteren werden explizite und spezifische Erläuterungen zwischen den resultierenden Kenngrößen und des chemischen sowie morphologischen Aufbaus der Polyoxymethylene aufgezeigt und ergründet.

Kategorie und Themenbeschreibung: Polymertribologie und -analyse

Zusätzliche Schlüsselwörter: Polymere, Polyoxymethylen, Polyacetale, Tribologie, Fördertechnik, Kunststofftechnik

### 1 EINLEITUNG

Der Einsatz technischer Kunststoffe im Bereich der Logistik (Förderung und Lagerung) gewinnt zunehmend an Bedeutung. Eigenschaften wie geringe Dichte, die Möglichkeit zum ungeschmierten Betrieb (Trockenlauf) und eine kostengünstige Serienfertigung unterstützen die vielfältigen Einsatzmöglichkeiten. So wird ein Großteil (40% Marktanteil in EU + USA [Fau10]) von Förderketten aus Kunststoffen hergestellt. Der Werkstoff Stahl wird zunehmend verdrängt, behält aber seine Berechtigung im höheren Lastsegment. Um die Einsatzgrenzen zu erweitern werden die Werkstoffe durch Additive tribologisch und mechanisch optimiert. Der Antrieb von sowohl Kunststoff-, als auch Stahlketten erfolgt durch Kettenräder aus Kunststoff. Als polymere Werkstoffe für Kettenräder werden häufig Polyamid (PA) und Polyoxymethylen (POM) verwendet [Ill13]. Für Gleitketten- und Mattenkettenwerkstoffe kommen

Autoren: Dipl.-Ing. (FH) R. Illek, Dipl.-Ing. (FH) M. Hering, F. Pohli, Prof.-Dr.-Ing. K. Faust  
Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt  
Fakultät Maschinenbau, Ignaz-Schön-Straße 11, 97421 Schweinfurt, Bayern, Deutschland  
Email: rene.illek@fhws.de, karsten.faust@fhws.de  
Telefon: +49 9721 940-9940, Fax: +49 9721 940-900, Website: www.fhws.de

Die Erlaubnis zur Kopie in digitaler Form oder Papierform eines Teils oder aller dieser Arbeit für persönlichen oder pädagogischen Gebrauch wird ohne Gebühr zur Verfügung gestellt. Voraussetzung ist, dass die Kopien nicht zum Profit oder kommerziellen Vorteil gemacht werden und diese Mitteilung auf der ersten Seite oder dem Einstiegsbild als vollständiges Zitat erscheint. Copyrights für Komponenten dieser Arbeit durch Andere als FHWS müssen beachtet werden. Die Wiederverwendung unter Namensnennung ist gestattet. Es andererseits zu kopieren, zu veröffentlichen, auf anderen Servern zu verteilen oder eine Komponente dieser Arbeit in anderen Werken zu verwenden, bedarf der vorherigen ausdrücklichen Erlaubnis.

heute bevorzugt Polyoxymethylen (POM), Polypropylen (PP), Polyamid (PA) und Polyethylen (hoher Dichte, high density) (PE-HD) zur Anwendung, wobei POM und PP am häufigsten Einsatz finden ([Fau10], [Ill13], [Rasch12]).

Polyoxymethylene (Polyacetale) sind technische Kunststoffe, welche sich durch hervorragende mechanische Eigenschaften (Festigkeit, Steifigkeit, Härte, Zähigkeit), tribologische Systemeigenschaften (Gleit- und Abriebverhalten) und sehr gute chemische Beständigkeit auszeichnen. In Abbildung 1 sind die wichtigsten technischen Thermoplaste aus TTS (Tribotechnische Systeme) von förderntechnischen Anwendungen gegenübergestellt. Aus dieser sind die Vorzüge der Polyoxymethylene anhand deren spezifischer Eigenschaften ersichtlich. Da die tribologischen Eigenschaften systemspezifische Parameter sind wurden in dieser Gegenüberstellung auf eine Angabe von Reibungszahlen und Verschleißangaben verzichtet. In tribologischen Paarungen hat jedes Material eigene Vorzüge und Anwendungsgebiete. Allerdings weist POM im Allgemeinen in tribologischen Systemen überragende Eigenschaften auf ([Rös05], [Raa13]). Für die hervorragenden mechanischen und tribologischen Kennwerte spielt die hohe Kristallinität der Polyoxymethylene eine entscheidende Rolle. Das außerordentlich gute Preis/Leistungs-Verhältnis dieses Polymerwerkstoffes ist von entscheidender Bedeutung für eine Vielzahl von Einsatzgebieten.

Eigenschaften	Werkstoffe	POM	PA 66	PBT	PP
<b>Mechanische Eigenschaften</b>		trocken/feucht	trocken/feucht	trocken/feucht	trocken/feucht
• Zugfestigkeit [MPa]		72	85 / 50	60	33
• Zug-E-Modul [MPa]		3100	3100 / 1100	2500	1450
• Kerbschlagzähigkeit [kJ/m <sup>2</sup> ]		9	5 / 20	6	5
• Kugeldruckhärte [MPa]		174	160 / 100	130	72
<b>Tribologische Eigenschaften</b>		Systemeigenschaften			
• Reibungszahl					
• Verschleiß					
<b>Dimensionsstabilität</b>					
• Wasseraufnahme (Normalklima) [%]		0,2	2,8	0,25	< 0,1
• Wasseraufnahme (Wasserlagerung) [%]		0,8	8,5	0,5	< 0,1
• Wärmeausdehnung [10 <sup>-6</sup> /K] längs/quer zur Fließrichtung (FR)		122	70 / 100	130 - 160	100 - 200
<b>Morphologische Eigenschaften</b>		teilkristallin	teilkristallin	teilkristallin	teilkristallin
• Kristallinität [%]		POM-H: 70 - 80% POM-C: 50 - 60%	30 - 45%	40 - 50 %	60 - 70%
• Wärmeleitfähigkeit [W/K·m]		0,37	0,23	0,27	0,22
• Oberflächenenergie [mN/m]		42,1	47,5	43,6	31,2
• Dichte [g/cm <sup>3</sup> ]		1,42	1,14	1,3	0,9
<b>Richtpreis</b>		2,60 - 2,90 €/kg	3,10 - 3,40 €/kg	3,00 - 3,20 €/kg	ca. 1,20 €/kg

Abbildung 1: Vergleich von Thermoplasten für PPS in der Fördertechnik ([Ehr11], [Kern14], [Hel10], [Czy13], [Sae13])

Auf Basis der zahlreichen Anwendungsgebiete in Automobil-, Elektro-, Konsumgüterindustrie, Feinwerktechnik und Materialflusstechnik ist die globale Nachfrage von POM im vergangenen Jahrzehnt deutlich gestiegen. Dies soll an dieser Stelle kurz anhand einer zusammenfassenden Darstellung in Abbildung 2 verdeutlicht werden. Hierbei ist anzumerken, dass über 90% des weltweiten POM-Verbrauchs im Spritzguss verarbeitet werden. ([Kirn13], [Vaa11], [Vaa10], [Sed08], [For07], [Kut05], [Disch02]).

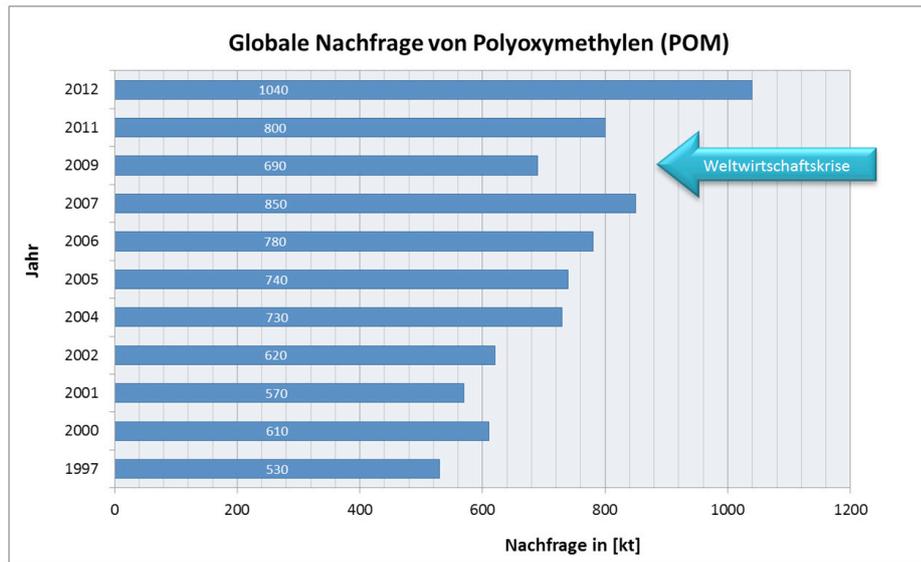


Abbildung 2: Globale Nachfrage von Polyoxymethylen (POM)

## 2 HERSTELLUNG UND AUFBAU DER POLYOXYMETHYLENE

Die Herstellung von Polyoxymethylen unterteilt sich grundlegend in die Erzeugung von Homopolymerisaten (POM-H) und Copolymerisaten (POM-C). Die nachfolgende Abbildung 3 stellt das grundsätzliche Herstellungsschema der Polyoxymethylenen dar und soll an dieser Stelle einen prinzipiellen Überblick über die Ausgangsprodukte (Edukte) und die einzelnen Verfahrensschritte geben. POM-Homopolymerisate werden durch anionische Polymerisation aus Formaldehyd oder alternativ aus Trioxan (Trimere des Formaldehyds) hergestellt. Dagegen werden POM-Copolymerisate durch kationische Ringöffnungspolymerisation von Trioxan unter Zugabe von 2 - 4 [Gew.-%] Ethylenoxid (Comonomer-Bausteine) erzeugt. Hierbei erfolgt die Verteilung dieser Comonomer-Bausteine statistisch entlang der Polymerkette, wobei zu Beginn des Polymerisationsvorgangs vorwiegend comonomerreiches POM-C entsteht, da die Comonomere schneller polymerisieren als Trioxan. Diese blockartigen Comonomersequenzen werden durch Transacetalisierung, welche parallel zur Wachstumsreaktion abläuft, weitgehend statistisch verteilt. Es treten bei den Copolymeren zusätzlich oftmals sogenannte Fines mit ca. 5 - 7 [Gew.-%] auf, d. h. ein erhöhter Anteil an niedermolekularen Kettenlängen die meist zyklisch verteilt sind. Dieser Fine-Anteil wirkt wie ein innerer Schmierstoff, führt aber gleichzeitig zu veränderten mechanischen Kennwerten. Die Produkte weisen durch die niedermolekularen Anteile (verkürzte Polymerketten - Fines) eine höhere Steifigkeit aber gleichzeitig etwas geringere Zähigkeit auf. ([Kai11], [Gee06], [Gna91], [Wit14]).

