ISSN 1580-2949
UDK 669+666+678+53 MTAEC9, 44(4)163–232(2010)

MATER. TEHNOL.	LETNIK VOLUME 44	ŠTEV. NO. 4	STR. P. 163–232	LJUBLJANA SLOVENIJA	JULY–AUG. 2010

VSEBINA – CONTENTS

UDK 533
Review article/Pregledni članek

ISSN 1580-2949
MTAEC9, 44(4)165(2010)

SURFACE MODIFICATION OF MATERIALS USING AN EXTREMELY NON-EQUILIBRIUM OXYGEN PLASMA

MODIFIKACIJA POVRŠINE MATERIALOV Z IZRAZITO NERAVNOVESNO KISIKOVO PLAZMO

Miran Mozetič

Odsek za tehnologijo površin in optoelektroniko, Institut "Jožef Stefan", Jamova cesta 39, 1000 Ljubljana, Slovenija,
miran.mozetic@ijs.si

Prejem rokopisa – received: 2010-01-04; sprejem za objavo – accepted for publication: 2010-02-05

Several technological processes based on the interaction of extremely non-equilibrium oxygen plasma are described. A plasma with a low kinetic temperature of heavy particles and an extremely high density of neutral oxygen atoms is created in glass plasma reactors by inductively coupled radio-frequency (RF) discharges. The density of the charged particles is kept low at around 10^{16} m^{-3}, while the density of the neutral oxygen atoms may exceed a value of 1×10^{22} m^{-3}. The neutral oxygen atoms are chemically active and readily react with solid materials, even at low temperatures. The interaction of the oxygen atoms with metal samples often causes a rapid nucleation of metal oxide isles and the spontaneous growth of one- or two-dimensional structures, such as nanowires and nanobelts. These "nanofeatures" are often monocrystalline. The exposure of different carbon-rich materials to an oxygen plasma is applied for the functionalization with oxygen functional groups as well as for the controlled oxidation of various materials. The superhydrophilicity of several polymers and composites can be achieved. The technique is suitable for the destruction of bacteria and thus the sterilization of delicate materials.

Keywords: plasma, oxygen, metal oxide nanoparticles, polymer functionalization, sterilization

Opisani so nekateri tehnološki postopki za obdelavo materialov z izrazito neravnovesno kisikovo plazmo. Stanje plinske plazme z nizko kinetično energijo težkih delcev in zelo veliko gostoto nevtralnih kisikovih atomov se vzpostavi v induktivno sklopljeni radiofrekvenčni plinski razelektritvi v steklenem plazemskem reaktorju. Gostota nabitih delcev v takšni plazmi je razmeroma majhna in navadno ne presega vrednosti 1×10^{16} m^{-3}, medtem ko lahko gostota nevtralnih kisikovih atomov preseže vrednost 1×10^{22} m^{-3}. Nevtralni kisikovi atomi so kemijsko izredno aktivni in reagirajo s površino trdnih snovi že pri nizki temperaturi. Interakcija kisikovih atomov s površino kovinskih materialov pogosto povzroči bliskovito nukleacijo otočkov kovinskega oksida, iz katerih spontano rastejo eno- ali dvodimenzionalne strukture, kot so nanožice in nanotrakovi. Nanomateriali so pogosto monokristalinični. Materiali, ki vsebujejo ogljik, prav tako reagirajo z atomskim kisikom. Že majhna doza atomov povzroči modifikacijo površine tovrstnih materialov z nastankom kisikovih funkcionalnih skupin. Tako obdelani polimerni materiali so včasih superhidrofilni. Podaljšana izpostava organskih materialov kisikovim atomom vodi k postopni oksidaciji in s tem počasnemu jedkanju materialov. Pojav izkoriščamo pri sterilizaciji občutljivih organskih materialov, ki ne prenesejo klasičnih postopkov.

Ključne besede: plazma, kisik, nanodelci kovinskega oksida, funkcionalizacija polimerov, sterilizacija

1 INTRODUCTION

The increasing demands on the quality as well as the miniaturization of products facilitate the development of new technological processes. These new technologies should ensure the good quality of products, low production costs and should be environmentally acceptable. Classical approaches to the treatment of materials have been almost exhausted and industry is looking for new ideas. The majority of breakthrough technologies that have appeared in recent years are based on the application of non-equilibrium environments, especially non-equilibrium gaseous plasma.[1] A variety of gases are used to create a plasma with suitable properties[1]. A plasma of particular interest is created with oxygen[1]. Molecular oxygen is passed through an electrical discharge. The fast electrons then excite the neutral molecules with the original Maxwell-Boltzmann (MB) distribution causing a dramatic shift to the non-equilibrium state. Apart from neutral molecules in low ro-vibrational states, substantial amounts of highly excited molecules, neutral atoms and ionized molecules and atoms are created.[2–3] While the distribution of neutral molecules over translational states is preserved close to MB due to the poor kinetic energy exchange between fast electrons and slow neutrals during elastic collisions, it is not true for internal energy distribution. Two excited states of molecular oxygen are metastable (at the excitation energy of approximately 1 eV and 2 eV) and their concentration is far above the values calculated using MB distribution at room temperature.[4] Furthermore, the dissociation degree is far from the equilibrium value and the same applies for the ionization fractions. Different temperatures that are suitable for a description of such a non-equilibrium state of oxygen depend on the particular conditions (discharge and plasma parameters).[5] The largest shift from MB distribution is often observed by passing the oxygen through the radio frequency of microwave discharges created in glass plasma reactors. At the neutral gas kinetic temperature close to room temperature, the other temperatures (T) are as follows: vibrational temperature 500–1000 K, dissociation temperature 100 000 K or even higher, ionization temperature 10 000–50 000 K, electron temperature

15 000–50 000 K.[6] Such a non-equilibrium gas often acts as an ordinary gas at room or slightly elevated temperature in terms of the heat dissipation on solid materials, while it acts as an extremely hot gas (well over 10,000 K) in terms of chemical activity. Obviously, heavily non-equilibrium oxygen is used for the rapid chemical modification of solid materials at low temperature.[7]

2 DETERMINATION OF THE KEY PLASMA PARAMETERS

Plasma is a partially ionized gas, and the most obvious plasma parameters are the density of the charged particles (free electrons and positive ions) and their temperatures (or the kinetic energy distribution functions in the case the particles that are away from the MB distribution). Apart from these, another couple of parameters are often stated, i.e., the plasma potential (V_s) and the Debye length (λ_D). In many cases, both are calculated with equations:[8]

$$V_s - V_f = \frac{-kT_e}{2e_0} \ln \frac{m_+}{2m_e} \quad (1)$$

and

$$\lambda_s = \sqrt{\frac{kT_e}{N_e e_0^2}} \quad (2)$$

In equations (1) and (2), V_s is the plasma potential (often called the space potential), V_f is the floating potential, e_0 is the elementary charge, k is the Boltzmann constant, T_e is the electron temperature, m_+ is the positive ion mass, m_e is the electron mass, N_e is the electron density, and ε_0 is the vacuum permittivity. These two equations are useful only as long as some requirements are fulfilled. Obviously, one requirement is the MB distribution of electrons. The set of parameters T_e, N_e, V_s, λ_D is useful for describing the electrical properties of the plasma, but tell little about its chemical reactivity. Apart from these classical parameters, many others are necessary for a decent characterization of the processing plasma, and they include the rotational and vibrational temperature of the neutral and charged molecules, the density of the neutral atoms, the density of the metastable molecules and atoms, and the ionization fractions of the molecules and atoms. In a plasma of electronegative gases, the density of the negatively charged heavy particles is not negligible, either. Complete plasma characterization is obviously a difficult task, so numerous methods have been developed and applied more or less successfully. Among them, electrical, magnetic and catalytic probes[9–18] are worth mentioning, along with more complicated techniques, such as a variety of optical emission[19–21] and absorption techniques,[22] the propagation of electromagnetic waves,[6] and chemical titrations.[23] A proper determination of the parameters is crucial for a calculation of the fluxes of the reactive plasma particles on the surface of the samples as well as the kinetic energy of the charged particles interacting with plasma facing materials. Although modern plasma laboratories are well equipped with several different tools, plasma is often characterized insufficiently.

3 SINTHESIS OF METAL OXIDE NANOPARTICLES

A novel technology based on the application of non-equilibrium oxygen plasma is the synthesis of large quantities of metal oxide nanoparticles.[24–32] Metallic foils are exposed to oxygen plasma with a dissociation fraction of the order of 10 % (corresponding to a dissociation temperature of 22 500 K).[24] Neutral oxygen atoms are, by far, more reactive than their parent molecules and interact with the surface of solid materials both chemically and physically. Chemical reactivity causes the formation of metal oxides, while the physical interaction is predominantly demonstrated by the accommodation (relaxation) of metastable molecules and the heterogeneous surface recombination of neutral oxygen atoms.[33–34] Furthermore, the bombardment of the metal surface with positive ions from the plasma is observed as well as the neutralization of charged particles. All the chemical and physical reactions are heavily exothermic, causing a localized disturbance of the surface atoms from their (quasi)equilibrium positions. The surface atoms become extremely mobile, so a disturbance of the crystalline structure occurs.[35–37] The extremely high mobility of the surface atoms allows for the creation of (meta)stable chains of metal oxides, stretching perpendicularly from the surface plain, as shown by molecular dynamics simulations.[38] An example of such a chain is shown in **Figure 1**. The chains act as nuclei for the growing of nanofeatures from the surface of the metal samples. As long as the mobility of the surface atoms is much higher in the preferred direction than in other orientations, the growth is one-dimensional, resulting in

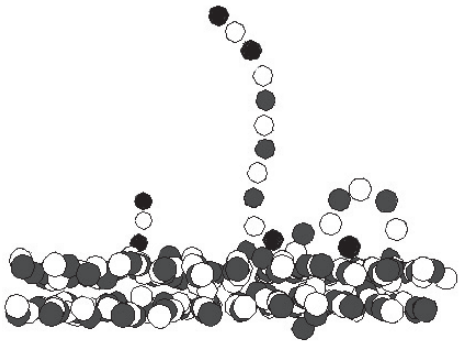


Figure 1: The results of a molecular dynamics simulation of the interaction of oxygen atoms added to the surface of nickel foil. The figure shows the formation of NiO chains. Dark and white particles stand for the O and Ni atoms, respectively.

Slika 1: Rezultati simulacije molekulske dinamike interakcije kisikovih atomov na površini nikljeve folije. Slika prikazuje tvorjenje verig NiO. Temni krogci prikazujejo kisikove atome, beli pa nikljeve.

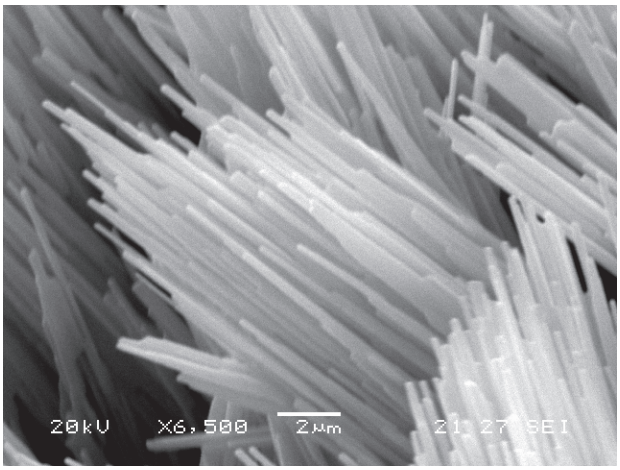


Figure 2: Nanowires of niobium pentoxide stretching from the surface of a niobium foil during exposure to an extremely non-equilibrium oxygen plasma with the following parameters: a neutral gas kinetic temperature of about 500 K, an electron temperature of 20 000 K, an ionization fraction of about 1×10^{-6} and a dissociation fraction of about 13 %.

Slika 2: Nanožičke niobijevega pentoksida, rastoče na površini niobijeve folije med izpostavo neravnovesni kisikovi plazmi z naslednjimi parametri: temperatura nevtralnega plina okoli 500 K, temperatura elektronov 20 000 K, delež ionizacije okoli 1×10^{-6} in delež disociacije okoli 13 %

long nanowires rather than other forms. **Figure 2** represents a SEM image of such nanowires, growing from the surface of niobium foil exposed to an extremely large flux of oxygen atoms of close to 1×10^{24} m^{-2} s^{-1}.[24] It is assumed that the nanowires are monocrystalline Nb_2O_5, the most stable form of niobium oxide. In the case of a lower dissociation fraction (lower O flux at about 3×10^{23} m^{-2} s^{-1}), other features grow from the surface of the metal foils. **Figure 3** shows nanobelts of rectangular shape growing from the surface of the niobium foils. Fluxes below about 1×10^{23} m^{-2} s^{-1} do not produce such interesting nanoparticles. Namely, at a low flux, the surface mobility of the atoms is too low to allow for the rapid growth of one- or two-dimensional structures, so the oxidation mechanism is close to that characteristic for a thermal (equilibrium) treatment. Similar results are obtained with many other metals, including Fe, V, Mo, Cr.[29]

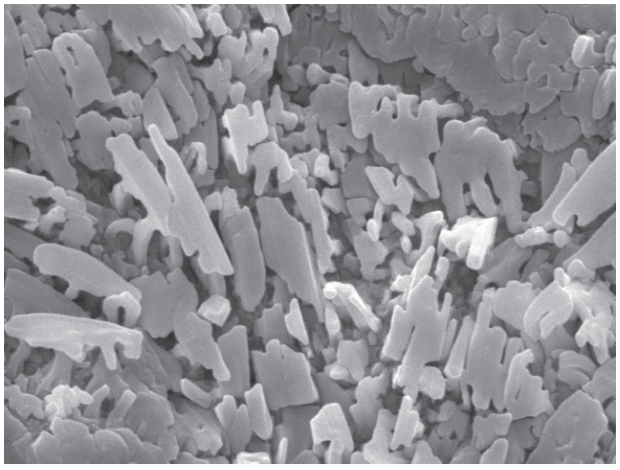


Figure 3: Rectangular nanoparticles of niobium pentoxide growing from the surface of niobium foil during exposure to highly non-equilibrium oxygen plasma with the following parameters: a neutral gas kinetic temperature of about 500 K, an electron temperature of 20 000 K, an ionization fraction of about 8×10^{-7} and a dissociation fraction of about 4 %.

Slika 3: Pravokotni nanodelci niobijevega pentoksida, rastoči na površini niobijeve folije med izpostavo neravnovesni kisikovi plazmi z naslednjimi parametri: temperatura nevtralnega plina okoli 500 K, temperatura elektronov 20 000 K, delež ionizacije okoli 8×10^{-7} in delež disociacije okoli 4 %.

4 SURFACE ACTIVATION OF POLYMER MATERIALS

A popular application of oxygen plasma is the surface functionalization of organic materials.[39–56] Polymer materials are usually hydrophobic with a fairly low surface energy. This hydrophobicity is usually the consequence of a poor concentration of polar functional groups on the surface of the polymers. In thermal equilibrium the concentration of the functional groups on the surface is the same as the bulk concentration. The surface properties of polymer materials are changed by the incorporation of different functional groups on the very surface.[39–40,50–51,54–55] Obviously, the new surface functional groups are not very stable and tend to decay spontaneously.[47] Several techniques for the functionalization of organic materials have appeared. Since the presence of foreign functional groups on the surface is not thermodynamically favorable, the obvious technique is the application of a non-equilibrium gaseous plasma. As mentioned above, non-equilibrium oxygen plasma with the kinetic temperature close to room temperature acts as an almost perfect medium for the surface functionalization with oxygen functional groups. Its extreme chemical reactivity allows for rapid functionalization at minimal thermal loading.[43] The applicability of this technology is demonstrated for the case of industrial LDPE (low-density polyethylene) polymer foil used for packaging (**Figures 4** and **5**).[57] The polymer is exposed to oxygen plasma at almost room neutral gas kinetic temperature, a dissociation fraction of about 10 % and a rather low ionization fraction of 1×10^{-6} (corresponding to the ionization temperature of about 12 000 K). The material heating due to the surface neutralization of charged particles is therefore negligible. Furthermore, the Debye length in such plasma is much larger than the mean free path of oxygen molecules or atoms. Therefore, the kinetic energy that the ions gain passing the potential difference between space and floating potentials (see equation (1)) is effectively distributed to neutral molecules or atoms before colliding with the polymer surface. The polymer material is therefore the subject of extremely poor heating, but high chemical modifications. Assuming an infinite planar geometry, the flux of ions (j_+) and atoms (j_a)

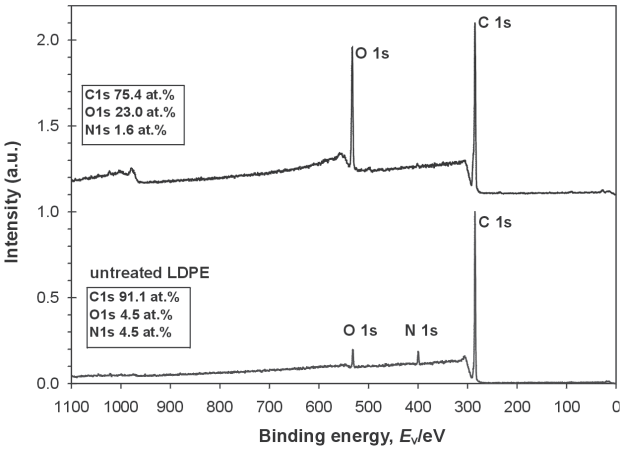


Figure 4: The XPS (X-ray photoelectron spectroscopy) survey spectra of untreated and plasma-treated industrial LDPE polymer foil. The oxygen concentration in the surface film of the polymer is increased to 23 % after treatment with plasma with the following parameters: a neutral gas kinetic temperature of about 400 K, an electron temperature of 20 000 K, an ionization fraction of about 1×10^{-6} and a dissociation fraction of about 10 %. The treatment time was 3 s.

Slika 4: Pregledni spekter XPS (rentgenske fotoelektronske spektroskopije) neobdelane in plazemsko obdelane polimerne folije LDPE. Koncentracija kisika na površini polimera je narasla na 23 % po obdelavi s plazmo z naslednjimi parametri: temperatura nevtralnega plina okoli 400 K, temperatura elektronov 20 000 K, delež ionizacije okoli 1×10^{-6} in delež disociacije okoli 10 %. Čas obdelave je bil 3 s.

onto the surface of the polymer does not depend on sheath properties, and is as follows:[58]

$$j_+ = ¼\, n_+ \langle v_+ \rangle \quad (3)$$

$$j_a = ¼\, n_a \langle v_a \rangle \quad (4)$$

Here, n_+ is the density of positively charged ions, $\langle v_+ \rangle$ is the average absolute random velocity of positively charged ions assuming an MB distribution, n_a is the neutral atom density, and $\langle v_a \rangle$ is the average absolute random velocity of neutral oxygen atoms, assuming an MB distribution.

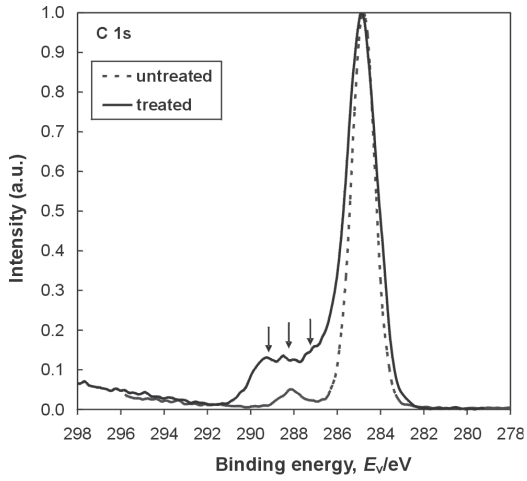


Figure 5. High-resolution C1s XPS peak for industrial LDPE polymer foil treated undert same conditions as in **Figure 4**

Slika 5: Visoko ločljivi spekter XPS vrha C1s industrijske polimerne folije LDPE, obdelane pri istih pogojih kot na **sliki 4**

There are few reasons for the acceleration of heavy charged particles (molecular and atomic ions) in high-frequency discharges. Namely, as the frequency of the electrical field approaches a few MHz, the local electromagnetic oscillations become too fast to be followed by heavy particles. Only the electrons can be heated (can gain kinetic energy) in radio-frequency fields. The distribution of heavy charged particles over translational states is thus the same as those for neutral particles – MB close to room temperature. The average velocity of the positively charged ions in equation (3) is therefore the same as for the neutral particles at room temperature (T): $\langle v_+ \rangle$ = 440 m/s for molecular ions and 630 m/s for atomic ions. The average velocity was calculated from equation $v = \sqrt{8kT / \pi m}$, where k is the Boltzmann constant and m is the ion mass.[58] Since the density of the molecules in the plasma is an order of magnitude larger than the density of the atoms, and since the dissociative ionization is energetically less favorable than the ordinary ionization of the molecules, the contribution of atomic ions is easily neglected and $\langle v_+ \rangle$ can be approximated to the value characteristic for molecules, i.e., 440 m/s. The resultant flux of ions (j_+) on the polymer surface exposed to plasma is thus:[58]

$$j_+ = ¼\, n_+ \langle v_+ \rangle = ¼\, 3 \times 10^{16}\ \mathrm{m^{-3}} \times 440\ \mathrm{m/s} = \\ = 3.3 \times 10^{18}\ \mathrm{m^{-2}s^{-1}} \quad (5)$$

Ions are recombined on the surfaces of all solid materials at a very high rate, close to 100 %. The thermal load per unit area caused by ion recombination is:[42]

$$P_+ = j_+ \mathrm{W_i} = 3.3 \times 10^{18}\ \mathrm{m^{-2}s^{-1}} \times 12 \times 1.6 \times 10^{-19}\ \mathrm{J} = \\ = 6\ \mathrm{W\ m^{-2}} \quad (6)$$

Here, P_+ is the power dissipated by the oxygen ion recombination per unit area and W_i is the ionization energy of the oxygen molecules, i.e., about 12 eV. The power P_+ is extremely low, so heating by ions is always negligible, as long as the ionization fraction is that low. The upper calculation is valid for the case when oxygen molecular ions are found in the ground state. In practice, this is never completely true, since there are always some excited ions in oxygen plasma. Fortunately, the density of ions with a high excitation energy in plasma with a low ionization fraction and a rather low electron temperature is usually negligible, so the upper calculation is a good approximation. The positive ions are accelerated in the sheath between the plasma and the surface (according to equation (1)) so the heating by ions is larger than the value given by (6). The maximum heating is obviously in the case of collision-less sheath when the ions gain a kinetic energy of $e_0 (V_s - V_f)$[8]. As explained above, however, the ions lose their kinetic energy in the gas phase as long as the Debye length is much larger than the mean free path of the molecules.

The heating of polymer materials by the recombination of neutral atoms is often also negligible. The thermal load per unit area due to this effect is calculated taking into account equation (4):[59]

$$P_a = 1/4\ n_a \langle v_a \rangle \gamma\ W_D/2 = 1/4 \times 6 \times 10^{20}\ m^{-3} \times \times 630\ m\ s^{-1} \times 10^{-4} \times 5.2 \times 1.6 \times 10^{-19}\ J / 2 = 4\ W\ m^{-2} \quad (7)$$

Here, γ is the recombination coefficient and W_D is the dissociation energy, i.e., $W_D = 5.2$ eV. The recombination coefficient for flat polymer materials is low, often of the order of 10^{-4}. Taking into account this value, the power dissipated on the polymer surface due to the recombination of atoms is approximately $P_a = 4\ W\ m^{-2}$. The total contributions of the neutral atom recombination (P_a) and the neutralization of ions (P_+) are $P = P_a + P_+$ and are therefore about 10 W m^{-2}. A polymer foil with a thickness d of 10 μm is therefore heated at:[58]

$$dT/dt = P/(\rho\ c_p\ d) = 10\ W\ m^{-2} / (10^3\ kg\ m^{-3} \times \times 1\ kJ\ kg^{-1}\ K^{-1} \times 10 \times 10^{-6}\ m) = 1\ K/s \quad (8)$$

This value, of course, depends on the properties of a particular polymer (the density ρ and the thermal capacity c_p) and is only a rough estimation. It depends also on the surface roughness (rough materials are heated at a faster rate) and the thickness (thicker materials are heated at lower rates). Since the flux of atoms on the surface is enormous (in the above calculation the value of $1 \times 10^{23}\ m^{-2}\ s^{-1}$ was taken into account), the surface of the polymer is saturated with oxygen functional groups in much less than a second, so far before any substantial heating of the material occurs.

5 STERILIZATION OF DELICATE MATERIALS

Since plasma treatment does not heat the materials much, oxygen plasma is used for the sterilization of delicate biocompatible materials.[60–65] Namely, oxygen plasma interacts with bacteria already at room temperature, causing the destruction of the bacterial cell wall and, therefore, sterilization. The sterilization effect is illustrated by **Figures 6** and **7**. **Figure 6** is an SEM image of the live bacterium *Bacillus Stearothermophilus*. The bacteria were deposited on substrates and treated with oxygen plasma with a neutral gas kinetic temperature of about 400 K, an electron temperature of 20 000 K, an ionization fraction of about 3×10^{-6} and a dissociation fraction of about 10 %.[62] The effect of the plasma treatment is well illustrated by **Figure 7**. The treatment time was 55 s. The reactive particles from the oxygen plasma (mainly neutral oxygen atoms) readily reacted with the bacteria. The capsule protecting the live bacteria disappeared and the bacterial cell wall is badly damaged. What remains from the bacteria after the plasma treatment is just the remnants of the bacterial cytoplasm. Obviously, some parts of the cytoplasm do not react with the plasma particles as much as the bacterial cell wall, let alone the capsule. The type of bacteria was not chosen accidentally. Bacteria *Bacillus Stearothermophilus* is famous for its resistance to prolonged heating at elevated temperature.[66] Even a half an hour treatment at 100 °C is not sufficient to ensure the sterility of objects contaminated with this type of bacteria. Plasma sterilization is therefore a promising method for the destruction of bacteria on biocompatible objects that do not stand standard autoclaving in humid air at 130 °C. Since the bacterium presented in **Figure 7** is very badly damaged it can be concluded that plasma treatment times shorter than a minute are sufficient for the sterilization by extremely non equilibrium oxygen plasma.

6 CONCLUSIONS

Some applications of extremely non-equilibrium oxygen plasma were presented. An oxygen plasma with a high degree of dissociation is created in a radio-frequency discharge. The neutral gas kinetic temperature remains close to room temperature, so the thermal loads on substrates are minimal. The chemical reactivity of such plasma, on the other hand, is extremely high. Such plasma is suitable for the modification of both inorganic and organic materials. The exposure of metal foils leads to the spontaneous growth of metal oxide nanoparticles.

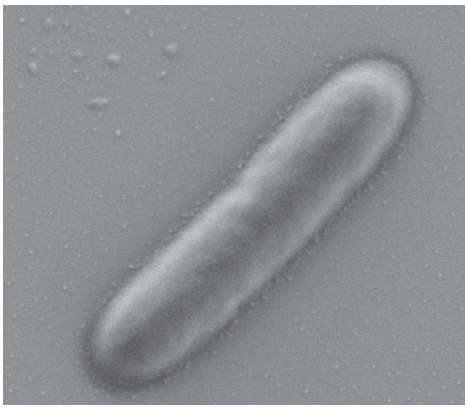


Figure 6: SEM image of *Bacillus Stearothermophilus* before plasma treatment

Slika 6: SEM slika bakterije *Bacillus Stearothermophilus* pred plazemsko obdelavo

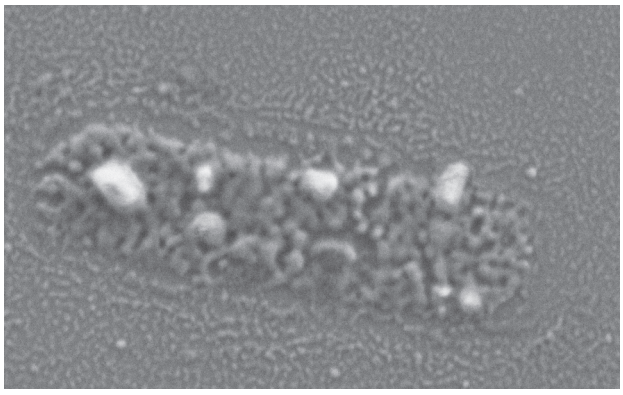


Figure 7: SEM image of *Bacillus Stearothermophilus* after treatment with plasma for 55 s. The plasma parameters were the same as for the treatment of the polymer foil (as in **Figure 4**).

Slika 7: SEM slika bakterije *Bacillus Stearothermophilus* po obdelavi v plazmi 55 s. Parametri plazme so bili isti kot pri obdelavi polimerne folije (kot na **sliki 4**).

As long as the flux of atoms on the surface of the samples is very large, an oxide film will grow in the form of long nanowires. At moderate flux, the oxide is often in the form of rectangular nanoparticles. The treatment of polymer materials with an oxygen plasma leads to a surface functionalization with oxygen-rich functional groups. Even a brief exposure results in a saturation of the surface with polar functional groups, causing a dramatic increase in the surface energy and thus the wettability. Treatment with oxygen plasma also causes the destruction of bacteria and therefore represents a suitable medium for the sterilization of delicate biocompatible objects.

7 REFERENCES

[1] R. d'Agostino, P. Favia, Y. Kawai, H. Ikegami, N. Sato, F. Arefi-Khonsari, Advanced Plasma Technology, Wiley, Weinheim 2008, 457

[2] Z. Lj. Petrovic, M. Suvakov, Z. Nikitovic, S. Dujko, O. Sasic, J. Jovanovic, G. Malovic, V. Stojanovic, Plasma Sources Sci. Technol., 16 (**2007**), S1–S12

[3] R. E. Robson, R. D. White, Z. Lj. Petrovic, Rev. Mod. Phys., 77 (**2005**) 4, 1303–1320

[4] A. Ricard, Reactive plasmas, SFV, Paris 1996, 156

[5] Z. Donko, P. Hartmann, K. Kutasi, Plasma Sources Sci. Technol., 15 (**2006**) 2, 178–186

[6] R. H. Huddlestone, S. L. Leonard, Plasma Diagnostic Techniques: Pure and Applied Physics Series Vol 21, Willey, New York 1965, 627

[7] T. Belmonte, C. Pintassilgo, T. Czerwiec, G. Henrion, V. Hody, J. M. Thiebaut, J. Loureiro, Surf. Coat. Tech., 200 (**2005**) 1–4, 26–30

[8] J. D. Swift, M. J. R. Schwar, Electrical Probes for Plasma Diagnostic, Iliffe Books, London 1969, 334

[9] A. Vesel, A. Drenik, M. Mozetic, M. Balat-Pichelin, Vacuum, 84 (**2010**) 7, 969–974

[10] A. Vesel, M. Mozetic, Vacuum, 61 (**2001**) 2–4, 373–377

[11] M. Mozetic, A. Vesel, V. Monna, A. Ricard, Vacuum, 71 (**2003**) 1–2, 201–205

[12] A. Drenik, U. Cvelbar, A. Vesel, M. Mozetic, Inf. MIDEM, 35 (**2005**), 85–91

[13] M. Mozetic, U. Cvelbar, A. Vesel, A. Ricard, D. Babic, I. Poberaj, J. Appl. Phys., 97 (**2005**) 10, 103308–1–103308–7

[14] M. Balat-Pichelin, A. Vesel, Chem. Phys., 327 (**2006**) 1, 112–118

[15] A. Drenik, U. Cvelbar, A. Vesel, M. Mozetic, Strojarstvo, 48 (**2006**) 1/2, 17–22

[16] M. Mozetic, A. Vesel, A. Drenik, I. Poberaj, D. Babic, J. Nucl. Mater., 363–365 (**2007**), 1457–1460

[17] A. Vesel, M. Mozetic, M. Balat-Pichelin, Vacuum, 81 (**2007**) 9, 1088–1093

[18] A. Drenik, A. Tomeljak, M. Mozetic, A. Vesel, D. Babic, M. Balat-Pichelin, Vacuum, 84 (**2010**) 1, 90–93

[19] N. Krstulovic, U. Cvelbar, A. Vesel, S. Milosevic, M. Mozetic, Mater. Tehnol., 43 (**2009**) 5, 245–249

[20] N. Krstulovic, I. Labazan, S. Milosevic, U. Cvelbar, A. Vesel, M. Mozetic, J. Phys. D: Appl. Phys., 39 (**2006**) 17, 3799–3804

[21] N. Krstulovic, I. Labazan, S. Milosevic, U. Cvelbar, A. Vesel, M. Mozetic, Mater. Tehnol., 38 (**2004**) 1, 51–54

[22] F. Gaboriau, U. Cvelbar, M. Mozetic, A. Erradi, B. Rouffet, J. Phys. D: Appl. Phys., 42 (**2009**) 5, 055204–1–055204–5

[23] A. Ricard, M. Gaillard, V. Monna, A. Vesel, M. Mozetic, Surf. Coat. Tech., 142–144 (**2001**), 333–336

[24] M. Mozetic, U. Cvelbar, M. K. Sunkara, S. Vaddiraju, Adv. Mater., 17 (**2005**) 17, 2138–2142

[25] U. Cvelbar, M. Mozetic, J. Phys. D: Appl. Phys., 40 (**2007**) 8, 2300–2303

[26] U. Cvelbar, K. Ostrikov, I. Levchenko, M. Mozetic, M. K. Sunkara, Appl. Phys. Lett., 94 (**2009**) 21, 211502–1–211502–3

[27] U. Cvelbar, K. Ostrikov, A. Drenik, M. Mozetic, Appl. Phys. Lett., 92 (**2008**) 13, 133505–1–133505–3

[28] Z. Chen, U. Cvelbar, M. Mozetic, J. He, M. K Sunkara, Chem. Mater., 20 (**2008**) 9, 3224–3228

[29] K. Ostrikov, Plasma Nanoscience: Basic Concept and Applications of Deterministic Nanofabrication, Wiley, New York 2008, 504

[30] A. Drenik, U. Cvelbar, K. Ostrikov, M. Mozetic, J. Phys. D: Appl. Phys., 41 (**2008**) 11, 115201–1–115201–7

[31] U. Cvelbar, K. Ostrikov, M. Mozetic, Nanotechnology, 19 (**2008**) 40, 405605–1–405605–7

[32] U. Cvelbar, Z. Chen, M. K. Sunkara, M. Mozetic, Small, 4 (**2008**) 10, 1610–1614

[33] M. Mozetic, A. Vesel, U. Cvelbar, A. Ricard, Plasma Chem. Plasma Process., 26 (**2006**) 2, 103–117

[34] M. Mozetic, U. Cvelbar, Plasma Sources Sci. Technol., 18 (**2009**) 3, 034002–1–034002–5

[35] D. Mariotti, K. Ostrikov, J. Phys. D: Appl. Phys., 42 (**2009**) 9, 092002–1–092002–4

[36] I. Levchenko, U. Cvelbar, K. Ostrikov, Appl. Phys. Lett., 95 (**2009**) 2, 021502–1–021502–3

[37] I. Levchenko, K. Ostrikov, K. Diwan, K. Winkler, D. Mariotti, Appl. Phys. Lett., 93 (**2008**) 18, 183102–1–183102–3

[38] A. Vesel, E. Vamvakopoulos, M. Mozetic, G. A. Evangelakis, Physica B, 324 (**2002**) 1–4, 261–267

[39] A. Asadinezhad, I. Novak, M. Lehocky, F. Bilek, A. Vesel, I. Junkar, P. Saha, A. Popelka, Molecules, 15 (**2010**) 2, 1007–1027

[40] U. Cvelbar, M. Mozetic, I. Junkar, A. Vesel, J. Kovac, A. Drenik, T. Vrlinic, N. Hauptman, M. Klanjsek-Gunde, B. Markoli, N. Krstulovic, S. Milosevic, F. Gaboriau, T. Belmonte, Appl. Surf. Sci., 253 (**2007**) 21, 8669–8673

[41] M.-J. Wang, Y.-I. Chang, F. Poncin-Epaillard, Surf. Interface Anal., 3, 348–355

[42] A. Vesel, M. Mozetic, A. Hladnik, J. Dolenc, J. Zule, S. Milosevic, N. Krstulovic, M. Klanjsek-Gunde, N. Hauptman, J. Phys. D: Appl. Phys., 40 (**2007**) 12, 3689–3696

[43] T. Vrlinic, A. Vesel, U. Cvelbar, M. Krajnc, M. Mozetic, Surf. Interface Anal., 39 (**2007**) 6, 476–481

[44] N. Médard, J.-C. Soutif, F. Poncin-Epaillard Langmuir, 18 (**2002**) 6, 2246–2253

[45] A. Vesel, M. Mozetic, A. Zalar, Surf. Interface Anal., 40 (**2008**) 3–4, 661–663

[46] G. Legeay, F. Poncin-Epaillard, C. R. Arciola, Int J. Artif. Organs., 29 (**2006**) 4, 453–461

[47] A. Vesel, I. Junkar, U. Cvelbar, J. Kovac, M. Mozetic, Surf. Interface Anal., 40 (**2008**) 11, 1444–1453

[48] M. Aouinti, P. Bertrand, F. Poncin-Epaillard, Plasmas Polym., 8 (**2003**) 4, 225–236

[49] A. Vesel, M. Mozetic, A. Zalar, Vacuum, 82 (**2008**) 2, 248–251

[50] M. Sowe, I. Novak, A. Vesel, I. Junkar, M. Lehocky, P. Saha, I. Chodak, Int. J. Polym. Anal. Ch., 14 (**2009**) 7, 641–651

[51] A. Vesel, Inf. MIDEM, 38 (**2009**), 257–265

[52] I. Junkar, U. Cvelbar, A. Vesel, N. Hauptman, M. Mozetic, Plasma Processes Polym., 6 (**2009**) 10, 667–675

[53] M. Gorjanc, V. Bukosek, M. Gorensek, A. Vesel, Tex. Res. J., (**2009**), 1–11, doi:10.1177/0040517509348330, in press

[54] A. Vesel, M. Mozetic, S. Strnad, Z. Peršin, K. Stana-Kleinschek, N. Hauptman, Vacuum, 84 (**2010**) 1, 79–82

[55] I. Junkar, A. Vesel, U. Cvelbar, M. Mozetic, S. Strnad, Vacuum, 84 (**2010**)1, 83–85

[56] C. Canal, F. Gaboriau, S. Villeger, U. Cvelbar, A. Ricard, Int. J. Pharm., 367 (**2009**) 1–2, 155–161

[57] T. Semenic, A. Vesel, M. Mozetic: Hydrophilization of LDPE and HDPE polymer by treatment in oxygen plasma, Proc of. 2nd Int. Conf. on Advanced Plasma Technol., Piran, 2009, 114–118

[58] J. Strnad, Fizika 1: Mehanika/Toplota, DMRS, Ljubljana 1989, 284

[59] M. Mozetic, Interakcija vodikove plazme s površinami trdnih snovi, Disertacija, Univerza v Mariboru 1997, 169

[60] D. Vujosevic, Z. Vratnica, A. Vesel, U. Cvelbar, M. Mozetic, A. Drenik, T. Mozetic, M. Klanjsek-Gunde, N. Hauptman, Mater. Tehnol., 40 (**2006**) 6, 227–232

[61] Z. Vratnica, D. Vujosevic, U. Cvelbar, M. Mozetic, IEEE Trans. Plasma Sci., 36 (**2008**) 4, 1300–1301

[62] D. Vujosevic, M. Mozetic, U. Cvelbar, N. Krstulovic, S. Milosevic, J. Appl. Phys., 101 (**2007**) 10, 103305-1-103305-7

[63] U. Cvelbar, D. Vujosevic, Z. Vratnica, M. Mozetic, J. Phys. D: Appl. Phys., 39 (**2006**) 16, 3487–3493

[64] O. Kylian, T. Sasaki, F. Rossi, Eur. Phys. J. Appl. Phys., 34 (**2006**) 2, 139–142

[65] U. Cvelbar, M. Mozetic, N. Hauptman, M. Klanjsek-Gunde, J. Appl. Phys., 106 (**2009**) 10, 103303–1–103303–5

[66] Available from World Wide Web: http://en.wikipedia.org/wiki/Bacillus_stearothermophilus

UDK 552.54(497.4):537.533.35
Review article/Pregledni članek

ISSN 1580-2949
MTAEC9, 44(4)173(2010)

MIDDLE TRIASSIC DRY-LAND PHASES IN SOUTHERN SLOVENIA

SREDNJETRIASNE KOPENSKE FAZE V JUŽNI SLOVENIJI

Stevo Dozet[1], Matjaž Godec[2]

[1]Geological Survey of Slovenia, Dimičeva ulica 14, 1000 Ljubljana, Slovenia
[2]Institute of Metals and Technology, Lepi pot 11, 1000 Ljubljana, Slovenia
matjaz.godec@imt.si

Prejem rokopisa – received: 2009-11-20; sprejem za objavo – accepted for publication: 2010-03-18

In the Middle Triassic lithostratigraphic sequence of southern Slovenia and neighbouring Gorski Kotar, Croatia, three dry-land phases have been recognized: 1 – between the Lower and Middle Triassic, 2 – in between the Anisian stratigraphic sequence, and 3 – between the Anisian and Ladinian beds. Traces of the Anisian volcanism have also been discovered. The dry-land episodes of the non-deposition and volcanic activity are ascribed to synsedimentary tectonic movements of the Alpine tectonic cycle. However, new data enable a new approach to the paleogeographical evolution of the area. In the Triassic "Dolomite" Complex, lenses of kaolinite deposits occur, representing an important correlation horizon for the basal Ladinian beds. They consist chiefly of kaolinite and quartz admixed in smaller quantities by hematite, goethite, nontronite, gibbsite, groutite and montmorillonite. The kaolinite deposits were formed in a marsh with weathering of the pyroclastic materials.

Key words: carbonate, clastic and volcanic rocks, dry-land phases, Middle Triassic, External Dinarides, Slovenia, SEM, EDS, XRD

V srednjetriasnem litostratigrafskem zaporedju južne Slovenije in sosednjega Gorskega Kotarja (Hrvaška) so ugotovljene tri okopnitve, in sicer: 1 – med spodnjim in srednjim triasom, 2 – v aniziju ter 3 – med anizijem in ladinijem. Odkriti so tudi sledovi anizijskega vulkanizma. Okopnitve in vulkanizem povezujemo s sočasnimi tektonskimi premiki Alpidskega tektonskega cikla. Novi podatki omogočajo nov pogled na paleogeografski razvoj obravnavanega ozemlja. V triasnem dolomitnem kompleksu se pojavljajo leče kaolinitnih usedlin, ki so pomemben korelacijski horizont za bazalne ladinijske plasti. Sestojijo v glavnem iz kaolinita in kremena, katerima so v manjših količinah primešani hematit, goethit, nontronit, gibbsit, groutit in montomorilonit. Kaolinitne usedline so nastale v močvirju pri razpadu piroklastičanega materiala.

Ključne besede: karbonatne, klastične in vulkanske kamnine, kopenske faze, srednji trias, Zunanji Dinaridi, Slovenija, SEM, EDS, XRD

1 INTRODUCTION

The territory that is in question, covers the southern Slovenia and the extreme northern parts of Gorski Kotar, which belongs to neighbouring Croatia. In the framework of the elaboration of the Basic Geologic Map on the scale of 1 : 100 000 the systematic regional geologic researches included among other Kočevje and Gorski Kotar [1] areas, White Carniola (Bela krajina) [2] and central Slovenia [3], that lie on the topographic Maps Grosuplje (No 45), Ribnica (No 54), Kočevje (No 55) and Črnomelj (No 56) on the scale of 1 : 25 000. In the last ten years the Geological Survey of Slovenia performed a detailed geological mapping for the elaboration of the Geologic Maps of Slovenia on the scales 1 : 25 000 and 1 : 50 000. The Map Sheet Grosuplje 1 : 25 000 consists of 25 sections on the scale of 1 : 10 000 embracing about 675 km^2. Also elaborated and digitalized is the Map Sheet Grosuplje on the scale 1 : 25 000. On the Map Sheet Grosuplje 1 : 25 000 lies among other Bloška planota that was the central part of our researches.

In southern Slovenia the dry-land phases and interruptions in sedimentation appear in different, and mostly in two, lithostratigraphic sequences: 1) in the variegated heterogeneous Triassic (Schytian-Carnian)

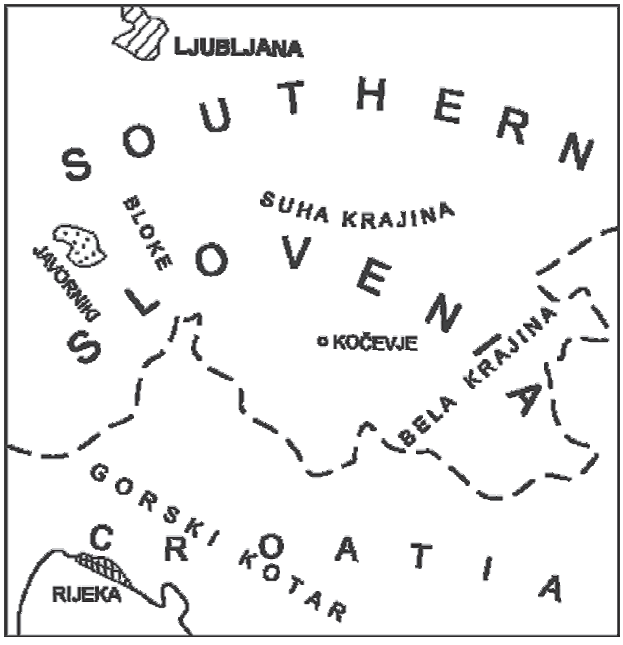


Figure 1: Location sketch map of the study area
Slika 1: Položajna skica obravnavanega ozemlja

chiefly clastic sedimentary succession of Kočevje, Carniola (Bela krajina) and Gorski Kotar, and 2) in the monotonous, light Middle Triassic "Dolomite" Complex on Bloke plateau.

The purpose of this work was to find answers and evidence for the following questions: What is the real age of the "Dolomite" Complex and its parts? Was the sedimentation of the "Dolomite" Complex continuous or interrupted? Are in the Triassic "Dolomite" Complex of Bloke plateau also present the Anisian, Ladinian and Cordevolian beds? How many interruptions occurred in the complex, and how long did they last? Finally, the aim of this paper is also to explain the particularities of the paleogeographical evolution of the study area during the Triassic period.

The listed problems required detailed stratigraphic investigations of the Middle Triassic "Dolomite" Complex and adjacent beds in the field and additional systematic laboratory examinations (**Figure 1**).

From the geographical point of view the investigated area belongs to the Lower and Inner Carniola (Dolenjska, Notranjska) [17,21], while, geologically and geotectonically, however, to the Slovenian and Dinaric Carbonate Platforms and the Lower and Inner Carniola Blocks [4,6,19].

2 EXPERIMENTAL PROCEDURES

In the past, geological researches in Slovenia were focused on systematic regional and detailed geological mappings of the considered area for the elaboration of the Basic Geological Map SFRY on the scale of 1 : 100 000 on the Map Sheet Ribnica (OGK I – RI). In the past ten years a detailed geological mapping for the Geological Map of Slovenia 1 : 50 000 on the Map Sheet Grosuplje 1 : 25 000 (GKS-GR) was achieved. It was accompanied by systematic laboratory examinations (a determination of the macro- and microfossils, XRD, geochemical and sediment-petrographic analyses and SEM/EDS analyses). The main mapping methods for the elaboration of the OGK I-RI were the mapping of all the exposures and the tracing of contacts, whereas the main methods for the elaboration of the GKS-GR were the mapping of motorway traces as well as stratimetric measurements and sedimentologic research of well-exposed representative cross-sections (stratimetric profiling).

Carbonate rocks were arranged according to FOLK's [23] and DUNHAM's [24] classifications, whereas the clastic sediments were defined regarding FOLK's [23] and PETTIJOHN's [25] classifications.

For the detailed analysis six selected rock samples were taken from three different regions in wider Bloke area, two from Puharje, two from Kaplanovo and two from Žgajnarji (**Table 1**). The elemental analyses of all six samples were performed using a field-emission scanning electron microscope JEOL JSM 6500F equipped with an EDS analyzer. Due to their very soft nature and the high crumbling tendency the samples were only carefully ground with 500 SiC paper. The accelerating voltage of the primary electron beam was 15 kV and the probe current was 0.2 nA.

Table 1: Designations of six selected samples that were taken at different places close to the Bloke region
Tabela 1: Oznake šestih izbranih vzorcev, ki so bili odvzeti na različnih mestih blizu področja Bloke

Sample designation	Name of the sample by the place where it was taken
1	Puharje 1
2	Puharje 2
3	Kaplanovo 1
4	Kaplanovo 2
5	Žgajnarji 1
6	Žgajnarji 2

With the aim to obtain information about the minerals present in the rocks, the phase analyses were performed using Philips Analytical X-ray diffraction (XRD) using $CuK_{\alpha 1}$ radiation (0.15406 nm) in the 2θ range from 0 to 70°.

The fraction of each mineral present in the six analysed samples was calculated from the XRD measurements and verified with EDS. The fraction of each mineral is expressed in approximate integer values, rounded up to 10 % or 5 %.

3 RESULTS

3.1 Geology

For the explanation of the paleogeographic development of the southern part of the Slovenian territory in the Middle Triassic period it is useful to recall the geological column of the clasticly developed Kočevje region and on the other side of the calcareously developed Bloke plateau (Bloška planota).

3.1.1 Geological column of the Kočevje region

The geological column of the Kočevje region consists of two parts: the lower clastic part and the upper carbonate part. The lower part is composed of the carbonate clastic Carboniferous, Permian and Triassic systems and the upper part involves the Upper Triassic series as well as the Jurassic and Cretaceous systems.

In the Kaptol-Mošenik cross-section south of Kočevje [26] the clastic formation of the Younger Paleozoic beds is subdivided into three lithostratigraphic units, which are in analogy with similar developments in Gorski Kotar and Ortnek [20] attributed to the Lower Permian (Trogkofelian). A relatively shallow sedimentation area, where the Lower Permian clastic rocks originated, began in the time of the ***Saalian movements*** to rise, so that the Kočevje and Gorski Kotar areas were in the Middle and Upper Permian dry land where in that

time no significant folding occurred. During a slow uplifting a relatively flat mainland originated, built of pretty clayey micaceous sediments with interbeds of compact quartz sandstones and conglomerates. The dry land was the product of the activity of the final strokes of the ***Hercynian orogeny***. The elevation of the considered area was accompanied by a fault tectonics in Gorski Kotar; however, with a weak magmatic activity as well [25].

The conclusions related to the dry land could also be deduced to the stratigraphic gap of the Scythian beds lying discordantly upon the Lower Permian bed. For the tectonic-erosional character of this boundary speaks also about 20-cm-thick limonite crust in the Gorski Kotar region [26–29] on the boundary between the Younger Paleozoic and Triassic sediments, that is a product of the strong oxidizing conditions. The elevation of the considered territory that began on the boundary between the Lower and Middle Permian in the time of the Saalian phase, continued uninterruptedly into the ***Pfälzian phase***, with which the Hercynian tectogenetic cycle in the Alps was ended. The Hercynian orogeny in the Dinarides area was followed by a common descending and afterwards the transgression of the Lower Triassic sea, which is according to U. Premru [16] closely connected with the extinguishing Herzynian orogenetic cycle. In the Triassic sea in the Kočevje, Gorski Kotar and Bela Krajina areas, a pretty variegated clastic-carbonate succession of sediments originated (**Figure 2**), of which the lower part is of Scythian and the upper one of Carnian age [10–13,30,31]. They are distinguished between themselves according to lithology, mineral composition, heavy minerals, geochemical composition, cement, diagenesis, source material, environment, Energy-index and climate.

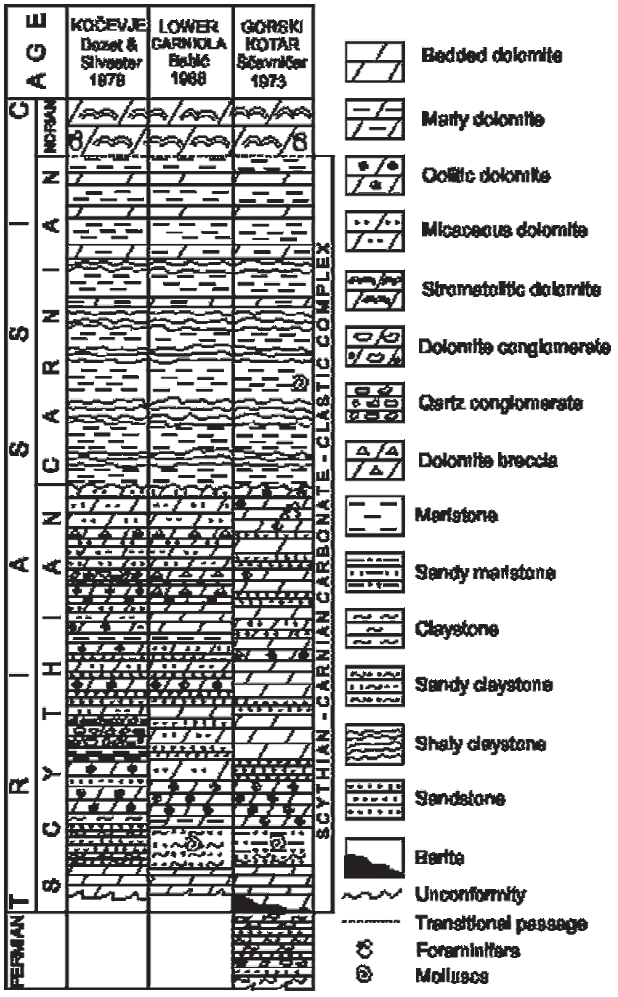


Figure 2: Correlation of the Scythian-Carnian carbonate-clastic succession of Kočevje, Gorski Kotar and Lower Carniola

Slika 2: Primerjava skitsko-karnijskega karbonatno-klastičnega zaporedja Kočevske, Gorskega Kotarja in Dolenjske

In the Scythian part of the clastic-carbonate succession oolitic and sandy dolomite prevail, and there is also less of the sandy dolomitic marlstone, claystone, micaceous sandstone and intraformational conglomerate. The Carnian (Julian and Tuvalian) beds are composed of red dolomitic marlstone containing interbeds of yellowish, greenish grey and grey dolomite as well as greenish grey claystone. Between heavy minerals turmaline prevails in the Scythian beds, and zircon in the Carnian beds.

The Scythian and Carnian sediments originated in a shallow sea and semiarid climate. The mainland in the background of the sedimentary basin was low. Toward the end of the Lower Triassic epoch the Kočevje region began slowly to uplift, as consequence of the ***Old Slovenian phase*** [32]. The elevation of the terrain is evidenced by interbeds of breccias and conglomerates in the Triassic sedimentary succession in the Kočevje region [9,32], the Gorski Kotar region [33–35] and the Lower Carniola (Dolenjska) region [34–37].

The continuous shallowing of the already relatively shallow Lower Triassic sea brought on a differentiation of the sedimentary area and finally originated the positive paleostructure comprising the Kočevje region, Bela krajina and Gorski Kotar [22]. The new dry land was gently anticlinally folded [38], and the folding was locally accompanied with weak radial tectonics. These events caused a reduction and later a complete interruption of the sedimentation in this part of the Dinarides, that lasted until the transgression of the Carnian (Julian) sea. The emersion in the Kočevje region lasted through the whole Anisian, Ladinian and Cordevolian epochs. With respect to the existence of dry land it is concluded with reference to the absence of the Middle Triassic and Lower Carnian (Cordevolian) rocks in this area, too.

3.1.2 "Dolomite" Complex on Bloke plateau

On the northeastern border of the Bloke plateau the Middle Triassic sedimentary succession occurs in a carbonate development, in which dolomite greatly prevails; therefore, it was denominated as the "Dolomite" Complex (**Figure 3**). In this Complex also more or less dolomitized limestones and even pure limestones occur,

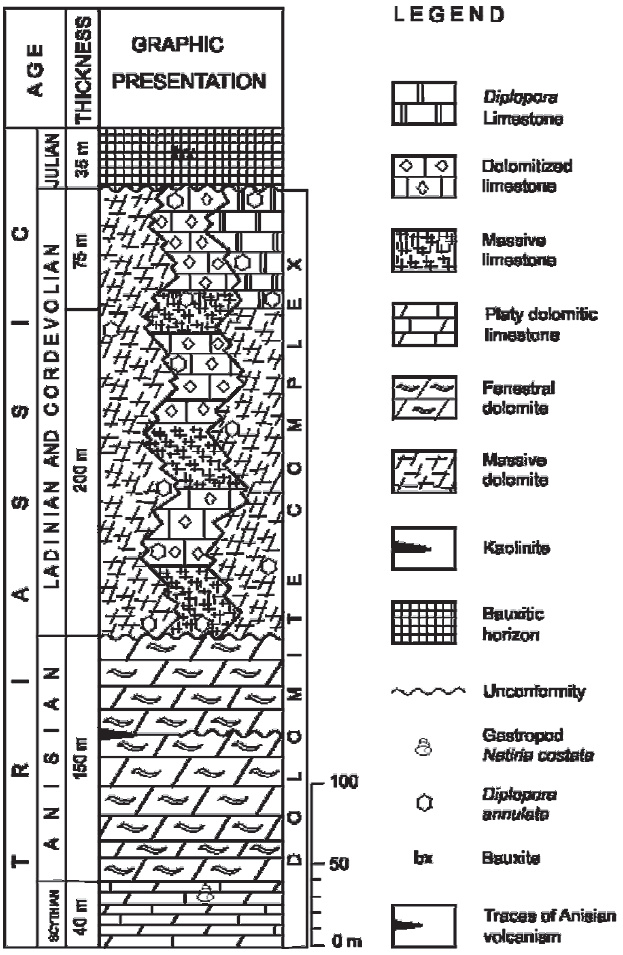


Figure 3: Geologic column of the "Dolomite Complex" with its footwall and hanging wall beds
Slika 3: Geološki stolpec "Dolomitnega" kompleksa s talninskimi in krovninskimi plastmi

that are due to the numerous dasycladal mostly rock-forming algae of the *Diplopora* genus, known under the name ***Diplopora Limestones*** and occurring in the complex above all in the uppermost part. In the Basic Geologic Map of SFRY on the scale 1 : 100 000 BUSER [3,39] ascribed the considered the "Dolomite" Complex to the Cordevolian substage, that is considered in Slovenia for several ten years already as belonging to the Lower Carnian substage. The dolomite or carbonate development of the Triassic beds also occurs in other parts of Slovenia and it is ascribed to its dissimilar age, from Anisian to Norian and Rhaetian [3,5,14,15,41–44].

3.1.3 Stratigraphic position

In the considered area the "Dolomite" Complex occurs mostly to an incomplete extent, since its uppermost or lowermost or both parts are tectonically cut off or eroded. In the valley of the Black Ravine (Črni potok) at Karlovica is exposed the footwall of the "Dolomite" Complex, represented by a several ten-meters-thick stratigraphic sequence of rosy, grey, platy dolomitic limestone, in which in the Škrilje and Podgozd areas the gastropod *Natiria costata* was found [3], that in Slovenia occurs in the uppermost Scythian beds. The conodont analysis of the samples from the beds at Podgozd was not positive [45].The hanging wall of the "Dolomite" Complex and their contacts are exposed on Veliki Vrh, where concordantly upon the "Dolomite" Complex is located a red pelitic and poorly oolitic iron-rich bauxite representing the basal horizon of the Julian beds. Concordantly upon the 30-meters-thick band of Carnian (Julian and Tuvalian) beds rests a dolomite in the Lofer development belonging to the Main Dolomite Formation. The "Dolomite" Complex thus lies concordantly between the Schythian and Julian-Tuvalian beds and represents, supposing that the sedimentation was uninterrupted, the Anisian, Ladinian (Fassanian and Langobardian) and Lower Carnian (Cordevolian) beds.

3.1.4 Lithology

The stratigraphic sequence of carbonate rocks denominated as the "Dolomite" Complex, originated on the Slovenian and Dinaric Carbonate Platforms, is not a homogeneous lithologic unit, since in some sections it includes, beside the dolomite, also, a strongly dolomitized limestone and bedded biostromal *Diplopora Limestone*. Also there, where the "Dolomite" Complex is built of pure dolomite, exists a certain heterogeneousity of the carbonate sequence as with regard to sedimentologic properties in the considered area, the lower and upper parts can be distinguished.

In the about 120-meters-thick interval of the lower part of stratigraphic sequence of the "Dolomite" Complex alternate light-grey to very-light-grey bedded fenestral, laminated and very fine-grained dolomites. The ***fenestral dolomite*** is a dolomite with numerous more or less parallel shrinkage pores. They may be an open space in the rock or completely or partly filled with secondarily introduced sediment or cement. The ***laminated dolomite*** is rare. In the rock the lamination is not produced by stromatolitic crusts, but by numerous, with stratification parallel, shrinkage pores filled with sparry calcite, limonite or sediments. The fine-grained dolomite (microdolosparite) does not show any porosity, whereas, the porosity of the laminated dolomite consists mostly of horizontal pores. The described carbonate sediments originated in a shallow lagune in an intertidal and supratidal environment. The types of carbonate rocks show that the Energy-index of the sea was low.

In the middle and upper part of the "Dolomite" Complex there is an about 350-to-400-meters-thick sedimentary succession built chiefly of massive and, here and there, bedded dolomite. The ***dolomite*** of the middle and upper part of the "Dolomite" Complex is mostly very light grey to almost white and sometimes it is bluish white, also the stratification of this sediment is rare. Almost always it is coarse-grained, respectively very coarse-crystallized, which speaks for its origin at a late

diagenesis of the *Diplopora* ***Limestone***. Usually, it is quite porous. The pores occur mostly in places that in former times were dasycladal algae. The pores and fine cracks are filled as a rule with fine dolomite crystals. The porosity of the rock is intragranular, sometimes, however, it is represented by isolated pores and fine cracks. The massive dolomite is compact and it is usually crumbling because of the coarse-grained texture. During weathering it disintegrates into a dolomitic sand, mylonite and ultramylonite.

The ***Diplopora Limestone*** is bedded and light grey, medium grey to grey or rosy. According to the texture it belongs to biosparite, bointrasparite, biolithite and rarely biomicrite. Commonly, it contains on the bed surfaces numerous round, oval and oblong sections of dasycladal algae of the genus *Diplopora*, that are mostly rock-forming. Diploporas lived on vast grasses on the bottom of the Cordevolian shallow and warm sea. More or less well-preserved *Diplopora* remains can also be found in the dolomite. Algae are better preserved in the limestones than in the dolomites. On the northern Borderland of the Bloke plateau we found in the *Diplopora Limestone* also an about 7.5-meters-thick interbed of bedded medium-grey stromatolitic limestone. The *Diplopora Limestone* passes laterally and vertically into the more or less dolomitized limestone or dolomitic limestone. In southern Slovenia in some places we come across larger complexes of dolomitic limestones. It is interesting and important that the dasycladacean are well-preserved in the dolomitic limestone as well. On the other hand, well-preserved dasycladacean remains are very rare due to diagenesis in the coarse-grained dolomite.

3.1.5 Tectonics

Our researches showed that in the Middle Triassic period there were three dry-land phases:

The first phase began already at the boundary between the Scythian and Anisian (Kočevje area and wider surrounding), the second one took place in the middle of the Anisian stage (Bloke plateau and environs) and the third phase was discovered somewhere in the middle part of the "Dolomite" Complex representing the boundary between the Anisian and Ladinian stage.

The first epeirogenic phase, denominated as ***Kočevje epeirogenic phase*** [9] is the largest phase especially with respect to the time of its lasting (from the end of the Scythian to the Julian). Besides the Kočevje region it also affected the White Carniola and Gorski Kotar area as well as some places in the Lower Carniola northwest of Kočevje. Unequal strongly intensified epeirogenic activity formed a positive structure with the centre in the Kočevje area. Upon the gently convexed antiform no sedimentation occurred. Evidence for the dry land is the absence of Anisian, Ladinian and Cordevolian rocks in the Kočevje region; however, primary exposures of the above-enumerated beds were not found in the White Carniola and Gorski Kotar as well. Anisian, Ladinian and Cordevolian rocks are also not found in the Triassic cross-sections of the Kočevje, White Carniola and Gorski Kotar regions. The Triassic columns consist of variegated (prevalently red) clastic sediments and more ore less frequent interbeds of dolomites and dolomitic marlstones. The lower part of the many-coloured clastic-carbonate succession belongs to Scythian; the upper one, however, to the Carnian. This is evidenced by macro- and microfossils as well as with petrographic-sedimentologic methods, heavy metals analysis and with XRD methods [28,30,35]. Since on the boundary between the Scythian and Carnian sequences there are no Anisian, Ladinian and Cordevolian rocks, and because the Carnian sedimentary succession starts nowhere with a thicker horizon of basal breccia or conglomerate, we believe that orogenetic forces and fault tectonics in that time were not present and that the newly formed dry land was flat and built of fine- and very fine-grained sediments.

Traces of the second Middle Triassic dry-land phase are exposed on the top of the lower third of the "Dolomite" Complex in the upper part of the sequence of light bedded fenestral dolomite, which is, according to its concordant position upon the rosy grey, thick-plated dolomitic limestone containing the gastropod *Natiria costata* and with reference to sedimentologic characteristics, attributed to the Anisian. The main evidence for this dry land phase, that was the shortest and the weakest at all three, is a 0.75-m- to 1.5-m-thick lense-like interbed of pale greenish kaolinite. The darker greenish tuff, however, speaks for volcanic activity [18] in the Anisian epoch. Otherwise, the Anisian volcanic rocks are more wide-spread in the Croatian territory [46].

The third Middle Triassic dry-land phase is discovered in the "Dolomite" Complex of Bloke plateau on the boundary between the underlying light fenestral dolomite, to which we ascribed the Anisian age, and the overlying very light grey to almost white massive coarse-grained porous dolomite, that in spots lies under the medium-grey to rosy bedded biostromal *Diplopora Limestone,* which is attributed in Slovenia to Cordevolian and Lower Carnian. The evidence for the Middle Triassic phase is based on finding 0.5- to 1.25-thick lense-like interbeds of kaolinite, bauxite and iron crust on the boundary between the Anisian light bedded fenestral dolomite and the Carnian-Ladinian white coarse-grained massive dolomite.

3.1.6 Paleogeography

The relatively shallow sedimentary environment in which the Lower Permian clastic rocks were deposited, began at the end of the Lower Permian during Saalian movements to elevate so that the Kočevje and Gorski Kotar area was in the Upper Permian a dry land. In this area in that time no important folding occurred. During the slow uplift a relatively flat dry land originated, built

of clayey sediments with interbeds of sandstones and conglomerates. This dry land was a consequence of the activity of the last strokes of the Hercynian orogeny [27,28].

After the last strokes of this orogeny (Pfälzian phase) the transgression of the Scythian sea takes place. The Scythian sediments originated in a very shallow sea, in between the intertidal band with erosional activity of the sea currents and waves on half-consolidated sediments of the sea bottom, that were mixed with a fine-grained dolomitic matrix. The Scythian intraformational conglomerate speaks, above all, for a shallow sea that from time to time flew off and was followed by drying out (desiccation) and cracking of the half consolidated mud material, of which fragments were afterwards rounded by sea currents. The material supply from the land was pretty considerable.

Wave ripple marks and cross-stratification in the Scythian sediments point to a turbulent shallow sea. The dry land in the vicinity of the sedimentary area had to be low and flat, since from this dry land came only fine and very fine rock material.

The absence of coarse-grained and non-sorted rock material speaks for the conclusion, that in the vicinity of the considered area in the Middle Triassic there were no mountains or higher elevated dry land. After the deposition of the Scythian sediments, the Kočevje region was elevated and a gently convexed antiform originated, upon which there was no sedimentation until the transgression of the Carnian sea.

The similar Scythian sedimentation on the Slovenian Carbonate Platform also originated on the Bloke plateau and in its surroundings, with a distinction that it was more calcareous.

In the Anisian epoch the carbonate platform in southern Slovenia became larger and more stable. The rock material supply from the land was insignificant. Monotonous sedimentation of light dolomites point at a stable shelf and sedimentologic properties of the carbonate rocks indicate a neritic sedimentation in the tidal area. In the middle of Anisian epoch the intensified epeirogenic movements assisted the forming of short-lasting local dry land with traces of volcanism, bauxite and kaolinite were preserved on the surface.

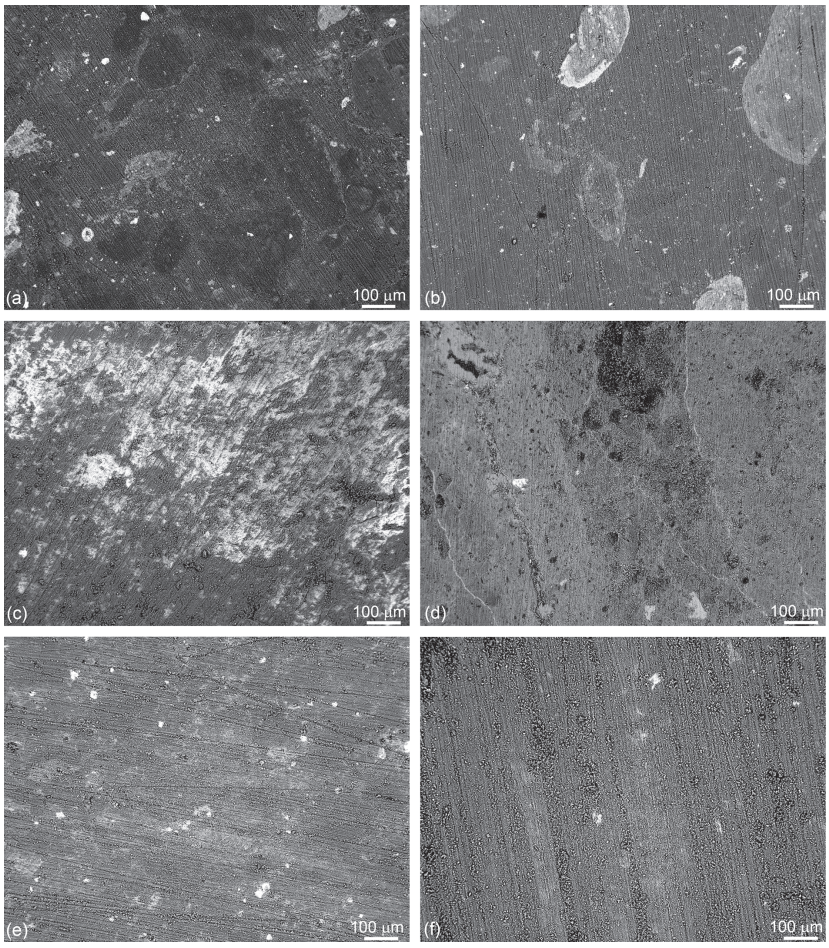


Figure 4: BE images of all six samples performed on an area 1.17 mm × 0.88 mm. The chemical inhomogeneity is shown by the lighter and darker areas in the micrographs

Slika 4: Slike, dobljene z odbitimi elektroni vseh šestih vzorcev, posnete na področju velikosti 1,17 mm × 0,88 mm. Kemijska nehomogenost se kaže v porazdelitvi temnejših in svetlejših področij na sliki

Greater activity of the Anisian volcanoes was recorded in the Croatian territory [46]. After the short-lasting dry land, that can be temporally correlated with the Monte Negro orogenic phase, repeatedly followed a neritic sedimentation on the carbonate platform, equal as before to the short-lasting interruption of sedimentation.

On the boundary between the Anisian and Ladinian stage followed a new, still more intense dry-land phase, which is evidenced by here and there preserved lenses of bauxite, kaolinite and iron crusts. In the whole area of southern Slovenia originated then on the Dinaric Platform carbonate sediments in a shallow sea. In the shallow and warm sea grew calcareous algae, forming numerous submarine lawns. With their coalcareous skeletons they contributed most to the origin of several hundreds of meters thick limestone sequence. Later the *Diplopora Limestone* changed at late diagenesis in part or completely into a coarse-grained (crystallized) dolomite [47].

The part of the "Dolomite " Complex, originated after the Anisian/Ladinian dry-land phase, the Slovenian geologists attributed to the Lower Carnian respectively to Cordevolian. In S. Buser's, and A. Ramovš's opinions [7] the algal species *Diplopora annulata* Schafhäutl occurring in southern Slovenia only in the Cordevolian beds, represents a guide fossil for this area. In our opinion, it was formed in the Langobardian and Cordevolian epoch.

The dry-land phases were detected and evidenced by the following signs and materials: erosion surfaces, barite-filled pockets, continental breccias and conglomerates, paleokarstic phenomena, red residual sediments, kaolinite deposits, as well as iron and kaolinite crusts.

Upon the size and thickness of the lense-like kaolinite deposits, that range normally from 4 to 9 meters and amount in spots even to 17.5 meters, influenced also the morphology of the solid rock. Namely, the kaolinite deposits were not formed or accumulated upon the elevated parts of the paleorelief. The paleorelief served only as a trap for the considered deposits [50]. The kaolinite deposits start in rare places with a thin layer of basal calcareous breccio-conglomerate with a kaolinite matrix.

The kaolinite deposits originated most probably from a pyroclastic material that was deposited into a marsh and there changed into the kaolinite [51–53].

3.2 Laboratory examinations

3.2.1 SEM/EDS analysis

Semi-quantitative EDS analyses of all six selected samples were obtained. BE images of the investigated samples are shown **Figure 4**. The rocks are not homogeneous and consist of areas with different chemical compositions. The higher intensity of the backscattered electrons represents a lighter colour in the BE image; therefore, it can be concluded that such areas consist of elements having higher atomic numbers. The overall elemental analyses of all six samples are shown in **Table 2**. All the samples have approximately mass fraction 60 % of oxygen, the exception being sample 4, where the content of oxygen is less than half, most probably due to the higher amount of iron. Silicon and aluminium are the next two most represented elements. In sample 1 the mass fraction of aluminium is almost 26 %, while in sample 2, 5 and 6 the content of aluminium is close to 17 % and in other samples it is of 8 %. The higher content of silicon is in sample 3 (27 %) followed by sample 4 and 5 (approximately 19 %) and there is near 10 % in samples 1 and 4. Iron gives the colour of the rocks; the higher content of iron and their oxides give a more reddish colour to the samples. The highest content of iron, of approximately 28 %, was found in sample 4, it was in sample 2 at 4.7 %, in samples 1 and 5 at 2.7 %, in sample 3 at 1.6 % and in sample 6 at 0.4 %. Some of the samples also contain a minor content of titanium, magnesium and some other elements presented in **Table 1**. The EDS spectra of all the examined samples are presented in **Figure 5**. In each sample a few spot

Table 2: Chemical composition of six samples determined with EDS on a field of view 1,17 mm × 0.88 mm; mass fractions *w*/%, mole fractions *x*/%

Table 2: Kemična sestava šestih vzorcev določena z EDS-analizo na vidnem polju 1,17 mm × 0,88 mm; masni deleži *w*/%, molski deleži *x*/%

		O	Mg	Al	Si	K	Ca	Ti	Fe
Sample 1	*w*/%	60.0	–	25.7	10.9	–	–	0.6	2.7
	x/%	72.8	–	18.5	7.5	–	–	0.3	0.9
Sample 2	*w*/%	59.5	–	17.1	18.6	–	–	–	4.7
	x/%	72.9	–	12.4	13.0	–	–	–	1.7
Sample 3	*w*/%	57.9	2.3	8.3	27.0	2.2	0.6	–	1.6
	x/%	71.2	1.9	6.0	18.9	1.1	0.3	–	0.6
Sample 4	*w*/%	51.9	–	8.4	9.4	–	–	2.0	28.3
	x/%	73.1	–	7.0	7.6	–	–	0.9	11.4
Sample 5	*w*/%	60.5	0.3	17.4	18.7	–	–	0.4	2.7
	x/%	73.3	0.2	12.5	12.9	–	–	0.2	0.9
Sample 6	*w*/%	61.2	–	17.9	20.0	–	–	0.5	0.4
	x/%	73.3	–	12.7	13.6	–	–	0.2	0.1

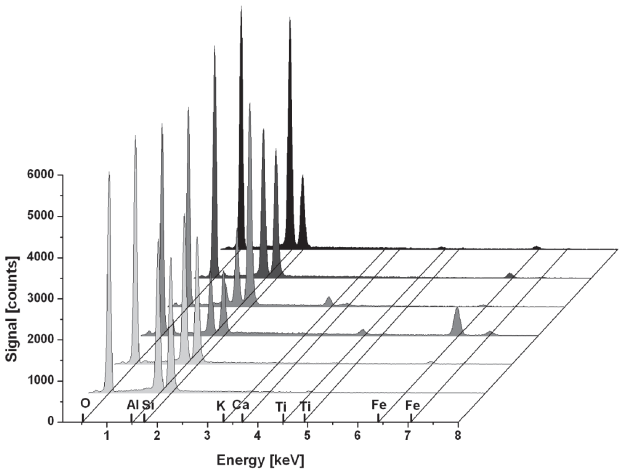


Figure 5: EDS spectra of all six samples taken for an area of 1.17 mm × 0.88 mm
Slika 5: EDS spektri šestih vzorcev posneti na področju 1,17 mm × 0,88 mm

analysis were performed, which clearly demonstrate the scattering of individual elements over the sample, especially iron, titanium and also silicon as well as aluminium, while oxygen is almost the same in all the examined samples. The results of the chemical composition of the individual small areas of sample 2 are shown in **Figure 6** and **Table 3**.