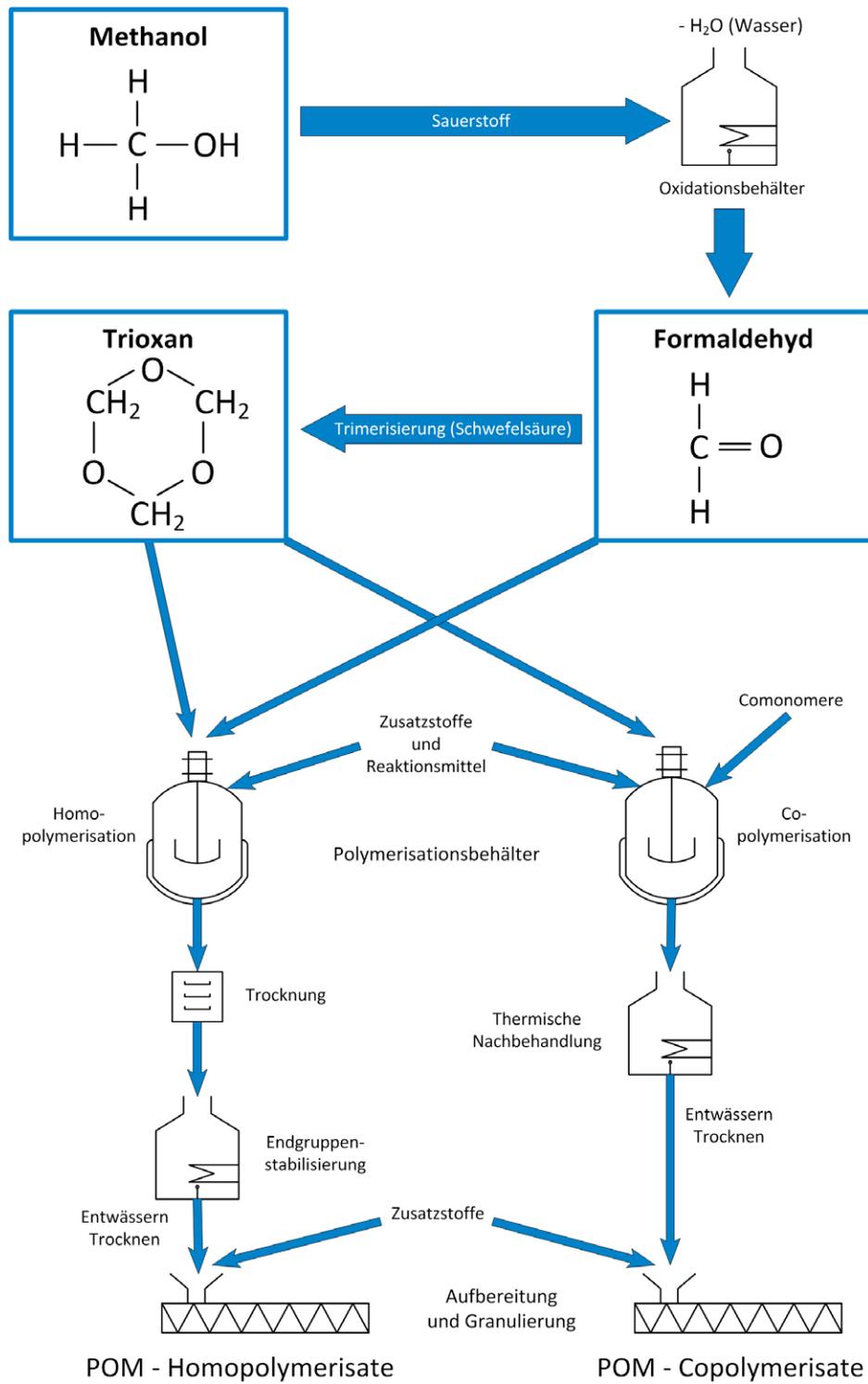
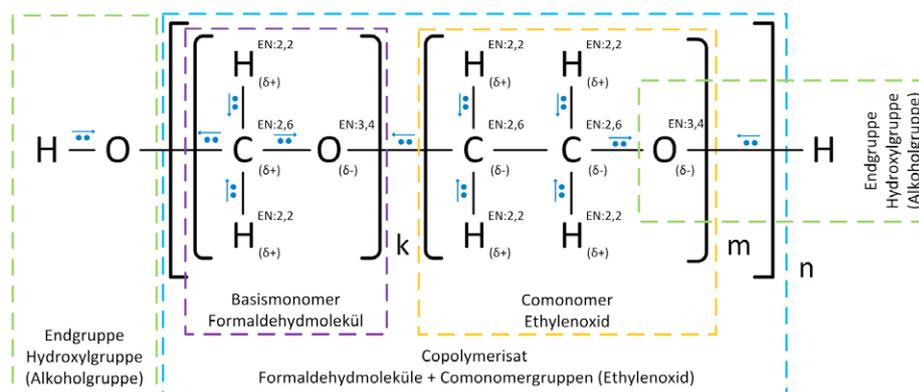


Abbildung 3: Grundsätzliches Herstellungsschema der Polyoxymethylene

Der strukturelle Aufbau von POM-Homopolymersisaten und POM-Copolymerisaten ist in der nachfolgenden Abbildung 4 dargestellt. In dieser sind außerdem die Elektronegativitäten EN der Atome und die aus den Differenzen der Elektronegativitäten resultierenden Dipole mit ihrer jeweiligen Partialladung ( $\delta^+$ ,  $\delta^-$ ) dargestellt. Die Dipole in Polymerketten bewirken Anziehungskräfte (Dipol-Dipol-Kräfte), auf die nächstliegenden Molekülketten. In den adhäsiven Kontaktstellen von tribologischen Polymerpaarungen bestimmen diese u.a. die adhäsiven Bindungskräfte beziehungsweise Kraftpotentiale zwischen den Reibpartnern.

### POM-C



### POM-H

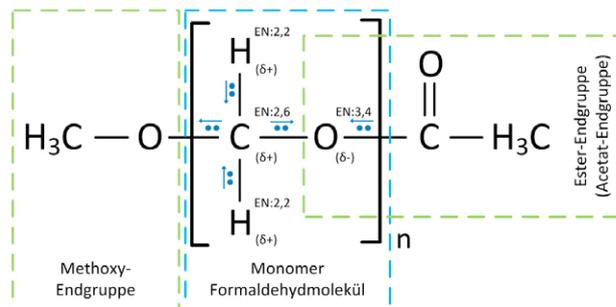


Abbildung 4: Strukturformeln von POM-C und POM-H

Die Comonomer-Bausteine sind in den POM-C-Kettenmolekülen, wie bereits einleitend erwähnt, statistisch verteilt (k und m statistisch verteilt). Es ist anzumerken, dass der Comonomeranteil (CMA) mit 2 - 4 [Gew.-%] gegenüber dem Basismonomeranteil (Formaldehydmolekülen) relativ gering ist. Trotzdem ergeben sich dadurch mannigfaltige Unterschiede zwischen POM-H und POM-C. Ein wesentlicher Unterschied ist die geringere Kristallinität von POM-C (siehe Abbildung 1). In Abhängigkeit vom Kristallisationsgrad ändern sich mechanische, tribologische und morphologische Eigenschaften, welche wiederum auf sehr komplexe Weise voneinander abhängig sind. Mit steigendem Kristallisationsgrad nehmen u.a. die Dichte, Elastizitätsmodul, Zugfestigkeit (Streckgrenze) und Härte zu. Gleichzeitig nehmen die Schlagzähigkeit, Bruchdehnung (Verstreckbarkeit), mechanische Dämpfung, Beständigkeit gegen Spannungsrissbildung, thermische Ausdehnung, Maßbeständigkeit, Quellung und Lichtdurchlässigkeit ab [Ehr11]. Unterschiedliche molare Massenverteilungen (Polymerisationsgradverteilungen) resultieren in unterschiedlichen Viskositätsklassen der Polymere. Eine

hochviskose Type ist demnach aus längeren Molekülketten (hohe molare Masse) aufgebaut, wobei eine niedrigviskose Type (niedrige molare Masse) kürzere Molekülketten enthält. Mit zunehmender Viskosität (Molekulargewicht) sinkt der Kristallisationsgrad, da die längeren Kettenmoleküle eine Kristallitbildung behindern bzw. erschweren [Ehr11]. Auf Basis dieser Grundlagen sollen die nachfolgenden wissenschaftlichen Untersuchungsergebnisse begründet werden.

### 3 VERGLEICH VON POM-HOMOPOLYMEREN UND -COPOLYMEREN

Aufgrund der enormen Bedeutung von POM für zahlreiche Anwendungen speziell in TTS der Fördertechnik, auf Basis von Polymer-Polymer-Paarungen, wurden unterschiedliche Polyoxymethylen-Typen einhergehend untersucht und gegenübergestellt, um grundlegende wissenschaftliche Erkenntnisse zu generieren und zu intensivieren. Es wurden sowohl Homopolymerisate als auch Copolymerisate betrachtet, wobei diese jeweils in unterschiedlichen Viskositätsklassen (Melt-Flow-Rate (MFR)) und weiteren Modifikationen (z.B. Typen mit Nukleierungsmittel (NM) oder Type mit reduziertem Comonomeranteil (R-CMA) analysiert wurden.

Es wurden die verschiedenen POM Werkstoffe anhand Ihrer mechanischen Kenngrößen (Zugfestigkeit, Zug-E-Modul, Schlagzähigkeit und Härte), tribologischen Kenngrößen (Reibungszahl, Reibkontakttemperatur und Verschleiß) sowie Morphologie (Oberflächenenergie, Kristallinität und Molare Massenverteilung) verglichen. Die Untersuchungen wurden anhand von Vielzweckprobekörpern für Kunststoffe nach [DIN 3167] und tribologischen Prüfkörpern (Unterproben (Platte: 200 x 80 x 3 mm), Oberproben (Platte: 15 x 10 x 3 mm)) durchgeführt. Die Herstellung der Prüfkörper erfolgte im Spritzgussverfahren unter definierten Verfahrensparametern nach den Empfehlungen der Rohstoffhersteller und Anpassung an das Spritzgusswerkzeug.

In diesem ersten Teil sollen die Ergebnisse der mechanischen Untersuchungen dargestellt und ergründet werden. In den nachfolgenden Abbildungen 5 und 6 sind die ausgewählten Polyoxymethylene anhand deren ermittelten Zugfestigkeit, Zug-E-Modul und Schlagzähigkeit gegenüber gestellt. Es ist hieraus ersichtlich, dass keine reproduzierbare Abhängigkeit zwischen der MFR bzw. der Viskosität und der Zugfestigkeit sowie dem Zug-E-Modul festgestellt werden konnte. Hingegen nimmt die Kerbschlagzähigkeit mit abnehmender MFR zu, demnach weisen hochviskose POM-Typen höhere Zähigkeiten auf als niedrigviskose. Dies ist begründet durch die stärkeren Verschlaufungen und die Abgleitbehinderung der längeren Molekülketten hochviskoser Typen. Durch die Anwendung von Nukleierungsmitteln zeigt sich eine Erhöhung der Zugfestigkeit, Zug-E-Modul und Kerbschlagzähigkeit. Dieser Anstieg der getesteten mechanischen Kennwerte resultiert aus der höheren Kristallinität dieser Typen durch die NM. Die höheren mechanischen Kennwerte der POM-H Typen im Vergleich zu den Copolymerisaten (POM-C) sind ebenfalls durch die höhere Kristallinität der Homopolymere begründet. Durch eine Reduzierung des Comonomeranteils (R-CMA) konnte keine wissenschaftlich vertretbare Verbesserung der untersuchten mechanischen Kenngrößen festgestellt werden. Dennoch können über den CMA die Eigenschaften der Polyoxymethylene verändert werden ([Kai11], [Gee06], [Gna91], [Wit14]). Auf Basis der vorliegenden Ergebnisse scheinen sich die Unterschiede zwischen POM Typen mit differenten CMA allerdings deutlich geringer auszuwirken als zunächst vermutet.