Table 3: Chemical composition of individual small areas of sample 2 determined with EDS; mass fractions *w*/%, mole fractions *x*/%
Table 3: Kemična sestava posameznih manjših področji vzorca 2 določena z EDS-analizo; masni deleži *w*/%, molski deleži *x*/%

		O	Al	Si	Ti	Fe
Spectrum 1	*w*/%	49.5	10.4	11.2	1.2	27.7
	x/%	70.4	8.8	9.0	0.5	11.3
Spectrum 2	wt.%	54.5	14.1	15.2	0.8	15.4
	x/%	71.5	11.0	11.4	0.3	5.8
Spectrum 3	*w*/%	60.8	17.6	18.7	0.2	2.7
	x/%	73.5	12.6	12.9	0.1	0.9
Spectrum 4	*w*/%	58.5	16.1	17.1	0.4	7.9
	x/%	73.1	7.0	7.6	0.9	11.4
Spectrum 5	*w*/%	54.5	13.4	14.4	0.7	17.0
	x/%	72.0	10.5	10.8	0.3	6.4

3.2.2 XRD analysis

Figure 7 to 12 shows the XRD spectra of six samples with labeled constituting minerals. The XRD results correspond very well with the EDS elemental analyses taking into account that there are some elements that cannot be analyzed, e.g., hydrogen, and that light elemental analysis is not very precise when using EDS. This analysis is difficult because of their low photon energy, what may lead to high absorption in the specimen and in the components of the spectrometer [48]. Sample 1 is dark-red oolitic pisolitic bauxite or boehmite bauxite with some minor minerals such as hematite, kaolinite, nantronite and berthierine (**Figure 7**). The

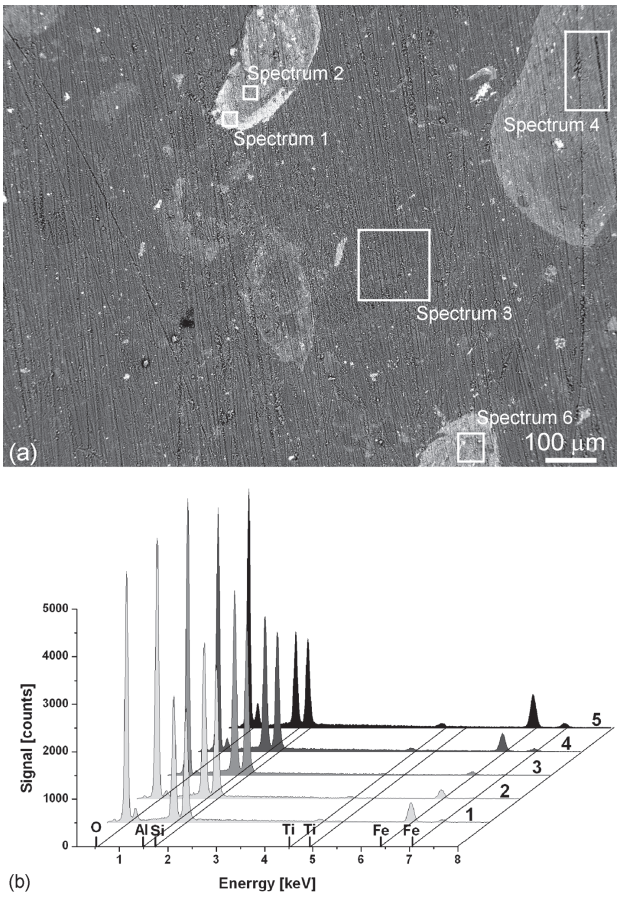


Figure 6: (**a**) BE image of sample 2 with marked areas of EDS analysis; (**b**) five EDS spectra performed on different areas
Slika 6: (**a**) Slika odbitih elektronov vzorca 2 z označenimi mesti EDS analize; (**b**) pet EDS spektrov, posnetih na različnih področjih

share of the boehmite phase is at 60 %, while the share of other minerals is approximately 10 % for each. Sample 2 is red pelitic bauxite or iron kaolinite (**Figure 8**). The ratio of the kaolinite mineral to the hematite is 8 to 2. Sample 3 is a quartz sandstone with a tuff admixture (**Figure 9**). The content of quartz is 80 %, while in the other two phases it is 10 % for each. Sample 4 is an iron-kaolinite-rich phase (**Figure 10**) where the

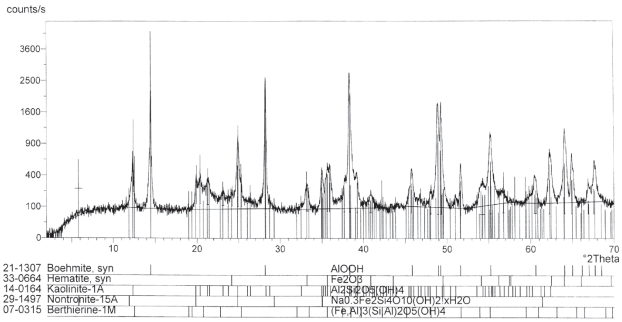


Figure 7: XRD-spectrum of sample 1 with labelled constituted minerals of analysed rock
Slika 7: XRD-spekter vzorca 1 z označenimi minerali, ki sestavljajo analizirane kamne

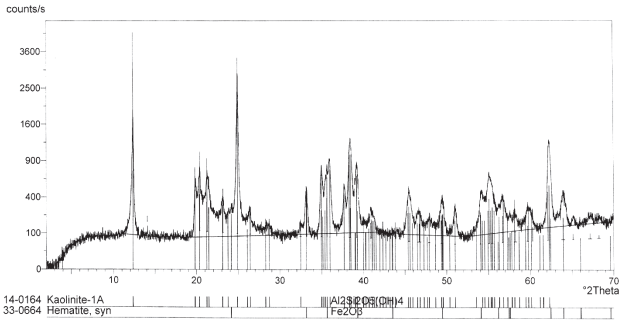


Figure 8: XRD-spectrum of sample 2 with labelled constituted minerals of analysed rock

Slika 8: XRD-spekter vzorca 2 z označenimi minerali, ki sestavljajo analizirane kamne

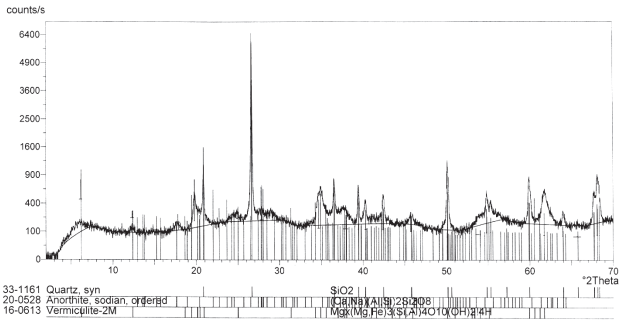


Figure 9: XRD-spectrum of sample 3 with labelled constituted minerals of analysed rock

Slika 9: XRD-spekter vzorca 3 z označenimi minerali, ki sestavljajo analizirane kamne

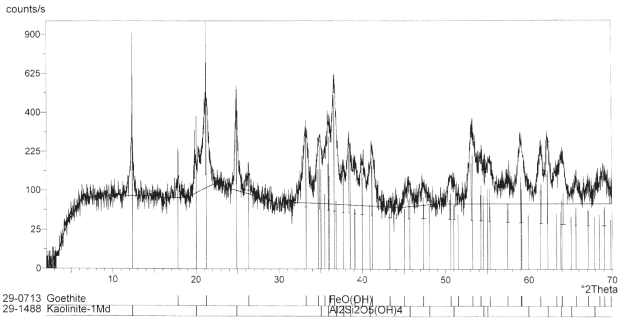


Figure 10: XRD-spectrum of sample 4 with labelled constituted minerals of analysed rock

Slika 10: XRD-spekter vzorca 4 z označenimi minerali, ki sestavljajo analizirane kamne

ratio of the goethite and kaolinite are 6 to 4. Sample 5 is kaolinite with the addition of iron in the goethite phase (**Figure 11**). The content of kaolinite in sample 5 is 70 %, while goethite and nontronite are represented each as 10 %, and gibbsite and groutite as 5 % each. Sample 6 is pure kaolinite (**Figure 12**).

4 CONCLUSIONS

In Southern Slovenia – which in the Scythian and Anisian epoch belonged to the Slovenian Carbonate

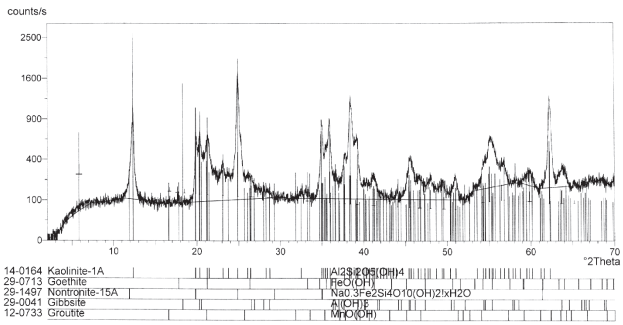


Figure 11: XRD-spectrum of sample 5 with labelled constituted minerals of analysed rock

Slika 11: XRD-spekter vzorca 5 z označenimi minerali, ki sestavljajo analizirane kamne

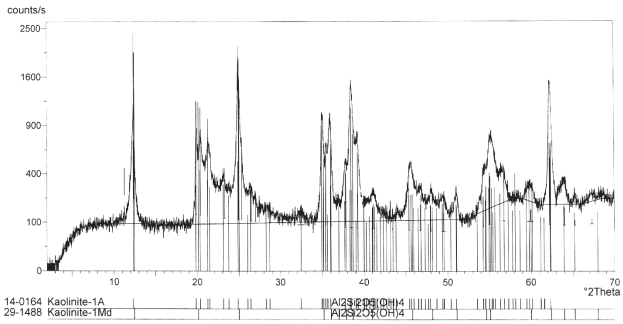


Figure 12: XRD-spectrum of sample 6 with labelled constituted minerals of analysed rock

Slika 12: XRD-spekter vzorca 6 z označenimi minerali, ki sestavljajo analizirane kamne

Platform including the whole Dinarides and from the Upper Cordevolian to the end of the Mesozoic period to the Dinaric Carbonate Platform – two Triassic sedimentary developments can be distinguished: the clastic and the carbonate. The clastic development occurs in the Kočevje and Bela krajina regions, while the carbonate appears in the Bloke plateau and in some other places in Slovenia. In the last three years we focused our research work on the Bloke plateau area, to the "Dolomite" Complex in particular.

With systematic regional and detailed geological and laboratory researches we came to the following conclusions:

1. The "Dolomite" Complex on the Bloke plateau is not a homogenous lithologic unit, but it consists of two, and in some places even of three or four, rock sequences.
2. In the Kočevje region, there was no sedimentation in the Anisian, Ladinian and Cordevolian epochs, since there are no Middle Triassic and Lower Carnian (Cordevolian) rocks exposed at the surface, and because these rocks do not crop out anywhere in the White Carniola and Gorski Kotar, as well.
3. The sedimentation in the "Dolomite" Complex of the Bloke plateau was not continuous; it was interrupted two times for a certain time: in the middle of the

Anisian epoch and on the boundary between the Anisian and Ladinian stage.

4. About S. Buser and A. Ramovš statement [7], i.e., that in southern Slovenia the dasycladal species *Diplopora annulata* Schafhäutl occurs only in the Cordevolian beds, we came to the conclusion that it was so only there, where the Ladinian is clasticly developed. In other places in Slovenia, however, we have also the Middle Triassic carbonate development ("Dolomite" Complex), where according to our data, the beds containing diploporas extend into the Ladinian stage as well. The considered dry-land phase started probably already in the topmost Anisian and lasted for some time into the Ladinian epoch (probably in the whole Fassanian). Afterwards, started the sedimentation of the *Diplopora Limestone* of the coarse-grained dolomites, that fill the Ladinian and Cordevolian interval of the Triassic lithostratigraphic column [15–17,49].
5. In the Anisian epoch volcanoes very probably also erupted.

5 REFERENCES

[1] D. Savić, S. Dozet, Tolmač za list Delnice (Explanation note of the Map Sheet Delnice), Osnovna geološka karta SFRJ, 1 : 100 000. – *Zvezni geološki zavod,* Beograd, (**1985**), 62

[2] J. Bukovac, M. Poljak, M. Šušnjar, M. Čakalo, Tumač za list Črnomelj (Explanation note of the Map Sheet Črnomelj), Osnovna geološka karta SFRJ 1 : 100 000. – *Zvezni geološki zavod*, Beograd (**1984**), 63

[3] S. Buser, Tolmač Lista Ribnica (Explanation note of the Map Sheet Ribnica). Osnovna geološka karta SFRJ, 1 : 100 000. – *Zvezni geološki zavod*, Beograd (**1974**), 60

[4] A. Melik, Posavska Slovenija (Posavje, Slovenia), II/3. – *Slovenska matica*, Ljubljana, (**1959**), 595

[5] S. Buser, Triassic beds in Slovenia. – *16th European micropaleont. colloq.*, Ljubljana (**1979**), 17–26

[6] S. Buser, Development of the Dinaric and the Julian Carbonate Platforms and of the intermediate basin (NW Jugoslavia). – *Mem. Soc. Geol. It.*, Trst 40 (**1989**), 313–320

[7] S. Buser, A. Ramovš, Razvoj triadnih skladov v slovenskih Zunanjih Dinaridih (Development of the Triassic beds in Slovenian Outer Dinarides). – *Prvi kolokvij o geologiji Dinaridov*, Ljubljana, 1 (**1968**), 33–42

[8] Ž. Đurđanović, *Geol. Vjesnik*, 20 (**1967**), 107–111

[9] S. Dozet, *Rud. met. zbornik;* 36 (**1989**), 663–673

[10] S. Dozet, *RMZ Materials and Geoenvironment*, 51 (**2004**), 2191–2208

[11] C. Germovšek, *Geologija*, 3 (**1955**), 116–135

[12] C. Germovšek, Razvoj mezozoika v Sloveniji (Die Entwicklung des Mesozoikums in Slowenien). – *Prvi jugosl. geol. kongr.*, Ljubljana, 1 (**1956**), 35–43

[13] M. Herak, B. Sokač, B. Šćavničar, Correlation of the Triassic in SW Lika, Paklenica and Gorski Kotar. – *Geol. Sbornik Slov. akad. Vied*, Bratislava, 18/2 (**1967**), 189–202

[14] M. Pleničar, Tolmač za list Postojna (Explanation note of the Map Sheet Postojna). Osnovna geološka karta SFRJ 1 : 100 000. – *Zvezni geološki zavod*, Beograd, (**1970**), 62

[15] U. Premru, Geološka zgradba južne Slovenije (Geological Structure of Southern Slovenia). – *Geologija*, 25 (**1982**) 1, 95–126

[16] U. Premru, Tektonika in tektogeneza Slovenije (Tectonics and Tectogenesis). – *Geološki zavod Slovenije*, Ljubljana, (**2005**), 518

[17] U. Premru, B. Ogorelec, L. Šribal, O geološki zgradbi Dolenjske (On Geological Structure of Dolenjska). – *Geologija*, 20 (**1977**), 167–192

[18] I. Rakovec, Triadni vulkanizem na Slovenskem (Triassic volcanism in Slovenia). – *Geografski vestnik*, 18 (**1946**), 129–171

[19] I. Rakovec, Pregled tektonske zgradbe Slovenije (A Survey of the Tectonic Structure of Slovenia). – *Prvi jugosl. geol. kongr.*, Ljubljana 1 (**1956**), 73–83

[20] A. Ramovš, Razvoj paleozoika na slovenskem (The Stratigraphic Development of Palaezoic in Slovenia). – *Prvi jugosl. geol. Kongr.*, Ljubljana, (**1956**), 149–151

[21] A. Ramovš, Biostratigrafske značilnosti triasa v Sloveniji (Biostratigraphishe Characteristik der trias in Slowenien). – *Geologija*, 16 (**1973**), 379–388

[22] B. Sokač, I. Velić, Triassic, Jurassic and Cretaceous of the Karst part of the Dinarides in the western Croatia. – *16th European micropaleont. Colloq.*, Ljubljana, (**1979**), 79–100

[23] R. I. Folk, Practical petrographic classification of limestones. – *Amer. Ass. Petrol. Geol. Bul.,l* Tulsa, (**1959**), 1–38

[24] R. J. Dunham, Classification of carbonate rocks according to depositional texture, In W.E. Ham. (ed): Classification of carbonate rocks. – *AAPG Memoir*, Tulsa, **1962**, 108–121

[25] F. H. Pettijohn, Sedimentary rocks. – *Harper and Row*, New York (**1975**), 628

[26] S. Dozet, M. Silvester, *Rud. met. Zbornik*, 31 (**1984**) 1, 5–19

[27] D. Savić, S. Dozet, M. Sarkotić, Odnos permskih i gornjetriaskih naslaga na području Gorskog Kotara (Relating between permian and Upper Triassic beds in the Gorski Kotar area). – *10 Kongr. geol.*, Budva, *SFRJ;* 1 (**1982**), 653–673

[28] D. Savić, S. Dozet, Tumač za list Delnice (Explanation note of the Map Sheet Delnice). Osnovna geološka karta SFRJ 1 : 100 000. – *Zvezni geol. Zavod*, Beograd, (**1985**), 66

[29] S. Dozet, Tolmač lista Delnice. Osnovna geološka karta SFRJ 1 : 100 000. – *Geološki zavod Slovenije*, (**1983**), 109

[30] S. Dozet, M. Silvester, Skitske in zgornjekarnijske kamenine na Kočevskem (Scythian and Upper Carnian rocks from the Kočevje region). – *Geologija*, 22 (**1979**) 2, 327–336

[31] S. Dozet, *Rud. met. Zbornik*, 37 (**1990**) 3, 391–408

[32] A. Ramovš, Tektonische Bewegungen in der Trias Sloweniens (NW Yugoslowien). – *Prvi simp. orog. faz. alp. Evrope* (*Beograd–Bor*), 1 (**1970**), 21–34

[33] B. Šćavničar, A. Šušnjara, *Geol. vjesnik*, 20 (**1967**), 82–106

[34] M. Babić, *Geol. Vjesnik*, 21 (**1968**), 10–18

[35] B. Šćavničar, Klastiti trijasa u Gorskom Kotaru (Clastic Sediments of the Triassic in the Gorski Kotar Region) – *Acta geologica Jug. ak. znan. umjet.*, 7 (**1973**) 3, 105–160

[36] S. Dozet, *Rud. met. Zbornik*, 32 (**1985**) 1–2, 27–49

[37] S. Dozet, *Geologija*, 42 (**2000**), 41–68

[38] B. Sokač, Paläostrukturen der Trias in dem Gebiete des Gorski Kotar und des Velebitgebirges. – *Bull. Sci. Cons. Acad. Yugosl.*, Zagreb, 14/5–6 (**1969**), 142–143

[39] S. Buser, Osnovna geološka karta SFRJ, list Ribnica 1:100 000 (Basic Geological Map of SFRY, Map Sheet Ribnica 1 : 100 000). – *Zvezni geološki zavod*, Beograd (**1969**)

[40] V. Jacobshagen, *N. Jb. Geol. Paläont.*, Stuttgart, 9 (**1961**), 477–483

[41] K. Grad, L. Ferjančič, Tolmač za list Kranj (Explanation note of the Map Sheet Kranj). Osnovna geološka karta SFRJ 1 : 100 000. – *Zvezni geološki zavod*, Beograd, (**1976**), 70

[42] M. Pleničar, U. Premru, Tolmač za list Novo Mesto (Explanation note of the Map Sheet Novo mesto). Osnovna geološka karta SFRJ 1 : 100 000. – *Zvezni geološki zavod*, Beograd, (**1977**), 61 pp

[43] K. Sikić, O. Basch, A. Šimunić, Tumač za list Zagreb (Explanation note of the Map Sheet Zagreb). Osnovna geološka karta SFRJ 1 : 100 000. – *Zvezni geološki zavod*, Beograd, (**1979**), 81

[44] U. Premru, *Geologija*, 17 (**1974**), 262–274

[45] S. Dozet, T. Kolar-Jurkovšek, *RMZ Materials and Geoenvironment*, 54 (**2007**) 3, 361–386

[46] A. Šimunić, Al. Šimunić, Triassic deposits of Hrvatsko Zagorje. – *Geol. Croatica*, Zagreb, 50/2 (**1997**)

[47] A. Ramovš, *Geologija*, 13 (**1970**), 159–173

[48] J. Goldstein, D. Newbury, D. Joy, C. Lyman, P. Pechlin, E. Lifshin, L. Sawyer, J. Michael, Scanning Electron Microscopy and X-Ray Microanalysis, third ed., Kluwer Academic, 2003, Chapter 10, Page 499

[49] B. Celarc – *Geologija*, 47 (**2004**) 2, 139–149

[50] S. Dozet, M. Godec, Mater. Tehnol., 43, 2 (**2009**), 97–102

[51] D. Stoffler, Neure Erkenntnisse in der Tonsteinfrage auf Grund sedimentpetrographischer und geoschemischer Untersuchungen in Flöz Wahlschied der Grube Ensdorf (Saar). – *Beitr. Miner. u. Petr.*, Göttingen, Heidelberg; 9 (**1963**)

[52] H. Füchtbauer, G. Müller, Sedimente und Sedimentgesteine. Sediment-Petrologie, Teil 2. *Schwatzbart'sche Verl.*, Stuttgart, (**1970**),p 784

[53] M. Drofenik, J. Čar, D. Strmole, Geologija, 18 (**1975**), 107–155

UDK 620.178.3:539.42
Original scientific article/Izvirni znanstveni članek

ISSN 1580-2949
MTAEC9, 44(4)185(2010)

ENHANCED FATIGUE ANALYSIS – INCORPORATING DOWNSTREAM MANUFACTURING PROCESSES

IZBOLJŠANA UTRUJENOSTNA ANALIZA – VKLJUČITEV IZDELAVNIH PROCESOV

Wilfried Eichlseder

University of Leobenm, Franz-Josef-strasse 18, A-8700 Leoben, Austria
wilfried.eichlseder@unileoben.ac.at

Prejem rokopisa – received: 2009-08-19; sprejem za objavo – accepted for publication: 2010-02-17

Most failures of engineering components can be traced back to cyclic loads which cause fatigue of the material. These cyclic loads can be either constant or variable. This influence has to be taken into account in the fatigue analysis, the aim of which is to assure a minimum required lifetime of a component without breakdown. Simultaneously, a lightweight design is intended by reducing the size of the component. The strength of materials under cyclic loading is essentially lower than those under static loading. However, the cyclic strength is of more importance in practical applications, as more components fail by cyclic loading than by static loads. Regrettably, the strength of cyclic loaded components was discovered relatively late and the first systematic investigations were performed in the 19th century. Cyclic loadings result from mechanic or thermal service loads, which can be superimposed to static loads. When looking at a truck, the chassis is loaded by the dead load of the vehicle. Due to loading and unloading procedures the load is changed into a quasi-static way. Additional cyclic loads emerge during driving operation in addition to the static loads due to braking, acceleration, cornering, or physical conditions just like bumps on the roadway or air drag. Additional loading is also due to resonance vibrations of engine, gearbox, fuel tank, or the spare wheel. These cyclic loads lead to fatigue of material hence overstraining them to a crack. These cracks propagate and finally lead to failure of the component.

Key words: fatigue life, complex structures, simulation, manufacturing processs, finite elements

Izračun trajnostne dobe za geomertično kompleksne strukture zahteva poznanje lokalne utrujenosti materiala. Po navadi raziskava lokalnega utrujenostnega vedenja komponent ni mogoča, ker je treba za študij vplivnih parametrov preveliko časa. Zato je potrebna metoda za izračun trajnostne dobe z upoštevanjem lokalnih vplivov. Lokalna utrujenostna trajnostna doba je odvisna od zarez, vrste obremenitve, temperature, višine in sekvence obremenitve itd., močan je tudi vpliv procesa izdelave. V tem članku so obravnavani procesi izdelave in njihov vpliv na mikrostrukturo in na utrujenostno trajnostno dobo. Posebej so raziskani vplivi litja, oblikovanja, varjenja in površinske obdelave. Različni procesi so bili simulirani, da bi bil dosežen lokalni učinek, kot so čas strjevanja, hitrost deformacije in rezidualne napetosti. Pripravljeni so bili preizkušanci z opredeljenimi procesnimi parametri, njihova utrujenostna trajnostna doba pa je bila določena eksperimentalno. Ti rezultati so bili podlaga za napoved utrujenostne trajnostne dobe geometrično kompleksnih oblik na podlagi rezultatov analize s končnimi elementi, ki je upoštevala lokalne napetosti, gradiente napetosti in proces izdelave. Simulacija procesa izdelave in razumevanje vpliva lokalnih razmer na utrujenostno trajnostno dobo omogoča dobro simulacijo zapiranja verige od procesa izdelave do napovedi te dobe.

Ključne besede: utrujenostna trajnostna doba, kompleksne strukture, simulacija, procesi izdelave, končni elementi

1 FATIGUE ANALYSIS

The science of fatigue analysis deals with the fatigue of components under cyclic loads as in their regular service life. The main goal is to obtain spatial dimensions of the components in a way that they will serve the minimum lifetime required without failure, with the focus also lying on weight reduction by using reduced amount of material. Basically dimensioning of components can be done in two ways.

Dimensioning by testing

In this case the expected loads are applied on prototypes. When the required fatigue life is not achieved, a modified prototype is built and tested again. This development cycle continues until the required specifications are reached. The weak points in components or complete construction could be easily detected as against oversized regions with no visible damage.

Dimensioning by simulation

Based on stresses, knowledge of material behaviour and the loading spectrum, a lifetime calculation is performed. At spots that do not achieve the required lifetime, the geometry or the material is changed or reinforced. Whereas less stressed spots provide scope for a reduction in weight by means of material reduction.

The latter has the advantage that it can be applied very early in the engineering phase and is also less time and cost intensive. However, in contrast to testing, the accuracy of simulation is lesser. Hence, in the initial phase of mass product development, simulations are performed and in the final stages, developed components are tested to evaluate their fatigue resistance.

1.1 Damage accumulation

In real life applications, dynamic loads have variable amplitudes instead of constant amplitudes that too with different frequencies and stochastic sequences. The dam-

age under cyclic load is based on the formation and motion of crystalline defects in the material which has not yet been measured in a quantitative way. Hence, the knowledge of an engineer depends on empirical models. In the last few decades numerous hypothesis have been established. One of the oldest with distinction from others and also most often used is that of Palmgren and Miner. This model, like the other models, concentrates on describing the two following observations:

a) There exists an additional damage D_i due to a number of load cycles n_i with a stress amplitude of σ_i, which is caused by the amplitude itself and the foregoing loadcycles and loads.
b) The accumulation of the single damage leads to a global damage D, at a critical damage parameter D_c failure occurs.

1.1.1 Palmgren Miner Method

The most famous rule to define the damage accumulation was already described by Palmgren in 1924 and Miner in 1945. Thereby, the service life at single constant amplitude is considered as the base for calculating the service life at variable amplitudes. The damage D_1 due to on single load cycle is given by,

$$D_1 = \frac{1}{N_i} \tag{1}$$

where N_i is the number of load cycles to fracture of the S-N curve at a load level σ_{ai}.

When, at the load level σ_{ai}, the load cycles appear with a frequency n_i, the damage D_i at that level is given by:

$$D_1 = \frac{n_i}{N_i} \tag{2}$$

By summing up the damage at each load cycle for a specific load spectrum, one gets the total damage

$$D = \sum_{i=1} \frac{n_i}{N_i} \tag{3}$$

wherein i is the spectrum level, n_i the number of load cycles on the i^{th} level and σ_{ai} the corresponding load amplitude. Experiments show that components fail below or above the damage sum of 1 whereas, components are to fail, as per definition, when the damage sum reaches a critical value of 1:

$$D = 1{,}0 \tag{4}$$

2 EXTENDING THE SIMULATION CHAIN – INTEGRATION OF THE MANUFACTURING PROCESS

For lifetime evaluation of complex geometric components, concepts based on local stresses or strains are required (**Figure 1**). For investigating local stresses, finite element method (FEM) has been well established. For

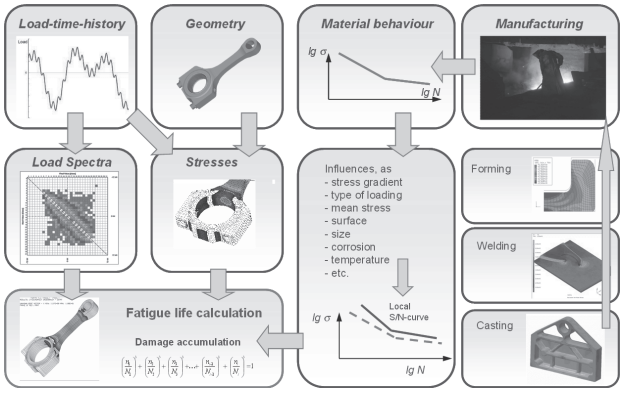


Figure 1: Lifetime prediction based on local stresses and strains
Slika 1: Napoved trajnostne dobe na podlagi lokalnih napetosti in deformacij

lifetime calculation, knowledge of local fatigue strength of material, described by the S-N curves of component and, the service conditions are needed, but these are not identical with the data of fatigue tests performed on idealised specimens. In an ideal case the strength of a component is determined by tests and can be used for lifetime calculation. Since no components exist in the construction phase, S-N curves are obtained by simulation.

These influences can either lead to an increase or decrease in stiffness and when occurring synchronous, these effects may boost or weaken each other. The experimental investigation of these effects on fatigue strength at a full scale is quite time consuming and also cost intensive whereby making it impractical. Hence, these tests are performed very selectively. The fatigue strength characteristics are not uniform all over the component and the S-N curve of the whole component is obtained by encompassing all the local S-N curves. The task is to transfer the S-N curve of the specimen with respect to the influencing factors into an S-N curve of the component.

In the following consideration, a few models for describing the influence of the production process are explained. These models facilitate lifetime calculation based on the results of finite element calculations.

2.1 Notches and load type

Notches, zones of bending stress, or components exposed to irregular force flow show irregular characteristics of the stress level. These irregularities can be derived from the stresses which gives the stress gradient χ or with reference to local stress, the relative stress gradient χ^* (**Figure 2**):

$$\chi^* = \frac{1}{\sigma_{max}} \left(\frac{d\sigma}{dx} \right) \tag{5}$$

As the stress gradient can be calculated easily with the help of finite element results, it is obvious to use it for the assessment of stresses and to describe the influences on the S-N curve in the case of notches as well as under tension/compression and bending loading.

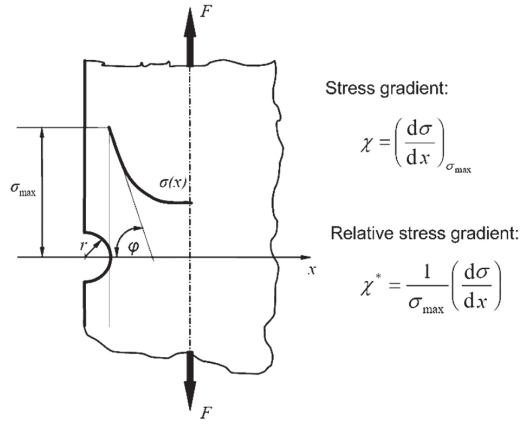


Figure 2: Stress gradient in notch root
Slika 2: Gradient napetosti na dnu zareze

2.1.1 Relative Stress Gradient Concept (RSG)

Notches are characterized by an increase in stress and formation of stress gradient within the component. Further, beams subjected to bending loading show a stress gradient. In both cases the material exhibits a locally higher fatigue stress limit than those under pure tensile loading on un-notched specimens which is a result of the support effect (**Figure 3**).

Two different fatigue limit parameters form the base for describing the influence of the stress gradient on the fatigue limit of any component:

– The fatigue limit σ_{zdw} of an un-notched specimen exposed to tension-compression load with the relative stress gradient $\chi^* = 0$ and
– The fatigue limit σ_{bw} of a bending specimen with thickness b and a relative stress gradient $\chi^* = 2/b$.

To describe the fatigue limit of components with arbitrary stress gradients one has to interpolate or extrapolate these values. Experience shows that the relationship between fatigue limit and stress gradient is not linearly proportional, with increasing stress gradient the growth of the fatigue limit is reduced (**Figure 4**).

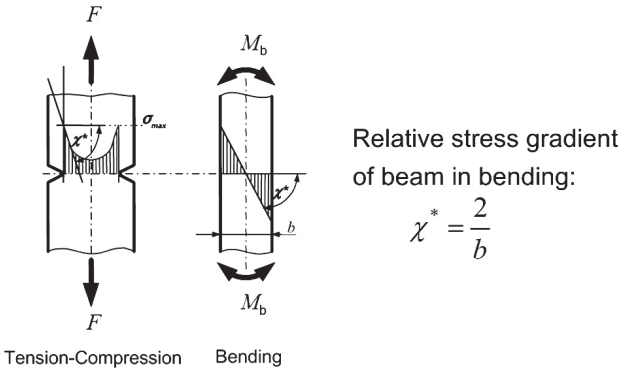
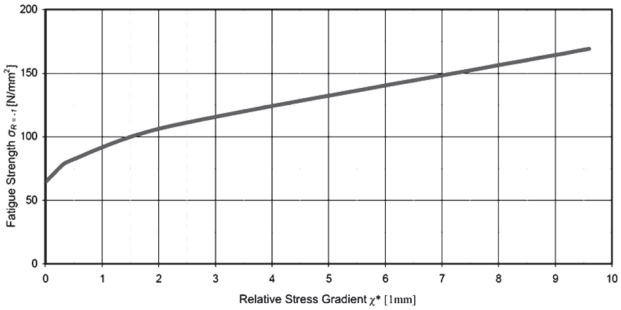


Figure 3: Stress gradient in notch for tension-compression and bending
Slika 3: Gradient napetosti za razteg-tlak in upogib



Figire 4: Fatigue limit at 10^7 load cycles depending on stress gradient
Slika 4: Meja utrujenosti pri 10^7 obremenilnih ciklih v odvisnosti od gradienta napetosti

To describe this relationship, an exponential curve is chosen, which is characterised by the exponent K_D:

$$\sigma_D = \sigma_{zdw^*}\left(1+\left(\frac{\sigma_{bw}}{\sigma_{zdw}}-1\right)\cdot\left(\frac{\chi^*}{2/b}\right)^{K_D}\right) \quad (6)$$

$$\sigma_D = \sigma_{zdw}\cdot n_\chi \quad (7)$$

For the fatigue life calculation with the help of S-N-curve according to (8) two additional parameters are necessary; the number of cycles at the fatigue limit N_D and the slope k of the S-N-curve.

$$N = N_D\cdot\left(\frac{\sigma_a}{\sigma_D}\right) \text{for } \sigma_a \geq \sigma_D \quad (8)$$

Generally speaking, with an increase in the sharpness of a notch the slope k gets steeper, while the number of cycles at the fatigue limit N_D decreases. To describe the trend of N_D and k of the S-N-curves, the corresponding minimum and maximum values at low and high gradient (χ^*) from test results are defined.

For the description of these values between these limits, exponential functions are chosen:

The values lg ($N_{D\min}$) and lg ($N_{D\max}$) describe the lowest and highest number of cycles at fatigue limit, while k_{min} and k_{max} refer to the minimum and maximum slope of the S-N-curve versus the stress gradient. The dependence of N_D and k as a function of n_χ is described by the exponents K_n and K_k, respectively.

$$\lg(N_D) = \lg(N_{D\min}) + \frac{\lg(N_{D\max}) - \lg(N_{D\min})}{n_\chi^{K_n}} \quad (9)$$

$$k = K_{min} + \frac{K_{max} - K_{min}}{n_\chi^{K_n}} \quad (10)$$

The equations (6) to (10) define the S-N curve for a component with notches and irregular stress distribution and can further be used for lifetime calculation of complex geometric components. This allows for an efficient analysis of S-N curves and an estimation of lifetime of finite element structures with more than a hundred thousand degrees of freedom in every node of the structure.

2.2 Statistical size effect

The statistical size effect supposes that the flaws in a component of large volume are more than those in a smaller component owing to to statistical reasons. Hence, bigger components show a higher probability for the occurance of critical local stress due to a flaw. This could probably lead to a crack which would grow and eventually lead to the failure of the component. For evaluation of the statistical size effect the stress integral for the high loaded volume (11) or the high loaded area (12) was proposed:

$$V_\sigma = \int_V \left(\frac{\sigma}{\sigma_{\max}} \right)^m \mathrm{d}V \tag{11}$$

$$A_\sigma = \int_A \left(\frac{\sigma}{\sigma_{\max}} \right)^m \mathrm{d}A \tag{12}$$

$\sigma_{\max}$ represents the notch root stress, σ is the stress of a volume or area element, m represents the Weibull exponent which characterizes the homogeneity of the material. For easier numerical calculation of highest loaded volume it is proposed in [3] that the volume $V_{90\,\%}$ which is loaded by at least 90 % of the maximum stress should be taken into account for evaluation. For calculation of a correction factor for fatigue strength the ratio between highest loaded volume of a reference specimen V_{ref} to the highest loaded volume V of the component is built.

$$n_{\mathrm{st}} = \left(\frac{V_{\mathrm{ref}}}{V} \right)^{\frac{1}{m}} \tag{13}$$

$$n_{\mathrm{st}} = \left(\frac{A_{\mathrm{ref}}}{A} \right)^{\frac{1}{m}} \tag{14}$$

2.3 Technological influences on forged components

Forging at different tool speeds primarily influences the microstructural properties of the material. In addition, a change in mechanical properties of the component like fatigue strength behaviour is also noticed, which is caused by a difference in grain size or grain size distribution.

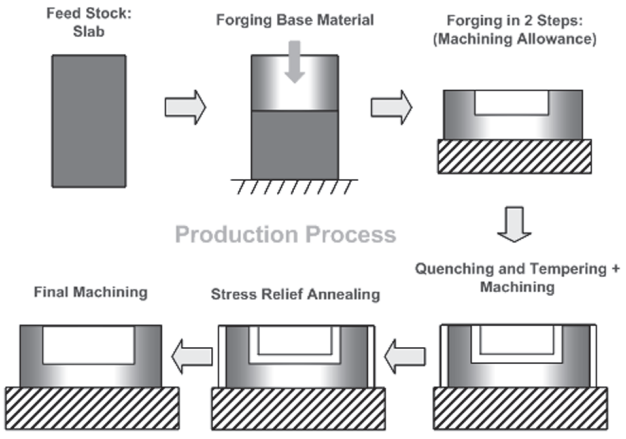


Figure 5: Principle chain of production process
Slika 5: Veriga postopkov procesa izdelave

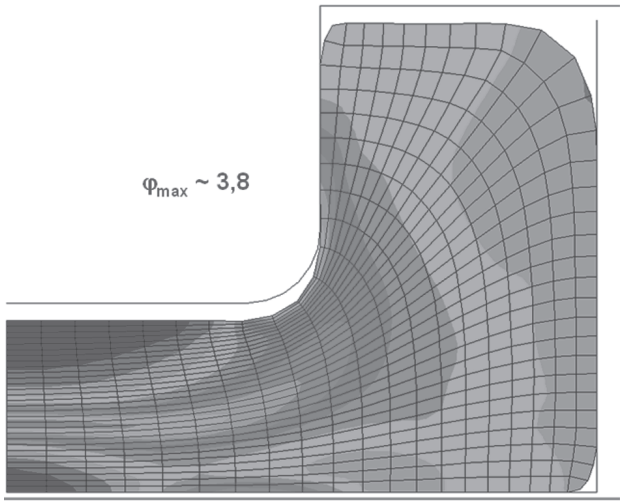


Figure 6: Simulation of degree of deformation
Slika 6: Simulacija stopnje deformacije

The influences on the fatigue limit are not uniform over the complete structure but dependent on local parameters like:

- Local degree of deformation
- Local deformation rate
- Forging temperature
- Grain Size (mean value, variation)
- Inclusions in the microstructure
- Segregation direction

In **Figure 5** the basic principle of a forging process is shown, while **Figure 6** depicts the distribution of degree of deformation of a component calculated by simulation. [4].

The S-N curve for influence of degree of deformation on fatigue strength of 16MnCr4 is depicted in **Figure 7**.

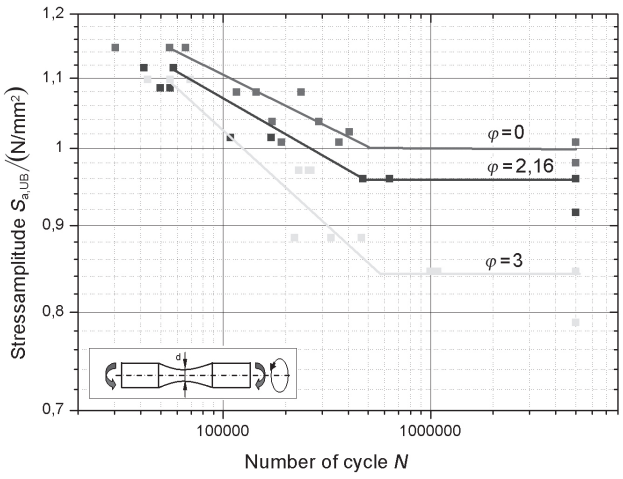


Figure 7: Influence of degree of deformation on performance of fatigue limit
Slika 7: Vpliv stopnje deformacije na velikost utrujenostne trdnosti

Rotating bending fatigue tests have been performed on un-notched specimens with local degrees of deformation $\varphi = 0$, $\varphi = 2.16$ and $\varphi = 3$. The fatigue limit shows a clear dependence on local degree of deformation. [4].

The number of cycles at fatigue limit nearly stays constant while the slope k of the S-N curve steepens with lowering fatigue strength. The reason for the decrease of strength is explained by metallographic analysis. Forged specimens show greater variation of grain size. This is ascribed to grain growth during the heating phase prior to the forging process. Hence, the influence of a chosen forging temperature on fatigue strength is a result of change in grain size and grain size distribution.

2.4 Influence of heat treatment on fatigue strength

Simulation of the cooling during the annealing process of a component is a big benefit for investigating the influence of heat treatment on fatigue strength. The temperature distribution during cooling calculated by simulation is shown in **Figure 8**. When combining local

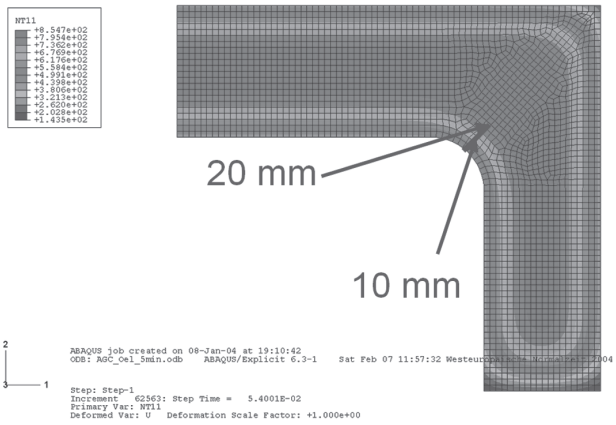


Figure 8: FE Simulation of annealing
Slika 8: Simulacija žarjenja s končnimi elementi

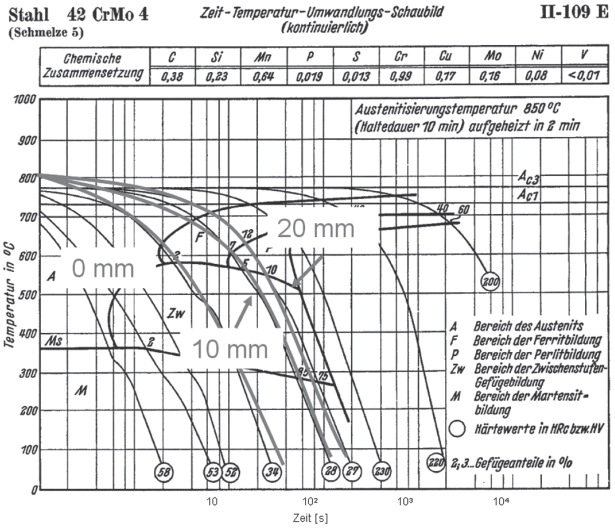


Figure 9: Continuous TTT curve of 42CrMo4
Slika 9: Kontinuirni TTT-diagram za 42CrMo4

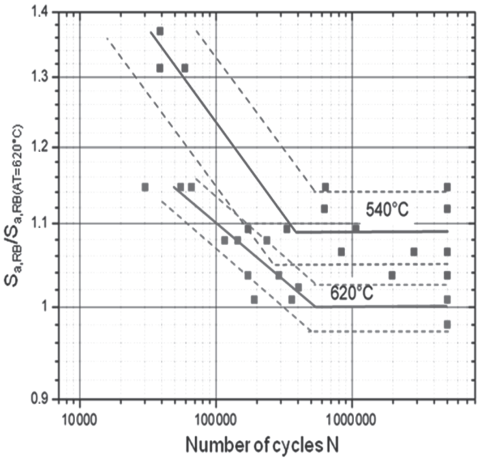


Figure 10: Influence of tempering temperature on fatigue strength
Slika 10: Vpliv temperature popuščanja na trdnost pri utrujenosti

temperatures in dependence of time with continuous TTT curves, the composition of local microstructure can be estimated and hence the local strength calculated [4] (**Figure 9**).

In the example shown below, the microstructure of the heat treated component shows a higher amount of martensite on the surface than 10 mm below the surface, which corresponds to the machining allowance.

Hence, on the surface of the machined component a considerably different microstructure with different strength characteristics exists. The influence of microstructural difference on the fatigue strength is shown in the S-N curve in **Figure 10**. For the investigated material the fatigue strength shows an increase of 7 % when tempered at 540 °C instead of 620 °C.

2.5 Technological influence on cast aluminum components

The technological influences, like casting process and refinement of microstructure, have an important effect on the fatigue behaviour of cast aluminum alloys. Furthermore, the local dendrite arm spacing (DAS) and, the distribution and size of pores can change the lifetime

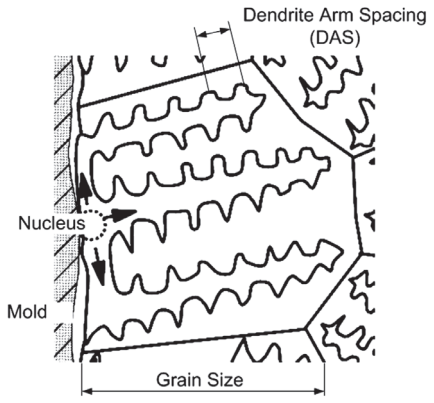


Figure 11: Definition of DAS [5]
Slika 11: Definicija DAS [5]

behaviour of the components. DAS and porosity are mainly affected by the freezing rate, which depends on the casting process, wall thickness, etc. Additionally, oxides and impurities affect fatigue strength.

Generally a dendrite grows from a single nucleus, which could be foreign particles or fragments of other grains as shown in **Figure 11**. The dendrite arm spacing is only affected by the local freezing rate or solidification time, whereas the grain size especially depends on the number of nuclei in the melt. Therefore a correlation between grain size and freezing rate cannot be shown.

The grain of a cast material can be built up by different dendrites of same origin. While grain size depends on number of apparent nuclei, DAS is only affected by local freezing rate and respective solidification time. In this way no correlation can be found between grain size and solidification time.

2.5.1 Die casting

For dimensioning dynamically loaded components with regard to fatigue lifetime, the knowledge of the local S-N-curve is necessary. These local S-N-curves, determined by the material, are essentially influenced by component specific effects, such as type of loading, size, stress gradient, temperature and the production process. In the case of cast components the technological influences can be described by DAS. DAS is not constantly spread within the component. Its distribution depends on the conditions under which the local solidification occurs. As models for the simulation of solidification are known and available local DAS can be determined in a very early stage of the development process. Based on the S-N-curves obtained by simulation and the local stress concept, which takes into account the local stress gradient, fatigue lifetime of geometrically complex, cast components can be calculated. [1]

For the die casting process DAS is a good parameter for characterizing the microstructure with respect to fatigue strength. Beside DAS, size and distribution of porosity also influence fatigue strength of cast aluminium alloys. Simulation of the freezing process allows the calculation of DAS with respect to local freezing rate. The S-N curves in **Figure 12** show the strong influence of DAS on fatigue strength of cast aluminium alloys. Specimen with DAS of 30 µm show a higher fatigue strength than the ones with 70 µm, additionally the slope of the S-N curve increases with greater DAS, and the number of cycles at fatigue limit decreases [6].

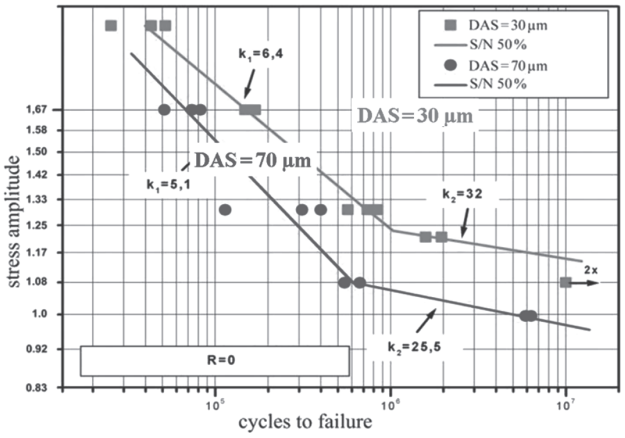


Figure 12: S-N curves for specimens with different DAS [6]
Slika12: S-N-krivulje za različen DAS[6]

2.5.2 High-Pressure die casting

While for die casting DAS is very important for fatigue strength, gas porosity plays a main role for high pressure die casting due to turbulent flow conditions which can't be avoided completely.

In [7] geometry and distribution of pores has been investigated using computer tomography (CT). CT analy-

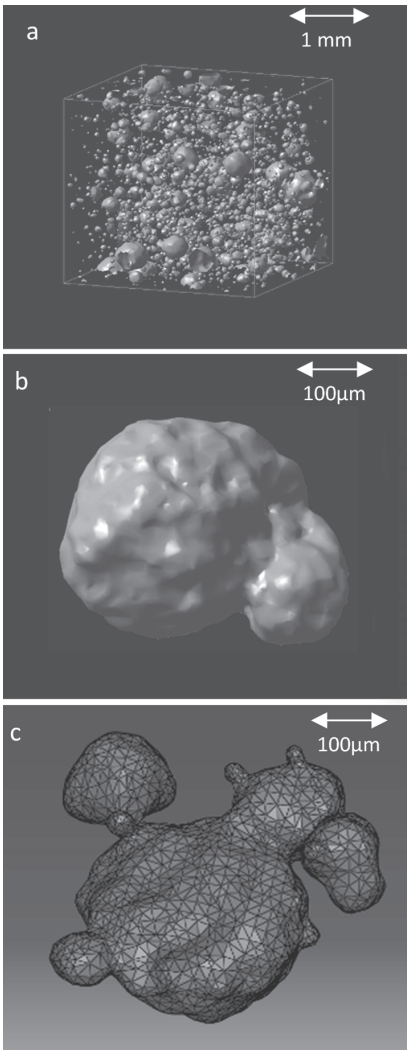


Figure 13: Results of computer tomography: a – Overview of tomographed cube; b – pore pressed by dendrites; c – meshed pore.
Slika 13: Rezultati računalniške tomografije: a – pogled na tomografirano kocko, b – pore, ki jih stiskajo dendriti, c – zamrežene pore

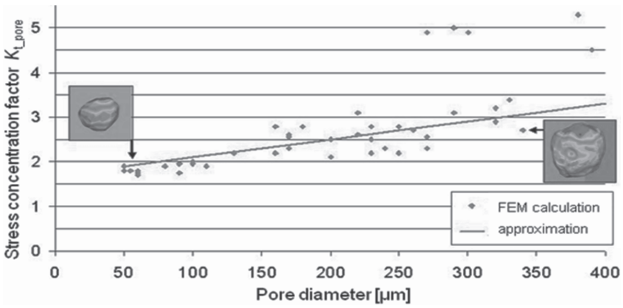


Figure 14: FEM calculation for tomographed spherical gas pores with varying diameters [7]
Slika 14: FEM-izračun za tomografirane sferične plinske pore z različnim premerom[7]

sis shows that the shape of gas pores obviously differs from a sphere as shown in **Figure 13**.

By finite element calculations, the impact of gas pores on the stress state in specimen has been examined. Local stress elevations which result out of irregular shape and notches show values in a range from 1,8 to 2,9. The larger the size of a pore, the stronger are the dendrites pressed into the surface, thus resulting in sharper notches.

Figure 14 shows the correlation between local stress elevation of approximately spherical gas pores and the pore diameter for tomographed pores with a range in diameter from 50 µm to 350 µm.

2.5.3 Calculation of pore distribution in high pressure die casting components

In [8] for calculation of pore distribution in a high pressure die casting component the use of the so called statistical porosity model is suggested. For a defined temperature range, post pressure and hold pressure, the corresponding Weibull distribution of porosity in the component is calculated by (15)

$$k_{\text{Weibull}} = (c_1 p_{\text{solid}}^2 + c_2 p_{\text{solid}} - c_3) T_{\text{rel}} + (c_4 p_{\text{solid}}^2 - c_5 p_{\text{solid}} + c_6)$$
$$d_{\text{Weibull}} = (c_7 p_{\text{solid}} - c_8) T_{\text{rel}} + (c_9 p_{\text{solid}} - c_{10}) \quad (15)$$

Exact position of a pore in a component is not predictable, but distribution of porosity within a defined domain is. The porosity distribution provides information regarding the percentage of elements within a domain, defined by the temperature distribution during the freezing process.

Pore diameter is an essential input in the fracture mechanics material model described in (8) in order to calculate the pore size dependence of fatigue strength. Hence, with the help of (16) the local porosity in the component has to be converted to a single pore of an equivalent diameter.

$$d_{\text{equ}} = \sqrt{\frac{4 A_{\text{element}} \text{Porosity}[\%]/100}{\pi}} \quad (16)$$

wherein, A_{element} corresponds to the area of the element, which originally was used for porosity calculation.

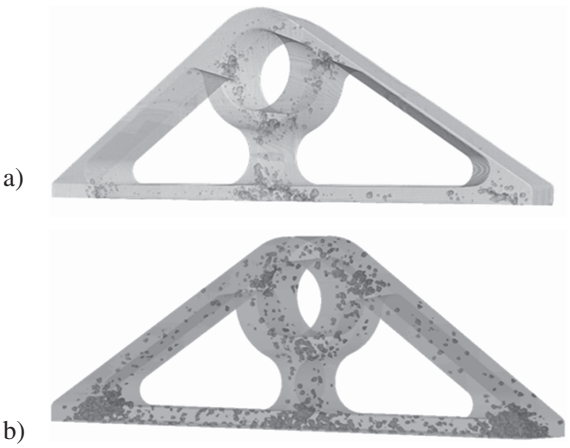


Figure 15: Pore distribution in real component (computertomography) (a), calculated pore distribution (b)
Slika 15: Porazdelitev por v realnem elementu (računalniška tomografija) (a), izračunana porazdelitev por (b)

Multiplying porosity with element area gives the area of the pore within the element. The area equivalent diameter d_{equ} can be calculated using this equation. If two small pores exist in an element in the original calculation the recalculation will result into one big pore which will further lead to conservative results. Calculated pore distribution in comparison to real pore distribution is presented in **Figure 15** and an obvious and pretty good correlation between both is observed.

The calculated pore distribution depends on post pressure and shows a different pore distribution for each calculation cycle as it relies on a random function. This reflects reality in a good way, as components even if they are of the same charge never show exactly the same pore distribution. But in a statistic point of view, the results can be compared by looking at the density and size of pores.

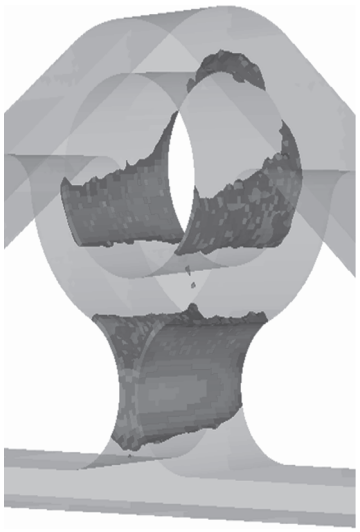


Figure 16: Distribution of safety against cyclic failure in the component
Slika 16: Porazdelitev varnosti priti ciklični obremenitvi komponente

2.5.4 Calculation of safety against cyclic failure

A fracture mechanic model defines the correlation between fatigue strength and pore size [8]. With this model, the local load capacity depending on the pore size can be calculated in each element. The load in the component is further calculated by FEM. Local safety for each element is due to the correlation between local load capacity and local load. The distribution of safety against cyclic failure for the above shown reference component is depicted in **Figure 16**.

3 CONCLUSION

For further optimization of components with the focus on higher utilisation of materials beside the knowledge of local stresses also knowledge of local strengths is needed. The later shows distinction on many different influences

- Kind of loading (Push/Pull, Bending , Torsion)
- Geometry and size
- Temperature
- Mean Stress
- Surface Layer (Surface topography, residual stress, microstructure, hardness)
- Load sequence
- Production process, Casting, Deforming, Cutting or Welding
- etc.

All these influences lead to higher or lower strength of the component. When occurring simultaneously the effects can either strengthen or attenuate each other. The experimental investigation of all these effects on fatigue strength in its whole sum is due to time and cost factors nearly impossible and can only be performed punctual, additional simulation is needed for calculation.

On the basis of stress calculations with FEM and simulation of casting and deforming process the dimensioning process of components was shown. These examples show how interdisciplinary work enhances the significance of strength calculations.

4 REFERENCES

[1] Bargel H. J., Schulze G.: Werkstoffkunde, Springer-Verlag Berlin Heidelberg, 2005

[2] Böhm J.: Zur Vorhersage von Dauerschwingfestigkeiten ungekerbter und gekerbter Bauteile unter Berücksichtigung des statistischen Größeneinflusses. Dissertation, TU München, 1980

[3] Sonsino, C. M.: Zur Bewertung des Schwingfestigkeitsverhaltens von Bauteilen mit Hilfe örtlicher Beanspruchungen. Konstruktion 45 (**1993**)

[4] Fröschl J.: Fatigue effects of forged components: Technological effects and multiaxial fatigue, Dissertation, Montanuniversität Leoben, 2006

[5] Altenpohl, D.: Aluminium von innen. Aluminium-Verlag, Düsseldorf, 1994

[6] Minichmayr, R., Eichlseder, W.: Lebensdauerberechnung von Gussbauteilen unter Berücksichtigung des lokalen Dendritenarmabstandes und der Porosität, Gießerei, (**2003**) 5

[7] Powazka D.: Einfluss der Porosität auf die Betriebsfestigkeit von Al-Druckgussbauteilen, Dissertation, Montanuniversität Leoben, 2009

[8] Oberwinkler Christian: Virtuelle betriebsfeste Auslegung von Aluminium-Druckgussbauteilen, Dissertation, Montanuniversität Leoben, 2009

UDK 54
Original scientific article/Izvirni znanstveni članek

ISSN 1580-2949
MTAEC9, 44(4)193(2010)

A NEW COMPLEX VECTOR METHOD FOR BALANCING CHEMICAL EQUATIONS

NOVA KOMPLEKSNA METODA ZA URAVNOTEŽENJE KEMIČNIH ENAČB

Ice B. Risteski

2 Milepost Place # 606, Toronto, Ontario, Canada M4H 1C7
ice@scientist.com

»Just as tensor analysis is the proper language of kinematics, so linear algebra is the proper language of stoichiometry.«
Aris, R.; Mah, R. H. S.; *Ind. Eng. Chem. Fundam*. 2 (**1963**), 90.

Prejem rokopisa – received: 2009-10-22; sprejem za objavo – accepted for publication: 2010-01-21

In this article, the author discovers a paradox of balancing chemical equations. The many counterexamples illustrate that the considered procedure of balancing chemical equations given in the paper[1] is inconsistent. A new complex vector method for paradox resolution is given too.

Keywords: paradox, balancing, chemical equations, vector space.

V članku avtor opisuje paradoks pri uravnoteženju kemijskih reakcij. Več primerov dokazuje, da je procedura uravnoteženja kemijskih reakcij v viru[1] inkonsistentna. Predstavljena je nova kompleksna vektorska metoda za rešitev paradoksa.

Ključne besede: paradoks, uravnoteženje, kemijske reakcije, vektorski prostor

1 INTRODUCTION

Balancing chemical equations is a basic matter of chemistry, if not one of its most important issues, and it plays a main role in its foundation. Indeed, it is a subtle question which deserves considerable attention.

Also, this topic was a magnet for a number of researchers around the world, because the general problem of balancing chemical equations was considered as one of the hardest problems in chemistry, as well as in mathematics.

In chemical literature there are a great number of papers which consider the problem of balancing chemical equations in different *chemical ways*, but all of them offer only some particular procedures for balancing simple chemical equations. Very often, these chemical procedures generate absurd results, because most of them are founded on *presumed chemical principles*, but not on genuine exact principles. And these *presumed chemical principles* generate paradoxes! The balancing chemical equations does not depend on *chemical principles*, it is a mathematical operation which is founded on true mathematical principles.

There is a number of paradoxes in chemistry about balancing chemical equations, and these will be systematized and studied in a special article of higher level by the same author which will appear in the near future. In this article only one paradox in balancing chemical equations, discovered in[1] is considered.

The author would like to emphasize very clearly, that *balancing chemical equations is not chemistry; it is just linear algebra.* Although it is true, that it is not chemistry, it is very important for chemistry! In this particular case the following question comes up: *If balancing chemical equations is not chemistry, then why is it considered in chemistry*? Or maybe more important is this question: *If it is linear algebra, then it should be studied in mathematics, so why bother chemists*?

The author will address the above questions assuming that: *the problem of balancing chemical equations was a multidisciplinary subject and for its solution both mathematicians and chemists are needed. The job of chemists is to perform reactions, while balancing their equations is a job for mathematicians. Mathematics, as a servant to other sciences, is a problem solver, but chemistry is a result user.*

The skepticism about balancing chemical equations by *chemical principles* appeared a long time ago. Let's quote the opinions of three chemists.

In 1926 the American chemist Simons[2] wrote: *The balancing of an equation is a mathematical process and independent of chemistry. The order of steps in the process is as essential as the order of steps in long division and the process is much simpler than is ordinary assumed.* Seventy-one years later, the Dutch chemist Ten Hoor[3], made a similar statement: *Balancing the equation of the reaction is a matter of mathematics only.*

Now is the right place to paraphrase the criticism of the American chemist Herndon[4]: *The major changes that have taken place over seven decades are substitutions of the terms 'change in oxidation number' and 'algebraic*

method' for the terms 'valence change' and 'method of undetermined coefficients', respectively. This assertion shows very clearly the picture of chemists' contributions to balancing chemical equations by chemical principles. Probably expressing his satisfaction with the chemists' contributions to balancing chemical equations, Herndon, the editor of the Journal of Chemical Education[4], decided that further discussion *of equation balancing will not appear in the Journal unless it adds something substantively new to what has already appeared.*

In view of the above assertions this question arises: *Are there in chemistry 'chemical principles' capable of offering solutions of the general problem of balancing chemical equations? Maybe more interesting is the question: What are 'chemical principles' – a rhetorical sophism of chemists or their hopes?*

More interesting for us is to give an answer to both questions from a scientific view point. Perhaps, the more appropriate short answer to the first question is: *'Chemical principles' are not defined entities in chemistry, and so this term does not have any meaning. They are not capable to provide solutions of the general problem of balancing chemical equations, because they are founded on an intuitive basis and they represent only a main generator for paradoxes. However, the necessary and sufficient conditions for a complete solution of the general problem of balancing chemical equations lie outside of chemistry, and we must look for them in an amalgamated theory of n-dimensional vector spaces, linear algebra, abstract algebra and topology. It is a very hard problem of the highest level in chemistry and mathematics, which must be considered only on a scientific basis.* Like this should look the answer to the first question.

The answer to the second question will be described in this way: *Actually, 'chemical principles' are a remnant of an old traditional approach in chemistry, when chemists were busy with the verification of their results obtained in chemical experiments. It is true, that until second half of 20th century there was no mathematical method for balancing chemical equations in chemistry, other than Bottomley's algebraic method[5]. Chemists balanced simple particular chemical equations using only Johnson's change in oxidation number procedure[6], Simons – Waldbauer – Thrun's partial reactions procedure[2,7] and other slightly different modifications derived from them. The 'chemical principles' were an assumption of traditional chemists, who thought before the appearance of Jones' problem[8], that the solution of the general problem of balancing chemical equations is possible by use of chemical procedures. But, practice showed that the solution of the century old problem is possible only by using contemporary sophisticated mathematical methods.*

These questions require a deeper elaboration than given here, and it will be a main concern of the author in his future research.

2 PRELIMINARIES

The exact statements about balancing chemical equations agree with the following well-known results.

Theorem 1. *Every chemical reaction can be reduced in a matrix equation* $\boldsymbol{Ax} = \boldsymbol{0}$*, where* $\boldsymbol{A}$ *is a reaction matrix,* $\boldsymbol{x}$ *is a column-vector of the unknown coefficients and* $\boldsymbol{0}$ *is a null column-vector.*

Proof. The proof of this theorem immediately follows from[9].

Remark 2. The coefficients satisfy three basic principles (corresponding to a closed input-output static model[10,11])

- the law of conservation of atoms,
- the law of conservation of mass, and
- the time-independence of the reaction.

Theorem 3. *Every chemical reaction can be presented as a matrix Diophantine equation* $\boldsymbol{Ax} = \boldsymbol{By}$*, where* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are matrices of reactants and products, respectively, and* $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are column-vectors of unknown coefficients.* The proof of this theorem is given in.[12,13]

This theorem is actually the Jones'[8] problem founded by virtue of Crocker's procedure for balancing chemical equations[14], and from this problem the formalization of chemistry began. This problem is a milestone in chemistry and mathematics as well, and just it opened the door in chemistry to enter a new mathematical freshness. Jones[8] with his problem transferred the general problem of balancing chemical equations from the field of chemistry into the field of mathematics and opened way to solve this problem with mathematical methods founded on principles of linear algebra.

Before the appearance of this problem, the approach of balancing chemical equations was intuitive and useful only for some elementary chemical equations. Now, some well-known results from the theory of complex n-dimensional vector spaces[15,16], for resolution of balancing chemical equations will be introduced. Here, by $\mathbb{C}$ is denoted the set of complex numbers and by $\mathbb{C}^n$ is denoted the Euclidian n-dimensional vector space with complex entries.

Definition 4. A vector space over the field $\mathbb{C}$ consists of a set V of objects called vectors for which the axioms for vector addition hold

(A_1) *If* $\boldsymbol{u}, \boldsymbol{v} \in V$*, then* $(\boldsymbol{u} + \boldsymbol{v}) \in V$,
(A_2) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$, $\forall \boldsymbol{u}, \boldsymbol{v} \in V$,
(A_3) $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$, $\forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$,
(A_4) $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u} = \boldsymbol{0} + \boldsymbol{u}$, $\forall \boldsymbol{u} \in V$,
(A_5) $-\boldsymbol{u} + \boldsymbol{u} = \boldsymbol{0} = \boldsymbol{u} + (-\boldsymbol{u})$, $\forall \boldsymbol{u} \in V$,

and the axioms for scalar multiplication

(S_1) *If* $\boldsymbol{u} \in V$*, then* $a\boldsymbol{u} \in V$, $\forall a \in \mathbb{C}$,
(S_2) $a(\boldsymbol{u} + \boldsymbol{v}) = a\boldsymbol{u} + a\boldsymbol{v}$, $\forall \boldsymbol{u}, \boldsymbol{v} \in V \wedge \forall a \in \mathbb{C}$,
(S_3) $(a + b)\boldsymbol{u} = a\boldsymbol{u} + b\boldsymbol{u}$, $\forall \boldsymbol{u} \in V \wedge \forall a,b \in \mathbb{C}$,
(S_4) $a(b\boldsymbol{u}) = (ab\boldsymbol{u})$, $\forall \boldsymbol{u} \in V \wedge \forall a,b \in \mathbb{C}$,
(S_5) $1\boldsymbol{u} = \boldsymbol{u}$, $\forall \boldsymbol{u} \in V$.

Remark 5. *The content of axioms* (A_1) *and* (S_1) *is described with thes assertion that V is closed under vector addition and scalar multiplication. The element* $\mathbf{0}$ *in axiom* A_4 *is called the zero vector.*

Definition 6. *If V is a vector space over the field* $\mathbb{C}$*, a subset U of V is called a subspace of V if U itself is a vector space over* $\mathbb{C}$*, where U uses the vector addition and scalar multiplication of V.*

Definition 7. *Let V be a vector space over the field* $\mathbb{C}$*, and let* $\boldsymbol{v}_i \in V$ $(1 \leq i \leq n)$*. Any vector in V of the form* $\boldsymbol{v} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_n\boldsymbol{v}_n$*, where* $a_i \in \mathbb{C}$*,* $(1 \leq i \leq n)$ *is called a linear combination of* $\boldsymbol{v}_i$*,* $(1 \leq i \leq n)$*.*

Definition 8. *The vectors* $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ *are said to span or generate V or are said to form a spanning set of V if* $V = \text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$*. Alternatively,* $\boldsymbol{v}_i \in V$ $(1 \leq i \leq n)$ *span V, if for every vector* $\boldsymbol{v} \in V$ *there exist scalars* $a_i \in \mathbb{C}$ $(1 \leq i \leq n)$ *such that*

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_n\boldsymbol{v}_n,$$

i. e., $\boldsymbol{v}$ is a linear combination of

$$a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_n\boldsymbol{v}_n.$$

Remark 9. *If* $V = \text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$*, then each vector* $\boldsymbol{v} \in V$ *can be written as a linear combination of the vectors* $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$*. Spanning sets have the property that each vector in V has exactly one representation as a linear combinations of these vectors.*

Definition 10. *Let V be a vector space over a field* $\mathbb{C}$*. The vectors* $\boldsymbol{v}_i \in V$ $(1 \leq i \leq n)$ *are said to be linearly independent over* $\mathbb{C}$*, or simply independent, if it satisfies the following condition: if*

$$s_1\boldsymbol{v}_1 + s_2\boldsymbol{v}_2 + \ldots + s_n\boldsymbol{v}_n = 0,$$

then

$$s_1 = s_2 = \ldots = s_n = 0.$$

Otherwise, the vectors that are not linearly independent, are said to be linearly dependent, or simply dependent.

Remark 11. *The trivial linear combination of the vectors* $\boldsymbol{v}_i$*,* $(1 \leq i \leq n)$ *is the one with every coefficient zero*

$$0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 + \ldots + 0\boldsymbol{v}_n.$$

Definition 12. *A set of vectors* $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ *is called a basis of V if it satisfies the following two conditions*

1° $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ *are linearly independent*,

2° $V = \text{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$.

Definition 13. *A vector space V is said to be of finite dimension n or to be n-dimensional, written dim V = n, if V contains a basis with n elements.*

Definition 14. *The vector space* $\{\mathbf{0}\}$ *is defined to have dimension* 0.

3 PARADOX APPEARANCE

Chemistry as other natural sciences is not immune of paradoxes. Unlike other natural sciences, in chemistry paradoxes appeared some time later, and it has only two, while in other sciences many such contradictions are met. These paradoxes are well-known Levinthal's paradox[17]: *The length of time in which a protein chain finds its folded state is many orders of magnitude shorter than it would be if it freely searched all possible configurations*, and Structure-Activity Relationship (SAR) paradox[18]: *Exceptions to the principle that a small change in a molecule causes a small change in its chemical behavior are frequently profound.*

However, these paradoxes are not alone and there are more, but now another will be mentioned, which appears in the balancing of chemical equations.

In the paper[1], the so-called formal balance numbers (FBN) are introduced, like this: *Formal balance numbers are an aid that may grossly facilitate the problem of balancing complex redox equations. They may be chosen as being equal to the traditional values of oxidation numbers, but not necessarily. An inspection of the redox equation may suggest the optimal values that are to be assigned to formal balance numbers. In most cases, these optimal values ensure that only two elements will 'change their state'* (*i. e. the values of the formal balance numbers*), *allowing the use of the oxidation number technique for balancing equations, in its simplest form. Just like for oxidation numbers, the algebraic sum of the formal balance numbers in a molecule/neutral unit is* 0, *while in an ion it is equal to its charge* (*the sum rule*).

Promptly, it was detected that the procedure given in[1] boils down to using of well-known unconventional oxidation numbers, which previously were advocated by Tóth[19] and Ludwig[20].

Consider this sentence from previous definition: *They may be chosen as being equal to the traditional values of oxidation numbers, but not necessarily.* It is a paradox! If the *formal balance numbers* can be the same as oxidation numbers or not, then the whole definition is illogical. This definition represents only a contradictory premise, which does not have any correlation with balancing chemical equations. The above definition does not speak anything about balancing chemical reactions in a chemical sense of the word, or their solution in a mathematical sense. *In order for a chemical equation to be balanced the first necessary and sufficient condition is its solvability,* but the above definition is far from it.

The so-called *formal balance numbers,* which actually are the same as the well-known oxidation numbers, *are not any criterion for balancing chemical equations.*

If the chemical equation is not formalized, then it generates only paradoxes. To support this assertion some ordinary counterexamples will be given.

1° Balancing chemical equations by using of change in oxidation number procedure has a limited usage! It holds only for some simple equations. *Is it possible to determine valence of elements in some complex organic molecules as they are* $C_{2952}H_{4664}N_{812}O_{832}S_8Fe_4$, $C_{738}H_{1166}N_{812}O_{203}S_2Fe$*, and so on*?

Not yet! For instance, one may show the powerlessness of that procedure by the following two counterexamples.

Example 1.

$$2993\ C_{2952}H_{4664}N_{816}O_{832}S_8Fe_4 + 9300568\ CsClO_3 + 408724\ Li_3PO_4 \rightarrow 11972\ Li_4Fe(CN)_6 + 8763504CO_2 + 2370456CsNO_3 + 1178284\ LiHCl_2 + 408724\ H_3PO_4 + 6944000\ Cs_{0.998}Cl + 23944\ H_2S + 5753504\ H_2O.$$

What is the valence of C, N, S *and* Fe?

Example 2.

$$249000\ C_{738}H_{1166}N_{812}O_{203}S_2Fe + 1098133767\ Pb(NO_3)_2 + 46812000\ HSiCl_3 + 7716178\ P_2O_5 + 23148534\ CoCl_2 \rightarrow 46812000\ SiCl_{3.989} + 7716178\ Co_3(PO_4)_2 + 124500\ H_2SO_4 + 549616500\ Pb_{1.998}O_3 + 124500\ Fe_2(SO_4)_3 + 15313500\ C_{12}H_{22}O_{11} + 2398455534\ NO_2.$$

The same previous question holds for this counterexample too. This means, the solution of the above equations is based on very sophisticated matrix methods[21–23], but in no case on the change in oxidation number procedure!

As consequence of that, same holds for so-called *formal balance numbers*. That procedure is useless to balance complex chemical equations as it does not give effects for balancing chemical equations. Still, there are many causes for arguing why that elementary procedure is useless, but the above mentioned two counterexamples are enough to show that it is ineffective.

2° Consider this simple reaction

Example 3.

$$x_1\ NaNO_2 + x_2\ FeSO_4 + x_3\ H_2SO_4 \rightarrow x_4\ NaHSO_4 + x_5\ Fe_2(SO_4)_3 + x_6\ NO.$$

In the above chemical reaction N and Fe change the valence, but chemical equation is impossible! Thus, is the procedure of *formal balance numbers* useful for balancing all chemical equations? Obviously, not! Since it does not have the capability to detect if some chemical equation is possible or not.

3° Now, one more counterexample will be given, where that particular procedure of *formal balance numbers* is impossible.

For instance, consider this chemical reaction

Example 4.

$$128147987406\ FeSO_4 + 27736706439\ PrTlTe_3 + 286573109604\ H_3PO_4 \rightarrow 128404797000\ Fe_{0.998}(H_2PO_4)_2{\cdot}H_2O + 13945051000\ Tl_{1.989}(SO_3)_3 + 13889187000\ Pr_{1.997}(SO_4)_3 + 83210119317\ TeO_2 + 14881757802\ P_2O_3 + 44645273406\ H_2S.$$

Is it possible to balance the above chemical equation by the procedure of *formal balance numbers*?

Not yet! Since it is not possible to determine the valence of Fe, Tl and Pr. Then, on what basis the author of the paper[1] states, that *his* procedure (there called *method*) *is proposed for fast and easy balancing of complex redox equations*? On top of all, he states that *the procedure* (there called *method*) *is probably the fastest of all possible methods*! Really, a *very modest* statement is offered by the author, which is wrong, not only because of today's current balancing methods view point, but also from an earlier view point, when that procedure was published.

Is it a *method* when somebody can find on every step counterexamples? Obviously, the answer is negative! It is merely a picture of the old chemical traditionalism. Or, maybe it is a lonely case lost in the newest progressive and contemporary mathematical generalism!

4 PARADOX RESOLUTION BY A NEW COMPLEX VECTOR METHOD

With the purpose of solution of the paradox, in this section a new complex vector method of balancing chemical equations will be developed. This method is founded on the theory of n-dimensional complex vector spaces.

Theorem 15. *Suppose that chemical equation*

$$x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \ldots + x_n\boldsymbol{v}_n = \boldsymbol{0}, \tag{1}$$

where $\boldsymbol{v}_i$ $(1 \le i \le n)$ *are the molecules and* x_i $(1 \le i \le n)$ *are unknown coefficients is a vector space* V *over the field* $\mathbb{C}$ *spanned by the vectors of the molecules* v_i $(1 \le i \le n)$. *If any set of* m *vectors of the molecules in* V *is linearly independent, then* $m \le n$.

Proof. Let be $V = \text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. We must show that every set $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ of vectors in V with $m > n$ fails to be linearly independent. This is accomplished by showing that numbers $x_1, x_2, \ldots, x_m$ can be found, not all zero, such that

$$x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \ldots + x_m\boldsymbol{u}_m = \boldsymbol{0}.$$

Since V is spanned by the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, each vector $\boldsymbol{u}_j$ can be expressed as a linear combination of $\boldsymbol{v}_i$

$$\boldsymbol{u}_j = a_{1j}\boldsymbol{v}_1 + a_{2j}\boldsymbol{v}_2 + \ldots + a_{nj}\boldsymbol{v}_n.$$

Substituting these expressions into the preceding equation gives

$$\boldsymbol{0} = \sum_{j=1}^{m} x_j \left(\sum_{i=1}^{n} a_{ij}\boldsymbol{v}_j \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij}x_j \right) \boldsymbol{v}_i .$$

This is certainly the case if each coefficient of $\boldsymbol{v}_i$ is zero, *i. e.*, if

$$\sum_{j=1}^{m} a_{ij}x_j = 0, \ (1 \le i \le n).$$

This is a system of n equations with m variables $x_1, x_2, \ldots, x_m$, and because $m > n$, it has a nontrivial solution.

This is what we wanted. Now we shall prove the following results.

Theorem 16. *Let U be a subset of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$. *Then U is a subspace of V if and only if it satisfies the following conditions*

$$\mathbf{0} \in U, \text{ where } \mathbf{0} \text{ is the zero vector of } V, \tag{2}$$

$$\text{If } \boldsymbol{u}_1, \boldsymbol{u}_2 \in U, \text{ then } (\boldsymbol{u}_1 + \boldsymbol{u}_2) \in U, \tag{3}$$

$$\text{If } \boldsymbol{u} \in U, \text{ then } a\boldsymbol{u} \in U, \forall a \in \mathbb{C}. \tag{4}$$

Proof. If U is a subspace of V of the chemical equation (1), it is clear by axioms (A_1) and (S_1) that the sum of two vectors in U is again in U and that any scalar multiple of a vector in U is again in U. In other words, U is *closed* under the vector addition and scalar multiplication of V. The converse is also true, *i. e.*, if U is closed under these operations, then all the other axioms are automatically satisfied. For instance, axiom (A_2) asserts that holds $\boldsymbol{u}_1 + \boldsymbol{u}_2 = \boldsymbol{u}_2 + \boldsymbol{u}_1$, $\forall \boldsymbol{u}_1, \boldsymbol{u}_2 \in U$. This is clear because the equation is already true in V, and U uses the same addition as V. Similarly, axioms (A_3), (S_2), (S_3), (S_4) and (S_5) hold automatically in U, because they are true in V. All that remains is to verify axioms (A_4) and (A_5).

If (2), (3) and (4) hold, then axiom (A_4) follows from (2) and axiom (A_5) follows from (4), because $-\boldsymbol{u} = (-1)\boldsymbol{u}$ lies in U, $\forall \boldsymbol{u} \in U$. Hence, U is a subspace by the above discussion. Conversely, if U is a subspace, it is closed under addition and scalar multiplication and this gives (3) and (4). If $\boldsymbol{z}$ denotes the zero vector of U, then $\boldsymbol{z} = 0\boldsymbol{z}$ in U. But, $0\boldsymbol{z} = \mathbf{0}$ in V, so $\mathbf{0} = \boldsymbol{z}$ lies in U. This proves (2).

Remark 17. *If U is a subspace of V of the chemical equation* (1) *over the field, then the above proof shows that U and V share the same zero vector. Also, if* $\boldsymbol{u} \in U$, *then* $-\boldsymbol{u} = (-1)\boldsymbol{u} \in U$, *i. e., the negative of a vector in U is the same as its negative in V.*

Proposition 18. *If V is any vector space of the chemical equation* (1) *over the field* $\mathbb{C}$, *then* $\{\mathbf{0}\}$ *and V are subspaces of V.*

Proof. $U = V$ clearly satisfies the conditions of the Theorem 16. As to $U = \{\mathbf{0}\}$, it satisfies the conditions because $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $a\mathbf{0} = \mathbf{0}$, $\forall a \in \mathbb{C}$.