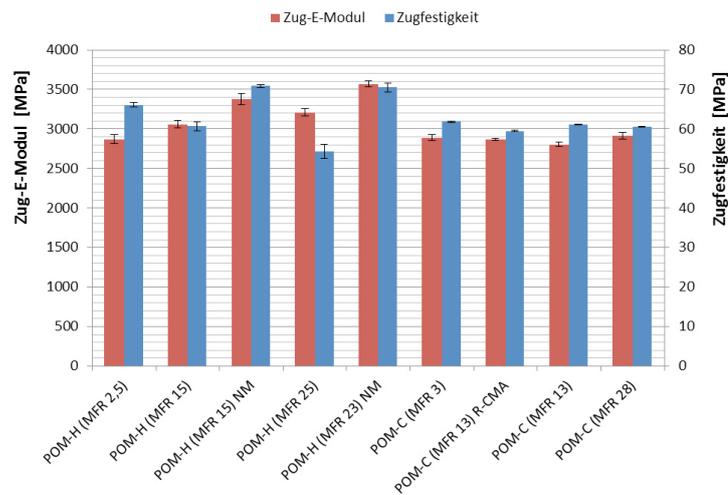


Abbildung 5: Zugfestigkeit und Zug-E-Modul von Polyoxymethylenen

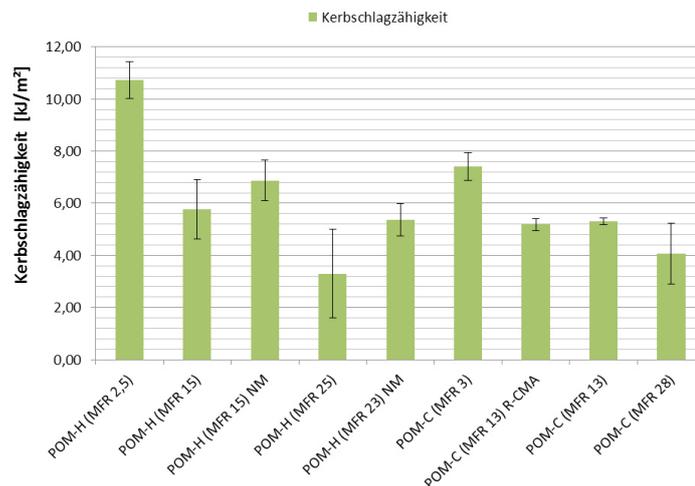


Abbildung 6: Kerbschlagzähigkeit von Polyoxymethylenen

Die Härtemessungen erfolgten mittels Kugeleindruckhärteprüfung nach [DIN 2039] an den tribologischen Probekörpern. Hierzu wurde auf jedem Probekörper eine Matrix an Messpunkten aufgezeichnet. Daraus wurde anschließend eine Härtepunktematrix erzeugt und ausgewertet. Es konnte zum einen nachgewiesen werden, dass sich die Härtewerte an den Rändern der Probekörper deutlich von den Härtewerten in der Probenmitte unterscheiden. Des Weiteren wurden auch Härteunterschiede zwischen angussnahen und angussfernen Messpunkten festgestellt. Diese Phänomene sind begründet durch die schnelleren Abkühlungsbedingungen an den Rändern der Prüfkörper und angussfernen Messpunkten. Aufgrund der langsameren Abkühlung in der Mitte der Probekörper resultiert eine höhere Kristallinität in diesen lokalen Bereichen. Aus diesem Grund weisen diese Segmente höhere Härtewerte auf. Der signifikante Einfluss der Geometrie auf die Härtewerte in geometrisch unterschiedlichen Systemen wird auch beim direkten Vergleich von Abbildung 7 und Abbildung 8 deutlich. Hier zeigen die wesentlich kleineren Oberproben, im Vergleich zu den größeren Unterproben, geringere Härtewerte. Wobei die Differenzen unterschiedlich ausgeprägt ausfallen. Beim Vergleich von POM-Homopolymeren und Copolymeren weisen letztere geringere Härtewerte auf, was

wiederum durch die höhere Kristallinität der Homopolymerisate begründet ist. Durch die Anwendung von Nukleierungsmitteln zeigt sich wiederum eine Erhöhung der Härtewerte, wobei dieser Effekt bei den Oberproben (Abbildung 8) deutlich ausgeprägter ausfällt als bei den Unterproben (Abbildung 7). Dieses Resultat stützt die These, dass sich durch die NM der Kristallisationsgrad erhöht und gerade bei schnellen Abkühlbedingungen während der Herstellung (Geringen Zykluszeiten) sich verbesserte mechanische Kenngrößen ergeben als im Vergleich zu Typen ohne NM. Des Weiteren zeigt sich, dass die Härte mit sinkender MFR abnimmt, d. h. hochviskose Typen lassen niedrigere Härtewerte erkennen als niedrigviskose. Auch dieser Effekt ist je nach geometrischem System unterschiedlich ausgeprägt (Vergleich Abbildung 7 und Abbildung 8). Der R-CMA zeigt wiederum keinen eindeutigen Einfluss sondern verhält sich bei Ober- und Unterprobe gegensätzlich, was potentiell auf die unterschiedlichen geometrischen Gegebenheiten zurückzuführen ist.

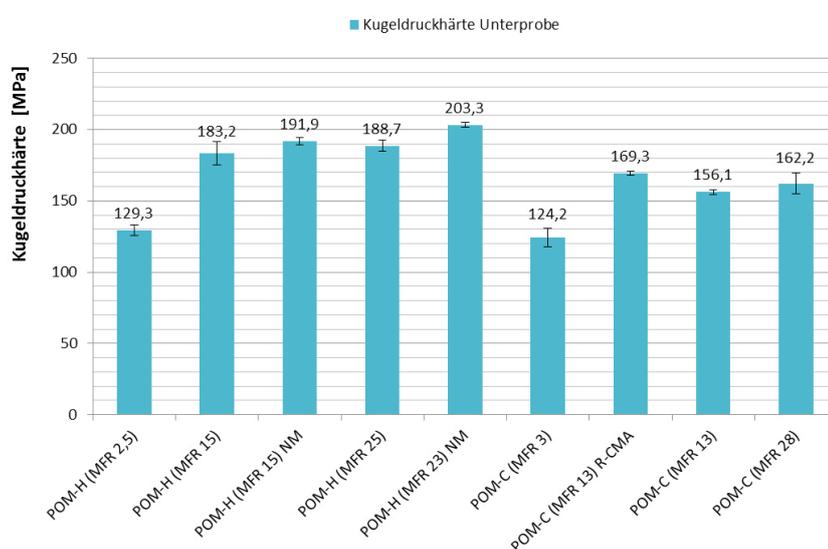


Abbildung 7: Kugeleindruckhärte der Unterproben (Tribologischer Grundkörper) von Polyoxymethylenen

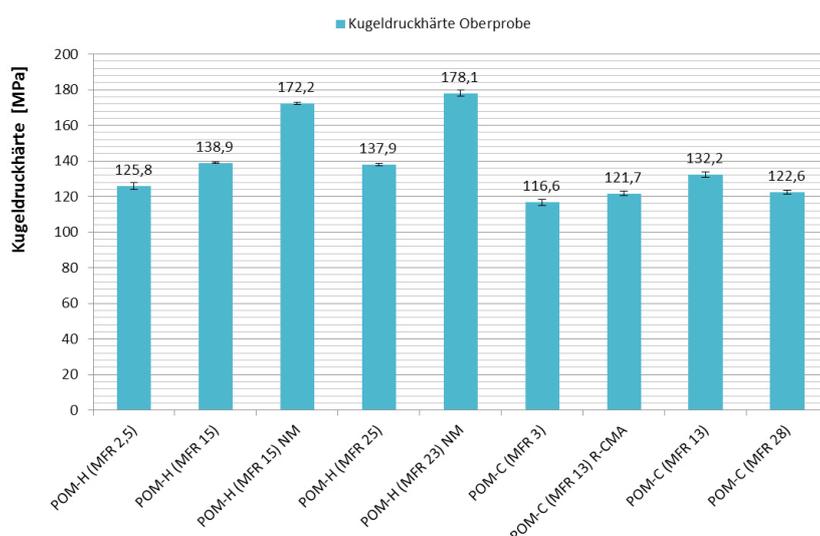


Abbildung 8: Kugeleindruckhärte der Oberproben (Tribologischer Gegenkörper) von Polyoxymethylenen

#### 4 ZUSAMMENFASSUNG

Es wurden verschiedene Polyoxymethylentypen systematisch ausgewählt und einander in puncto mechanischen, tribologischen und morphologischen Gesichtspunkten gegenübergestellt. Als POM-Typen wurden sowohl Homopolymerisate als auch Copolymerisate betrachtet, welche jeweils in unterschiedlichen Viskositätsklassen (Melt-Flow-Rate (MFR)) und weiteren Modifikationen (z.B. Typen mit Nukleierungsmittel (NM) oder Type mit reduziertem Comonomeranteil (R-CMA)) analysiert wurden.

Im ersten Teil dieser Veröffentlichung wurden die mechanischen Kenngrößen (Zugfestigkeit, Zug-E-Modul, Kerbschlagzähigkeit und Härte) der POM Typen ausgeführt. Gerade bei Typen mit NM wurden bessere mechanische Kenngrößen ermittelt. Eine Type mit R-CMA hat dagegen keine eindeutigen Verbesserungen aufgezeigt. Die unterschiedlichen Viskositäten (MFR) haben keine Differenzierung der Zugfestigkeit und des Zug-E-Modul in deren Abhängigkeit zugelassen. Allerdings wurden deutliche Unterschiede bei Kerbschlagzähigkeit und Härte aufgezeichnet. Es hat sich bei den Untersuchungen herausgestellt, dass eine Härteanalyse in Form einer Härtepunktmatrix von sehr zentraler Wichtigkeit ist, gerade auch für Weiterführende tribologische Untersuchungen, um sinnvolle und repräsentative Prüfbereiche auf den Probekörpern festzulegen. Des Weiteren wurde aufgezeigt, dass Härteanalysen bei geometrisch differenten Systemen sehr unterschiedlich ausfallen können, die Tendenzen jedoch durchaus vergleichbar sind.

In den nächsten Teilen der Veröffentlichung sollen die Ergebnisse von morphologischen und tribologischen Untersuchungen dargestellt und anhand der kunststofftechnischen Grundlagen ergründet werden.

## REFERENZEN

- [Czy13] Czyborra L.: Fachseminar - Tribologische Eigenschaften von thermoplastischen Kunststoffen: Süddeutsches Kunststoff-Zentrum (SKZ), Würzburg, 2013.
- [Disch02] Disch S.; Szneczana J.; Nenad K.: Polyoxymethylen (POM): Kunststoffe - Werkstoffe, Verarbeitung, Anwendung (S. 86 - 93). Carl Hanser Verlag, München, 10/2002.
- [DIN 2039] DIN EN ISO 2039-1: Kunststoffe - Bestimmung der Härte - Kugeleindruckversuch: Beuth Verlag, Berlin, Ausgabe Juni 2003.
- [DIN 3167] DIN EN ISO 3167: Kunststoffe - Vielzweckprobekörper: Beuth Verlag, Berlin, Ausgabe Dezember 2003.
- [Ehr11] Ehrenstein Gottfried W.: Polymer Werkstoffe - Struktur, Eigenschaften, Anwendung: Carl Hanser Verlag, München, 3. Auflage, 2011.
- [Fau10] Faust Karsten: Neue polymere Werkstoffkonzepte für Fördergleitketten und Systemanalyse der Korrelation von Reibungskoeffizienten: Dissertation, TU Chemnitz, 2010.
- [For07] Forschler B.; Geiser Ch. Polyoxymethylen (POM): Kunststoffe - Werkstoffe, Verarbeitung, Anwendung (S. 112 - 114). Carl Hanser Verlag, München, 10/2007.
- [Hel10] Hellerich W.; Harsch G.; Baur E.: Werkstoff-Führer Kunststoffe: Carl Hanser Verlag, München, 10. Auflage, 2010.
- [Ill13] Illek R., Kellner S., Faust K.: Kombination von Produktentwicklung, Kunststofftechnik und Tribologie für neue Werkstoffkonzepte in fördertechnischen Anwendungen: FHWS Science Journal, Jahrgang 1 - Ausgabe 2, Scheinfurt - Würzburg, 2013.
- [Kai11] Kaiser Wolfgang: Kunststoffchemie für Ingenieure: Carl Hanser Verlag, München, 3. Auflage, 2011.
- [Kern14] Kern: KERN RIWETA Material Selector 4.0: Werkstoffdatenbank für Kunststoffe, www.kern.de, 2014.
- [Kirn13] Kirner H.-H.; Palakaya F.: Polyoxymethylen (POM): Kunststoffe - Werkstoffe, Verarbeitung, Anwendung (S. 86 - 94). Carl Hanser Verlag, München, 10/2013.
- [Kut05] Kutta S.: Polyoxymethylen (Polyacetale, POM): Kunststoffe - Werkstoffe, Verarbeitung, Anwendung (S. 101 - 106). Carl Hanser Verlag, München, 10/2005.
- [Gee06] Geerißen H.: Skript Kunststoffchemie - Einführung in die Allgemeine und organische Chemie: Hochschule Darmstadt, 11. Auflage, 2006.
- [Gna91] Gnauck B.; Fründt P.: Einstieg in die Kunststoffchemie: Carl Hanser Verlag, München, 3. Auflage, 1991.
- [Raa13] Raab Peter: POM ein überraschend vielseitiger Werkstoff: Vortrag und Präsentation, Fachkolloquium Kunststoff-Gleitketten und Tribologie, TU Chemnitz, 2013.