Remark 19. *The vector space* $\{\mathbf{0}\}$ *is called the zero subspace of V of the chemical equation* (1) *over the field* $\mathbb{C}$. *Since all zero subspaces look alike, we speak of the zero vector space and denote it by* 0. *It is the unique vector space containing just one vector.*

Proposition 20. *If* $\boldsymbol{v}$ *is a vector of some molecule in a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$, *then the set* $\mathbb{C}\boldsymbol{v} = \{a\boldsymbol{v}, \forall a \in \mathbb{C}\}$ *of all scalar multiplies of* $\boldsymbol{v}$ *is a subspace of V.*

Proof. Since $\mathbf{0} = 0\boldsymbol{v}$, it is clear that $\mathbf{0}$ lies in $\mathbb{C}\boldsymbol{v}$. Given two vectors $a\boldsymbol{v}$ and $b\boldsymbol{v}$ in $\mathbb{C}\boldsymbol{v}$, their sum $a\boldsymbol{v} + b\boldsymbol{v} = (a + b)\boldsymbol{v}$ is also a scalar multiple of $\boldsymbol{v}$ and so lies in $\mathbb{C}\boldsymbol{v}$. Therefore $\mathbb{C}\boldsymbol{v}$ is closed under addition. Finally, given $a\boldsymbol{v}$, $r(a\boldsymbol{v}) = (ra)\boldsymbol{v}$ lies in $\mathbb{C}\boldsymbol{v}$, so $\mathbb{C}\boldsymbol{v}$ is closed under scalar multiplication. Now, if we take into account the Theorem 16, immediately follows the statement of the proposition.

Theorem 21. *Let* $U = \text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *in a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$. *Then,*

U is a subspace of V containing each of $\boldsymbol{v}_i$ $(1 \le i \le n)$, (5)

U is the smallest subspace containing these vectors in the sense that any subspace of V that contains each of v_i $(1 \le i \le n)$, *must contain U.* (6)

Proof. First we shall proof (5). Clearly

$$\mathbf{0} = 0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 + \ldots + 0\boldsymbol{v}_n$$

belongs to U. If

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_n\boldsymbol{v}_n$$

and

$$\boldsymbol{w} = b_1\boldsymbol{v}_1 + b_2\boldsymbol{v}_2 + \ldots + b_n\boldsymbol{v}_n$$

are two members of U and $a \in U$, then

$$\begin{gathered}\boldsymbol{v} + \boldsymbol{w} \\ = (a_1 + b_1)\boldsymbol{v}_1 + (a_2 + b_2)\boldsymbol{v}_2 + \ldots + (a_n + b_n)\boldsymbol{v}_n, \\ a\boldsymbol{v} = (aa_1)\boldsymbol{v}_1 + (aa_2)\boldsymbol{v}_2 + \ldots + (aa_n)\boldsymbol{v}_n,\end{gathered}$$

so both $\boldsymbol{v} + \boldsymbol{w}$ and $a\boldsymbol{v}$ lie in U. Hence, U is a subspace of V. It contains each of $\boldsymbol{v}_i$ $(1 \le i \le n)$. For instance,

$$\boldsymbol{v}_2 = 0\boldsymbol{v}_1 + 1\boldsymbol{v}_2 + 0\boldsymbol{v}_3 + \ldots + 0\boldsymbol{v}_n.$$

This proves (5).

Now, we shall prove (6). Let W be subspace of V that contains each of $\boldsymbol{v}_i$, $(1 \le i \le n)$. Since W is closed under scalar multiplication, each of $a_i\boldsymbol{v}_i$ $(1 \le i \le n)$ lies in W for any choice of a_i $(1 \le i \le n)$ in $\mathbb{C}$. But, then $a_i\boldsymbol{v}_i$ $(1 \le i \le n)$ lies in W, because W is closed under addition. This means that W contains every member of U, which proves (6).

Theorem 22. *The intersection of any number of subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$ *is a subspace of V.*

Proof. Let $\{W_i: i \in \boldsymbol{I}\}$ be a collection of subspaces of V and let $W = \cap (W_i: i \in \boldsymbol{I})$. Since each W_i is a subspace, then $\mathbf{0} \in W_i$, $\forall i \in \boldsymbol{I}$. Thus $\mathbf{0} \in W$. Assume $\boldsymbol{u}, \boldsymbol{v} \in W$. Then, $\boldsymbol{u}, \boldsymbol{v} \in W_i$, $\forall i \in \boldsymbol{I}$. Since each W_i is a subspace, then $(a\boldsymbol{u} + b\boldsymbol{v}) \in W_i$, $\forall i \in \boldsymbol{I}$. Therefore $(a\boldsymbol{u} + b\boldsymbol{v}) \in W$. Thus W is a subspace of V of the chemical equation (1).

Theorem 23. *The union* $W_1 \cup W_2$ *of subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$ *need not be a subspace of V.*

Proof. Let $V = \mathbb{C}^2$ and let $W_1 = \{(a, 0): a \in \mathbb{C}\}$ and $W_2 = \{(0, b): b \in \mathbb{C}\}$. That is, W_1 is the x-axis and W_2 is the y-axis in $\mathbb{C}^2$. Then W_1 and W_2 are subspaces of V of the chemical equation (1). Let $\boldsymbol{u} = (1, 0)$ and $\boldsymbol{v} = (0, 1)$. Then the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ both belong to the union $W_1 \cup W_2$, but $\boldsymbol{u} + \boldsymbol{v} = (1, 1)$ does not belong to $W_1 \cup W_2$. Thus $W_1 \cup W_2$ is not a subspace of V.

Theorem 24. *The homogeneous system of linear equations obtained from the chemical equation* (1), *in n*

unknowns $x_1, x_2, \ldots, x_n$ *over the field* $\mathbb{C}$ *has a solution set W, which is a subspace of the vector space* $\mathbb{C}^n$.

Proof. The system is equivalent to the matrix equation $\boldsymbol{Ax} = \boldsymbol{0}$. Since $\boldsymbol{A0} = \boldsymbol{0}$, the zero vector $\boldsymbol{0} \in W$. Assume $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors in W, *i. e.*, $\boldsymbol{u}$ and $\boldsymbol{v}$ are solutions of the matrix equation $\boldsymbol{Ax} = \boldsymbol{0}$. Then $\boldsymbol{Au} = \boldsymbol{0}$ and $\boldsymbol{Av} = \boldsymbol{0}$. Therefore, $\forall a, b \in \mathbb{C}$, we have $\boldsymbol{A}(a\boldsymbol{u} + b\boldsymbol{v}) = a\boldsymbol{Au} + b\boldsymbol{Av} = a\boldsymbol{0} + b\boldsymbol{0} = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$. Hence $a\boldsymbol{u} + b\boldsymbol{v}$ is a solution of the matrix equation $\boldsymbol{Ax} = \boldsymbol{0}$, *i. e.*, $a\boldsymbol{u} + b\boldsymbol{v} \in W$. Thus W is a subspace of $\mathbb{C}^n$.

Theorem 25. *If S is a subset of the vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then*

1° *the set* $\operatorname{span}\{S\}$ *is a subspace of V of the chemical equation* (1) *over the field* $\mathbb{C}$ *which contains S.*

2° $\operatorname{span}\{S\} \subseteq W$*, if W is any subspace of V of the chemical equation* (1) *over the field* $\mathbb{C}$ *containing S.*

Proof. 1°. If $S = \varnothing$, then $\operatorname{span}\{S\} = \{\boldsymbol{0}\}$, which is a subspace of V containing the empty set $\varnothing$. Now assume $S \neq \varnothing$. If $\boldsymbol{v} \in S$, then $1\boldsymbol{v} = \boldsymbol{v} \in \operatorname{span}\{S\}$, therefore S is a subset of $\operatorname{span}\{S\}$. Also, $\operatorname{span}\{S\} \neq \varnothing$ because $S \neq \varnothing$. Now assume $\boldsymbol{v}, \boldsymbol{w} \in \operatorname{span}\{S\}$; say

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + \ldots + a_m\boldsymbol{v}_m$$

and

$$\boldsymbol{v} = b_1\boldsymbol{w}_1 + \ldots + b_n\boldsymbol{w}_n,$$

where $\boldsymbol{v}_i, \boldsymbol{w}_j \in S$ and a_i, b_j are scalars.

Then

$$\boldsymbol{v} + \boldsymbol{w} = a_1\boldsymbol{v}_1 + \ldots + a_m\boldsymbol{v}_m + b_1\boldsymbol{w}_1 + \ldots + b_n\boldsymbol{w}_n$$

and for any scalar k,
$k\boldsymbol{v} = k(a_1\boldsymbol{v}_1 + \ldots + a_m\boldsymbol{v}_m) = ka_1\boldsymbol{v}_1 + \ldots + ka_m\boldsymbol{v}_m$ belong to $\operatorname{span}\{S\}$ because each is a linear combination of vectors in S. Thus $\operatorname{span}\{S\}$ is a subspace of V of the chemical equation (1) over the field $\mathbb{C}$ which contains S.

2°. If $S = \varnothing$, then any subspace W contains S, and $\operatorname{span}\{S\} = \{\boldsymbol{0}\}$ is contained in W. Now assume $S \neq \varnothing$ and assume $\boldsymbol{v}_i \in S \subset W$ $(1 \leq i \leq m)$. Then all multiples $a_i\boldsymbol{v}_i \in W$ $(1 \leq i \leq m)$ where $a_i \in \mathbb{C}$, and therefore the sum $(a_1\boldsymbol{v}_1 + \ldots + a_m\boldsymbol{v}_m) \in W$. That is, W contains all linear combinations of elements of S. Consequently, $\operatorname{span}\{S\} \subseteq W$, as claimed.

Proposition 26. *If W is a subspace of V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then* $\operatorname{span}\{W\} = W$.

Proof. Since W is a subspace of V of the chemical equation (1) over the field $\mathbb{C}$, W is closed under linear combinations. Hence, $\operatorname{span}\{W\} \subseteq W$. But $W \subseteq \operatorname{span}\{W\}$. Both inclusions yield $\operatorname{span}\{W\} = W$.

Proposition 27. *If S is a subspace of V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then* $\operatorname{span}\{\operatorname{span}\{S\}\} = \operatorname{span}\{S\}$.

Proof. Since $\operatorname{span}\{S\}$ is a subspace of V, the above Proposition 26 implies that $\operatorname{span}\{\operatorname{span}\{S\}\} = \operatorname{span}\{S\}$.

Proposition 28. *If S and T are subsets of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, such that* $S \subseteq T$*, then* $\operatorname{span}\{S\} \subseteq \operatorname{span}\{T\}$.

Proof. Assume $\boldsymbol{v} \in \operatorname{span}\{S\}$. Then

$$\boldsymbol{v} = a_1\boldsymbol{u}_1 + \ldots + a_r\boldsymbol{u}_r,$$

where $a_i \in \mathbb{C}$, $(1 \leq i \leq r)$ and $\boldsymbol{u}_i \in S$ $(1 \leq i \leq r)$. But $S \subseteq T$, therefore every $\boldsymbol{u}_i \in T$ $(1 \leq i \leq r)$. Thus $v \in \operatorname{span}\{T\}$. Accordingly, $\operatorname{span}\{S\} \subseteq \operatorname{span}\{T\}$.

Proposition 29. *The* $\operatorname{span}\{S\}$ *is the intersection of all the subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$ *which contains S.*

Proof. Let $\{W_i\}$ be the collection of all subspaces of a vector space V of the chemical equation (1) containing S, and let $W = \cap\, W_i$. Since each W_i is a subspace of V, the set W is a subspace of V. Also, since each W_i contains S, the intersection W contains S. Hence $\operatorname{span}\{S\} \subseteq W$. On the other hand, $\operatorname{span}\{S\}$ is a subspace of V containing S. So $\operatorname{span}\{S\} = W_k$ for some k. Then $W \subseteq W_k = \operatorname{span}\{S\}$. Both inclusions give $\operatorname{span}\{S\} = W$.

Proposition 30. *If* $\operatorname{span}\{S\} = \operatorname{span}\{S \cup \{\boldsymbol{0}\}\}$*, then one may delete the zero vector from any spanning set.*

Proof. By Proposition 28, $\operatorname{span}\{S\} \subseteq \operatorname{span}\{S \cup \{\boldsymbol{0}\}\}$. Assume $\boldsymbol{v} \in \operatorname{span}\{S \cup \{\boldsymbol{0}\}\}$, say

$$\boldsymbol{v} = a_1\boldsymbol{u}_1 + \ldots + a_n\boldsymbol{u}_n + b\cdot 0$$

where $a_i, b \in \mathbb{C}$ $(1 \leq i \leq n)$ and $\boldsymbol{u}_i \in S$ $(1 \leq i \leq n)$.

Then $\boldsymbol{v} = a_1\boldsymbol{u}_1 + \ldots + a_n\boldsymbol{u}_n$, and so $\boldsymbol{v} \in \operatorname{span}\{S\}$. Thus $\operatorname{span}\{S \cup \{\boldsymbol{0}\}\} \subseteq \operatorname{span}\{S\}$. Both inclusions give $\operatorname{span}\{S\} = \operatorname{span}\{S \cup \{\boldsymbol{0}\}\}$.

Proposition 31. *If the vectors* $\boldsymbol{v}_i \in V$ $(1 \leq i \leq n)$ *span a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then for any vector* $\boldsymbol{w} \in V$*, the vectors* $\boldsymbol{w}, \boldsymbol{v}_i$ $(1 \leq i \leq n)$ *span V.*

Proof. Let $\boldsymbol{v} \in V$. Since the $\boldsymbol{v}_i$ $(1 \leq i \leq n)$ span V, there exist scalars a_i $(1 \leq i \leq n)$ such that $\boldsymbol{v} = a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n + 0\boldsymbol{w}$. Thus $\boldsymbol{w}, \boldsymbol{v}_i$ $(1 \leq i \leq n)$ span V.

Proposition 32. *If* $\boldsymbol{v}_i$ $(1 \leq i \leq n)$ *span a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, and for* $k > 1$*, the vector* $\boldsymbol{v}_k$ *is a linear combination of the preceding vectors* $\boldsymbol{v}_i$ $(1 \leq i \leq k-1)$*, then* $\boldsymbol{v}_i$ *without* $\boldsymbol{v}_k$ *span V, i. e.,* $\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n\} = V$.

Proof. Let $\boldsymbol{v} \in V$. Since the $\boldsymbol{v}_i$ $(1 \leq i \leq n)$ span V, there exist scalars a_i $(1 \leq i \leq n)$ such that

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n.$$

Since $\boldsymbol{v}_k$ is a linear combination of $\boldsymbol{v}_i$ $(1 \leq i \leq k-1)$, there exist scalars b_i $(1 \leq i \leq k-1)$ such that

$$\boldsymbol{v}_k = b_1\boldsymbol{v}_1 + \ldots + a_{k-1}\boldsymbol{v}_{k-1}.$$

Thus

$$\begin{aligned}\boldsymbol{v} &= a_1\boldsymbol{v}_1 + \ldots + a_k\boldsymbol{v}_k + \ldots + a_n\boldsymbol{v}_n\\ &= a_1\boldsymbol{v}_1 + \ldots + a_k(b_1\boldsymbol{v}_1 + \ldots + b_{k-1}\boldsymbol{v}_{k-1}) + \ldots + a_n\boldsymbol{v}_n\\ &= (a_1 + a_kb_1)\boldsymbol{v}_1 + \ldots + (a_{k-1} + a_kb_{k-1})\boldsymbol{v}_{k-1}\\ &\quad + a_{k+1}\boldsymbol{v}_{k+1} + \ldots + a_n\boldsymbol{v}_n.\end{aligned}$$

Therefore,

$$\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n\} = V.$$

Proposition 33. *If W_i, $(1 \leq i \leq k)$ are subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, for which* $W_1 \subset W_2 \subset \ldots \subset W_k$ *and*

$$W = W_1 \cup W_2 \cup \ldots \cup W_k,$$

then W is a subspace of V.

Proof. The zero vector $\mathbf{0} \in W_1$, hence $\mathbf{0} \in W$. Assume $\boldsymbol{u}, \boldsymbol{v} \in W$. Then, there exist j_1 and j_2 such that $\boldsymbol{u} \in Wj_1$ and $\boldsymbol{v} \in Wj_2$. Let $j = \max(j_1, j_2)$. Then $Wj_1 \subseteq W_j$ and $Wj_2 \subseteq W$, and so $\boldsymbol{u}, \boldsymbol{v} \in W_j$. But W_j is a subspace. Therefore $(\boldsymbol{u} + \boldsymbol{v}) \in W_j$ and for any scalar s the multiple $s\boldsymbol{u} \in W_j$. Since $W_j \subseteq W$, we have $(\boldsymbol{u} + \boldsymbol{v})$, $s\boldsymbol{u} \in W$. Thus W is a subspace of V.

Proposition 34. *If W_i $(1 \leq i \leq k)$ are subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$ *and S_i $(1 \leq i \leq k)$ span W_i $(1 \leq i \leq k)$, then* $S = S_1 \cup S_2 \cup \ldots \cup S_k$ *spans W.*

Proof. Let $\boldsymbol{v} \in W$. Then there exists j such that $\boldsymbol{v} \in W_j$. Then $\boldsymbol{v} \in \text{span}\{S_j\} \subseteq \text{span}\{S\}$. Therefore $W \subseteq \text{span}\{S\}$. But $S \subseteq W$ and W is a subspace. Hence $\text{span}\{S\} \subseteq W$. Both inclusions give $\text{span}\{S\} = W$, *i. e.*, S spans W.

Theorem 35. *Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a linearly independent set of vectors in a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then the following conditions*

1° $\{\boldsymbol{v}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *is a linearly independent set,*

2° $\boldsymbol{v}$ *does not lie in* $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$,

are equivalent for a vector $\boldsymbol{v}$ *in V.*

Proof. Assume 1° is true and assume, if possible, that $\boldsymbol{v}$ lies in $\text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$, say,

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n.$$

Then

$$\boldsymbol{v} - a_1\boldsymbol{v}_1 - \ldots - a_n\boldsymbol{v}_n = \mathbf{0}$$

is a nontrivial linear combination, contrary to 1°. So 1° implies 2°. Conversely, assume that 2° holds and assume that

$$a\boldsymbol{v} + a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n = \mathbf{0}.$$

If $a \neq 0$, then

$$\boldsymbol{v} = (-a_1/a)\boldsymbol{v}_1 + \ldots + (-a_n/a)\boldsymbol{v}_n,$$

contrary to 2°. So $a = 0$ and

$$a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n = \mathbf{0}.$$

This implies that

$$a_1 = \ldots = a_n = 0,$$

because the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is linearly independent. By this is proved that 2° implies 1°.

Proposition 36. *Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be linearly independent in a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then $\{a_1\boldsymbol{v}_1, a_2\boldsymbol{v}_2, \ldots, a_n\boldsymbol{v}_n\}$, such that the numbers a_i $(1 \leq i \leq n)$ are all nonzero, is also linearly independent.*

Proof. Suppose a linear combination of the new set vanishes

$$s_1(a_1\boldsymbol{v}_1) + s_2(a_2\boldsymbol{v}_2) + \ldots + s_n(a_n\boldsymbol{v}_n) = \mathbf{0},$$

where s_i $(1 \leq i \leq n)$ lie in $\mathbb{C}$.

Then

$$s_1a_1 = s_2a_2 = \ldots = s_na_n = 0$$

by the linear independence of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. The fact that each $a_i \neq 0$ $(1 \leq i \leq n)$ now implies that $s_1 = s_2 = \ldots = s_n = 0$.

Proposition 37. *No linearly independent set of vectors of molecules can contain the zero vector.*

Proof. The set $\{\mathbf{0}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ cannot be linearly independent, because

$$1\mathbf{0} + 0\boldsymbol{v}_1 + \ldots + 0\boldsymbol{v}_n = \mathbf{0},$$

is a nontrivial linear combination that vanishes.

Theorem 38. *A set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ of vectors of molecules in a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$ *is linearly dependent if and only if some $\boldsymbol{v}_i$ is a linear combination of the others.*

Proof. Assume that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is linearly dependent. Then, some nontrivial linear combination vanishes, *i. e.*,

$$a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_n\boldsymbol{v}_n = \mathbf{0},$$

where some coefficient is not zero. Suppose $a_1 \neq 0$.

Then

$$\boldsymbol{v}_1 = (-a_2/a_1)\boldsymbol{v}_2 + \ldots + (-a_n/a_1)\boldsymbol{v}_n,$$

gives $\boldsymbol{v}_1$ as a linear combination of the others. In general, if $a_i \neq 0$, then a similar argument expresses $\boldsymbol{v}_i$ as linear combination of the others.

Conversely, suppose one of the vectors is a linear combination of the others, *i. e.*,

$$\boldsymbol{v}_1 = a_2\boldsymbol{v}_2 + \ldots + a_n\boldsymbol{v}_n.$$

Then, the nontrivial linear combination $1\boldsymbol{v}_1 - a_2\boldsymbol{v}_2 - \ldots - a_n\boldsymbol{v}_n$ equals zero, so the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is not linearly independent, *i. e.*, it is linearly dependent. A similar argument works if *any* $\boldsymbol{v}_i$ $(1 \leq i \leq n)$ is a linear combinations of the others.

Theorem 39. *Let $V \neq 0$ be a vector space of the chemical equation* (1) *over the field* $\mathbb{C}$*, then*

1° *each set of linearly independent vectors is a part of a basis of V,*

2° *each spanning set V contains a basis of V,*

3° *V has a basis and* $\dim V \leq n$.

Proof. 1° Really, if V is a vector space that is spanned by a finite number of vectors, we claim that *any* linearly independent subset $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ of V is contained in a basis of V. This is certainly true if $V = \text{span}\{S\}$ because then S is *itself* a basis of V. Otherwise, choose $\boldsymbol{v}_{k+1}$ outside $\text{span}\{S\}$. Then $S_1 = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}\}$ is linearly independent by Theorem 35. If $V = \text{span}\{S_1\}$, then S_1 is the desired basis containing S. If not, choose $\boldsymbol{v}_{k+2}$ outside $\text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}\}$ so that $S_2 = \{\boldsymbol{v}_1, \boldsymbol{v}_2,$

…, $\boldsymbol{v}_k$, $\boldsymbol{v}_{k+1}$, $\boldsymbol{v}_{k+2}$} is linearly independent. Continue this process. Either a basis is reached at some stage or, if not, arbitrary large independent sets are found in *V*. But this later possibility cannot occur by the Theorem 15 because *V* is spanned by a finite number of vectors.

2° Let $V = \text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$, where (as $V \neq 0$) we may assume that each $\boldsymbol{v}_i \neq \boldsymbol{0}$. If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ is linearly independent, it is itself a basis and we are finished. If not, then according to the Theorem 38, one of these vectors lies in the span of the others. Relabeling if necessary, it is assumed that $\boldsymbol{v}_1$ lies in $\text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ so that $V = \text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$. Now repeat the argument. If the set $\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ is linearly independent, we are finished. If not, we have (after possible relabeling) $V = \text{span}\{\boldsymbol{v}_3, \ldots, \boldsymbol{v}_m\}$. Continue this process and if a basis is encountered at some stage, we are finished. If not, we ultimately reach $V = \text{span}\{\boldsymbol{v}_m\}$. But then $\{\boldsymbol{v}_m\}$ is a basis because $\boldsymbol{v}_m \neq \boldsymbol{0}$ ($V \neq 0$).

3° *V* has a spanning set of *n* vectors, one of which is nonzero because $V \neq 0$. Hence 3° follows from 2°.

Corollary 40. *A nonzero vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$ *is finite dimensional only if it can be spanned by finitely many vectors.*

Theorem 41. *Let V be a vector space of the chemical equation* (1) *over the field* $\mathbb{C}$ *and* $\dim V = n > 0$*, then*

1° *no set of more than n vectors in V can be linearly independent,*

2° *no set of fewer than n vectors can span V.*

Proof. *V* can be spanned by *n* vectors (any basis) so 1° restates the Theorem 15. But the *n* basis vectors are also linearly independent, so no spanning set can have fewer than *n* vectors, again by Theorem 15. This gives 2°.

Theorem 42. *Let V be a vector space of the chemical equation* (1) *over the field* $\mathbb{C}$ *and* $\dim V = n > 0$*, then*

1° *any set of n linearly independent vectors in V is a basis* (*that is, it necessarily spans V*),

2° *any spanning set of n nonzero vectors in V is a basis* (*that is, they are necessarily linearly independent*).

Proof. 1° If the *n* independent vectors do not span *V*, they are a part of a basis of more than *n* vectors by property 1° of the Theorem 39. This contradicts Theorem 41.

2° If the *n* vectors in a spanning set are not linearly independent, they contain a basis of fewer than *n* vectors by property 2° of Theorem 39, contradicting Theorem 41.

Theorem 43. *Let V be a vector space of dimension n of the chemical equation* (1) *over the field* $\mathbb{C}$ *and let U and W denote subspaces of V, then*

1° *U is finite dimensional and* $\dim U \leq n$,

2° *any basis of U is a part of a basis for V,*

3° *if* $U \subseteq W$ *and* $\dim U = \dim W$*, then* $U = W$.

Proof. 1° If $U = 0$, $\dim U = 0$ by Definition 14. So assume $U \neq 0$ and choose $\boldsymbol{u}_1 \neq \boldsymbol{0}$ in *U*. If $U = \text{span}\{\boldsymbol{u}_1\}$, then $\dim U = 1$. If $U \neq \text{span}\{\boldsymbol{u}_1\}$, choose $\boldsymbol{u}_2$ in *U* outside $\text{span}\{\boldsymbol{u}_1\}$. Then $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, is linearly independent by Theorem 35. If $U = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, then $\dim U = 2$. If not, repeat the process to find $\boldsymbol{u}_3$ in *U* such that $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is linearly independent and continue in this way. The process must terminate because the space *V* (having dimension *n*) cannot contain more than *n* independent vectors. Therefore U has a basis of at most *n* vectors, proving 1°.

2° This follows from 1° and Theorem 39.

3° Let $\dim U = \dim W = m$. Then any basis $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ of *U* is an independent set of *m* vectors in *W* and so is a basis of *W* by Theorem 42. In particular, $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ spans *W* so, because it also spans *U*, $W = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\} = U$. By this is proved 3°.

Proposition 44. *If U and W are subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then U + W is a subspace of V.*

Proof. Since *U* and *W* are subspaces, $\boldsymbol{0} \in U$ and $\boldsymbol{0} \in W$. Hence $\boldsymbol{0} = \boldsymbol{0} + \boldsymbol{0} \in U + W$. Assume $\boldsymbol{v}, \boldsymbol{v}' \in U + W$. Then there exist $\boldsymbol{u}, \boldsymbol{u}' \in U$ and $\boldsymbol{v}, \boldsymbol{v}' \in W$ such that $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ and $\boldsymbol{v}' = \boldsymbol{u}' + \boldsymbol{w}'$. Since *U* and *W* are subspaces, $\boldsymbol{u} + \boldsymbol{u}' \in U$ and $\boldsymbol{w} + \boldsymbol{w}' \in W$ and for any scalar *k*, $k\boldsymbol{u} \in U$ and $k\boldsymbol{w} \in W$. Accordingly, $\boldsymbol{v} + \boldsymbol{v}' = (\boldsymbol{u} + \boldsymbol{w}) + (\boldsymbol{u}' + \boldsymbol{w}') = (\boldsymbol{u} + \boldsymbol{u}') + (\boldsymbol{w} + \boldsymbol{w}') \in U + W$ and for any scalar *k*, $k\boldsymbol{v} = k(\boldsymbol{u} + \boldsymbol{w}) = k\boldsymbol{u} + k\boldsymbol{w} \in U + W$. Thus $U + W$ is a subspace of *V*.

Proposition 45. *If U and W are subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then U and W are contained in U + W.*

Proof. Let $\boldsymbol{u} \in U$. By hypothesis *W* is a subspace of V and so $\boldsymbol{0} \in W$. Hence $\boldsymbol{u} = \boldsymbol{u} + \boldsymbol{0} \in U + W$. Accordingly, *U* is contained in $U + W$. Similarly, *W* is contained in $U + W$. By this the proof is finished.

Proposition 46. *If U and W are subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then U + W is the smallest subspace of V containing U and V, i. e.,* $U + W = \text{span}\{U, W\}$.

Proof. Since $U + W$ is a subspace of *V* containing both *U* and *W*, it must also contain the linear span of *U* and *W*, *i. e.*, $\text{span}\{U, W\} \subseteq U + W$.

On the other hand, if $\boldsymbol{v} \in U + W$, then $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w} = 1\boldsymbol{u} + 1\boldsymbol{w}$, where $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in W$. Hence, $\boldsymbol{v}$ is a linear combination of elements in $U \cup W$ and so belongs to $\text{span}\{U, W\}$. Therefore, $U + W \subseteq \text{span}\{U, W\}$. Both inclusions give us the required result.

Proposition 47. *If W is a subspace of a vector space V of the chemical equation* (4. 2) *over the field* $\mathbb{C}$*, then* $W + W = W$.

Proof. Since *W* is a subspace of *V*, we have that *W* is closed under vector addition. Therefore, $W + W \subseteq W$. By Proposition 45, $W \subseteq W + W$. Thus, $W + W = W$. By this the proof is finished.

Proposition 48. *If U and W are subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, such that* $U = \text{span}\{S\}$ *and* $W = \text{span}\{T\}$*, then* $U + W = \text{span}\{S \cup T\}$.

Proof. Since $S \subseteq U \subseteq U + W$ and $T \subseteq W \subseteq U + W$, we have $S \cup T \subseteq U + W$. Hence $\text{span}\{S \cup T\} \subseteq U + W$.

Now assume $\boldsymbol{v} \in U + W$. Then $\boldsymbol{v} = \boldsymbol{u} + W$, where $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in W$. Since $U = \text{span}\{S\}$ and $W = \text{span}\{T\}$, $\boldsymbol{u} = a_1\boldsymbol{u}_1 + \ldots + a_r\boldsymbol{u}_r$ and $\boldsymbol{w} = b_1\boldsymbol{w}_1 + \ldots + b_s\boldsymbol{w}_s$, where $a_i, b_j \in \mathbb{C}$, $\boldsymbol{u}_j \in S$, and $\boldsymbol{w}_i \in T$. Then $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w} = a_1\boldsymbol{u}_1 + \ldots + a_r\boldsymbol{u}_r + b_1\boldsymbol{w}_1 + \ldots + b_s\boldsymbol{w}_s$. Thus, $U + W \subseteq \text{span}\{S \cup T\}$. Both inclusions yield $U + W = \text{span}\{S \cup T\}$.

Proposition 49. *If U and W are subspaces of a vector space V of the chemical equation* (1) *over the field* $\mathbb{C}$*, then* $V = U + W$ *if every* $\boldsymbol{v} \in V$ *can be written in the form* $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$*, where* $\boldsymbol{u} \in U$ *and* $\boldsymbol{w} \in W$.

Proof. Assume, for any $\boldsymbol{v} \in V$, we have $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ where $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in W$. Then $\boldsymbol{v} \in U + W$ and so $V \subseteq U + W$. Since U and V are subspaces of V, we have $U + W \subseteq V$. Both inclusions imply $V = U + W$.

By this, in whole is given the skeleton of the complex vector method.

5 AN APPLICATION OF THE MAIN RESULTS

In this section the above complex vector method will be applied on some chemical equations for their balancing. All chemical equations balanced here appear for the first time in professional literature and they are chosen with an intention to avoid all well-known to date chemical equations which were repeated many times in the chemical journals for explanation of certain particular techniques for balancing of some chemical equations using only atoms with integer oxidation numbers.

1° First, we shall consider the case when the chemical reaction is infeasible.

Example 5. Consider this chemical reaction

$$z_1\,K_4Fe(CN)_6 + z_2\,K_2S_2O_3 \rightarrow z_3\,CO_2 + z_4\,K_2SO_4 + z_5\,NO_2 + z_6\,FeS,\quad z_i \in \mathbb{C},\ (1 \le i \le 6).$$

From the following scheme

	$\boldsymbol{v}_1 = K_4Fe(CN)_6$	$\boldsymbol{v}_2 = K_2S_2O_3$	$\boldsymbol{v}_3 = CO_2$	$\boldsymbol{v}_4 = K_2SO_4$	$\boldsymbol{v}_5 = NO_2$	$\boldsymbol{v}_6 = FeS$
K	4	2	0	2	0	0
Fe	1	0	0	0	0	1
C	6	0	1	0	0	0
N	6	0	0	0	1	0
S	0	2	0	1	0	1
O	0	3	2	4	2	0

follows this vector equation

$$z_1\boldsymbol{v}_1 + z_2\boldsymbol{v}_2 = z_3\boldsymbol{v}_3 + z_4\boldsymbol{v}_4 + z_5\boldsymbol{v}_5 + z_6\boldsymbol{v}_6,$$

i. e.,

$$z_1\,(4, 1, 6, 6, 0, 0)^{\mathrm{T}} + z_2\,(2, 0, 0, 0, 2, 3)^{\mathrm{T}} = z_3\,(0, 0, 1, 0, 0, 2)^{\mathrm{T}} + z_4\,(2, 0, 0, 0, 1, 4)^{\mathrm{T}} + z_5\,(0, 0, 0, 1, 0, 2)^{\mathrm{T}} + z_6\,(0, 1, 0, 0, 1, 0)^{\mathrm{T}},$$

or

$$(4z_1 + 2z_2,\ z_1,\ 6z_1,\ 6z_1,\ 2z_2,\ 3z_2)^{\mathrm{T}} = (2z_4,\ z_6,\ z_3,\ z_5,\ z_4 + z_6,\ 2z_3 + 4z_4 + 2z_5)^{\mathrm{T}}.$$

From the system of linear equations

$$\begin{gathered} 4z_1 + 2z_2 = 2z_4,\\ z_1 = z_6,\\ 6z_1 = z_3,\\ 6z_1 = z_5,\\ 2z_2 = z_4 + z_6,\\ 3z_2 = 2z_3 + 4z_4 + 2z_5, \end{gathered}$$

one obtains the contradictions $z_2 = 3z_1$ and $z_2 = 44z_1/3$, that means that the system is inconsistent. According to Definition 8, the vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$, $\boldsymbol{v}_4$, $\boldsymbol{v}_5$ and $\boldsymbol{v}_6$ of the molecules of the chemical reaction (5. 1) do not generate a vector space V, and according to the Definition 10 they are linearly independent, *i. e.*, we have only a trivial solution $z_i = 0$, $(1 \le i \le 6)$, that means that the chemical reaction is infeasible.

2° Next, we shall consider the case when the chemical reaction is feasible and it has a unique solution.

This type of chemical equations really is the most appropriate for study the process of balancing chemical equations, because it gives an excellent opportunity for application of the group theory.

At once, we would like to emphasize here, that by application of groups theory one may determine Sylow subgroups, conjugacy classes of maximal subgroups, proper normal subgroups, and so one. The main reason why we confined ourselves to the next group of calculations is limitation of the size of the work.

Example 6. Consider this chemical reaction

$$z_1\,Fe_2(SO_4)_3 + z_2\,PrTlTe_3 + z_3\,H_3PO_4 \rightarrow z_4\,Fe_{0.996}(H_2PO_4)_2{\cdot}H_2O + z_5\,Tl_{1.987}(SO_3)_3 + z_6\,Pr_{1.998}(SO_4)_3 + z_7\,Te_2O_3 + z_8\,P_2O_5 + z_9\,H_2S,\ z_i \in \mathbb{C},\ (1 \le i \le 9).$$

According to the scheme given below

	$\boldsymbol{v}_1 = Fe_2(SO_4)_3$	$\boldsymbol{v}_2 = PrTlTe_3$	$\boldsymbol{v}_3 = H_3PO_4$	$\boldsymbol{v}_4 = Fe_{0.996}(H_2PO_4)_2{\cdot}H_2O$	$\boldsymbol{v}_5 = Tl_{1.987}(SO_3)_3$	$\boldsymbol{v}_6 = Pr_{1.998}(SO_4)_3$	$\boldsymbol{v}_7 = Te_2O_3$	$\boldsymbol{v}_8 = P_2O_5$	$\boldsymbol{v}_9 = H_2S$
Fe	2	0	0	0.996	0	0	0	0	0
S	3	0	0	0	3	3	0	0	1
O	12	0	4	9	9	12	3	5	0
Pr	0	1	0	0	0	1.998	0	0	0
Tl	0	1	0	0	1.987	0	0	0	0
Te	0	3	0	0	0	0	2	0	0
H	0	0	3	6	0	0	0	0	2
P	0	0	1	2	0	0	0	2	0

one obtains the following vector equation

$$z_1\boldsymbol{v}_1 + z_2\boldsymbol{v}_2 + z_3\boldsymbol{v}_3 = z_4\boldsymbol{v}_4 + z_5\boldsymbol{v}_5 + z_6\boldsymbol{v}_6 + z_7\boldsymbol{v}_7 + z_8\boldsymbol{v}_8 + z_9\boldsymbol{v}_9,$$

i. e.,

$$z_1\,(2,3,12,0,0,0,0,0)^{\mathrm{T}} + z_2\,(0,0,0,1,1,3,0,0)^{\mathrm{T}} + z_3\,(0,0,4,0,0,0,3,1)^{\mathrm{T}} = z_4\,(0.996,0,9,0,0,0,6,2)^{\mathrm{T}} + z_5\,(0,3,9,0,1.987,0,0,0)^{\mathrm{T}} + z_6\,(0,3,12,1.998,0,0,0,0)^{\mathrm{T}} + z_7\,(0,0,3,0,0,2,0,0)^{\mathrm{T}} + z_8\,(0,0,5,0,0,0,0,2)^{\mathrm{T}} + z_9\,(0,1,0,0,0,0,2,0)^{\mathrm{T}},$$

from where follows this system of linear equations

$$\begin{gathered} 2z_1 = 0.996z_4, \\ 3z_1 = 3z_5 + 3z_6 + z_9, \\ 12z_1 + 4z_3 = 9z_4 + 9z_5 + 12z_6 + 3z_7 + 5z_8, \\ z_2 = 1.998z_6, \\ z_2 = 1.987z_5, \\ 3z_2 = 2z_7, \\ 3z_3 = 6z_4 + 2z_9, \\ z_3 = 2z_4 + 2z_8, \end{gathered}$$

From the last system one obtains the required solution. This show that the vectors $\boldsymbol{v}_i$ $(1 \leq i \leq 9)$ generate a vector space V and they are linearly dependent. The balanced reaction has this form

$$\begin{gathered} 17839883133\ \mathrm{Fe_2(SO_4)_3} + 12843034110\ \mathrm{PrTlTe_3} + 81542933266\ \mathrm{H_3PO_4} \rightarrow 35823058500\ \mathrm{Fe_{0.996}(H_2PO_4)_2{\cdot}H_2O} + 6463530000\ \mathrm{Tl_{1.987}(SO_3)_3} + 6427945000\ \mathrm{Pr_{1.998}(SO_4)_3} + 19264551165\ \mathrm{Te_2O_3} + 4948408133\ \mathrm{P_2O_5} + 14845224399\ \mathrm{H_2S}. \end{gathered}$$

3° Next, the case when the chemical reaction is non-unique will be considered, *i. e.*, when it has infinite number of solutions.

Example 7. Consider double reduction of titanium dioxide with carbon and chlorine given by the reaction

$$z_1\,\mathrm{TiO_2} + z_2\,\mathrm{C} + z_3\,\mathrm{Cl_2} \rightarrow z_4\,\mathrm{TiCl_4} + z_5\,\mathrm{CO} + z_6\,\mathrm{CO_2}.$$

This reaction plays an important role in theory of metallurgical processes, especially in processes of direct reduction of metal oxides. Sure, that this reaction is not unique, and there are many other reactions of that kind, which may to be used for analysis of this particular case.

From the following scheme

	$\boldsymbol{v}_1 = \mathrm{TiO_2}$	$\boldsymbol{v}_2 = \mathrm{C}$	$\boldsymbol{v}_3 = \mathrm{Cl_2}$	$\boldsymbol{v}_4 = \mathrm{TiCl_4}$	$\boldsymbol{v}_5 = \mathrm{CO}$	$\boldsymbol{v}_6 = \mathrm{CO_2}$
Ti	1	0	0	1	0	0
O	2	0	0	0	1	2
C	0	1	0	0	1	1
Cl	0	0	2	4	0	0

follows this vector equation

$$z_1\boldsymbol{v}_1 + z_2\boldsymbol{v}_2 + z_3\boldsymbol{v}_3 = z_4\boldsymbol{v}_4 + z_5\boldsymbol{v}_5 + z_6\boldsymbol{v}_6,$$

i. e.,

$$(z_1, 2z_1, 0, 0)^{\mathrm{T}} + (0, 0, z_2, 0)^{\mathrm{T}} + (0, 0, 0, 2z_3)^{\mathrm{T}} = (z_4, 0, 0, 4z_4)^{\mathrm{T}} + (0, z_5, z_5, 0)^{\mathrm{T}} + (0, 2z_6, z_6, 0)^{\mathrm{T}},$$

or

$$(z_1, 2z_1, z_2, 2z_3)^{\mathrm{T}} = (z_4, z_5 + 2z_6, z_5 + z_6, 4z_4)^{\mathrm{T}},$$

i. e., immediately follows this system of linear equations

$$\begin{gathered} z_1 = z_4, \\ 2z_1 = z_5 + 2z_6, \\ z_2 = z_5 + z_6, \\ 2z_3 = 4z_4, \end{gathered}$$

which general solution is $z_3 = 2z_1$, $z_4 = z_1$, $z_5 = -2z_1 + 2z_2$, $z_6 = 2z_1 - z_2$, where z_1 and z_2 are arbitrary complex numbers.

Now, balanced general chemical reaction has this form

$$z_1\,\mathrm{TiO_2} + z_2\,\mathrm{C} + 2z_1\,\mathrm{Cl_2} \rightarrow z_1\,\mathrm{TiCl_4} + (-2z_1 + 2z_2)\,\mathrm{CO} + (2z_1 - z_2)\,\mathrm{CO_2}, \quad (\forall z_1, z_2 \in \mathbb{C}).$$

According to the Definition 8, the vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$, $\boldsymbol{v}_4$, $\boldsymbol{v}_5$ and $\boldsymbol{v}_6$ of the molecules of the chemical reaction generate infinite number of vector spaces V_∞, and according to the Definition 10 they are linearly dependent, *i. e.*, we have an infinite number of solutions $(z_1, z_2, 2z_1, z_1, -2z_1 + 2z_2, 2z_1 - z_2)$, $(\forall z_1, z_2 \in \mathbb{C})$, that means that the this chemical reaction is non-unique.

6 DISCUSSION

In his previous work[22], the author announced a cleaning of chemistry from barren intuitionism and its substitution by an elegant formalism from one side, and substitution of the old chemical traditionalism by a new mathematical generalism, from other side. This announcement is realized in this work that gives a new contribution to the theory as well as practice of balancing chemical equations.

The complex vector method of balancing chemical equations augmented the research field in chemistry and made obsolete the old traditional approach, and gave reliable results for paradox resolution.

By this work will begin consideration of paradoxes in chemistry as a serious object, and it will increase researchers' carefulness to avoid appearance of paradoxes.

7 CONCLUSION

The new complex mathematical method of balancing chemical equations, which was used for the solution of a paradox is farewell to the chemical tradition, which still

respects the composers of *general chemistry* textbooks, that affirm that chemistry is a science which studies the structure of substances, how they react when combined or in contact, and how they behave under different conditions. These subjects include a great part of the matter to which chemistry was applied.

In this work the *foundation of chemistry* is enriched by one more new topic, and a contribution to a new formalization of chemistry founded by virtue of a new complex vector method of balancing chemical equations is offered. This work opens doors for the next research in chemistry for diagnostic of paradoxes and their resolution. It will accelerate the newest mathematical research in chemistry and it will surmount the barriers hampering the development of chemistry.

This work is a critical survey that requires changes of chemical thinking. Hence, it must be distinguished from the uncritical penetration, in which chemistry itself is developed sometimes.

ACKNOWLEDGEMENT

Author would like to thank Prof. Dr Valery C. Covachev from the Bulgarian Academy of Sciences and to the Dutch chemist Marten J. Ten Hoor for reading the article and for their remarks.

In addition, author would like to thank Prof. Dr Franc Vodopivec from the University of Ljubljana for his helpful suggestions.

8 REFERENCES

[1] Petruševski, V. M.; *Glas. hem. tehnol. Makedonija 17* (**1998**), 141
[2] Simons, J. H.; *J. Chem. Educ. 3* (**1926**), 1305
[3] Ten Hoor, M. J.; *J. Chem. Educ. 74* (**1997**), 1367
[4] Herndon, W. C.; *J. Chem. Educ. 74* (**1997**), 1359
[5] Bottomley, J.; *Chem. News 37* (**1878**), 110
[6] Johnson, O. C.; *Chem. News* 42 (**1880**), 51
[7] Waldbauer, L. J.; Thrun, W. E.; *J. Chem. Educ. 3* (**1926**), 1430
[8] Jones, M.; *SIAM Rev. 13* (**1971**), 571
[9] Alberty, R. A.; *J. Chem. Educ. 68* (**1991**), 984
[10] Zeggeren, V. F.; Storey, S. H.; *The Computation of Chemical Equlibria,* Cambridge Univ. Press: London, 1970
[11] Smith, W. R.; Missen, R. W.; *Chemical Reaction Equilibrium Analysis*: *Theory and Algorithms,* Wiley: New York, 1982
[12] Risteski, I. B.; *SIAM Problems & Solutions*, 2007, 1–10
[13] Risteski, I. B.; *J. Korean Chem. Soc. 52* (**2008**), 223
[14] Crocker, R.; *J. Chem. Educ. 45* (**1968**), 731
[15] Halmos, P.; *Finite Dimensional Vector Spaces,* Springer Vlg: Berlin, 1998
[16] Kurepa, S.; *Konačno dimezionalni vektorski prostori i primjene,* 5-to izd., Sveučilisna naklada Liber: Zagreb, 1990
[17] Levinthal, C.; *J. Chim. Phys. 65* (**1968**), 44
[18] Patani, G. A.; LaVoie, E. J.; *Chem. Rev. 96* (**1996**), 3147
[19] Tóth, Z.; *J. Chem. Educ. 74* (**1997**), 1270
[20] Ludwig, O.; *J. Chem. Educ. 74* (**1997**), 1270
[21] Risteski, I. B.; *Internat. J. Math. Manuscripts 1* (**2007**), 180
[22] Risteski, I. B.; *Bol. Soc. Quím. Méx. 2* (**2008**), 104
[23] Risteski, I. B.; *J. Chin. Chem. Soc. 56* (**2009**), 65

UDK 621.791:669.715
Original scientific article/Izvirni znanstveni članek

ISSN 1580-2949
MTAEC9, 44(4)205(2010)

APPLICATION OF GREY RELATION ANALYSIS (GRA) AND TAGUCHI METHOD FOR THE PARAMETRIC OPTIMIZATION OF FRICTION STIR WELDING (FSW) PROCESS

UPORABA GREYJEVE ANALIZE (GRA) IN TAGUCHIJEVE METODE ZA PARAMETRIČNO OPTIMIZACIJO VARJENJA Z VRTILNO-TORNIM PROCESOM (FSW)

Hakan Aydin[1], Ali Bayram[2], Ugur Esme[3*], Yigit Kazancoglu[4], Onur Guven[5]

[1,2]Uludag University, Faculty of Engineering and Architecture, Department of Mechanical Engineering, 16059, Gorukle-Bursa/Turkey
[3]Mersin University Tarsus Technical Education Faculty Department of Mechanical Education, 33480, Tarsus-Mersin/Turkey
[4]Izmir University of Economics, Department of Business Administration, 35330, Balcova-Izmir/Turkey
[5]Mersin University, Engineering Faculty, Department of Mechanical Engineering, 33400, Mersin/Turkey
uguresme@gmail.com

Prejem rokopisa – received: 2010-02-15; sprejem za objavo – accepted for publication: 2010-02-25

This study focused on the multi-response optimization of friction stir welding (FSW) process for an optimal parametric combination to yield favorable tensile strength and elongation using the Taguchi based Grey relational analysis (GRA). The objective functions have been selected in relation to parameters of FSW parameters; rotating speed, welding speed and tool shoulder diameter. The experiments were planned using Taguchi's L_8 orthogonal array. Multi-response optimization was applied using Grey relation analysis and Taguchi approach to solve the problem. The significance of the factors on overall quality characteristics of the welding process has also been evaluated quantitatively by the analysis of variance (ANOVA) method. Optimal results have been verified through confirmation experiments. This study has also showed the application feasibility of the Grey relation analysis in combination with Taguchi technique for continuous improvement in welding quality.

Keywords: Friction stir welding, Grey relation analysis, Taguchi method, optimization

Cilj raziskave je bila večodgovorna optimizacija procesa varjenja z vrtilnim trenjem (FSW) za kombinacijo parametrov za dosego ugodnih raztržne trdnosti in raztezka z uporabo Taguchi-Greyjeve racionalne analize. Primerne funkcije so bile izbrane v povezavi s FSW-parametri: hitrost vrtenja, hitrost varjenja in premer ramen orodja. Preizkusi so bili izvršeni z uporabo Taguchijeve ortogonalne mreže L_8. Odgovori so bili optimizirani z uporabo Greyjeve analize s Taguchijevim približkom. Pomen dejavnikov splošnih značilnosti procesa varjenja je bil kvantitativno analiziran z analizo variance (ANOVA). Optimalni rezultati so bili preverjeni s preizkusi. Raziskava je tudi pokazala uporabnost Greyjeve analize v povezavi s Taguchijevo tehniko za stalno izboljšanje tehnike varjenja.

Ključne besede: vrtilno torno varjenje, Greyjeva analiza odvisnosti, Taguchijeva metoda, optimizacija

1 INTRODUCTION

In today's manufacturing world, quality is of vital importance. Quality can be defined as the degree of customer's satisfaction as provided by the procured product. The product quality depends on the desired requirements gained in the product that suits its functional requirements in various areas of application.[1]

In the field of welding, weld quality mainly depends on the welding type, mechanical properties of the weld metal and heat affected zone (HAZ), which in turn is influenced by metallurgical characteristics and chemical compositions of the weld.[1] Moreover, these mechanical-metallurgical features of the weldment directly related to welding process parameters. In other words, weld quality depends on welding process parameters.[1]

The welding of aluminum and its alloys has always represented a great challenge for designers and technologists.[2] Friction stir welding (FSW) is a welding technique, patented in 1991 by TWI.[3,4]

As a solid-state process, FSW can avoid the formation of solidification cracking and porosity associated with fusion (FSW) welding processes and significantly improve the weld properties of aluminum alloys.[4,5] As illustrated in **Figure 1**, this technique involves a non-con-

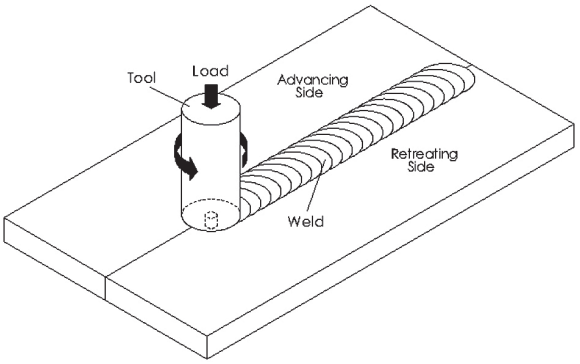


Figure 1: Schematic representation of FSW
Slika 1: Shematična predstavitev FSW

sumable, cylindrical, rotating tool (usually hardened steel) which moves between the seam of two butted plates and stirs them together.[2–6] The effect of friction stir welding on the material is both on heat flow and plastic strain. The heat is generated by friction between the tool shoulder and the top of the sheets.

When compared to traditional welding techniques, FSW strongly reduces the presence of distortions and residual stresses.[8–11] The FSW process is a solid state process and therefore a solidification structure is absent in the weld. A detailed description of the FSW process is present in the literature.[7–16] The process can be easily monitored and replicated. In addition, it does not produce any major safety hazards, such as fume or radiation.[17] This process is used to bond metals without fusion or filler materials.[18] FSW of aluminum has several advantages over fusion welding processes. Problems arising from fusion welding of aluminum alloys, such as solidification cracking, liquation cracking and porosity, are eliminated with FSW, due to its solid-state nature.[17–23]

The Taguchi method is very popular for solving optimization problems in the field of production engineering.[24,25] The method utilizes a well-balanced experimental design (allows a limited number of experimental runs) called orthogonal array design, and signal-to-noise ratio (*S/N* ratio), which serve the objective function to be optimized (maximized) within experimental domain.[24] However, traditional Taguchi method cannot solve multi-objective optimization problem. To overcome this, the Taguchi method coupled with Grey relational analysis has a wide area of application in manufacturing processes. This approach can solve multi-response optimization problem simultaneously.[26,27]

Planning the experiments through the Taguchi orthogonal array has been used quite successfully in process optimization by Chen and Chen,[28] Fung and Kang,[29] Tang et al,[30] Vijian and Arunachalam[31] as well as Zhang et al.[32] Therefore, this study applied a Taguchi L_8 orthogonal array to plan the experiments on FSW welding process. Three controlling factors including rotating speed (*w*), welding speed (*V*) and shoulder diameter (*d*) were selected. The Grey relational analysis is then applied to examine how the welding process factors influence the tensile strength (*TS*) and percent elongation (*e*). An optimal parameter combination was then obtained. Through analyzing the Grey relational grade matrix, the most influential factors for individual quality targets of FSW welding process can be identified. Additionally, the analysis of variance (ANOVA) was also utilized to examine the most significant factors for the tensile strength and elongation in FSW welding process.

2 GREY RELATIONAL ANALYSIS (GRA)

2.1 Data Preprocessing

In Grey relational analysis, experimental data i.e., measured features of quality characteristics are first normalized ranging from zero to one. This process is known as Grey relational generation. Next, based on normalized experimental data, Grey relational coefficient is calculated to represent the correlation between the desired and actual experimental data. Then overall Grey relational grade is determined by averaging the Grey relational coefficient corresponding to selected responses.[26] The overall performance characteristic of the multiple response process depends on the calculated Grey relational grade. This approach converts a multiple response process optimization problem into a single response optimization situation with the objective function is overall Grey relational grade. The optimal parametric combination is then evaluated which would result highest Grey relational grade. The optimal factor setting for maximizing overall Grey relational grade can be performed by Taguchi method.[26,33]

In Grey relational generation, the normalized E corresponding to the smaller-the-better (SB) criterion which can be expressed as:

$$x_i(k) = \frac{\max y_i(k) - y_i(k)}{\max y_i(k) - \min y_i(k)} \tag{1}$$

TS should follow the larger-the-better (LB) criterion, which can be expressed as:

$$x_i(k) = \frac{y_i(k) - \min y_i(k)}{\max y_i(k) - \min y_i(k)} \tag{2}$$

where $x_i(k)$ is the value after the Grey relational generation, min $y_i(k)$ is the smallest value of $y_i(k)$ for the k^{th} response, and max $y_i(k)$ is the largest value of $y_i(k)$ for the k^{th} response.[26] An ideal sequence is $x_0(k)$ (k = 1, 2, 3....., 8 for the responses. The definition of Grey relational grade in the course of Grey relational analysis is to reveal the degree of relation between the 16 sequences [$x_0(k)$ and $x_i(k)$, i = 1, 2, 3.....,8]. The Grey relational coefficient # can be calculated as:

$$\xi_i(k) = \frac{\Delta_{\min} - \psi\Delta_{\max}}{\Delta_{0i}(k) + \psi\Delta_{\max}} \tag{3}$$

where $\Delta_{0i} = \left\| x_0(k) - x_i(k) \right\|$ = difference of the absolute value $x_0(k)$ and $x_i(k)$; ψ is the distinguishing coefficient $0 \le \psi \le 1$; $\Delta_{\min} = \forall j^{\min} \in i \forall k^{\min} \left\| x_0(k) - x_j(k) \right\|$ = the smallest value of Δ_{0i}; and $\Delta_{\max} = \forall j^{\max} \in i \forall k^{\max} \left\| x_0(k) - x_j(k) \right\|$ = largest value of Δ_{0i}. After averaging the Grey relational coefficients, the Grey relational grade γ_i can be computed as:

$$g_i = \frac{1}{n} \sum_{k=1}^{n} \xi_i(k) \tag{4}$$

where n is the number of process responses. The higher value of Grey relational grade corresponds to intense relational degree between the reference sequence $x_0(k)$ and the given sequence $x_i(k)$. The reference sequence $x_0(k)$ represents the best process sequence; therefore,

Table 1: Chemical and mechanical properties of AA1050 aluminum alloy
Tabela 1: Kemična sestava in mehanske lastnosti zlitine AA1050

Chemical composition *w*/%	Al	Mg	Si	Mn	Zn	Fe	Ti	Sn
	Balance	0.007	0.18	0.05	0.033	0.30	0.009	0.182
Mechanical properties	Yield strength (MPa)		Tensile strength (MPa)		Elongation (%)		Vickers Hardness (*HV*)	
	155		175		4		50	

Table 2: Process parameters and their limits
Tabela 2: Parametri in limite procesa

Parameters	Notation	Unit	Levels of factors			
			1	2	3	4
Rotating speed	*w*	r/min	740*	1070	1520	2140
Welding speed	*V*	mm/min	80*	224	–	–
Shoulder diameter	*d*	mm	15*	20	–	–

*Initial factor settings

Table 3: Orthogonal array L_8 of the experimental runs and results
Tabela 3: Ortogonalna mreža L_8 eksperimentalnih varkov in rezultatov

Run no	Experimental results					
	w	*V*	*d*	*TS*/ MPa	*E*/%	Fracture location HAZ: Heat affected zone TMAZ: Thermo-mechanically affected zone NZ: Nugget zone BM: Base metal
1	1	1	1	93	14.8	The interface between HAZ and TMAZ on the retreating side
2	1	2	2	65	5.50	NZ
3	2	1	1	90	17.3	BM
4	2	2	2	89	13.5	HAZ on the retreating side
5	3	1	2	92	18.3	BM
6	3	2	1	93	14.5	HAZ on the advancing side
7	4	1	2	94	19.1	BM
8	4	2	1	92	14.1	HAZ on the advancing side

higher Grey relational grade means that the corresponding parameter combination is closer to the optimal. The mean response for the Grey relational grade with its grand mean and the main effect plot of Grey relational grade are very important because optimal process condition can be evaluated from this plot.[26]

3 EXPERIMENTAL PROCEDURE AND TEST RESULTS

3.1 Experimental Details

AA1050-H22 aluminum alloy material was used as a workpiece material with the thickness of 4 mm. The workpieces were machined out in 360 mm lengths and 200 mm widths. The mechanical properties and percent composition of workpiece material is listed in **Table 1**.

1.2367 (X38CrMoV5-3) hardened and threaded (left screw with 0.8 mm pitch) pins with the shoulder diameters of 15 mm and 20 mm were used as welding tools. The dimensions of the welding tools are shown in **Figure 2**.

The pre-machined aluminum plates were fixed rigidly on the table of the vertical semiautomatic milling machine for lap joint of FSW as shown in **Figure 3**.

The rotating tool was fixed to the spindle of the milling machine and then the spindle of the milling machine was adjusted at an angle of 2–3° away from the spindle travel path. To generate the required pre-frictional heating, the shoulder of the rotating tool was held in its ini-

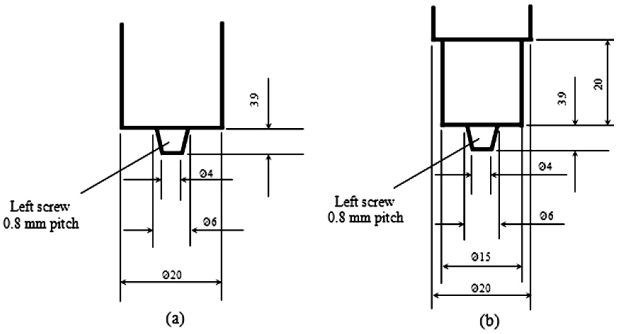


Figure 2: Dimensions of the welding tools used in the experiments: a) 20 mm shoulder diameter, b) 15 mm shoulder diameter
Slika 2: Dimenzije pri preizkusih uporabljenih varilnih orodij: a) premer ramena 20 mm, b) premer ramena 15 mm

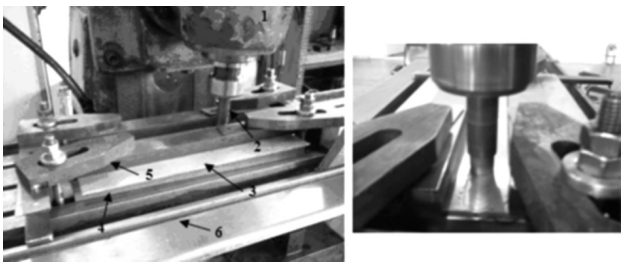


Figure 3: FSW applications on conventional vertical milling machine; 1 milling head, 2 welding tool, 3 aluminum plates, 4 Steel backing plate, 5 clamping setup, 6 machine table
Slika 3: Uporaba FSW na pokončnem vrtalnem stroju; 1 vrtalna glava, 2 varilno orodje, 3 aluminijevi plošči, 4 jeklena oporna plošča, 5 prijemno orodje, 6 delovna miza stroja

tial position for 30 s rubbing with the surface of the workpiece.

Figure 4 shows the dimensions of the tensile test specimens prepared according to TS138 EN10002-1 standard. The tensile tests were carried out at a room temperature and crosshead speed of 10 mm/min using using a ZWICK Z-050 tensile testing machine. Each tensile test is the average of four specimens cut from the same joint.

3.2 Process Parameters and Test Results

In full factorial design, the number of experimental runs exponentially increases as the number of factors as well as their level increases. This results huge experimentation cost and considerable time.[26] So, in order to compromise these two adverse factors and to search the optimal process condition through a limited number of experimental runs, Taguchi's L_8 orthogonal array consisting of 8 sets of data has been selected to optimize the multiple performance characteristics of FSW. Experiments have been conducted with the process parameters given in **Table 2**, to obtain butt welding on AA1050-H22 aluminum 4 mm thickness with (360 × 200) mm dimensions by FSW welding process.

Table 3 shows the selected design matrix based on Taguchi L_8 orthogonal array consisting of 8 sets of coded conditions and the experimental results for the responses of TS and E. All these data have been utilized for analysis and evaluation of optimal parameter combination required to achieve desired quality weld within the experimental domain.

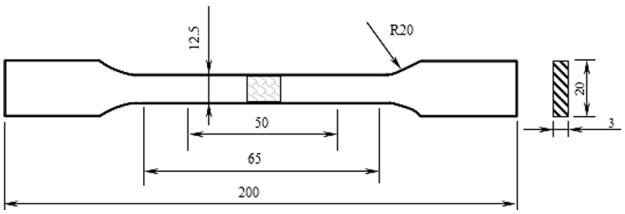


Figure 4: Dimensions of tensile test specimens
Slika 4: Dimenzije raztržnega preizkušanca

4 PARAMETRIC OPTIMIZATION OF FSW PROCESS

4.1 Evaluation of Optimal Process Condition

First, by using Eqs. (1) and (2), experimental data have been normalized to obtain Grey relational generation.[26] The normalized data and Δ_{0i} for each of the responses have been furnished in **Table 4** and **Table 5** respectively. For *TS* larger-the-better (LB) and for *E* smaller-the-better (SB) criterion has been selected.

Table 4: Grey relational generation of each performance characteristics
Tabela 4: Generiranje Greyjeve odvisnosti za značilnosti vsakega preizkusa

Run no	*TS*	*E*
	Larger-the-better	Smaller-the-better
Ideal sequence	1	1
1	0.966	0.463
2	0.000	1.000
3	0.862	0.132
4	0.828	0.412
5	0.931	0.059
6	0.966	0.338
7	1.000	0.000
8	0.931	0.368

Table 5: Evaluation of Δ_{0i} for each of the responses
Tabela 5: Ocena Δ_{0i} za vsak odgovor

Run no	*R*a	*HV*
Ideal sequence	1	1
1	0.034	0.537
2	1.000	0.000
3	0.138	0.868
4	0.172	0.588
5	0.069	0.941
6	0.034	0.662
7	0.000	1.000
8	0.069	0.632

Table 6 shows the calculated Grey relational coefficients (with the weights of $\psi_{TS} = 0.7$ and $\psi_E = 0.3$) of each performance characteristic using Eq. (3).

Table 6: Grey relational coefficient of each performance characteristics ($\psi_{TS} = 0.7$, $\psi_E = 0.3$)
Tabela 6: Greyjevi koeficienti odvisnosti za značilnosti vsakega preizkusa ($\psi_{TS} = 0.7$, $\psi_E = 0.3$)

Run no	TS	E
Ideal sequence	1	1
1	0.953	0.359
2	0.412	1.000
3	0.829	0.564
4	0.795	0.741
5	0.906	0.531
6	0.951	0.685
7	1.000	0.507
8	1.000	0.706

The Grey relational coefficients, given in **Table 7**, for each response have been accumulated by using Eq. (4) to evaluate Grey relational grade, which is the overall representative of all the features of FSW quality. Thus, the multi-criteria optimization problem has been transformed into a single equivalent objective function optimization problem using the combination of Taguchi approach and Grey relational analyses. Higher is the value of Grey relational grade, the corresponding factor combination is said to be close to the optimal[26].

Table 7: Grey relational grade
Tabela 7: Stopnje Greyjeve odvisnosti

Run no	Grey relational grade	Rank
1	0.7747	6
2	0.5882	8
3	0.7493	7
4	0.7787	5
5	0.7936	4
6	0.8710	2
7	0.8520	3
8	0.9119	1

Table 8 shows the *S/N* ratio based on the larger-the-better criterion for overall Grey relational grade calculated by using Eq. (5).

$$S/N = -10\lg\left[\frac{1}{n}\sum_{i=1}^{n}\frac{1}{y_i^2}\right] \quad (5)$$

where n is the number of measurements, and y_i is the measured characteristic value.

Graphical representation of S/N ratio for overall Grey relational grade is shown in **Figure 5**. The dashed line is the value of the total mean of the S/N ratio.

As indicated in **Figure 5**, the optimal condition for the FSW of aluminum alloy becomes $w_4V_1d_1$. **Table 9** shows the mean Grey relational grade ratio for each level of the process parameters.

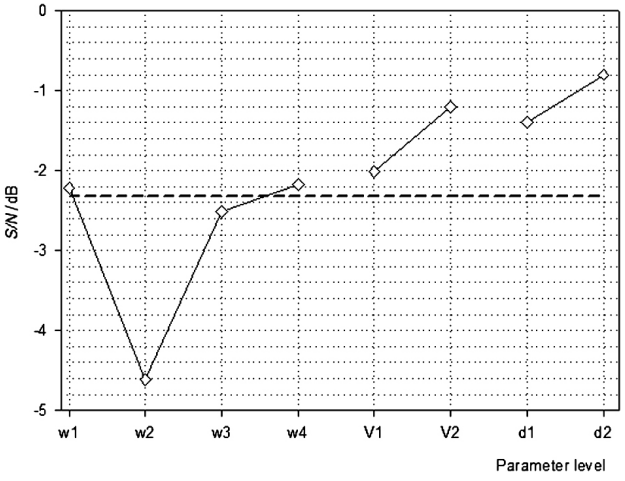


Figure 5: *S/N* ratio plot for the overall Grey relational grade
Slika 5: Razmerje *S/N* za splošno Greyjevo odvisnost

Table 8: *S/N* ratio for overall Grey relational grade
Tabela 8: Razmerje *S/N* za splošno Greyjevo stopnjo

Run no	S/N
1	–2.22
2	–4.61
3	–2.51
4	–2.17
5	–2.01
6	–1.20
7	–1.39
8	–0.80

Table 9: Response table for the mean Grey relational grade
Tabela 9: Odgovori za povprečno stopnjo po Greyu

Factors	Grey relational grade				
	Level 1	Level 2	Level 3	Level 4	Max-Min
w	0.68	0.76	0.83	0.88	0.20
V	0.79	0.78	–	–	0.01
d	0.83	0.75	–	–	0.08

Total mean Grey relational grade = 0.79

4.2 Analysis of Variance (ANOVA)

The purpose of the analysis of variance (ANOVA) is to investigate which welding parameters significantly affect the performance characteristic.[26,33,34] This is accomplished by separating the total variability of the grey relational grades, which is measured by the sum of the squared deviations from the total mean of the grey relational grade, into contributions by each welding parameters and the error.[26,34] Thus

$$SS_T = SS_F + SS_e \quad (6)$$

where

$$SS_T = \sum_{j=1}^{p}(\gamma_j - \gamma_m)^2 \quad (7)$$

and

SS_T	Total sum of squared deviations about the mean
γ_j	Mean response for j^{th} experiment
γ_m	Grand mean of the response
p	Number of experiments in the orthogonal array
SS_F	Sum of squared deviations due to each factor
SS_e	Sum of squared deviations due to error

In addition, the F test was used to determine which welding parameters have a significant effect on the

Table 10: ANOVA results of FSW process
Tabela 10: ANOVA-rezultati FSW-procesa

Parameter	Degree of Freedom	Sum of Square	Mean Square	*F*	Contribution (%)
w	3	0.050	0.020	0.88	65.61
V	1	0.015	0.002	0.62	19.68
D	1	0.010	0.010	0.47	13.12
Error	1	0.0012	0.010		1.58
Total	6	0.0762			100

performance characteristic. Usually, the change of the welding parameter has a significant effect on the performance characteristic when the F value is large. ANOVA for overall Grey relational grade is shown in **Table 10**.

4.3 Confirmation Test

After evaluating the optimal parameter settings, the next step is to predict and verify the enhancement of quality characteristics using the optimal parametric combination. The estimated Grey relational grade $\hat{\gamma}$ using the optimal level of the design parameters can be calculated as:

$$\hat{\gamma} = \gamma_m + \sum_{i=1}^{o} (\bar{\gamma}_j - \gamma_m) \quad (8)$$

where γ_m is the total mean Grey relational grade, $\bar{\gamma}_i$ is the mean Grey relational grade at the optimal level, and o is the number of the main design parameters that affect the quality characteristics.[26] **Table 11** indicates the comparison of the predicted tensile strength and elongation with that of actual by using the optimal welding conditions. Good agreement between the actual and predicted results has been observed (improvement in overall Grey relational grade was found to be as 0.20).

Table 11: Results of confirmation test
Tabela 11: Rezultati preizkusov preverjanja

	Initial factor settings	Optimal process condition	
		Prediction	Experiment
Factor levels	$w_1V_1d_1$	$W_4V_1d_1$	$W_4V_1d_1$
TS	93		96
E	14.8		12.3
S/N ratio of overall Grey relational grade	–2.22	–0.58	–1.80
Overall Grey relational grade	0.72	0.89	0.92

Improvement in Grey relational grade = 0.20

In Taguchi method, the only performance feature is the overall Grey relational grade; and the aim should be to search a parameter setting that can achieve highest overall Grey relational grade.[26,33] The Grey relational grade is the representative of all individual performance characteristics. In the present study, objective functions have been selected in relation to parameters of tensile strength and elongation. The weight calculations were done by using Analytic Hierarchy Process (AHP) and the weights were found to be as 0.70 and 0.30 for the responses of tensile strength and elongation respectively.

The results showed that using optimal parameter setting ($w_4V_1d_1$) caused lower elongation with higher tensile strength.

5 CONCLUSION

Taguchi method is a very effective tool for process optimization under limited number of experimental runs. Essential requirements for all types of welding processes are higher tensile strength with lower elongation. This study has concentrated on the application of Taguchi method coupled with Grey relation analysis for solving multi criteria optimization problem in the field of friction stir welding process. Experimental results have shown that tensile strength and elongation of welded AA1050-H22 aluminum alloy are greatly improved by using Grey based Taguchi method.

6 REFERENCES

[1] L. Kukielka, Journal of Mechanical Technology, 19 (**1989**), 319–356
[2] H. Aydin, A. Bayram, A. Uguz, S. K. Akay, Mater Des 30 (**2009**), 2211–2221
[3] W. M. Thomas, Friction Stir Butt Welding International Patent Application, No. PCT/GB92 Patent Application No. 9125978.8, (1991)
[4] C. J. Dawes, W. M. Thomas, Welding Journal 75 (**1996**), 41–45
[5] Y. C. Chen, H. J. Liu, J. C. Feng, Journal of Materials Science 41 (**2006**), 297–299
[6] M. Boz, A. Kurt, Mater Des 25 (**2004**), 343–347
[7] R. A. Murr, Scr Mater, 45 (**2001**), 75–80
[8] G. Bussu, P. E. Irving, Int J Fatigue, 25 (**2003**), 77–88
[9] R. John, K.V. Jata, K. Sadananda, Int J Fatigue 25 (**2003**), 939–948
[10] K. V. Jata, K. K. Sankaran, J. Ruschau, Metall Mater Trans, 31A (**2000**), 2181–2192
[11] M. K. Kulekci, A. Sik, E, Kaluc, Int J Adv Manuf Technol, doi10.1007/s00170-006-0901-z
[12] W. B. Lee, S. B. Jung, Mater Lett, 6 (**2005**), 1041–1046
[13] W. B. Lee, M. Y. Yeon, S. B. Jung, Scr Mater, 49 (**2005**), 423–428
[14] S. Lim, S. Kim, C. G. Lee, C. D. Yim, S. G. Kim, Metals Materials Transactions, 36A (**2005**), 1609–1612
[15] R. S. Mishra, Z.Y. Ma, Mater Sci Eng, R50 (**2005**), 1–78
[16] M. P. Miles, B. J. Decker, T. W. Nelson, Metall Mater Trans A, 35A (**2004**), 3461–3468
[17] W. M. Thomas, P. L. Threadgil, E. D. Nicholas, Sci Technol Weld Join, 4 (**1999**), 365–372
[18] M. K. Kulekci, F. Mendi, I. Sevim, O. Basturk, Metalurgija, 44 (**2005**), 209–213
[19] G. Çam, V. Ventzke, J. F. Dos Santos, M. Koçak, G. Jennequin, P. Gontier-Maurin, Sci Technol Weld Join, 4 (**1999**), 317–323
[20] O. V. Flores, C. Kennedy, L. E. Murr, D. Brown, S. Pappu, B. M. Nowak, J. C. McClure, Scr Mater 38 (**1998**), 703–708
[21] H. Uzun, C. D. Donne, A. Argagnotto, T. Ghidini, C. Gambaro, Mater Des, 26 (**2005**), 41–46
[22] K. Colligan, Welding Research, 6 (**1999**), 229
[23] S. Benavides, Y. Li, L. E. Murr, D. Brown, J.C. McClure, Scr Mater, 41 (**1999**), 809–815
[24] W. H. Yang, Y. S. Tarng, J Mater Process Technol, 84 (**1998**), 122–129
[25] H. Rowlands, J. Antony, G. Knowles, The TQM Magazine, 12 (**2000**), 78–83
[26] S. Datta, A. Bandyopadhyay, P. K. Pal, Int J Adv Manuf Technol, 39 (**2008**), 1136–1143
[27] S. H Lim, C. M Lee, WJ. Chung, International Journal of Precision Engineering and Manufacturing, 7 (**2006**), 18–23
[28] D. C Chen, C. F. Chen, J Mater Process Technol, 190 (**2007**), 130–137

[29] C. P. Fung, P. C. Kang, J Mater. Process Technol, 170 (**2005**), 602–610
[30] S. H Tang, V. J Tan, S. M Sapuan, S Sulaiman, N. Ismail, R. Samin, J Mater Process Technol, 182 (**2007**), 418–426
[31] P. Vijian, V. P. Arunachalam, J Mater Process Technol, 180 (**2006**), 161–166
[32] J. Z. Zhang, J. C. Chen, E. D. Kirby, J Mater Process Technol, 184 (**2007**), 233–239
[33] U. Esme, M. Bayramoglu, Y. Kazançoglu, S. Özgun, Mater. Tehnol., 43 (**2009**), 143–149
[34] U. Esme, A. Sagbas, F. Kahraman, M.K. Kulekci, Mater. Tehnol., 42 (**2008**), 215–219

UDK 536:543.57:546.92 57 28
Original scientific article/Izvirni znanstveni članek

ISSN 1580-2949
MTAEC9, 44(4)213(2010)

PHASE RELATIONS IN THE Pt–Ag17.5–Si4.5 TERNARY ALLOY

FAZNA ODVISNOST V TERNARNI ZLITINI Pt–Ag17.5–Si4.5

Grega Klančnik, Primož Mrvar, Jožef Medved

University of Ljubljana, Faculty of Natural Science and Engineering, Department of Materials and Metallurgy, Aškerčeva 12, SI-1000 Ljubljana, Slovenia
grega.klancnik@ntf.uni-lj.si

Prejem rokopisa – received: 2009-12-01; sprejem za objavo – accepted for publication: 2010-04-22

An investigation of the Pt–Ag17.5–Si4.5 ternary alloy was carried out. Thermodynamic data on platinum systems are scarce and are mostly based on binary systems. No accurate thermodynamic study of the Pt–Ag–Si ternary system has been published so far. Thermodynamic calculations were used as an indication of the possible phase composition for the investigated platinum-based alloy. The results are based on differential scanning calorimetry (DSC) tests, scanning electron microscopy with energy-dispersive X-ray spectroscopy (SEM – EDS) and X-ray diffraction measurements (XRD). The analyses revealed the presence of two phases that were found after homogenization at 900 °C: a high-temperature phase, Pt_2Si, and a Ag (Pt) solid solution.

Keywords: Pt–Ag–Si alloys, thermal analysis, thermodynamics

V tem prispevku je opisana raziskava, ki je osredinjena na ternarno zlitino Pt-Ag17.5–Si4.5. Za zlitinske sisteme, ki so osnovani na platini, so termodinamski podatki skopi in temeljijo predvsem na binarnih faznih sistemih. Nobena termodinamska študija še ni bila opravljena na ternarnem sistemu Pt–Ag–Si. Za izbiro ustrezne kemijske sestave smo uporabili termodinamski izračun. Prvi rezultati so bili dobljeni z diferenčno vrstično kalorimetrijo (DSC), rastrskim elektronskim mikroskopom z energijsko disperzivno rentgensko spektroskopijo (SEM – EDS) in rentgensko analizo (XRD). Analize so potrdile obstoj dveh binarnih faz, dobljeni po homogenizaciji na 900 °C. Najdeni fazi sta visokotemperaturna faza Pt_2Si in trdna raztopina na osnovi srebra Ag (Pt).

Ključne besede: Pt-Ag-Si zlitine, termična analiza, termodinamika

1 INTRODUCTION

In this study we have investigated an alloy from the ternary Pt–Ag–Si system with the aim to determine the phase composition of the low-melting platinum-based alloy. Using thermodynamic equilibrium calculations of the phase diagram, an indication of a suitable chemical composition for the alloy can be made. From the analysis of the ternary alloy the investigated new information on the extension of the binary phases into the ternary system was expected. Differential scanning calorimetry, scanning electron microscopy with EDS analyses, and X-ray diffraction measurements were used for the investigation.