[Rasch12] Rasch Frank: Reibungsminderung an Stütz- und Führungselementen für Kunststoffketten. Dissertation, TU Chemnitz, 2012.

[Rös05] Rösler Jan: Zur Tragfähigkeitssteigerung thermoplastischer Zahnräder mit Füllstoffen: Dissertation, TU Berlin, 2005.

[Sae13] Saechtling: Kunststoff Taschenbuch: Carl Hanser Verlag, München, 31. Auflage, 2013.

[Sed08] Sedler F.: Polyoxymethylen (Polyacetale, POM): Kunststoffe - Werkstoffe, Verarbeitung, Anwendung (S. 160 - 165). Carl Hanser Verlag, München, 10/2008.

[Vaa11] Vaahs Tilo: Polyoxymethylen (POM): Kunststoffe - Werkstoffe, Verarbeitung, Anwendung (S. 64 - 68). Carl Hanser Verlag, München, 10/2011.

[Vaa10] Vaahs Tilo; Smeets D.: Polyoxymethylen (POM): Kunststoffe - Werkstoffe, Verarbeitung, Anwendung (S. 124 - 127). Carl Hanser Verlag, München, 10/2010.

[Wit14] Witan K.: Fachbereich Maschinenbau und Kunststofftechnik, Hochschule Darmstadt, Gesprächs- und Emailnotizen vom Juni 2014.



## Verschleißquantifizierungsverfahren für Polymer-Polymer-Gleitpaarungen

Dipl.-Ing.(FH) Rene Illek, Frank Pohli, Stefan Kellner, Prof. Dr.-Ing. Karsten Faust  
Fakultät Maschinenbau, Bereich Kunststofftechnik, Produktentwicklung und Tribologie  
Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt

Aktuelle wissenschaftliche Forschungsarbeiten des Fachbereiches befassen sich u.a. mit der Verschleißquantifizierung von Kunststoff-Kunststoff-Gleitpaarungen in tribotechnischen Systemen (TTS) aus dem Bereich der Fördertechnik. Im ersten Schritt wurden, aufbauend auf durchgeführten Marktrecherchen, verschiedene Analyse- und Messverfahren anhand von Testmessungen detailliert betrachtet, bewertet und ggf. validiert. Im Weiteren wurde ein gravimetrisches Verschleißquantifizierungsverfahren für Polymer-Polymer-Gleitpaarungen ausgearbeitet. Zur Validitätsprüfung wurden die Einflussfaktoren/Störgrößen dieses Verfahrens ermittelt, bewertet und ggf. durch entsprechende Abhilfemaßnahmen oder Korrekturfaktoren berücksichtigt. Ein exemplarischer Überblick, zur gravimetrischen Verschleißquantifizierung, soll die neuen Möglichkeiten und Grenzen dieses Verschleißanalyseverfahrens aufzeigen.

Kategorie und Themenbeschreibung: Polymertribologie und -analyse

Zusätzliche Schlüsselwörter: Verschleiß, Verschleißquantifizierung, Polymere, Tribologie, Fördertechnik, Kunststofftechnik

### 1 EINLEITUNG

Die Verschleißbewertung und die Beurteilung der tribologischen Eignung von Polymeren in Kunststoff-Kunststoff Gleitpaarungen wird bisher mit dem an der TU Chemnitz entwickelten Verschleißbewertungsverfahren durch Verschleißkennwerte KV und dem Tribokennwert KT durchgeführt [Sum11]. Hierbei werden die Verschleißkennwerte der tribologischen Probekörper anhand der vorliegenden Verschleißerscheinungsformen visuell begutachtet und durch subjektive Zuordnung, anhand von Referenzbildern, in bewertete Kategorien klassifiziert. Je nach Kategorie der Verschleißintensität wird den einzelnen Probekörpern (Grundkörper und Gegenkörper) ein Verschleißkennwert KV zugeordnet. Anhand dieses Kennwertes wird der Verschleiß der Polymere charakterisiert. Der Tribokennwert vereint den Verschleißkennwert und den Reibungskoeffizienten  $\mu$  mittels mathematischer Operatoren zur Beurteilung der Gleitpaarungseignung [Sum11]. Diese Vorgehensweise kann nicht als Ersatz für eine Verschleißmessung durch physikalische Größen herangezogen werden. Die subjektive Einteilung des Verschleißes durch den Verschleißkennwert KV entspricht einem

Autoren: Dipl.-Ing.(FH) R. Illek, F. Pohli, S. Kellner, Prof. Dr.-Ing. K. Faust  
Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt  
Fakultät Maschinenbau, Ignaz-Schön-Straße 11, 97421 Schweinfurt, Bayern, Deutschland  
E-Mail: rene.illek@fhws.de, karsten.faust@fhws.de  
Telefon: 09721 940-9940, Fax: 09721 940 900, Website: www.fhws.de

Die Erlaubnis zur Kopie in digitaler Form oder Papierform eines Teils oder aller dieser Arbeit für persönlichen oder pädagogischen Gebrauch wird ohne Gebühr zur Verfügung gestellt. Voraussetzung ist, dass die Kopien nicht zum Profit oder kommerziellen Vorteil gemacht werden und diese Mitteilung auf der ersten Seite oder dem Einstiegsbild als vollständiges Zitat erscheint. Copyrights für Komponenten dieser Arbeit durch Andere als FHWS müssen beachtet werden. Die Wiederverwendung unter Namensnennung ist gestattet. Es andererseits zu kopieren, zu veröffentlichen, auf anderen Servern zu verteilen oder eine Komponente dieser Arbeit in anderen Werken zu verwenden, bedarf der vorherigen ausdrücklichen Erlaubnis.

Verschleißbewertungsverfahren und keinem Verschleißmessverfahren. Dennoch können durch dieses Bewertungsverfahren, vorzugsweise bei höheren und/oder deutlich differenten Verschleißintensitäten, die vorliegenden Tendenzen valide berücksichtigt und aufgezeigt werden. Nach [Gre13] und [Fein13] werden Verschleißmessmethoden in drei Hauptgruppen eingeteilt. Innerhalb der Vorabrecherchen und ersten Testmessungen im Rahmen eines Industrieprojektes wurden die Verschleißmessverfahren dieser einzelnen Gruppierungen erfasst und bewertet. Aufgrund der hieraus generierten Ergebnisse wurde der Fokus für detailliertere Betrachtungen auf die Verfahren der Hauptgruppe I (Erfassung der verschleißbedingten Maß- und/oder Gewichtsänderungen von Bauteilen/Prüfkörpern) gerichtet. Innerhalb der Recherchen wurden Offline- und Online-Verschleißmessverfahren betrachtet. Im Weiteren wurden allerdings potentielle Online-Messverfahren verworfen, da diese gerade bei Polymer-Polymer-Gleitpaarungen eine Vielzahl von unzureichenden Schwächen aufweisen. Wesentlich sind hierbei vor allem, dass sich dem Online-Verschleißsignal zahlreiche Störgrößen überlagern, wie beispielsweise Wärmeausdehnungen, Materialübertragung und Kriechvorgänge der Polymere. Diese Störgrößen sind messtechnisch und in der späteren Auswertung nicht voneinander zu differenzieren ([GfT11], [Gre13], [Fein13]). Zentral ist auch die Thematik, dass bei einer Online-Verschleißmessung nur der Systemverschleiß im Gesamten oder der eines Reibpartners (Voraussetzung: Vernachlässigbarer Verschleiß eines Reibpartners) quantifiziert werden können. Zur Verschleißanalyse eines TTS ist es nach [Czi10] erforderlich den Verschleiß des Grund- und Gegenkörpers explizit zur erfassen und zu bewerten (Ganzheitliche Systembetrachtung). Gerade aufgrund der Systemeigenschaften von tribologischen Kenngrößen ist dies notwendig, um Werkstoffpaarungen und speziell Polymer-Polymer-Paarungen im Gesamten tribologisch beurteilen zu können.

## 2 ALLGEMEINE VERSCHLEIßQUANTIFIZIERUNGSVERFAHREN

Im Vordergrund der Verfahrensrecherchen standen Messverfahren zur Erfassung direkter Verschleißmessgrößen an gleitmodifizierten und / oder faserverstärkten Polymeren in Kunststoff-Kunststoff-Gleitpaarungen, an den Probekörpern des eingesetzten Tribometers (TriboSPEEDSTER 2.0). Die direkten und bezogenen Verschleißmessgrößen bzw. Verschleißkennwerte wurde in Anlehnung an das Arbeitsblatt Nr. 7 der Gesellschaft für Tribologie [DFS02] und [Czi10] definiert und benannt. Als direkte Verschleißmessgrößen wurden somit der massenmäßige Verschleißbetrag  $W_m$  [kg] und der volumetrischen Verschleißbetrag  $W_v$  [m<sup>3</sup>] herangezogen. Aus diesen wurden im Anschluss, unter Einbeziehung der Versuchsparameter, die bezogenen Verschleißkenngrößen abgeleitet. Im Fokus standen hierbei die spezifische Verschleißintensität  $W_s$  [m<sup>3</sup>/Nm], die volumetrische Verschleißgeschwindigkeit  $W_{vt}$  [m<sup>3</sup>/h] sowie das volumetrische Verschleiß-Weg-Verhältnis  $W_{vs}$  [m<sup>3</sup>/m].

Innerhalb einer detaillierten Marktanalyse wurden die potentiell geeigneten Verfahren zur Verschleißquantifizierung nach Art und Ausführung der Messinstrumente in die Kategorien Mikroskopie, Sensorik, Spektroskopie, Massenspektroskopie, Indirekte Messverfahren (Analyse mechanischer, physikalischer und elektrischer Eigenschaften) sowie sonstige Verfahren strukturiert. Zusätzlich wurden die Verfahren hinsichtlich der analysierbaren Werkstoffe gekennzeichnet. Die Spezifikationen der einzelnen Messinstrumente wurden im Bezug zur Aufgabenstellung methodisch beurteilt und zusammengestellt. Hierbei haben sich die Messverfahren der Mikroskopie, Sensorik sowie eine gravimetrische Verschleißquantifizierung durchgesetzt und wurden zur vollständigen Erfüllung der Zielvorgabe nochmals detaillierter gegenübergestellt und bewertet. Die Weiteren Methoden wurden als ungeeignet eingestuft, da eine Erfassung der verschleißbedingten Maß- oder Masseänderungen zur Generierung direkter Verschleißmessgrößen nicht oder nur durch einen erheblichen Zusatzaufwand durchführbar

ist. Dies erfordert beispielsweise bei indirekten Verschleißmessmethoden (Hauptgruppe III nach [Gre13] und [Fein13]) Referenzskalen/Kennlinien zur Umrechnung der physikalischen Messgrößen. In Abbildung 1 und Abbildung 2 sind die Ergebnisse der Marktrecherchen der Kategorien Mikroskopie und Sensorik exemplarisch dargestellt.

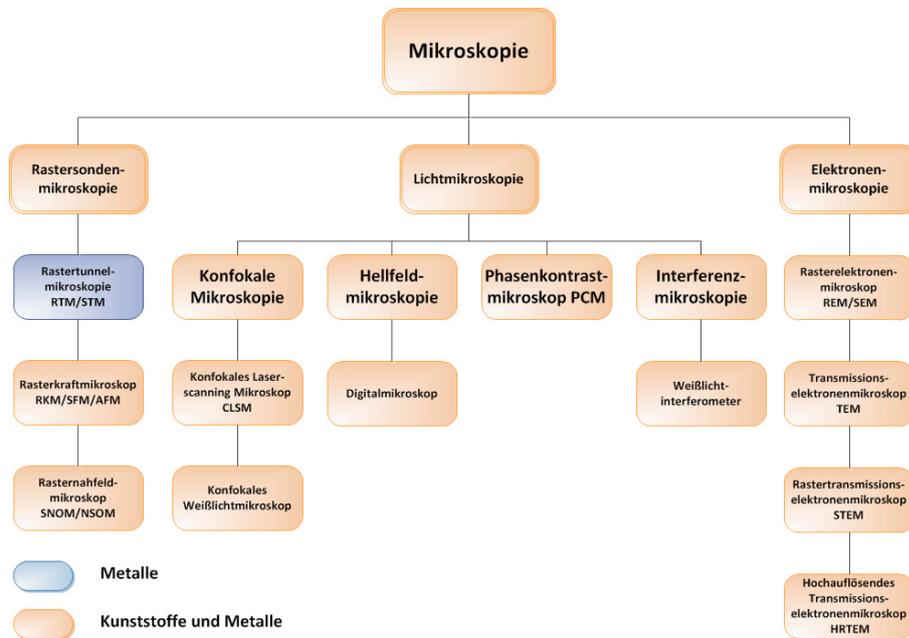


Abbildung 1: Zusammenfassung der Ergebnisse zur Marktanalyse - Mikroskopie

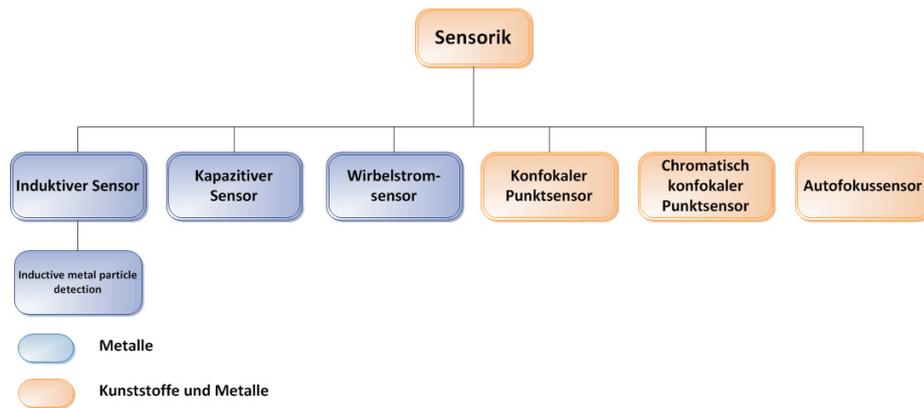


Abbildung 2: Zusammenfassung der Ergebnisse zur Marktanalyse - Sensorik

Eine detaillierte Bewertung der ausgewählten Verfahren anhand der Systematik und Werkzeuge eines Produktentwicklungsprozesses reduzierten die potentiellen Lösungsmöglichkeiten auf eine optische, taktile oder gravimetrische Verschleißquantifizierung [VDI97], [Con10]. Zu den optischen Messinstrumenten zählten hierbei optisch, konfokale Multisensorsysteme mit konfokalen oder chromatisch-konfokalen Punktsensoren und/oder optisch, konfokale Mikroskope mit entsprechender Objektivausstattung. Als taktiles Verfahren wurde ein hochauflösendes Profilometer mit unterschiedlichen Tastarmperepherien herangezogen. In Zusammenarbeit mit verschiedenen Herstellern und Fachbereichen für optische und taktile Messsysteme

wurden praxisbezogene Messungen durchgeführt und Testberichte zusammengestellt. Die Probemessungen resultierten in einer konkreten, anwendungsorientierten Ermittlung der direkten Verschleißmessgrößen. Dies beinhaltet sowohl den Einsatz der in der Praxis zu analysierenden Probekörper als auch eine Bewertung von Störgrößen auf das Messergebnis einer optischen und taktilen Verschleißquantifizierung. Zu den Störgrößen zählen hierbei Einflussfaktoren aus der direkten Umgebung, wie beispielweise Vibrationen sowie geometrische und werkstoffspezifische Störgrößen der Probekörper, wie Formabweichungen (Verzug, Quellung) oder Inhomogenitäten. Die gravimetrische Verschleißquantifizierung wurde durch eine Semimikrowaage umgesetzt und unter Berücksichtigung von Auslösung, Reproduzierbarkeit, Handhabung und Störgrößen detailliert analysiert und ausgearbeitet.

### 3 GRAVIMETRISCHE VERSCHLEIßQUANTIFIZIERUNG

Im nachfolgenden wird die Methodik der gravimetrischen Verschleißquantifizierung an Polymer-Polymer-Gleitpaarungen durch eine Semimikrowaage mit integriertem Ionisator in Auszügen dargestellt und erläutert. Die im Gesamten untersuchten Polymere wurden aus zuvor durchgeführten Polymerwerkstoffrecherchen in fördertechnischen Anwendungen (Förderketten, Gleitschienen und Kettenrädern) gezielt ausgewählt. Zur Validierung des gravimetrischen Verfahrens wurden umfangreiche Testwägungen an den polymeren Probekörpern sowie zahlreiche Literaturrecherchen zu potentielle Einflussfaktoren auf das Messergebnis durchgeführt. Diese unterteilten sich in Einflussfaktoren aus der direkten Umgebung des Messinstrumentes und aus den werkstoffspezifischen Eigenschaften der Probekörper. Die Auftriebseinflüsse der Umgebungsluft und deren Schwankungen durch Änderungen der Umgebungsbedingungen (relative Feuchtigkeit und Temperatur der Umgebungsluft) oder externe Schwingungen wurden den Einflussfaktoren aus der direkten Umgebung zugeordnet [Mett08]. Die werkstoffspezifischen Einflussfaktoren der Probekörper beinhalten eine positive bzw. negative elektrostatische Aufladung, Materialübertragungen, Feuchtigkeitsaufnahme (Hygroskopie) bzw. Feuchtigkeitsabgabe und die Diffusion von Additiven [Sart13]. Die Ermittlung reproduzierbarer und valider Verschleißmessergebnisse bei Polymerwerkstoffen erforderte die Generierung von Korrekturfaktoren für die jeweiligen Einflussfaktoren. In der nachfolgenden Abbildung 3 sind die Einflussfaktoren/Störgrößen der gravimetrischen Verschleißquantifizierung von Polymer-Polymer-Paarungen zusammenfassend dargestellt.



Abbildung 3: Einflussfaktoren/Störgrößen bei der gravimetrischen Verschleißquantifizierung

Die Korrekturfaktoren  $\Delta m_{\text{rel. Luftfeuchtigkeit}}$  und  $\Delta m_{\text{Lufttemperatur}}$ , welche die Auftriebskräfte bei der Wägung durch die Umgebungsbedingungen berücksichtigen, wurden anhand eines Berechnungsschemas für die Umgebungsbedingungen am Standort der Waage quantifiziert, beurteilt und zu einem Korrekturfaktor  $\Delta m_{\text{Auftrieb}}$  zusammengefasst. Die Korrekturfaktoren der elektrostatischen Aufladung  $\Delta m_{\text{E+}}$  bzw.  $\Delta m_{\text{E-}}$  wurden aus der waagenspezifischen Fachliteratur übernommen und durch eigene Messungen eingehend untersucht [Sart02]. Die Neutralisierung einer elektrostatischen Aufladung der Probekörper mithilfe des in der Waage integrierten Ionisators wurde an verschiedenen POM-C und POM-H sowie einer PE-UHMW Type untersucht. In Kooperation mit dem Waagenhersteller wurden hierbei die benötigten Einstellparameter Intensität und Einwirkdauer einer Ionisatorbetätigung festgelegt. Auf Basis dieser Parameter wurden im Weiteren die benötigten Betätigungsintervalle des Ionisators durch explizite Messreihen an den verschiedenen Polymerwerkstoffen ermittelt. Die Standardabweichung der Messreihen konnte durch den gezielten Einsatz des Ionisators bei jedem Polymerwerkstoff um bis zu 200% (bezogen auf die Standardabweichung ohne Ionisator) reduziert werden. Die Korrekturfaktoren  $\Delta m_{\text{H+}}$  bzw.  $\Delta m_{\text{H-}}$  einer Feuchtigkeitsaufnahme bzw. Feuchtigkeitsabgabe sowie der Korrekturfaktor für Diffusionsvorgänge von Additiven  $\Delta m_{\text{Diff}}$  wurden anhand von mehrwöchigen Wägungen der Probekörper für die jeweilige Polymertypen quantifiziert und werden während der tribologischen Prüfung durch Referenzprobekörper erfasst. Durch die Wägeregebnisse konnte der Effekt einer Feuchtigkeitsaufnahme (Hygroskopie) sowie Feuchtigkeitsabgabe und Diffusion der Probekörper im Austausch mit der Umgebungsluft nachgewiesen werden. Zudem hat sich herausgestellt, dass sich eine Trocknung der Probekörper vor dem tribologischen Versuch kontraproduktiv auf die Ermittlung aussagekräftiger und reproduzierbarer Wägeregebnisse auswirkt. Des Weiteren wurde bewiesen, dass die Generierung von einheitlichen Korrekturfaktoren  $\Delta m_{\text{H+}}$  bzw.  $\Delta m_{\text{H-}}$  und  $\Delta m_{\text{Diff}}$  nicht valide ist und diese demnach für jede Polymertypen und Prüfkörpergeometrie explizit ermittelt werden müssen. Im Rahmen der Testmessungen stellte sich heraus, dass die Diffusion von Additiven die Standardabweichungen der Messreihen der Vorher-Wägungen erheblich beeinflusst. Durch eine sondierte Lagerung der einzelnen Probekörper konnten dieser Effekt reduziert und gleichzeitig die Standardabweichungen verbessert werden.

Die Validierung des gravimetrischen Lösungskonzeptes erforderte die Erstellung eines detaillierten, chronologischen Versuchsablaufplans zur gravimetrischen Verschleißquantifizierung. Dieser beinhaltet die Versuchsvorbereitung, den Versuchsablauf sowie die einzelnen Arbeitsschritte der Vorher- und Nachher-Wägungen. Zusätzlich wurden potentielle Fehlerquellen ermittelt, Abhilfemaßnahmen und Hilfsmittel getestet und ggf. umgesetzt. Der Korrekturfaktor  $\Delta m_{\text{Auftrieb}}$  konnte aufgrund konstanter Umgebungsbedingungen vernachlässigt werden. Dieser kann aber bei entsprechenden Gegebenheiten mit berücksichtigt werden. Die Neutralisation der elektrostatischen Aufladung wird, innerhalb des gravimetrischen Messablaufes, durch den Einsatz des Ionisators mit den zuvor ermittelten Parametern für Intensität, Wirkdauer und Einschaltintervall sichergestellt. Zur Ermittlung der Korrekturfaktoren  $\Delta m_{\text{H+}}$  bzw.  $\Delta m_{\text{H-}}$  und  $\Delta m_{\text{Diff}}$  werden, parallel zur Verschleißphase der tribologischen Probekörper, Referenzprobekörper herangezogen. Diese bilden den Raumprozess durch Feuchteinfluss und Diffusionsvorgänge ab, wobei sich die Feuchtigkeitseinflüsse und Diffusionsvorgänge stets überlagern und nicht getrennt erfasst werden können. Es hat sich innerhalb der Untersuchungen herausgestellt, dass die Haupteinflussfaktoren der gravimetrischen Verschleißquantifizierung die Hygroskopie, Feuchtigkeitsabgabe und überlagerte Diffusionsvorgänge der Additive darstellen. Folglich sind diese zwingend durch die Referenzprobekörper und Probenkörperkonditionierungen zu erfassen und in Form der vorgestellten Korrekturfaktoren zu berücksichtigen.