2 THEORY

The binary boundary systems of the ternary Pt–Ag–Si system according to the literature are shown in **Figure 1**.[2,3] The Pt–Si binary system is rather complex with several intermetallic phases (Pt_3Si, $Pt_{12}Si_5$, Pt_2Si, Pt_6Si_5 and PtSi). The phases of the binary system Pt–Si are presented in **Table 1**. The Ag–Pt binary system is a

Table 1: The binary phases taken from the Pt–Si system
Tabela 1: Faze v binarnem sistemu Pt–Si

Phase	Space group	Structure type	Lattice parameter /nm			Ref
			a	b	c	
Pt3Si	C12/m1	Pt3Ge	0.7724(2)	0.7767(2)	0.5390(2)	5
HT-Pt3Si	Pnma	Fe3C	0.5579	0.7697	0.5520	6
Pt12Si5	I4/m	Ni12P5	0.9607	0.9607	0.5542	6
HT-Pt12Si5*	P4/nZ	/	0.13404	0.13404	0.5451	7
Pt2Si	I4/mmm	ThH2	0.3933	0.3933	0.5910	8
HT-Pt2Si	P6-2M	Fe2P	0.6440	0.6440	0.3573	8
PtSi	Pbnm	FeB	0.5932(1)	0.5595(1)	0.3603(1)	9
PtSi	Pnma	MnP	0.5595	0.3603	0.5932	9
Pt6Si5*	P121/m1	/	0.15308	0.348	0.6120	10
Pt2Si3	P63/mmc	Sc2O2S	0.3841(1)	0.3841	1.1924(5)	11

* – no data present, HT – high-temperature phase

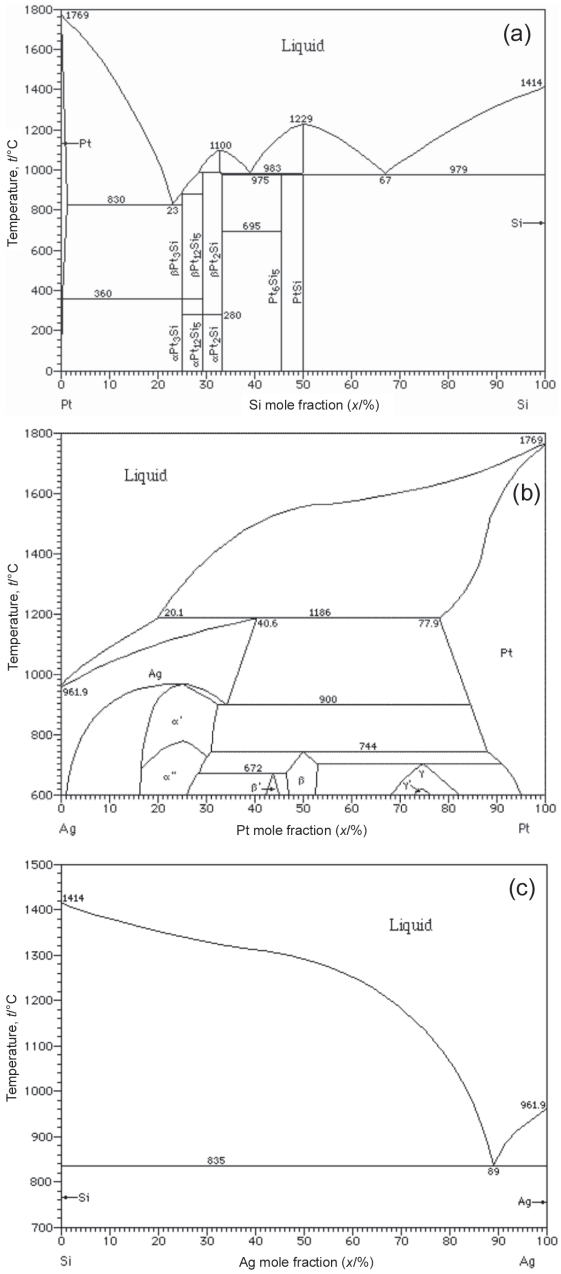


Figure 1: Phase diagrams of the binary systems: Pt–Si[2] **(a)**, Ag–Pt[3] **(b)** and Si–Ag[4] **(c)**

Slika 1: Fazni diagrami binarnih sistemov: Pt–Si[2] **(a)**, Ag–Pt[3] **(b)** and Si–Ag[4] **(c)**

peritectic system with several reactions in the solid state. The Ag–Pt phase diagram assessed by Karakaya and Thompson[3] identified a superstructure at low temperature and a miscibility gap between the silver and the platinum-rich corner.[1] The intermediate phases are known as β (PtAg), γ and γ' (Pt_3Ag). Other phases, such as α', α'' ($PtAg_3$) and β', are also possible. Invariant reactions other than the peritectic are uncertain, because of the lack of conclusive studies. The phase diagram of the Si–Ag binary system is a eutectic type, without intermediate or intermetallic phases.

3 EXPERIMENTAL

The chemical composition of the alloy was selected on the basis of the results of thermodynamic calculations of the liquidus surface for the isothermal section of the ternary system Pt–Ag–Si using Thermo Calc 4 for Windows (TCW 4) with the SSOL4 database.

The composition of the alloy is presented in **Table 2**.

The alloy was melted in a high-purity Al_2O_3 crucible in an argon atmosphere from elements of high-purity grade (99.99). The melting was performed with a high-temperature furnace, where the platinum was heated up to 1800 °C. The temperature control was established with a type-D thermocouple (W- 3 % Re – W- 25 % Re). After adding silver and silicon, the melt was mixed to achieve homogeneity. The alloy was then re-melted and slowly cooled to room temperature. The melting crucible inside the high-temperature furnace is presented in **Figure 2**.

The alloy was investigated with thermal analyses (STA – 449 Jupiter, Netzsch) and X-ray diffraction (XRD) patterns. These X-ray diffraction patterns were recorded at room temperature using Cu Kα1 and Kα2 radiation with a Bruker X8 APEX II CCD diffractometer. The diffraction patterns were compared with those

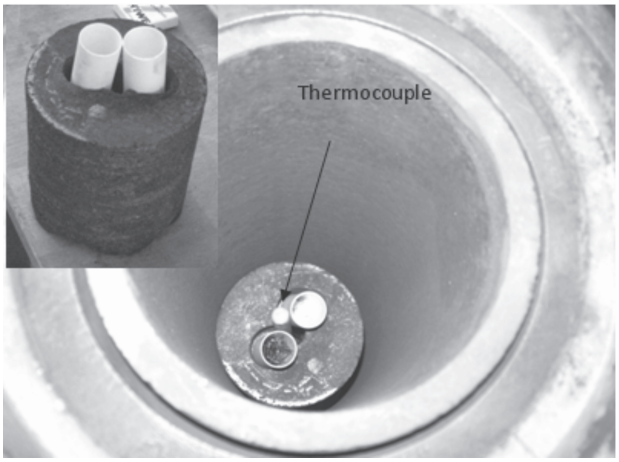


Figure 2: High-temperature furnace and position of the type-D thermocouple

Slika 2: Visokotemperaturna peč in položaj termoelementa tipa D

Table 2: Chemical composition of the investigated alloy

Tabela 2: Kemična sestava preiskovane zlitine

Sample	mass fraction (w/%)			mole fraction (x /%)			mass, m/g		
	Pt	Ag	Si	Pt	Ag	Si	Pt	Ag	Si
1	77.952	17.522	4.526	55.312	22.436	22.277	1.957	0.439	0.113

generated for known compounds and used to calculate the refined unit-cell parameters. The second recording of the X-ray diffraction patterns, using the same radiation, for the aged sample was carried out using PANalytical X-pert PRO MPD for samples with flat surfaces.

The maximum temperature for the STA 449 Jupiter – Netzsch device was 1550 °C for the as-cast sample, with a 20 K/min heating rate up to 1200 °C for the homogenized sample. The measurements were performed in an inert atmosphere of argon, applying the empty crucible as a reference. The heating program for the homogenized sample was as follows: heating at 25 K/min to the homogenization temperature (900 °C), holding for 30 min, and then further heating at 5 K/min up to 1200 °C.

The microstructure was examined on polished specimens using a Jeol 5610 scanning electron microscope (SEM) equipped for energy-dispersive X-ray spectroscopy (EDS). Several EDS point analyses were obtained for each phase.

4 RESULTS

4.1 Thermodynamic calculation

The calculation of the phase diagrams with TCW4 is an important tool for the design of materials and it significantly decreases the amount of experimental work required. The purpose of the thermodynamic calculations was to predict the liquid region inside the ternary Pt–Ag–Si system. The last liquid region served for a determination of the nominal alloy composition with a predominating platinum component (with a minimum content of the mass fraction 50 %). Below a temperature of 1089 °C the thermodynamic calculation predicts the crossing of the liquidus surfaces in the ternary system. **Figure 3** presents the isothermal section of Pt–Ag–Si at 1089 °C, which is the lowest liquidus temperature for an alloy with a content of platinum and the mass fraction of Ag 40 % and 6 % of Si. The composition of the investigated alloy is inside the region of this composition.

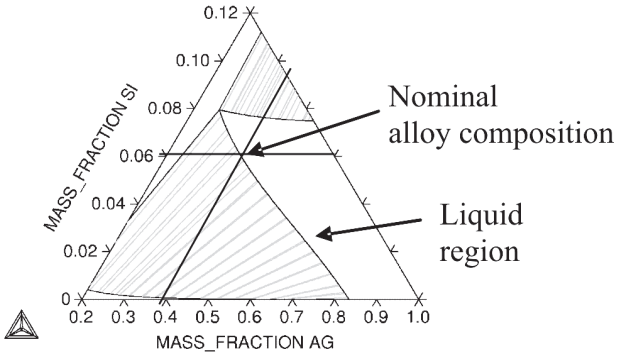


Figure 3: Isothermal section of the Pt–Ag–Si ternary system: liquid region at 1089 °C and crossing of the liquidus surfaces

Slika 3: Izotermni prerez ternarnega sistema Pt–Ag–Si: področje taline pri 1089 °C in prerez liquidusnih ploskev

4.2 Differential scanning calorimetry (DSC)

The aim of the DSC analyses was to determine the characteristic temperatures of possible reactions in the alloy in the temperature interval between 25 °C and 1550 °C. The DSC heating curve is presented in **Figure 4a**.

Two peaks were found on the DSC heating curve in **Figure 4a**, with a total heat of fusion of 89.147 J/g, which is the contribution of two different main reactions. **Figure 4a** represents the melting point at 964 °C with a heat of fusion of 5.887 J/g, with an additional reaction at 969.2 °C. The second main reaction is found at 1044.3 °C, with a heat of fusion of 83.26 J/g. The dislocated first peak away from the second one is assumed to be the result of a minute content of the first melting phase and a distinct temperature difference between the first and second reactions. The presence of a minute content of the first melting phase is also discernable from the values of the enthalpy of fusion (**Figure 4a**). Two peaks were also found on the DSC cooling curve in **Figure 4b**, which are also the contribution of two different main reactions, with a total heat of solidification of 102.726 J/g. The liquidus temperature determined from the cooling curve is at 1019.9 °C, and the solidification occurs with a high undercooling, with a solidification enthalpy of 97.45 J/g for the first reaction. With further cooling to 832 °C an exothermic peak with a small initiation at 853.2 °C and an enthalpy of solidification of 5.276 J/g was found. The high undercooling is discussed in relation to the theoretical liquidus determined from the heating curve. High undercooling usually indicates a lack of nucleation sites inside the melt at the theoretical solidification temperature determined during the heating

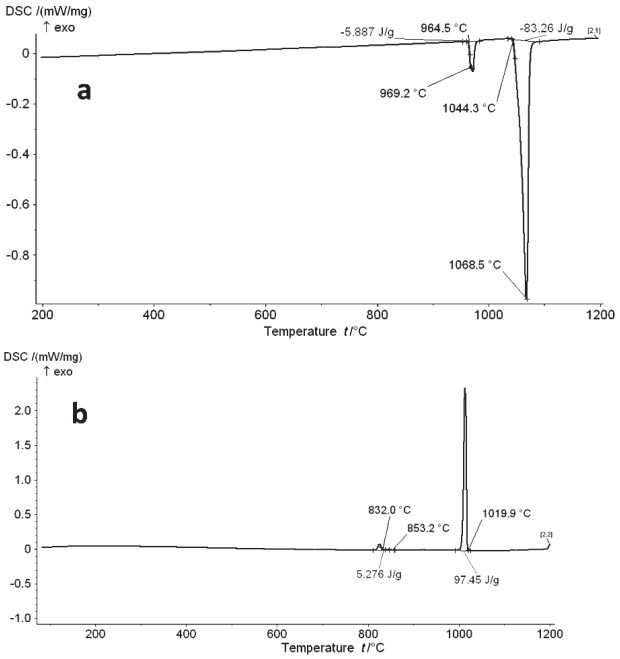


Figure 4: DSC analyses of the alloy: heating **(a)** and cooling curve **(b)**

Slika 4: DSC analiza zlitine: segrevalna **(a)** in ohlajevalna krivulja **(b)**

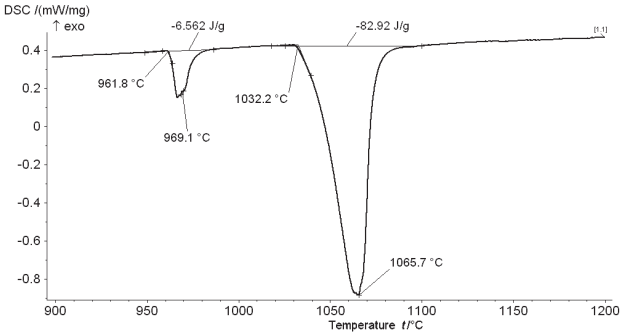


Figure 5: The DSC heating curve of the homogenized sample
Slika 5: DSC-segrevalna krivulja homogeniziranega vzorca

of the same alloy. The nucleation of the liquid during the melting is facilitated by the large number of microstructural defects, such as grain boundaries, especially segregating impurities, and makes the determination of the liquidus more reliable.

The homogenization temperature is selected below the first reaction. For the alloy homogenized at 900 °C, the heating curve between 900 °C and 1200 °C is shown in **Figure 5**. It shows a melting point at 961.8 °C. A reaction was also recorded at 969.1 °C, and the second strong effect was recorded at 1032 °C. The liquidus temperature determined from the DSC heating curve was 1065.7 °C.

4.3 Energy-dispersive X-ray spectroscopy (EDS)

The points of EDS analyses are presented in **Figure 6**. The results for the homogenized sample, gathered in **Table 3**, show a possible intermetal phase matrix based on platinum and silicon, which has been observed already.[2] The mole fraction of platinum in solution in the silver phase is about 2 %. Small deviations in the platinum content inside the silver-based phase are in relation to the reaction at 969.1 °C (spectra 1 and 2 in **Figure 6**). It is confirmed later in this paper that silicon only appears in the Pt–Si system (**Figure 7c**).

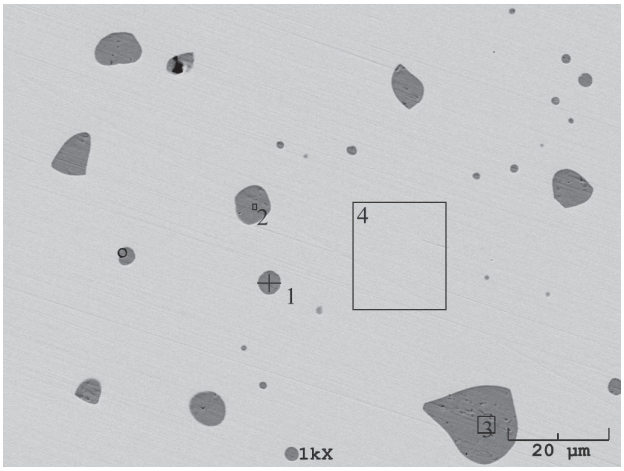


Figure 6: SEM micrograph of the microstructure after homogenization at 900 °C
Slika 6: SEM-mikroskopska slika mikrostrukture po homogenizaciji na 900 °C

Table 3: EDS analyses
Tabela 3: EDS analiza

Specter	Element	mole fraction *x* / %	mass fraction *w* / %
1	Ag Pt	97.51 2.49	95.58 4.41
2	Ag Pt	97.55 2.44	95.66 4.33
3	Ag Pt	98.43 1.56	97.19 2.80
4	Si Ag Pt	26.21 2.57 71.21	4.93 1.86 93.19

4.4 X-ray diffraction phase analysis (XRD)

The XRD analyses of the homogenized sample confirmed the presence of a small content of a silver-based phase and a larger amount of HT-Pt_2Si (high-temperature) metastable phase (**Figure 7b,c**). The recorded

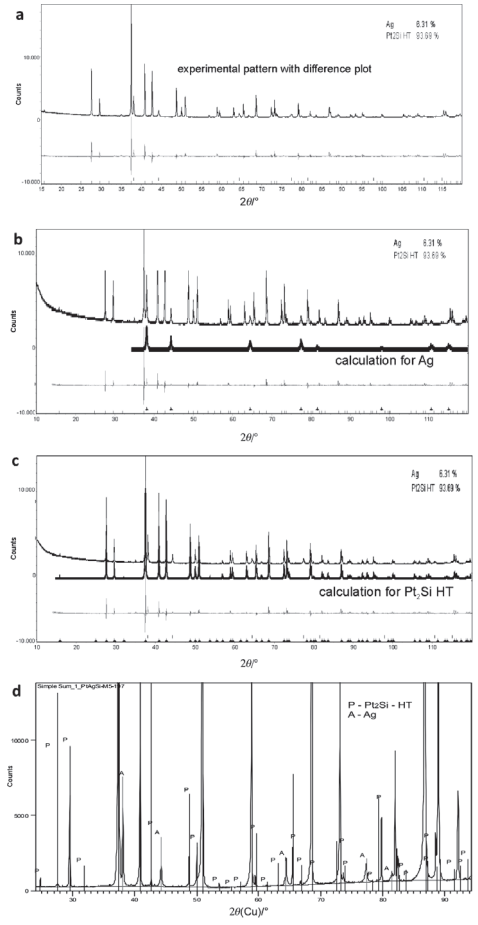


Figure 7: X-ray powder diffraction pattern: experimental pattern (a), silver-based phase (b), HT–Pt_2Si phase (c) and diffraction pattern of the aged sample at room temperature (d)
Slika 7: Posnetki rentgenska praškovne difrakcije: eksperimentalni posnetek (a), srebrova faza (b), faza HT–Pt_2Si (c) in difrakcijski posnetek staranega vzorca na sobni temperaturi (d)

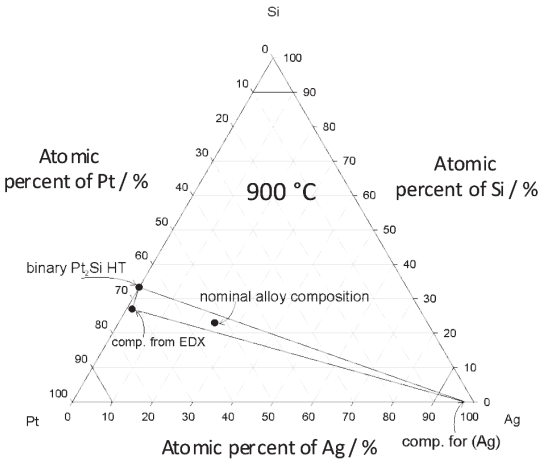


Figure 8: Isothermal section of the Ag–Pt–Si system at 900°C
Slika 8: Izotermni prerez sistema Ag–Pt–Si pri 900 °C

X-ray diffraction pattern of the alloy is shown in **Figure 7a**. The X-ray diffraction pattern recorded again after 2 months of aging at room temperature for possible changes in the HT Pt_2Si phase revealed the presence of a high-temperature phase HT-Pt_2Si and a silver-based phase (**Figure 7d**).

5 DISSCUSSION

After homogenization at 900 °C the melting point was at 961.8 °C, and near to the melting point of silver. The EDS analyses confirmed the presence of the silver phase with a small content of platinum. The EDS analyses also revealed a phase with a high content of platinum. The X-ray diffraction patterns after homogenization at 900°C revealed the existence of two phases: a silver-based phase and a HT-Pt_2Si phase. The determined lattice constant for the silver-based phase is 0.408364(18) nm. The lattice parameter for silver is at 0.40857 nm, with an Fm-3m space group. The determined HT–Pt_2Si phase has a lattice parameter a = 0.6460310(45) nm and b = 0.3577385(45) nm with a P6-2m space group. The literature, shows that the HT–Pt_2Si or the β–Pt_2Si is a congruent phase with the following lattice parameters: a = 0.6436 nm and c = 0.3569 nm or a = 0.6440 nm and c = 0.3573 nm. Thus, similar values to those in our investigation. From **Figure 8** it can be concluded that the nominal alloy composition in mole fractions (55.36 % Pt, 22.43 % Ag, 22.27 % Si) is really a two-phase region of the silver-based phase and the HT-Pt_2Si phase inside the Gibbs concentration triangle between the silver-rich corner and the binary Pt_2Si phase.

6 CONCLUSIONS

After homogenization at 900 °C the Pt–Ag17.5–Si4.5 ternary alloy was characterized by the presence of two compounds: the HT–Pt_2Si phase and a silver-based phase. It can be concluded from the X-ray diffraction patterns, the EDS and the DSC that the alloy with mass fractions 17.5 % of silver and 4.5 % of silicon has a melting point at 961.8 °C as a result of the melting of the silver-based phase. The effect at 1032 °C, determined from the DSC heating curve, is related to the melting of the HT-Pt_2Si phase and the liquidus temperature of the alloy is at 1065.5 °C.

7 LITERATURE

[1] M. H. F. Suiter, C. Colinet, A. Pasturel, Ab initio calculation of the phase stability in Au-Pd and Ag-Pt alloys, Physical review, B 73 (**2006**) 17, 174204
[2] L. E. Tanner, H. Okamoto, The Pt-Si system (Platinum-Silicon), Phase Diagram Evaluation; Journal of Phase Equilibria, 12 (**1991**) 5, 571
[3] I. Karakaya, W. T,Thompson, The Ag-Pt system, Royal military college of Canada, 8 (**1987**) 4, 334–430
[4] ASM handbook, Alloy phase diagrams, ASM international, 3 (**1992**), 410
[5] A. Gribanov, A. Grytstiv, E. Royanian, P. Rogl, E. Bauer, G. Giester, Y. Seropegin, On the system cerium-platinum-silicon; Journal of solid state chemistry; 181 (**2008**), 2964–2975
[6] R. Phal Ram, S. Bhan, On the constitution of platinum-silicon alloys; Zeitschrift fuer Metallkunde, 69 (**1978**), 524–529
[7] W. Gold, K. Schubert, Kristallstruktur von Pt12Si5, Zeitschrift fuer Kristallographie, Kristallgeometrie, Kristallphysik, Kristallchemie, 128 (**1969**), 406–413
[8] R. Gohle, K. Schubert, Zum aufbau des systems platin-silizium; Zeitschrift fuer metallkunde, 55 (**1978**), 503–511
[9] H. Pfeisterer, K. Schubert, Neuen phasen vom Mn P (B-31) – typ; Zeitschrift fuer metallkunde, 41 (**1950**), 358–367
[10] R. Gohle, K. Schubert, Journal of physics and Chemistry of Solids, 46 (**1985**), 631–641
[11] B. Y. Tsaur, J. W. Mayer, K. N. Tu, Ion beam induced metastable Pt2Si3 phase, Journal of Applied Physics, 51 (**1980**), 5326–5333

UDK 539.3:534.1
Original scientific article/Izvirni znanstveni članek

ISSN 1580-2949
MTAEC9, 44(4)219(2010)

A NEW SOLUTION OF THE HARMONIC FUNCTIONS IN THE THEORY OF ELASTICITY

NEKATERE ZNAČILNOSTI HARMONIČNIH FUNKCIJ V TEORIJI O ELASTNIČNOSTI

Valeryy Victorovich Chygyryns'kyy[1], Vladimir Grigorovich Shevchenko[1], Ilija Mamuzic[2], Sergey Borisevich Belikov[1]

[1]Zaporozhskyy National Technical University, Str. Zukovsky 64, Zaporižž'a, Ukraine
[2]University of Zagreb, Faculty of Mettalurgy Sisak, Str. A. N. heroja 3, Sisak, Croatia
mamuzic@simet.hr

Prejem rokopisa – received: 2008-11-21; sprejem za objavo – accepted for publication: 2009-08-17

A new approach to the solution of a plane problem of the theory of elasticity with the use of two harmonic functions with a Cauchy-Riemann analytical link is developed. The analysis of the harmonic functions shows that some allow a new approach to the solution of problems of the theory of elasticity. For the solution of linear differential equations a fundamental substitution is used, written in the general form $\psi(x,y) = y = C_\sigma \cdot \exp\theta$, with $\theta = \theta(x,y)$ as a function of the strain centre.

The transformations are explained with the properties of harmonic functions, where the Cauchy-Riemann relations can be used. The considered variants extend the possibilities for solutions and, if necessary, to obtain suitable functions for predetermined tasks. The new method is universal and can be effectively used when the fields of stresses and strains are described with trigonometric expressions.

Key words: theory of elasticity, harmonic functions, Cauchy-Riemann expressions.

Razvit je bil nov način reševanja ravninskega problema teorije elastičnosti s Cauchy-Riemanovo analitično zvezo. Analiza harmoničnih funkcij pokaže, da nekatere omogočajo nov način za rešitve problemov iz teorije elastičnosti. Za rešitev linearnih diferencialnih enačb se uporablja temeljna substitucija, zapisana v splošni obliki z $\psi(x,y) = y = C_\sigma \cdot \exp\theta$, z $\theta = \theta(x,y)$ kot funkcijo središča deformacije.

Transformacije smo razložili z lastnostmi harmoničnih funkcij, pri katerih je dovoljena uporaba Cauchy-Riemanovih povezav. Upoštevane variante razširjajo možnost rešitev in, če je potrebno, omogočijo, da dobimo rešitve za vnaprej načrtovano uporabo. Nova metoda je univerzalna in se lahko učinkovito uporabi, če so polja napetosti in deformacij opisana s trigonometričnimi odvisnostmi.

Ključne besede: teorija elestičnosti, harmonične funkcije, Cauchy-Riemanovi izrazi

1 INTRODUCTION AND FORMULATION OF THE TASK

The analysis of the peculiarities of the harmonic functions shows that some of them allow new approaches to the solution of problems of the theory of elasticity. Let us consider a plane problem of this theory. We have a set of equilibrium equations[1].

$$\frac{\partial\sigma_x}{\partial x}+\frac{\partial\tau_{xy}}{\partial y}=0;\ \frac{\partial\tau_{xy}}{\partial x}+\frac{\partial\sigma_y}{\partial y}=0 \qquad (1)$$

The equation of joint strains

$$\nabla^2(\sigma_x+\sigma_y)=0 \qquad (2)$$

The stresses' boundary conditions

$$\tau_n=\frac{\sigma_x-\sigma_y}{2}\cdot\sin 2\alpha-\tau_{xy}\cdot\cos 2\alpha \qquad (3)$$

Applying these expressions, the harmonic law of the distribution of contact stresses is determined[2], which formally coincides with that in[3]:

$$\tau_n=-\psi(x,y)\cdot\sin(A\Phi-2\alpha)$$

where $\psi(x,y)$ is the coordinate function of the strain centre; A is the constant determining the elastic state of a deformable medium; Φ is the coordinate function characterizing the allocation of contact shearing stresses; α is the slope angle of an element.

In place of equations (1) and (2), the biharmonic equation (4) can be applied:

$$\nabla^4\varphi=\frac{\partial^4\varphi}{\partial x^4}+2\frac{\partial^4\varphi}{\partial^2 x\cdot\partial^2 y}+\frac{\partial^4\varphi}{\partial y^4} \qquad (4)$$

with φ as a stress function.

The expression fulfils the boundary conditions (3)

$$\tau_{xy}=\psi(x,y)\cdot\sin(A\Phi) \qquad (5)$$

The stress difference in (3) is determined with

$$\sigma_x-\sigma_y=2\cdot\psi(x,y)\cdot\cos(A\Phi) \qquad (6)$$

2 SOLUTION OF THE TASK

The fundamental substitution is often used during the solution of linear differential equations[4], which can be written in the following general form

$$\psi(x,y)=\psi=C_\sigma\cdot\exp\theta \qquad (7)$$

with $\theta = \theta(x,y)$ being the unknown coordinate function of the strain centre.

Let us examine the harmonic functions $A\Phi$ and θ. The analytical link between them is admitted by the Cauchy-Riemann expressions [4,5]

$$\theta_x = \pm A\Phi_y \quad \theta_y = \pm A\Phi_x \tag{8}$$

After the derivation of equation (5), consideration of equation (7) and substitution in the equilibrium equations we obtain

$$\frac{\partial \sigma_x}{\partial x} + C_\sigma \cdot \theta_y \cdot \exp\theta \cdot \sin(A\Phi) + \\ + C_\sigma \cdot A\Phi_y \cdot \exp\theta \cdot \cos(A\theta) = 0$$

$$\frac{\partial \sigma_y}{\partial y} + C_\sigma \cdot \theta_x \cdot \exp\theta \cdot \sin(A\Phi) + \\ + C_\sigma \cdot A\Phi_x \cdot \exp\theta \cdot \cos(A\Phi) = 0$$

with θ_y, $A\Phi_x$ as the partial derivatives of the appropriate functions of the coordinates y and x. Passing from one variable to the other with the help of (8), we obtain, after integration and simplifications, the normal and shearing stresses

$$\begin{aligned} \sigma_x &= C_\sigma \cdot \exp\theta \cdot \cos(A\Phi) + f(y) + C \\ \sigma_y &= C_\sigma \cdot \exp\theta \cdot \cos(A\Phi) + f(x) + C \\ \tau_{xy} &= C_\sigma \cdot \exp\theta \cdot \sin(A\Phi) \end{aligned} \tag{9}$$

Substituting $f(y) = f(x) = 0$ in (9), we obtain the relation (6) that fulfils the boundary conditions for equation (3).

Considering (9), the sum of the stresses is

$$\sigma_x + \sigma_y = 2C$$

and the equation of joint strains (2) is automatically fulfilled. It is interesting that during the evaluation of the Laplacian for each value $C_\sigma \cdot \exp\theta \cdot \cos(A\Phi)$ and the substitution (8) the identity $0 \equiv 0$ is obtained. Using this peculiarity, the sum of stresses can be expressed as a product of the functions

$$\sigma' = \sigma_x + \sigma_y = n \cdot C \cdot \exp\theta \cdot \cos(A\Phi) \tag{10}$$

with n as the number that defines the influences of hydrostatic pressure on the medium of the stressed state in the strain zone.

By substituting (10) in (2) we obtain

$$\nabla^2(\sigma_x + \sigma_y) = \nabla^2\left[n \cdot C_\sigma \cdot \exp\theta \cdot \cos(A\Phi)\right] = \\ = n \cdot C_\sigma \cdot \exp\theta \left\{ \left[\theta_{xx} + (\theta_x - A\Phi_y)(\theta_x + A\Phi_y) + \right.\right. \\ \left. + \theta_{yy} + (\theta_y - A\Phi_x)(\theta_y + A\Phi_x) \right] \cdot \cos(A\Phi) - \\ \left. - \left[2\theta A\Phi_x + 2\theta_y A\Phi_y + A\Phi_{xx} + A\Phi_{yy} \right] \cdot \sin(A\Phi) \right\} \tag{11}$$

It is clear from the analysis of the differential equation (11), that it turns to identity under the condition of

$$\theta_x = \pm A\Phi_y \quad \theta_y = \pm A\Phi_x$$

This is the relation (8), which was introduced as an assumption during the solution of the equilibrium equations. Differentiating further, we obtain

$$\theta_{xx} = \pm A\Phi_{yx} \quad \theta_{yy} = \pm A\Phi_{xy}$$
$$\theta_{xy} = \pm A\Phi_{yy} \quad \theta_{yx} = \pm A\Phi_{xx}$$

The last relations convert equation (11) into identity.

The last expressions show that the indicated functions are harmonic, i.e.

$$\theta_{xx} + \theta_{yy} = 0 \quad A\Phi_{xx} + A\Phi_{yy} = 0$$

It is remarkable that the operators of the trigonometrical functions are equal to zero. This peculiarity shows that the function (10) also fulfils the biharmonic equation (4). Considering equation (10), the Laplace of the equation has the form

$$\nabla^2\varphi = \nabla^2\left[C_\sigma \cdot \exp\theta \cdot \cos(A\Phi)\right] = \\ = C_\sigma \cdot \exp\theta \cdot \left[\begin{pmatrix} \theta_{xx} + (\theta_x^2 - A\Phi_y^2 + \\ + \theta_{yy} + (\theta_y^2 - A\Phi_x^2 \end{pmatrix} \cdot \cos(A\Phi) - \right. \\ \left. - \begin{pmatrix} 2 \cdot \theta_x A\Phi_x + 2\theta_y A\Phi_y + \\ + A\Phi_{xx} + A\Phi_{yy} \end{pmatrix} \cdot \sin(A\Phi) \right] = 0$$

Let us introduce the symbolisms

$$L(x,y) = L = \theta_{xx} + \theta_x^2 - A\Phi_y^2 + \theta_{yy} + \theta_y^2 - A\Phi_x^2$$
$$M(x,y) = M = 2 \cdot \theta_x \cdot A\Phi_x + 2\theta_y \cdot A\Phi_y + A\Phi_{xx} + A\Phi_{yy}$$

Then, with consideration of the symbolisms, the accurate form of the Laplace equation is obtained

$$\alpha(x,y) = \alpha = L \cdot \cos(A\Phi) - M \cdot \sin(A\Phi) = 0$$

If the factors in the trigonometrical functions are equal to zero, also the operators $L = M = 0$. For a more integrated analysis let us write a Laplacian for the function $\alpha(x,y)$

$$\nabla^4\varphi = \frac{\partial^4\varphi}{\partial x^4} + 2\frac{\partial^4\varphi}{\partial^2 x \cdot \partial^2 y} + \frac{\partial^4\varphi}{\partial y^4} = \\ = \nabla^4\alpha = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial^2 y} \right) \cdot \left(L \cdot \cos(A\Phi) - M \cdot \sin(A\Phi) \right) = \\ = (L_{xx} - L \cdot A\Phi_x^2 - 2M_x \cdot A\Phi_x - M \cdot A\Phi_{xx} + L_{yy} - L \cdot A\Phi_y^2 - \\ - 2M_y \cdot A\Phi_y - M \cdot A\Phi_{yy}) \cdot \cos(A\Phi) - (2L_x \cdot A\Phi_x + \\ + L \cdot A\Phi_{xx} + M_{xx} - M \cdot A\Phi_x^2 + 2L_y \cdot A\Phi_y + L \cdot A\Phi_{yy} + \\ + M_{yy} - M_y \cdot A\Phi_y^2) \cdot \sin(A\Phi)$$

It is expected that the partial derivatives from zero functions are equal to zero, thus, $\nabla^4\alpha \equiv 0$. Let us write the partial derivatives of separate operators and track the mechanism of the turning into identity of the harmonic functions

$$L_x = (\theta_{xxx} + \theta_{yyx}) + (2\theta_x\theta_{xx} - 2A\Phi_y A\Phi_{yx} + \\ + (2\theta_y\theta_{yx} - 2A\Phi_x A\Phi_{xx})$$

Following (8), we have

$$\theta_x = -A\Phi_y \quad \theta_y = A\Phi_x$$
$$\theta_{xx} = -A\Phi_{yx} \quad \theta_{yy} = A\Phi_{xy}$$
$$\theta_{xy} = -A\Phi_{yy} \quad \theta_{yx} = A\Phi_{xx}$$
$$\theta_{xxx} = -A\Phi_{yxx} \quad \theta_{yyy} = A\Phi_{xyy}$$

$$\theta_{xyy} = -A\Phi_{yyy} \quad \theta_{yxx} = A\Phi_{xxx} \tag{12}$$

Let us substitute (12) in the operator L_x.

$$L_x = \frac{\partial}{\partial x}(\theta_{xx} - \theta_{yy}) + \left[2\theta_x\theta_{xx} - 2(\theta_x)(\theta_{xx})\right] + \left[2\theta_y\theta_{yx} - 2(\theta_y)(\theta_{yx})\right] \equiv 0$$

We have obtained the identity, "quod erat demonstrandum".

The same approaches take place during the evaluation of the operator L_{xx}. Let us write it as

$$L_{xx} = (\theta_{xxxx} - \theta_{yyxx}) + (2\theta_{xx}\theta_{xx} - 2A\Phi_{yx}A\Phi_{yx}) + (2\theta_x\theta_{xxx} - 2A\Phi_y A\Phi_{yxx}) - (2A\Phi_{xx}A\Phi_{xx} - 2\theta_{yx}\theta_{yx}) - (2A\Phi_x A\Phi_{xxx} - 2\theta_y\theta_{yxx})$$

Substituting (12) into the expression for L_{xx} and factoring out a flexion on x from the first brackets, after conversion the identity $0 \equiv 0$ is obtained.

Thus, the operators

$$L_{xx} = L_{yy} = M_{xx} = M_{yy} = L_x = L_y = M_x = M_y = L = M = 0$$

demonstrate that the function (10) fulfils the biharmonic equation (4) and it can be used for the evaluation of the components of the stress tensor. It is necessary to ensure that the field of stresses and the stress function are described, as a matter of fact, with identical expressions (9) and (10) linked analytically[6] with

$$\sigma_x = \frac{\partial^2\varphi}{\partial y^2} \qquad \sigma_y = \frac{\partial^2\varphi}{\partial x^2} \qquad \tau_{xy} = \frac{\partial^2\varphi}{\partial y \partial x}$$

Such schemes of transformations are explained with the properties of harmonic functions where the Cauchy-Riemann relations can be applied. The considered variants allow us to extend the possibility of solutions and, if necessary, to obtain suitable functions for the development of a predetermined result.

Let us return to the expressions for the stress tensor components and consider the equilibrium equations in the components of the stress deviator. Let us introduce the symbols

$$\sigma_x'' = \sigma_x - \sigma - f(y) - C$$
$$\sigma_y'' = \sigma_y - \sigma - f(x) - C \tag{13}$$

with σ being the mean stress.

Considering (13), the equilibrium equation (1), can be rewritten in the form[6]

$$\frac{\partial^2\sigma_x''}{\partial x} + \frac{\partial^2\tau_{xy}}{\partial y} = 0 \qquad \frac{\partial\tau_{xy}}{\partial x} + \frac{\partial\sigma_y''}{\partial y} = 0$$

By analogy with (9) and with integration and simplifications we obtain the stress tensor components

$$\sigma_x = C_\sigma \cdot \exp\theta \cdot \cos(A\theta) + \sigma + f(y) + C$$
$$\sigma_y = -C_\sigma \cdot \exp\theta \cdot \cos(A\theta) + \sigma + f(x) + C$$
$$\tau_{xy} = C_\sigma \cdot \exp\theta \cdot \sin(A\theta) \tag{14}$$

With: $\theta_x = \pm A\Phi_y$, $\theta_y = \pm A\Phi_x$

It follows from expressions (14) that their deviator part for the normal stresses $C_\sigma \cdot \exp\theta \cdot \cos(A\Phi)$ coincides with the shifting part σ in (10). Considering (13), (6) is fulfilled and the boundary conditions (3) are satisfied.

The outcome (14) can be generalized. The analytical link of the functions with the opposite signs is obtained in relations (8) and different signs of an index in an exponential curve result can be obtained. Therefore, the index of an exponential curve in a solution will be not unique. The exponential function can be written in the form of a sum with the use of the hyperbolic cosine or sine in the general form

$$\sigma_x = \left[C_1 \cdot ch(\theta) \pm C_2 \cdot sh(\theta)\right]\cos(A\Phi) + \sigma + f(y) + C$$
$$\sigma_y = -\left[C_1 \cdot ch(\theta) \pm C_2 \cdot sh(\theta)\right]\cos(A\Phi) + \sigma + f(x) + C$$
$$\tau_{xy} = +\left[C_1 \cdot ch(\theta) \pm C_2 \cdot sh(\theta)\right]\sin(A\Phi) \tag{15}$$

In these expressions it is assumed that the arguments of the trigonometric and exponential functions can be represented in the form of a series of harmonic functions interlinked with the Cauchy-Riemann relations.