In Abbildung 4 ist der Einfluss eines Massenverlustes infolge von Feuchtigkeitsabgabe oder Diffusion auf die Ermittlung des massenmäßigen Verschleißbetrages  $W_m$  sowie die Berechnung des korrigierten massenmäßigen Verschleißbetrages  $W_{m/K}$  allgemein dargestellt. Analog ist hierzu in Abbildung 5 der Einfluss einer Massenzunahme infolge von Feuchtigkeitsaufnahme auf die Ermittlung des massenmäßigen Verschleißbetrages  $W_m$  sowie die Berechnung des korrigierten massenmäßigen Verschleißbetrages  $W_{m/K}$  allgemein dargestellt.

#### Einfluss einer Feuchtigkeitsabgabe oder Diffusion auf den massenmäßigen Verschleißbetrag

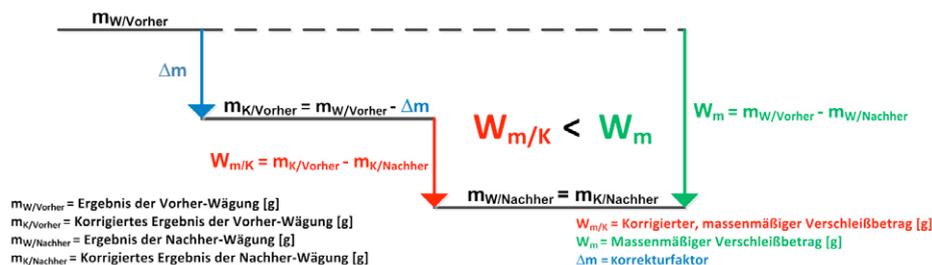


Abbildung 4: Einfluss und Korrektur der Feuchtigkeitsabgabe oder Diffusion

#### Einfluss einer Feuchtigkeitsaufnahme auf den massenmäßigen Verschleißbetrag

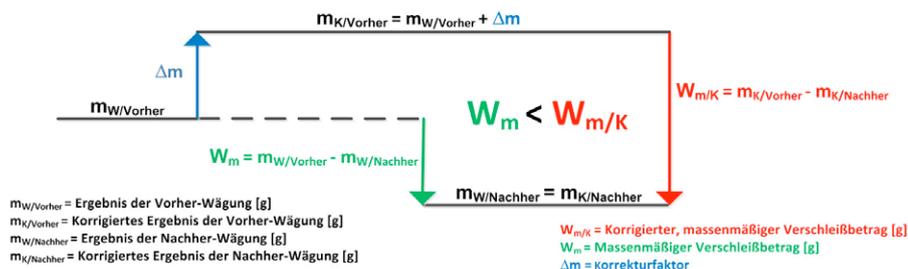


Abbildung 5: Einfluss und Korrektur der Hygroskopie (Feuchtigkeitsaufnahme)

Zur weiteren Validierung des Messkonzeptes wurden umfangreiche Versuchsreihen mit verschiedenen Grundpolymertypen (POM-C, POM-H, PP, PE-UHMW) und mit Modifizierungen (Inkorporierte Schmierstoffe oder Fasern) in verschiedenen Gleitpaarungen durchgeführt. Die Versuchsdurchführung orientierte sich konkret an dem zuvor generierten Versuchsplan. In Abbildung 6, Abbildung 7 und Abbildung 8 sind die Ergebnisse (Indirekte Verschleißkenngrößen) der gravimetrischen Verschleißquantifizierung durch ein exemplarisches Auswertungsbeispiel dargestellt. Bei diesen Abbildungen wurde, aufgrund der beispielhaften Darstellung, auf eine Angabe der konkreten Versuchsparameter verzichtet. In den Abbildungen sind u.a. die Verschleißkenngrößen der vier Messbahnen (Laufbahnen), welche am tribologischen Prüfstand immer parallel getestet werden, ersichtlich. Diese abgeglichenen Messwerte dienen anschließend der statistischen Auswertung der generierten Ergebnisse. Als Kriterien zur Validitätsprüfung des ausgearbeiteten Verfahrens wurde u.a. die Vergleichbarkeit der vier Messbahnen sowie ein Abgleich durch in Reihe geschaltete, gleichartige Versuche herangezogen.

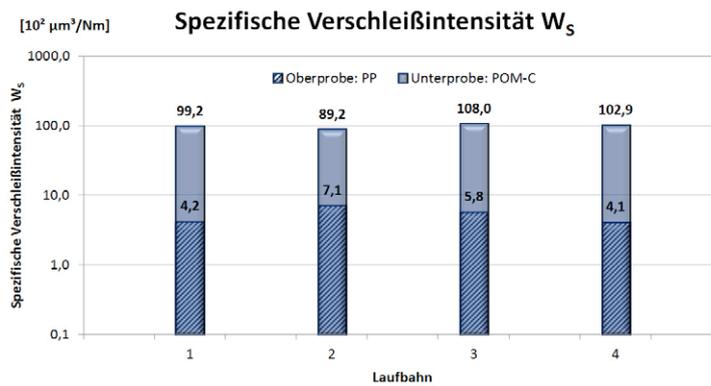


Abbildung 6: Auswertungsbeispiel - Spezifische Verschleißintensität  $W_s$

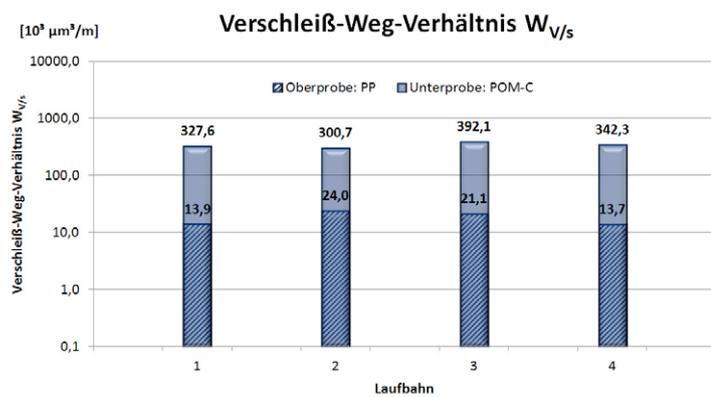


Abbildung 7: Auswertungsbeispiel - Verschleiß-Weg-Verhältnis  $W_{V/s}$

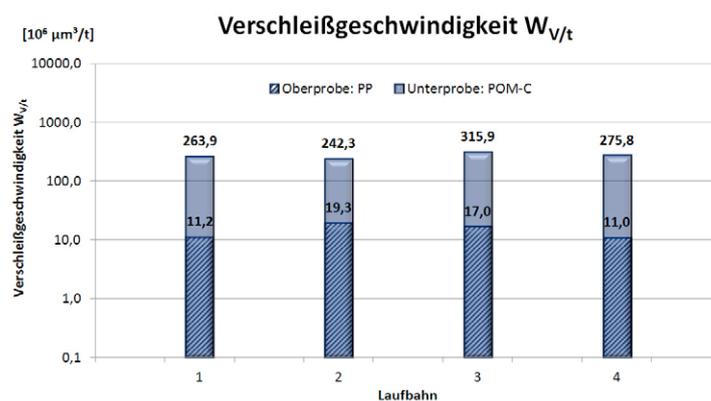


Abbildung 8: Auswertungsbeispiel - Verschleißgeschwindigkeit  $W_{V/t}$

Die Versuchsreihen zur Validierung des Messkonzeptes lieferten für die getesteten Polymere reproduzierbare, valide und aussagekräftige Ergebnisse. Bei Werkstoffen mit einer stärker ausgeprägten Hygroskopie oder Feuchtigkeitsabgabe, wie beispielsweise bei Polyamiden (PA), ist

diese Messmethode nur bedingt valide anwendbar und bedarf einer detaillierteren Betrachtung in Form von mehreren Referenzproben zur ausführlichen Betrachtung und Ermittlung der Korrekturfaktoren. Aus diesem Grund ist dieses Verfahren für derartige Kunststoffe weniger praktikabel und die Korrekturfaktoren bzw. Verschleißkennwerte weisen deutlich größere Messunsicherheiten auf. Vorzugsweise sollte das Verfahren hier nur bei höheren Verschleißintensitäten herangezogen werden, wobei sichergestellt werden sollte, dass der Verschleißbetrag größer ist als die Feuchtigkeitsaufnahme/-abgabe bzw. diese valide durch die Referenzproben erfasst und anschließend durch die Korrekturfaktoren berücksichtigt werden. In jedem Fall können aber sogenannte „Worst-Case-Statements“ zum Verschleißverhalten getroffen werden. Die vorzugsweise anzuwendende Verschleißquantifizierungsmethode (gravimetrisch, taktil oder optisch) sollte anhand des spezifischen Eigenschaftsprofils des zu analysierenden Polymers und dem zu erwartenden Verschleißbetrag ausgewählt werden. Bei sehr geringen Verschleißbeträgen und/oder hoher Feuchtigkeitsaufnahme/-abgabe sollten eine taktile oder optische Verschleißquantifizierung Anwendung finden. Hier liegt der wesentliche Vorteil einer optischen oder taktilen Verschleißquantifizierung durch eine hohe und variable Volumenauflösung. Zur direkten Gegenüberstellung der umsetzbaren räumlichen Auflösungen wurde das Kriterium der Verschleißvolumenauflösung eingeführt. Dieses stellt den kleinsten, messbaren Volumenunterschied der Messinstrumente dar. Die Verschleißvolumenauflösung eines optischen Verfahrens mittels eines chromatisch-konfokalen Punktsensors ist im Vergleich zur Semimikrowaage um einen Auflösungsfaktor von ca. 6.000.000 größer. Ein weiterer zentraler Vorteil eines optischen und taktilen Verfahrens ist die Unabhängigkeit der Verschleißvolumenauflösung von der Dichte des Werkstoffes, da bei diesen Verfahren der volumetrische Verschleißbetrag  $W_v$  direkt ermittelt wird. Durch das gravimetrische Verfahren können stets nur der Gesamtverschleiß der Probekörper erfasst werden, wobei durch optisch oder taktile Methoden auch Analysen von einzelnen Teilbereichen und ggf. Verschleißpartikeln durchführbar sind. Allerdings stellt dies gleichermaßen einen Vorteil des gravimetrischen Verfahrens bei Probekörper mit einem Eingriffsverhältnis von  $\epsilon = 1$  (kontinuierlicher Eingriff) dar, da hier keine vom Verschleiß unbehafte Referenzflächen, zur Verschleißvolumenermittlung mittels optischem oder taktilem Messverfahren, vorhanden sind. Demnach erfordert dies einen Vorher-Nachher-Scan derartiger Prüfkörper, was wiederum mit zusätzlichen Störgrößen verbunden ist. Dennoch ist ein wesentlicher Nachteil des gravimetrischen Verfahrens gegenüber einer taktilen oder optischen Verschleißquantifizierung der mannigfaltige Einfluss der zahlreichen Störgrößen sowie der erhöhte Bedienerinfluss (Handling) auf das Messergebnis. Die hohen Anschaffungskosten eines optischen oder taktilen Messinstrumentes stellen den wesentlichen Nachteil dieser Messmittel gegenüber dem gravimetrischen Verfahren dar.

#### 4 ZUSAMMENFASSUNG

Es wurde eine Methode der gravimetrischen Verschleißquantifizierung zur Ermittlung des Verschleißes an Polymeren in Kunststoff-Kunststoff-Gleitpaarungen erarbeitet, umgesetzt und validiert. Dies beinhaltete eine vorangestellte Ermittlung potentiell geeigneter Messverfahren und Methoden zur Verschleißquantifizierung, als auch eine methodische Bewertung der einzelnen Varianten. Innerhalb der Ausarbeitung und Umsetzung der gravimetrischen Verschleißquantifizierung wurden Korrekturfaktoren für die zuvor ermittelten Störeinflüsse generiert und in einen Ablaufplan zur Versuchsdurchführung und das abschließende Auswertungsschema integriert. Das neue Verschleißquantifizierungsverfahren bildet die Grundlage zur Analyse und Entwicklung verschleißoptimierter Polymerwerkstoffe, zur Ermittlung von Verschleißkennlinien sowie zur Untersuchung der Korrelationen zwischen Verschleiß und verschiedenen Materialeigenschaften oder anderen Ausführungen von Tribometern und

Bauteilprüfständen. Abschließend können systemspezifische Verschleißkenngrößen, durch eine kombinierte Betrachtung der Verschleiß-Messgrößen und Verschleißerscheinungsformen, definiert und beschrieben werden. Verschleiß-Kennlinien aus Modelversuchen und Verschleiß-Versagens-Diagramme stellen einen ersten Ansatz zur Festlegung von verschleißbedingten Wartungsintervallen und Lebensdauerprognosen von fördertechnischen Bauteilen dar. Im Weiteren werden verschiedene taktile und optische Verfahren mit höherer Verschleißvolumenauflösung detaillierter ausgearbeitet und validiert. Außerdem werden unterschiedliche Konzepte zur Ermittlung von Verschleißkennlinien an Modellprüfkörpern auf Basis der vorgestellten Verschleißquantifizierungsverfahren untersucht und umgesetzt.

## REFERENZEN

- [Sum11] Sumpf J.; Schuhmann A.; Weise S.; Nendel K.; Eichhorn S.: Neues Prüfverfahren zur Reibungs- und Verschleißbewertung von Kunststoff-Gleitpaarungen: Institut für Fördertechnik und Kunststoffe (ifk) - TU Chemnitz, Tribologie + Schmierung, 58. Jahrgang, 2011.
- [Gre13] Grebe M.: Fachseminar - Tribometrie: Technische Akademie Esslingen (TAE), Ostfildern - Nellingen, 2013.
- [GfT11] GfT: 52. Tribologie-Fachtagung - Reibung, Schmierung und Verschleiß - Forschung und praktische Anwendungen: Gesellschaft für Tribologie e.V. (GfT), Göttingen, 2011.
- [Fein13] Feinle P.: Fachseminar - Tribologie der Kunststoffe: Technische Akademie Esslingen (TAE), Ostfildern - Nellingen, 2013.
- [Czi10] Czichos Horst: Tribologie-Handbuch - Tribometrie, Tribomaterialien, Tribotechnik: Vieweg + Teubner Verlag, Wiesbaden, 3. Auflage, 2010.
- [DFS02] Deters, Ludger; Fischer, Alfons; Santer, Erich; Stolz Ulrich: Gesellschaft für Tribologie: Arbeitsblatt 7, 2002.
- [VDI97] Verein Deutscher Ingenieure: VDI2222, Konstruktionsmethodik- Methodisches Entwickeln von Lösungsprinzipien, Beuth Verlag, Berlin, Ausgabe 06/1997.
- [Con10] Conrad, Klaus-Jörg: Grundlagen der Konstruktionslehre – Methoden und Beispiele für den Maschinenbau, Hanser Verlag München, Auflage 5, 2010.
- [Mett08] Mettler-Toledo AG (Hrsg.): Wägefibel, Schweiz, 2008.
- [Sart13] Sartorius Weighing Technology GmbH (Hrsg.): Correct Use and Handling of Analytical and Microbalances, Publication No.: W-0135-a11064, Order No.: 98648-014-11, o.O., 2013. [Sart02] Sartorius Weighing Technology GmbH (Hrsg.): WägeRaum Nr.4, Göttingen, Ausgabe 4, 2002.

## Correlation between Airlines' use of foreign languages and their country of origin's acceptance of foreigners

Jerome Dumetz, Rainer Wehner  
Plekhanov Russian Economic University, Moscow  
Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt  
Moskau, Würzburg

The Airline sector is probably one of the most international industries nowadays. Thus, one should expect that airline companies are also international orientated and show their open-mindedness through all the services they offer. Using the renowned cross-cultural Bennett scale to rate the degree of ethnorelativity, this empirical study aims at analysing the link between the use of foreign language in inflight magazines and the attitude towards foreigners in countries the airlines originates from. The authors discovered that there exists a strong correlation between the use of foreign language in inflight magazines offered to the passengers and the degree welcoming foreigners in the country of origine of the airline.

Categories and Subject Descriptors: Social and Human Sciences

Additional Key Words and Phrases: Intercultural Differences, Ethnorelativity, Cross-cultural Communication, Airline Industry

### 1 INTRODUCTION

The simingly ever growing number of pasengers travelling around the globe by plane, the increasing number of airlines or simply the multiplication of online travelers' blogs should question anyone as to whether the world is becoming more inclusive in its diversity, or at least, more tolerant towards foreigners. In parallel, news reports display troubling, sometimes shocking features of intolerance in general, witnessing an exclusive nationalism in many countries or regions around the world. In fact, a travellers experience differs greatly depending where you land.

The authors, both working in the international field of their Universities and frequent travelers around the world with different airlines, realized that some services international airlines offer cannot be seen as a handsome gesture towards foreigners. For example: If you travel by plane, it is always nice to gain some information about the country of your destination in the in-flight magazine. If the in-flight magazine is only printed in the language of the country the airline is origine and you cannot read this language, the in-flight magazine is useless. The fact

Author's address: Dumetz, Jerome, Plekhanov Russian Economic University, Moscow, Russia, [www.rea.ru](http://www.rea.ru);  
Wehner, Rainer, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Würzburg, Germany, [www.fhws.de](http://www.fhws.de)

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than FHWS must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission.

that most of the airlines in the world are state owned can lead to the conclusion that some countries are more than others interested in been seen as friendly towards foreigners or different cultures.

Having this in mind, the authors collected data about the percentage of English or non-native language featured in in-flight magazines of international airlines. They connected the figures to the Bennett scale, also called the DMIS (for Developmental Model of Intercultural Sensitivity) and then compared the results with the Travel & Tourism Competitiveness Report presented at the World Economic Forum 2013 (WEF, 2013) in Davos.

The aim of this paper is to verify if the ethnorelativity concept can be verified in the airline industry, using empirical research correlated to academic findings.

The authors have chosen the Bennett scale because in comparison to other authors like Pierre Bourdieu or Mitchel R. Hammer, who invented the IDI (Intercultural Development Inventory), Milton Bennetts concept is more detailed dealing with intercultural sensitivity (BENNETT, 1986) than Bourdieu's or Hammer's, whose models are simpler or mostly closer connected to single persons (HAMMER, BENNETT, WISEMAN 2003).

A great number of models exist in the field of cross-cultural communication. Academic citations show that Hofstede's is the most cited author. Trompenaars' model and the GLOBE project are also frequently quoted. Among practitioners, no survey exist about the professionals' preferences, however, it is likely that the same models are used, besides proprietary ones that require royalties payments (DUMETZ, 2014).

All those models, however, are designed and used to compare cultures among themselves, using a series of behavioral vectors also called cultural dimensions. Such models allow to creating cultural profiles of national cultures (or individuals and even companies) that resemble personality types. While very useful in other cases, only the Bennett scale focuses exclusively with the degree of acceptance to other cultures, a concept called ethnorelativity.

In this article, the authors did not seek to compare airlines using cultural dimensions but to assess their level of ethnorelativity, and then to correlate the results with another existing survey that analyse the level of acceptance of foreigners of countries, a concept close to ethnorelativity.

## 2 MILTON BENNETT'S CONCEPT OF ETHNORELATIVITY

Milton Bennett (BENNETT, 1998) developed a cross-cultural model that describes different ways how people react to cultural differences. The model is organized in six stages. The first three stages are describing an ethnocentric behavior starting at a level of denying other cultures and seeing ones own culture as the only existing one. Moving towards up the scale to a more and more ethnorelative attitude, the individual is recognising its own culture in the context of other cultures and tends to adapt to them.

1. Denial: The individual does not recognize the existence of cultural differences
2. Defense: The individual does recognizes some differences exist, but they are often seen as negative

3. Minimization: The individual is unaware of the projection of its own cultural values; those values are mostly seen as superior
4. Acceptance: The individual shifts perspectives in order to understand that a „ordinary“ behavior can have different meanings in different cultures
5. Adaptation: The individual can evaluate the other’s behaviour from their frame of reference and can adapt its behaviour to fit the norms of the other culture
6. Integration: The individual easily shifts its frame of reference to fully adapt to the other culture.



Figure 1: The Milton Bennett scale

The Bennett scale is a convenient model to rate the degree of ethnorelativity from low (the individual display ethnocentric behaviours) to high (the individual display full integration to another culture) (BENNETT, 2011).

The DMIS is a combination of running intercultural workshops and doing consciousness studies. This was an early attempt to define intercultural competence. He'd define the end-state as integration – the ability to shift from one state to another – whether it be bicultural or multicultural, what he now calls an expanded repertoire of worldview.

Some of his early studies at Portland were around cults: a little model of an ethnocentric culture. The cult-people think their view of reality is completely and uniquely central to reality. It's this experience of the own culture as being central and it moves through this forming-feeling process that he names stages. The DMIS is based on subjective organization.

### 3 THE APPLICATION TO THE AIRLINE INDUSTRY

The airline industry is a natural choice to conduct this analysis due to its role and implication in the effects of globalization. Indeed, the airline industry participated in the development of globalization helping people travel faster and at a lower cost than before. In parallel, the industry clearly benefited of the globalization, as showed by the significant increase of airline companies around the world, now several thousands, and the ever growing number of passengers over the years reaching 2,8 billion passengers carried by airlines in 2011 with a forecasted 5,3% growth per annum between 2012 and 2016 according to IATA (IATA press 2013).

The great number of airlines in the world in general and in Europe in particular made possible the collect of data in sufficient number (30 airlines surveyed) to make adequate analysis.

The authors have gathered a number of in-flight magazines presented by airlines to their passengers (Table 1). All the magazines were collected or are intended for international flights,

ensuring the intercultural aspect of the flying experience. Because English has become the lingua franca of the travel industry, selected airlines were exclusively from non-English speaking countries, hence allowing the authors to focus on the linguistic aspect of the ethno-relativity.

	Magazine/ airline	Date	Country of origin	Language(s) used	Part of foreign language	Level of ethnorelativity/ remarks
1	Rossiya	Feb. 2014	Russia	Russian and English	3 pages/173	Low
2	Czech Airlines	3-4/2014	Czech Republic	Czech, English and Russian	1/3, 1/3, 1/3	Middle/High
3	Transaero	May 2013	Russia	Russian	None	Low
4	Aeroflot	July 2013	Russia	Russian and English	20 pages /300	Low
5	Air France	May 2013	France	French and English	50/50	Middle
6	Lufthansa	June 2013	Germany	German and English	50/50	Middle
7	Alitalia	June 2013	Italy	Italian and English	50/50	Middle
8	Aeroflot Premium	July 2013	Russia	Russian and English	80/250	Low
9	Turkish Airlines	Aug. 2013	Turkey	Turkish and English	50/50	Middle
10	SAS	July 2013	Denmark/Norway/Sweden	English	100%	High
11	KLM	May 2014	Netherlands	English	100%	High
12	S7	March 2013	Russia	Russian and English	2 pages/80	Low
13	Bulgaria	Aug. 2013	Bulgaria	Bulgarian and English	50/50	Middle
14	Emirates	Sept. 2013	UAE	English	100%	High
15	Aegean	April 2014	Greece	Greek and English	50/50	Middle
16	Tarom	April 2014	Romania	Romanian and English	50/50	Middle
17	AeroSvit	May 2014	Ukraine	Ukrainian and English	Roughly 30%	Low to Middle
18	Finnair	May 2014	Finland	English	100%	High
19	Malev	July 2011	Hungary	Hungarian and English	50/50	Middle
20	Brussels	May 2014	Belgium	English, French and Dutch	80/10/10	High
21	Austrian	Feb. 2014	Austrian	German and English	50/50	Middle
22	Aerpostal	April 2014	Venezuela	Spanish	None	Low
23	Moldova	June 2013	Moldova	Moldovan, Russian and English	1/3, 1/3, 1/3	Middle/high
24	Air Algérie	Feb. 2013	Algeria	French	None	Low
25	IranAir	May 2014	Iran	Persian	None	Low
26	China-Southern	May 2014	China	Chinese and English	50/50	Middle
27	ChinaEastern	May 2014	China	Chinese and English	50/50	Middle
28	Air China	May 2014	China	Chinese and English	50/50	Middle
29	Hainan Airlines	May 2014	China	Chinese	None	Low
30	TAM	2014	Brazil	Portugese/English and Portugese/ Spanish	50/50 and 50/50	Middle

Table 1: Surveyed airlines and data

### Hypothesis

Using this primary data, two hypotheses may be formulated to verify the correlation between Airlines' ethno-relativity and their country of origin's acceptance of foreigners

H1: Airlines wishing to show their open-mindedness to the world, or, at least a tolerance towards non-native passengers, should be providing information in a foreign language.

H2: Airlines originating from ethnocentric cultures should not provide magazines displaying foreign language text.

#### 4 COMPARING THE RESULTS WITH CROSS-CULTURAL RESEARCH

The fact that most of the large airlines are state owned leads to the assumption that there might be a correlation between the ethnorelativity of the country and the cultural artefacts, like the in-flight magazines, offered to their passengers.

Unfortunately there is no up-to-date source of an ethnorelativity score of each country. Therefore the authors decided to use the Travel & Tourism Competitiveness Report 2013, which was presented on the World Economic Forum (WEF, 2013) in Davos 2013.

The ranking on pillar 12, figure 12.02 (WEF, 2013, 455) about the attitude of the population toward foreign visitors moves into the same direction and shows, how foreign visitors are welcome in the surveyed countries.

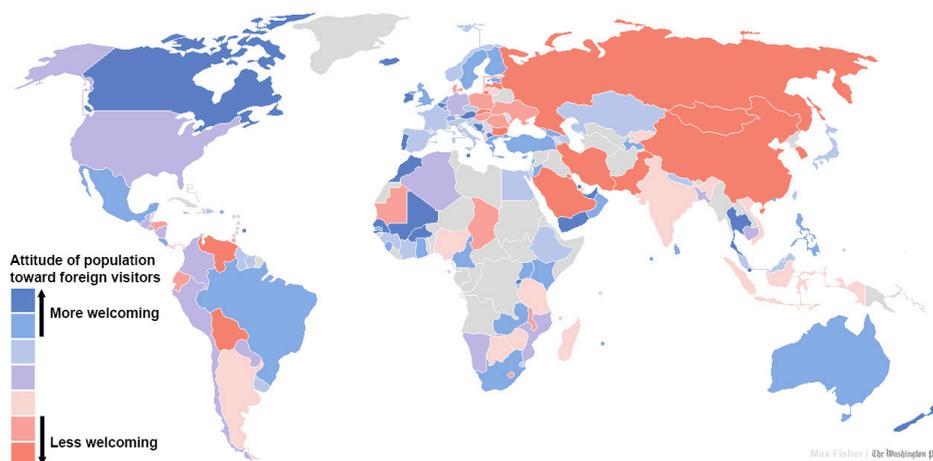


Figure 2: Map by Max Fisher, from *The Washington Post*<sup>1</sup>

Max Fisher, from *The Washington Post*, applied the WEF data, gathered between late 2011 to late 2012, from 140 countries into a map showed in Fig.2 (FISCHER 2013).

According to the data the red countries are less welcoming towards foreign visitors, and, the blue countries are more welcoming towards them.

The top three most welcoming countries for foreigners are, in order: Iceland, New Zealand and Morocco. Other high-ranking countries are Ireland, Canada, Austria and a few tourist havens like Thailand or United Arab Emirates. Those data seem to correlate in a way with the gathered data of the airline magazines.

The three countries least welcoming to foreigners are, in order: Bolivia, Venezuela and Russia. Other poorly ranked countries are located in the greater Middle East like Iran, Pakistan, Saudi Arabia but also China and South Korea.

This low ranking might have different reasons. The troubles in the Middle East states might had an influence on the votes. Another reason for the low ranking could be an existing or growing nationalism in the mentioned countries that could leads towards behavior of exclusion

(as opposed to inclusion) of the “others”. If we look at Russia for example, the country always displayed strong patriotic feelings, but in the last few years the number of popular events of nationalist inspiration has become more public. Topping the table, Iceland is known to be a liberal country, with various protections against discriminations of all kind.

## 5 FINDINGS

In order to calculate a possible correlation, the authors have set an ethnorelativity score from 0 (low ethnorelativity) to 1 (high ethnorelativity). The WEF ranking was transformed as a percentage ( $x/14$ ). The ranking was then inverted in order to have nations open to foreigners with a high score and nations not open to foreigners a low score.

	Country of origin (Airline)	Ethnorelativity	Ethnorating	WEF ranking /140	WEF /100	Neg WEF /100
1	Russia	Low	0	138	98,57	1,43
2	Czech Republic	Middle/High	0,8	66	47,14	52,86
3	Russia	Low	0	138	98,57	1,43
4	Russia	Low	0	138	98,57	1,43
5	France	Middle	0,5	80	57,14	42,86
6	Germany	Middle	0,5	83	59,29	40,71
7	Italy	Middle	0,5	79	56,43	43,57
8	Russia	Low	0	138	98,57	1,43
9	Turkey	Middle	0,5	53	37,86	62,14
10	Denmark/Norway/Sweden	High	1	66	47,14	52,86
11	Netherlands	High	1	52	37,14	62,86
12	Russia	Low	0	138	98,57	1,43
13	Bulgaria	Middle	0,5	132	94,29	5,71
14	UAE	High	1	15	10,71	89,29
15	Greece	Middle	0,5	37	26,43	73,57
16	Romania	Middle	0,5	122	87,14	12,86
17	Ukraine	Low to Middle	0,3	127	90,71	9,29
18	Finland	High	1	31	22,14	77,86
19	Hungary	Middle	0,5	116	82,86	17,14
20	Belgium	Middle to High	0,8	19	13,57	86,43
21	Austria	Middle	0,5	5	3,57	96,43
22	Venezuela	Low	0	139	99,29	0,71
23	Moldova	Middle to High	0,8	93	66,43	33,57
24	Algeria	Low	0	94	67,14	32,86
25	Iran	Low	0	135	96,43	3,57
26	China	Middle	0	130	92,86	7,14
27	China	Middle	0,5	130	92,86	7,14
28	China	Middle	0,5	130	92,86	7,14
29	China	Low	0	130	92,86	7,14
30	Brazil	Middle	0,5	43	30,71	69,29

Table 2: Correlation between ethnorating and WEF ranking

The correlation coefficient is 0,73 between the use of language in inflight magazines and the WEF ranking surveying the attitude of nationals towards foreigners.

$$\text{Correlation Coefficient (X, Y)} = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}}$$

The relatively high correlation between foreigners feeling unfriendly treated and use of foreign languages in national airline inflight magazines is remarkable. Reasons for that phenomenon could be various, ranging from politics, ethics or the society's open-mindedness in general.

While outside the scope of this article, it is interesting to notice the apparent link towards the ethnocentricity of the airlines's nations with local treatment of homosexuality. The ILGA Report on State-Sponsored Homophobia (ILGA, 2013) analyses the legal answers to homosexuality (positive or negative) that shows the level of acceptance of gays in a given society.

While Iceland, New Zealand and Morocco top the WTE list, Ireland, Canada, Austria can also be counted among the most welcoming countries for foreigners. All of them, except Morocco (a touristic destination where the authorities turn a blind eye on homosexuality) score high on the gay-right report.

At the bottom of the WTE list, Bolivia, Venezuela and Russia; but also Iran, Pakistan, Saudi Arabia or China are least welcoming towards foreigners. Russia isn't too welcoming or open towards gays. Bolivia, Venezuela are not very open to others, but neutral towards gays. Iran, Pakistan, Saudi Arabia are all three not open towards foreigners and not tolerant towards gays.

## 6 CONCLUSION AND FURTHER DEVELOPEMENTS

Both hypothesis are validated:

H1: Airlines wishing to show their open-mindedness to the world, or, at least a tolerance towards non-native passengers, tend to provide an inflight magazine in a foreign language.

H2: Airlines originating from ethnocentric cultures tend not to provide an inflight magazine in a foreign language.

As a result, the ethnocentricity theory, based on cultural artefacts such as airline magazines, delivers significant insight to analyse the degree of welcoming of foreigners in a country. This simple example shows the validity of the practical application of cross-cultural management models and theory.

Therefore, while the focus of this article is the correlation between the airline's magazine use of languages and how much the airline's country of origin is open to foreigners, it appears the ethnocentricity ranking created for this research (table 2) could be also used to lay the foundation for a value-based cultural dimension ranking of countries. The behavioural value would be intolerance towards foreigners.

## REFERENCES

BENNETT, MILTON J. (1998) Intercultural communication: A current perspective. Basic concepts of intercultural communication: selected readings. Yarmouth, ME: Intercultural Press, 3-17

BENNETT, MILTON J. (2011) SIETAR Europa Journal, "Milton Bennett talks about the DMIS" - March, 2011

BENNETT, MILTON J (1986) A developmental approach to training for intercultural sensitivity, in: International Journal of Intercultural Relations, Volume 10, Issue 2

DUMETZ, J., VISHNIAKOVA A., SOSNAKOVA A. (2014), "The European Influence on Cross-Cultural Models, a Historical Survey", Human Capital, N°2 (10), 2014

FISCHER, MAX (2013): <http://www.washingtonpost.com/blogs/worldviews/wp/2013/03/21/a-fascinating-map-of-countries-color-coded-by-their-openness-to-foreigners/> May 2014

HAMMER, MITCHEL R. / BENNETT, MILTON J. / WISEMAN, RICHARD L. (2003): Measuring intercultural sensitivity: The intercultural development inventory, in: International Journal of Intercultural Relations, Volume 27, Issue 4 2003

International Air Transport Association (IATA) press release N° 50, Dec. 2012, <http://www.iata.org/pressroom/pr/pages/2012-12-06-01.aspx>

ILGA – International Lesbian Gay Bisexual Trans and Intersex Association (2013), State Sponsored Homophobia Report [www.ilga.org](http://www.ilga.org)

WORLD ECONOMIC FORUM (2013): [http://www3.weforum.org/docs/WEF\\_TT\\_Competitiveness\\_Report\\_2013.pdf](http://www3.weforum.org/docs/WEF_TT_Competitiveness_Report_2013.pdf)

## IMPRESSUM

Herausgeber und V. i. S. d. P.  
Prof. Dr. Robert Grebner  
Hochschule für angewandte Wissenschaften  
Würzburg-Schweinfurt

Münzstraße 12  
97070 Würzburg  
Telefon +49 931 3511-6002  
Fax +49 931 3511-6044

Ignaz-Schön-Straße 11  
97421 Schweinfurt  
Telefon +49 9421 3511-6002  
Telefon +49 9421 3511-6002

Ansprechpartner  
Prof. Dr. Karsten Huffstadt  
Nadja Schaefer  
scj@fhws.de



ISSN 2196-6095

Bibliographische Information der Deutschen Bibliothek:  
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind immer im Internet über <http://dnb.ddb.de> abrufbar.