3 COMPARISON TO OTHER SOLUTIONS

The solutions of a plane problem with the help of a trigonometric series are often used. For example, the following combination of functions is often met[3]:

$$\varphi = \sin(\alpha \cdot x) \cdot \left[C_1 \cdot \exp(\alpha \cdot y) + C_2 \cdot \exp(-\alpha \cdot y)\right] \tag{16}$$

Let us ascertain whether the Cauchy-Riemann relation exists between the arguments of trigonometric and exponential functions

$$A\Phi = \alpha \cdot x \qquad \theta = \pm\alpha \cdot y \qquad A\Phi_x = \alpha$$
$$A\Phi_y = 0 \qquad \theta_y = \pm\alpha \qquad \theta_x = 0$$

The obtained relations take place for the functions

$$\theta_x = \mp A\Phi_y = 0 = 0 \qquad \theta_y = \pm A\Phi_x = \pm\alpha = \alpha$$

The peculiarity of these solutions is that they are common and do not contradict known partial solutions. The arguments $A\Phi$ and θ are harmonic functions of the coordinates x and y. They can be rather complicated and cannot be determined from the linear dependences for one coordinate.

Let us analyze the possibilities of the solution (14). The elementary variant of a harmonic function of two variables is $A\Phi = A \cdot x \cdot y$. Applying the relations (8), it is written as

$$\theta = \mp\frac{1}{2} \cdot A(x^2 - y^2)$$

Thus,

$$\theta_x = \mp A\Phi_y = \mp A \cdot x = A \cdot x \qquad \theta_y = \pm A\Phi_x = \pm A \cdot y = A \cdot y$$

Each function fulfils the Laplace equations.

4 CONCLUSION

A new approach to the solution of a plane problem of the theory of elasticity based on the use of two harmonic functions with the Cauchy-Riemann analytical link is developed. The new method is universal and can be effectively used when the fields of stresses and strains are described with trigonometric expressions.

5 REFERENCES

[1] Bezuchov N. I. Osnovy teoriy uprugosty, plastichnosty i polzuchesty (Basics of the theory of elasticity, plasticity and creep). Vysshaja shkola, 1968, 498 s

[2] Malinin N. N. Prikladnaja teorija plastichnosty i polzuchesty (Practical theory of plasticity and creep), Mashinostroenie, 1975, 399 s

[3] Chigirins'kyy V. V., Mazur V. L., Legotkin G.I., Slepynin A. G. and al. Proizvodstvo vysokoeffektivnogo metalloprokata (High efficient production of rolling stock), Dnepropetrovsk, PBA «Dnepro-VAl», 2006, 261 s

[4] Tihonov A. N., Samarskiy A. A. Uravnenija matematicheskoy fiziki (Equations of mathematical physics), Nauka, 1977, 735 s

[5] V. V. Chygyryns'kyy, I. Mamuzic, G. V. Bergeman. Analysis of the State of Stress of a Medium under Conditions of Inhomogeneous Plastic Flow, Metalurgija, 43 (**2004**), 87–93

[6] Nikiforov S. N. Teorija uprugosty i plastichnosty (Theory of elasticity and plasticity), Gosudarstvennoe izdatelstvo literatury po stroitelstvu i architecture, 1958, 283 s

UDK 669.18
Original scientific article/Izvirni znanstveni članek

ISSN 1580-2949
MTAEC9, 44(4)223(2010)

INFLUENCE OF VACUUM PROCESSING ON THE CONTENT OF SOME ELEMENTS IN MOLTEN METAL

VPLIV VAKUUMSKEGA PROCESIRANJA NA VSEBNOST NEKATERIH ELEMENTOV V KOVINSKIH TALINAH

Derviš Pihura, Derviš Mujagić

University of Zenica, Metallurgical institute "Kemal Kapetanović" Zenica, Travnićka c. 7, 72000 Zenica, Bosnia and Herzegovina
dervis.pihura@gmail.com

Prejem rokopisa – received: 2008-09-22; sprejem za objavo – accepted for publication: 2010-03-15

Residual and trace elements can affect physical, chemical and mechanical properties, especially those of high-alloy materials and steels. With a treatment of the molten metal in a vacuum-induction furnace and an electron-beam furnace the content of residual elements is diminished, depending on the treatment time and the vacuum pressure.

Key words: vacuum treatment, trace elements, electron-beam furnace, vacuum-induction furnace

Rezidualni elementi in elementi v sledovih vplivajo na kemične in mehanske lastnosti, posebno pri visokolegiranih zlitinah in jeklih. Pri obdelavi taline v vakuumski indukcijski peči in v peči z elektronskim curkom se vsebnost rezidualnih elementov zmanjša v odvisnosti od trajanja obdelave in od tlaka.

Ključne besede: vakuumska obdelava, elementi v sledovih, peč z elektronskim curkom, vakuumska indukcijska peč

1 INTRODUCTION

A vacuum is a powerful tool for improving the physical and chemical properties of ferrous and nonferrous alloys with vacuum melting, re-melting and treatment. Depending on the vacuum process used, the physical and chemical properties of the alloys can be improved. For this reason, it is important to select the vacuum treatment process, keeping in mind the required change in the properties of the alloy.

2 VACUUM PROCESSES

Vacuum melting is applied for the melting of raw materials in coreless induction furnaces, and is particularly suited to the melting and casting of different high-alloyed materials because it enables an exact control of the temperature and the chemical composition.[1–6] In this way, the physical characteristics and properties of the melted alloys can be tailored. The charged components are melted in vacuum or a protective atmosphere (Ar, N or H) and then vacuumed with a partial pressure of down to 10^{-4} mbar. Since the alloy may contain elements with a high vapour pressure, by selecting the proper vacuum partial pressure the chemical composition of the alloy must be kept in mind. With a vacuum treatment, the alloy is refined and its physical, chemical and mechanical properties as well as cleanliness may be changed. It is possible, in a vacuum furnace, to also treat the alloy with vacuum de-oxidation using appropriate additions or, through the effect of the vacuum itself, to decrease the content of the gasses N_2 and H_2 as well as to change the content of elements such as Sb, As, Bi, Sn, Cu, etc. By decreasing the content of some elements, non-metallic inclusions and gases, physical and mechanical properties that are better suited to a particular application are obtained.

Arc furnaces and electron-beam furnaces are mostly used for the re-melting of previously produced alloys applying a non-vacuum process. The best, which must be used for such an operation, are the electric arc furnace and electron-beam furnace, which is of particular interest for the aim of this paper. The electron-beam vacuum furnace is applied on the pilot and industrial scales for ingots up to 20 t. In this furnace, droplets of the molten alloy, or electrode, solidify rapidly in a water-cooled copper mould. The heat for melting is obtained with 3 to 12 electron beams targeted on the re-melted electrode. Depending on the energy consumption, the local electrode temperature may be increased to a few thousand degrees and, in this way, the conditions for evaporation of the high-partial-pressure components of the alloys are achieved. The high local temperature makes it difficult to measure the exact temperature of the molten drops. The evaporation may be considered as one of the advantages of the re-melting process, especially as the melt droplets form in a vacuum of different grades down to $1 \cdot 10^{-6}$ mbar. In this way, the molten metal is exposed to vacuum for a sufficient time and the extraction of the gases and the evaporation of high-vapour-pressure components is ensured. After re-melting, the alloy density and cleanliness are improved and the physical properties are substantially improved. The re-melted alloys can then be used for high-temperature and high-stress applications, as well as for the parts of jet engines.

A good technology for the vacuum treatment of molten metal is the circulating or RH process that is applied only on an industrial scale. In this process the molten metal is lifted from the container in the vacuum chamber, processed for the determined time and re-flown in the melting furnace. As part of the process the surface of the alloy surface in direct contact with the vacuum is increased and the rate and the extent of extraction is also increased. This process is also suited to the highest quality of steel grades.

3 EFFECT OF VACUUM TREATMENT

In a vacuum some of the physical properties of a molten metal are changed. Dissolved gases behave in accordance with Sieverts' and Nernst's laws and their extraction becomes possible. Nitrogen and hydrogen are extracted directly, while the extraction of oxygen occurs through vacuum de-oxidation with carbon and the extraction of CO.

Specific interactions between the elements present in the melt may help to displace some elements from the solution in the molten mother metal. The interaction coefficients are a measure of the mutual bonding energy of the dissolved X and Y atoms relative to the X-Fe and Y-Fe bonding. For dilute solutions, the bonding energies are independent of each other and additive, whereas in more concentrated solutions, Z atoms may affect the X-Y bonding. For multi-components solutions, the total activity coefficient for each element will depend on the mutual effect of each component. For a diluted solution the interaction can be described with a simple relation.

In the extraction process, the effect of the partial pressure of elements in the melt is of essential importance. For a number of elements, the effect of temperature on the partial pressure is shown in **Figure 1**. Generally, elements with a higher partial pressure are extracted more quickly from the melt; however, the evaporation also depends on the geometry, the vacuuming time and the ferro-static pressure in the reacting part of the melt. At the melting temperature of iron, the vapour pressure of the impurities and the alloying elements decreases in the order Ba > Pb > Sn > Mn > Si > Fe = 0.84 > 0.45 > 0.45 > 0.035 > $8 \cdot 10^{-4} > 2 \cdot 10^{-4}$ bar.[5,6]

4 RESULTS

In this work the results of investigations of the effect of the vacuum treatment of several alloys – low carbon iron and steel, alloys with about 60 % Ni and 18 % Cr, with 50 % Ni, steel with 18 % Cr and 10 % Ni and a low-alloy steel – in a vacuum-induction furnace and an electron-beam furnace are presented. In **Figure 1** the dependence of the partial vapour pressure on temperature is shown for different elements, as well as for elements found as impurities in the investigated alloys.

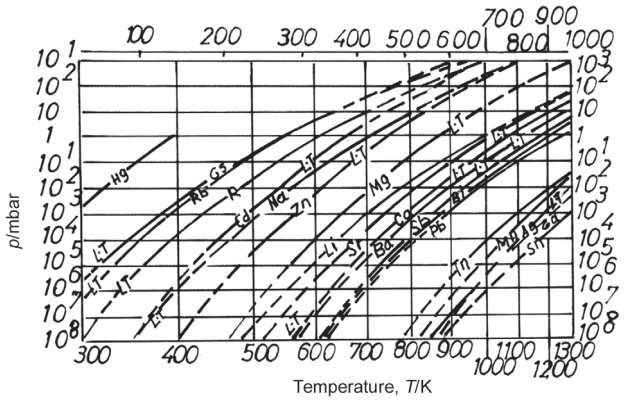


Figure 1: Dependence of the vapour or partial pressure for different elements on the temperature[5,6]

Slika 1: Odvisnost parcialnega parnega tlaka od temperature za različne elemente[5,6]

In **Figures 2 and 3** the effect of electron-beam melting on the content of some elements are shown – impurities in low-carbon iron, low-carbon steel, nickel or nickel chromium alloy types such as perm-alloy, Cr/Ni – 18/8 or 18/10, Cr/Ni – 18/60.

Figure 2 presents the results of the content change of tin, arsenic, copper and antimony, while in **Figure 3** the effects of the melting rate (kg/min) and the power ((kW h)/kg) are shown for arsenic and different groups of alloys. The content of all the investigated elements is diminished by maintaining the melt in vacuum; however, the extent and the kinetics of evaporation are different for the different alloys.

For the three elements with greater evaporation rates, two phases are seen on the kinetics curve: an initial phase of rapid evaporation down to a determined critical content and a phase with a much lower evaporation rate. For antimony the evaporation rate is 5.5 times lower in the second than in the first phase. The shape of the kinetics curves in **Figure 2** is related to the activity in the elements in solution in the molten iron. It is assumed that the evaporation rate is greater down to the equilibrium

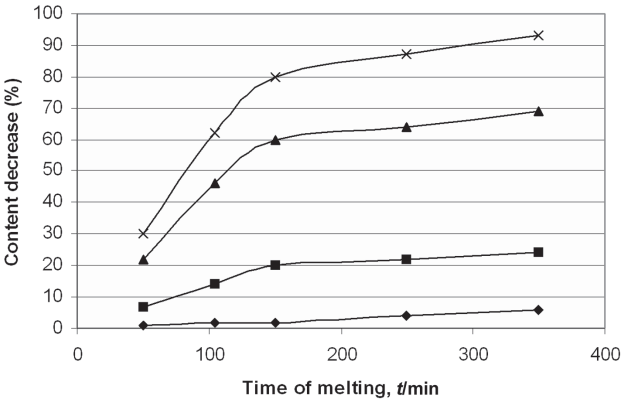


Figure 2: Decrease of some residual elements during electron-beam melting: Sn (◊), As (□), Cu (Δ) and Sb (x) in low-carbon iron

Slika 2: Zmanjšanje vsebnosti nekaterih elementov pri taljenju z elektronskim curkom: Sn (◊), As (□), Cu (Δ) in Sb (x)

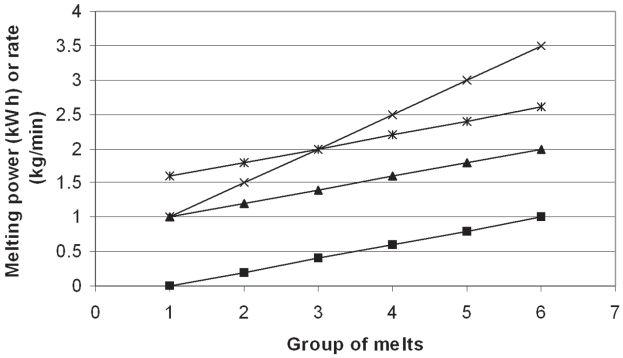


Figure 3: Loss of 22 % of arsenic by electron-beam melting for different group of melts and different melting rate MIZ (*) [6], As (□)[4] & lit. (Δ)5; low-carbon iron (1), steel (2), Cr/Ni – 18/8 (3) or 18/10 (4), Cr/Ni – 18/60 (5), perm-alloy (6) elemente

Slika 3: Izguba 22 % arzena pri taljenju v peči z elektronskim curkom za različne skupine zlitin pri različni hitrosti taljenja; MIZ (*) [6], As (□)[4], ref.5 (Δ). Maloogljično železo (1), jeklo (2), Cr/Ni 18-8 (3 in 18-10(4), Cr/Ni 18/60 (5), perm. zlitina (6)

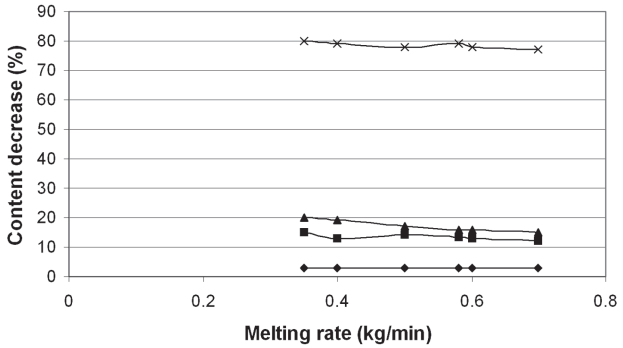


Figure 5: Decrease of some elements during melting in an electron-beam vacuum furnace: P (◊), Si (□), S (Δ) and Mn (×) depending on the melting rate of Cr/Ni – 18/8 (3), 18/10 (4) and perm-alloy6

Slika 5: Zmanjšanje vsebnosti nekaterih elementov pri taljenju v peči z elektronskim curkom; P (◊), Si (□), Cu (Δ) in Mn (×) v odvisnosti od hitrosti taljenja Cr/Ni – 18/8 (3), 18/10 (3) in perm. zlitine[6]

activity, which depends on the bath's chemical composition and the temperature.

The evaporation of antimony was the largest, but it is also large for copper. It is low for arsenic, while it is scarcely discernible for tin. In **Figure 3** the effect of the melting power and rate is shown for arsenic for the groups of melts of low-carbon iron (1), steel (2), perm-alloy (6), Cr/Ni – 18/8 (3) or 18/10 (4) and Cr/Ni – 18/60 (5). It indicates that an average decease of arsenic of 22 % for different groups of materials depends on the content of the alloying elements and on the influence of their physical properties on each material. The dependences are based on our own experimental findings[6] and reference.[5] The comparison of melting characteristics indicates a simple relation between the melting rate and the power consumption for different materials for the same average loss of arsenic.

The effect of melting in the vacuum induction furnace is, in principle, similar to the electron-beam furnace, with the higher evaporation rate for antimony and the lowest for tin (**Figure 4**). The evaporation rate increases with the increasing content of the element in

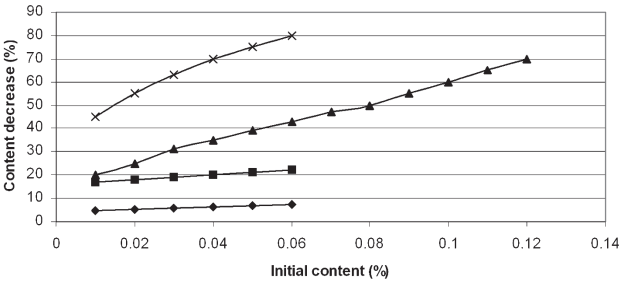


Figure 4: Decrease of some residual elements during the melting of low-carbon steels and alloys[1,2,6] in a vacuum induction furnace: Sn (◊), As (□), Cu (Δ) and Sb (×)

Slika 4: Zmanjšanje vsebnosti nekaterih rezidualnih elementov pri taljenju malooglјičnih jekel in zlitin[1,2,6] v vakuumski indukcijski peči: Sn (◊), As (□), Cu (Δ) in Mn(×)

the alloy and it is quite different for the same content of different elements in the alloy. Compared with the evaporation rate of tin, the evaporation rate is 2.5 times greater for arsenic, five times greater for copper and 10 times greater for antimony. The difference is very probably related to the difference of the activity of the elements in solution in the molten metal and their partial pressure and depends on the interaction coefficient of some element also, as the activity of the impurities in molten metal is affected by the presence of other components. To describe this phenomenon it is convenient to maintain the same standard state as with the binary melts, and to introduce an activity coefficient (f_x), which describes the effect of the third component. This coefficient depends on the interaction coefficients, which differ greatly,[5] and is the product of the interaction coefficients and the content of the third component in the melt.

It is assumed that with a vacuum treatment on the industrial scale the relative differences in the evaporation of elements would be similar to those presented in this work, while the absolute rate of evaporation and the change of element contents would depend for all elements on the conditions specific to the processing, especially the presence of the slag layer on the surface of the molten alloy.

The evaporation rates of sulphur, silicon and phosphorus by melting in an electron-beam vacuum furnace is low and are virtually unaffected by the melting rate, while the evaporation of manganese is large and again independent on the melting rate (**Figure 5**). Again, the difference is very probably related to the activity of the elements in solution in the iron bath. With respect to the loss of elements during other methods of vacuum processing of iron alloys, for phosphorus, sulphur, silicon and manganese a similar conclusion seems to be justified, as was earlier suggested for tin, arsenic, copper and antimony. It should also be considered that the slag could have a positive effect on the change of the content

of phosphorus and particularly sulphur in the melted iron alloy.

In vacuum, the gases are extracted from the metal bath according to Sieverts' law. The degassing intensity also depends on the metal bath temperature and the chemical composition, which determine the coefficients of interaction. This is confirmed by the data in **Table 1**, showing that the degassing constant is different for alloys with different chemical compositions and temperatures.

Table 1: Constant nitrogen-degassing rate ($K_N \cdot 10^{-4}$)[1,2,6]
Tabela 1: Konstanta hitrosti degazacije dušika ($K_N \cdot 10^{-4}$)[1,2,6]

Temperature T/°C	Material				
	Cr/Ni – 18/8	Cr/Ni – 18/60	Perm alloy	LAS	Cr/Ni – 18/10
1450	0.41	2.51	2.91	–	–
1500	1.11	1.82	–	1.41	0.81
1550	1.45	1.12	1.09	1.35	1.39
1600	1.82	0.75	0.81	1.21	1.83

5 CONCLUSION

The experimental data presented show the loss of elements as a result of evaporation in vacuum is very different; it is below the detection limits for some elements, while it is significant for other elements present in the iron bath. The evaporation rate is particularly high for antimony, copper and manganese, while it is virtually null for tin and low sulphur, phosphorus and silicon. The findings show that the evaporation rate depends on the interaction coefficient of the elements in solution in the metal bath and the temperature. When extrapolating the experimental findings to other vacuum processes for iron alloys, it should be kept in mind that the evaporation process also depends on the geometry of the vacuum exposure of the melt, and remembering that the presence of the slag on the metal bath surface would significantly affect the evaporation loss in two ways: by preventing a contact between the bath surface and the atmosphere and by a direct chemical reaction of the slag with some elements in solution in the molten bath, especially sulphur and phosphorus.

6 REFERENCES

[1] Pihura D. at al. Promjena sadržaja dušika u vakuumskoj peć (Change of nitrogen content in the vacuum furnace), Zenica, 2008
[2] Pihura D. at al. Osvajanje proizvodnje legura tipa nimonic i nitronic u vakuumskoj peći (Manufacturing of nimonic and nitronic type alloys in a vacuum furnace), Zenica, 2007
[3] Pihura D. Kazanska metalurgija (Ladle metallurgy) – VI-VIII dio; Met. Inst., Zenica, 1986/9
[4] Pihura D. Vacuum Melting Technology Monogr., UNIDO, Vienna, 1980
[5] Winkler O, Bakhish Vacuum metallurgy, Elsevier, London 1971
[6] Pihura D. Vakuumske tehnike i tehnologije (Vacuum technique and technology), Met. Inst., Zenica, 1979

UDK 621.039.568:669.14.018:539.42
Professional article/Strokovni članek

ISSN 1580-2949
MTAEC9, 44(4)227(2010)

INFLUENCE OF δ-FERRITE ON THE FATIGUE RESISTANCE OF BLADE MATERIALS

VPLIV δ-FERITA NA UTRUJENOSTNO TRDNOST JEKEL ZA LOPATICE

Pavlína Hájková[1], Jiří Janovec[1], Jan Siegl[2], Bohumil Smola[3], Michaela Vyroubalová[1], Daniela Tůmová[1]

[1]CTU in Prague – Faculty of Mechanical Engineering, Karlovo nám. 13, 121 35 Praha 2, Czech Republic
[2]CTU in Prague – Faculty of Nuclear Sciences and Physical Engineering, Trojanova 13, 121 35 Praha 2, Czech Republic
[3]Charles University in Prague – Faculty of Mathematics and Physics, KE Karlovu 5, 121 35 Praha 2, Czech Republic
pavlina.hajkova@fs.evut.cz

Prejem rokopisa – received: 2009-11-16; sprejem za objavo – accepted for publication: 2010-02-21

The low-pressure parts of a TG 1000 MW turbine, for the third wheels, are made from modified 12 % Cr martensitic steel AK1 TD.9, and the fourth, are made from X2CrNiMo13-4 steel. The fracture of the third turbine wheel at the point of the connection part of the blade has a high-cycle fatigue character with multiple initiation sites. An analysis of the microstructure has proved the influence of both the δ-ferrite content and the arrangement on the initiation point of the fatigue process. Also, an inspection of the fatigue crack's kinetics in relationship to the operating mode of the power plant has been carried out.

Key words: low-pressure turbine, steel, fatigue fracture, microstructure, δ ferrite

Nizkotlačni deli turbine TG 1000 MW iz tretjega venca so iz modificiranega 12-odstotnega Cr martenzitnega jekla AK TF.9, iz četrtega venca pa iz jekla X2CrNiMo 13-4. Prelom tretjega turbinskega venca v točki vpetja ima značaj visoke ciklične utrujenosti z mnogimi začetnimi mesti. Analiza mikrostrukture je pokazala vpliv vsebine in porazdelitve δ-ferita na začetno točko procesa utrujenosti. Ugotovljena je bila tudi povezava kinetike utrujenostne razpoke z načinom dela turbine.

Ključne besede: nizkotlačna turbina, jeklo, utrujenostna razpoka, mikrostruktura, δ-ferit

1 INTRODUCTION

In October 2008 the blade of the third turbine wheel (3TW) placed on the third low-pressure part (3LP) of the TG1 broke. Consequently, the broken blade damaged the turbine wheels numbers 3 and 4 on the LP1, LP2 and LP3 (see **Figure 1**). The break in the blade occurred near the connection part (**Figure 2**) during the blade's operation. Cracks in the connection-part regions within the turbine wheels of other low-pressure parts were discovered with defectoscopic analysis. The damaged connection part is shown in **Figure 2**. The chemical composition, mechanical properties, fractography and time-dependent estimation of the crack propagation of the removed blades were tested. The scheme of the damaged blades within the whole turbine vessel is shown in **Figure 3**.

2 EXPERIMENTAL PROCEDURE

2.1 Materials

The investigated blades (see **Table 1**) were drop forged, and a new steel, 1.4939, was investigated as a possible substitute material for the 3.TW.

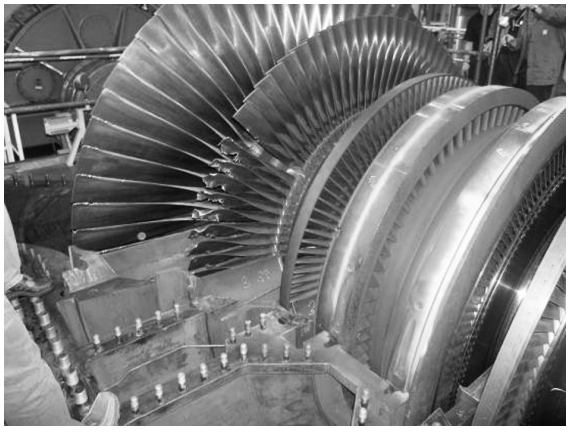
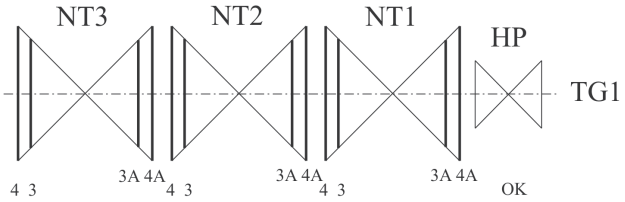


Figure 1: Damage to the 3TW on the TG1
Slika 1: Poškodba 3TW na TG 1

Figure 2: Damage to the connection part
Slika 2: Poškodba na mestu vpetja



Low-pressure part	Turbine Wheel number	Blade number
3LP	3. TW	152, 156, 106
1LP	4. TW	344
3LP	3. TW (suggestion)	1.4939

Figure 3: Scheme of broken blades within the whole turbine vessel
Slika 3: Shema prelomljenih lopatic v turbinskem ohišju

Table 1: Investigated samples
Tabela 1: Vzorci za preiskave

Blade No.	Low-pressure part	State	Material
152	3NT	Destroyed	modified 12 % Cr martensitic steel AK1 TD.9
156	3NT	Laboratory broken	modified 12 % Cr martensitic steel AK1 TD.9
106	3NT	Additionally identified	modified 12 % Cr martensitic steel AK1 TD.9
New material 1.4939		Identified	1.4939 + QT
344	NT 4a	OK	X2CrNiMo13-4 steel (WEV 400)

2.2 Investigation methods

2.2.1 Evaluation of the chemical and microchemical composition of the turbine-blade material

Three different devices were used for the chemical analysis: a ARL 34600 OE quantimeter, a Belec Vario Lab spectrometer, and an Electron Microanalysis Camebax MICRO CAMECA with a KAVEX energy-dispersion analyzer. These various devices were used because of the great variability in the chemical composition within the different melts.

2.2.2 Verification of the mechanical properties of the turbine blades

The mechanical properties were determined with tensile, Charpy impact and Brinell hardness tests. Two different types of samples with respect to the turbine-blade orientation were extracted. The first were extracted from the connection part perpendicular (PN) to the blade axis, the second samples extracted from the blade base in a parallel orientation (PO) to the blade axis.

Tensile tests were carried out on an INSTRON 100kN Series IX Automated Materials Testing System according to the testing conditions of the standard ČSN EN 10002-1. A Padostroj Charpy 300 J was used for the Charpy impact tests, according to the standard (with a test temperature of 20 °C and specimen dimensions of (10 × 10 × 55) mm with the V-notch) ČSN EN 10045-1. The Brinell hardness was determined with an EMCO-TEST M4C universal hardness tester according to the standard ČSN EN 6506-1.

2.2.3 Light-microscopy examination of metallographic samples

Qualitative and quantitative microstructural analyses of the base material and of the fracture areas were carried out for the perpendicular (PN) and parallel (PO) samples cut out from the blades. The samples were taken from the fracture zone and from the middle parts of the blade bodies. For the examination a Zeiss-Neophot 32 light microscope with suitable magnifications was used. For the X2CrNiMo13-4 steel, a solution of 5 mL HCl, 1 mL TNP (2,4,6-trinitrofenol) and 95ml of ethyl alcohol was used as an etching agent. The 12% Cr martensitic steel AK1 TD. 9 was etched in a water solution with 2.5 mL of HNO_3.

2.2.4 Electron-microscopy examination of metallographic samples

The samples for examination with transmission electron microscopy were prepared with spark erosion and cut parallel to the crack-propagation plane. From the thin plates, discs of 2.7 mm in diameter were prepared with spark-erosion cutting and mechanically thinned to 70–80 μm. Thin foils for the electron-microscopy observation were prepared with electrolytic polishing in TENUPOL equipment, applying a 6 % solution of $HClO_4$ in methanol, a temperature of 35 °C and a current of 150 mA. The samples were examined in a JEOL JEM2000FX transmission electron microscope and the energy-dispersion microanalyzer of an RTG emission LINK AN10000 was used for the chemical analyses.

3 RESULTS AND DISCUSSION

3.1 Chemical composition and mechanical properties

The results of the chemical analysis and the measurements of the mechanical properties (yield strength, ultimate strength, ductility, reduction of cross-section area, Charpy impact strength, hardness) are summarized in **Table 2**, **Table 3** and **Table 4**. It is evident that the requirements of the standards AK1 TD.9 and WEV 400 are fulfilled almost completely. Only the lower Charpy impact strength was determined for the material of the blades 152, 156, 106. This can be due to the operation of the components for a long time or the non-standard operating conditions.

3.2 Optical microscopy

The microstructure of the base material was typical for modified 12 % Cr martensitic steel. The microstructure of the heat-treated (quenched and annealed) steel is shown in **Figure 4**. The large amount of δ-ferrite present in the microstructure was aligned in rows, oriented parallel with the initial crack growth and may have

Table 2: Chemical composition (*w*/%)
Tabela 2: Kemična sestava jekel (*w*/%)

Sample	C	Si	Mn	P	S	Cr	Ni	Mo	V	W
Specification AK1 TD.9	0.10–0.16	< 0.60	0.60	< 0.025	< 0.015	10.5–12	1.5–1.8	0.35–0.5	0.18–0.30	1.6–2.0
152	0.142 ±0.004	0.371 ±0.002	0.46 ±0.004	0.022 ±0.002	0.01 ±0.001	11.71 ±0.14	1.853 ±0.022	0.448 ±0.004	0.3 ±0.006	1.77 ±0.142
156	0.143 ±0.002	0.374 ±0.002	0.462 ±0.004	0.021 ±0.002	0.011 ±0.002	11.62 ±0.022	1.843 ±0.012	0.451 ±0.006	0.302 ± 0.006	1.768 ±0.044
106	0.16 ±0.004	0.31 ±0.006	0.41 ±0.006	0.024 ±0.004	0.010 ±0.002	11.44 ±0.014	1.90 ±0.042	0.44 ±0.004	0.29 ±0.008	1.59 ±0.022
Specification 1.4939	0.15	< 0.35	0.9	0.020	0.015	12.50	3.00	2.00	0.40	–
New mat. 1.4939+QT	0.314 ±0.004	0.102 ±0.004	0.829 ±0.018	0.011 ±0.002	0.007 ±0.001	11.57 ±0.20	2.498 ±0.076	1.568 ±0.014	0.345 ±0.008	0.222 ±0.036
Specification WEV 400	max. 0.05	0.15–0.35	0.20–0.80	max. 0.020	max. 0.025	12.0–14.0	3.5–4.5	0.30–0.50	–	–
344	0.013 ±0.002	0.29 ±0.002	0.76 ±0.004	0.015 ±0.008	0.002 ±0.001	12.88 ±0.012	4.14 ±0.022	0.46 ±0.003	0.02 ±0.002	0.017 ±0.014

Table 3: Mechanical properties
Tabela 3: Mehanske lastnosti

Sample (blade No.)	Yield Strength $R_{p\,0.2}$/MPa	Ultimate strength R_m/MPa	Ductility *A*/%	Contraction *Z*/%
152/1 PN	845	957	20.0	55.4
152/2 PN	860	969	20.0	55.4
156/1 PO	830	948	18.3	55.4
156/2 PO	835	966	16.7	55.4
156/3 PN	820	962	16.7	55.4
106/1 PO	860	965	16.7	50.9
106/2 PO	849	957	17.3	55.4
106/3 PN	865	976	16.7	50.9
344/1 PO	951	989	18.3	81.2
344/2 PO	950	984	18.0	79.7
344/3 PN	952	987	18.3	79.7
Specification AK1 TD.9	min. 700	850–1000	min. 14	min. 40
Specification WEV 400	min. 950	1000–1100	min. 12	min. 45

Table 4: Charpy impact test and hardness [HBW]
Tabela 4: Charpyjeva žilavost in trdota

Temperature, *T*/°C		+ 20 °C					Average value HBW 2.5/187.5
Blade No.	Sample	*KV*/J			*KCV*/(J/cm^3)	Required *R*/(J/cm^3)	
152	1PO-2PN-3PN	27	28	28	35	50	255
156	1PO-2PO-3PN	41	28	27	40		255
106	1PO-2PO-3PN	36	31	32	41		264
344	1PO-2PO-3PN	206	201	196	251		295

greatly influenced the propagation of the fatigue cracks (see **Figure 5**).

The microstructure of the base steel of X2CrNiMo13-4 consisted of sorbite, and was typical for the heat-treated (quenched and annealed) steel. No δ-ferrite inserts were found within the structure, even at a magnification of 400 (see **Figure 6**). The quantitative analysis of the amount of δ-ferrite is shown in **Figure 7**. The size of the original austenitic grains was evaluated according to the standards ČSN ISO 643 and ASTM E 112. A combination of Nital 3 % and Vilella-Bain etching agents was used to highlight the original austenitic grains. The obtained values showed that its size varied between 4.5 and 5.5 for both steel types (modified 12 % Cr martensitic steel and X2CrNiMo13-4 steel).

3.3 TEM

Figure 9 shows a grain of δ-ferrite in the modified 12 % Cr martensitic steel (**Figure 8**). The TEM revealed that the microstructure of blade No. 106 was similar to the microstructures of blades No. 152 and 156. Accord-

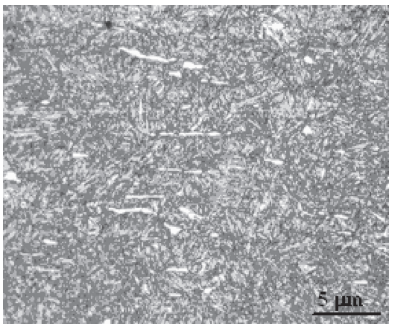


Figure 4: Blade No. 152 (PN), 200x, microstructure of base steel
Slika 4: Lopatica št. 152 (PN), mikrostruktura osnovnega jekla (200-kratna povečava)

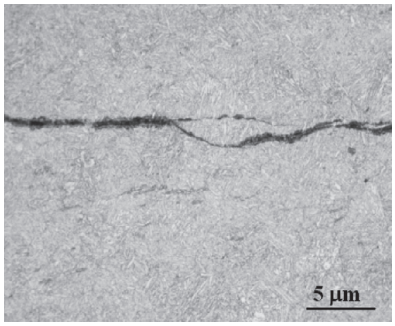


Figure 5: Blade No. 106, 200x, longitudinal crack and aligned inserts of δ-ferrite
Slika 5: Lopatica št. 106, podolžna razpoka in usmerjeni vložki δ-ferita (200-kratna povečava)

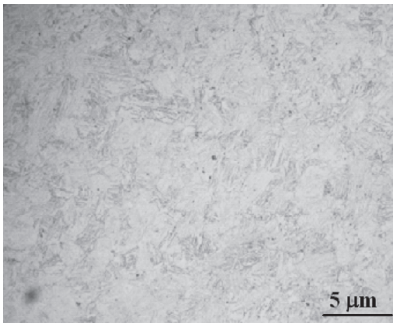


Figure 6: Blade No. 344 (PO), 200x, microstructure of base steel
Slika 6: Lopatica št. 344 (PO), mikrostruktura osnovnega jekla (200-kratna povečava)

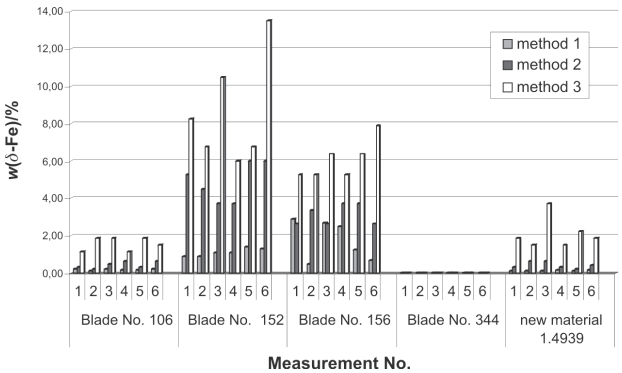


Figure 7: Content of δ-ferrite for different materials determined with different methods
Slika 7: Vsebnost δ-ferita v različnih jeklih, določena z različnimi metodami

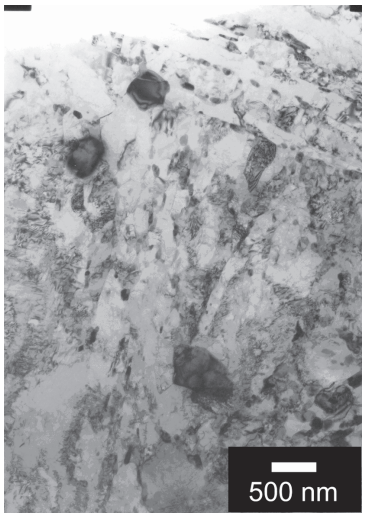


Figure 8: Blade No. 152, martensitic microstructure of base steel
Slika 8: Lopatica št. 152, martenzitna mikrostruktura osnovnega jekla

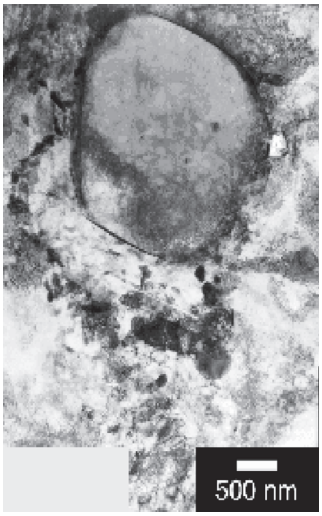


Figure 9: Blade No. 152, detail of δ-ferrite particle
Slika 9: Lopatica št. 152, detajl delca δ-ferita

Figure 10: Blade No. 344, martensitic microstructure
Slika 10: Lopatica št. 344, martenzitna mikrostruktura

Figure 11: Blade No. 344, fine precipitates displayed in dark field
Slika 11: Lopatica št. 344, izločki, prikazani v temnem polju

ing to the electron diffraction results, the carbide particles were of the $M_{23}C_6$ type and the microstructure was practically without δ-ferrite.

Blade No. 344 of the steel X2CrNiMo13-4 (**Figure 10**) contained very fine precipitates of the intermetallic phase Fe_2Mo (**Figure 11**). Similar fine particles were also found in the newly developed steel 1.4939, which was free of coarser particles typical for the other analysed steel. Nevertheless, the local occurrence of a very limited quantity of these carbide particles cannot be excluded.

3.4 Fractographical analysis of damaged blades (modified 12 % Cr martensitic steel)

The fracture surfaces were covered with a thick layer of corrosion products that complicated the investigation of the micromorphology in the region of the fatigue cracks. These layers were removed with ultrasonic cleaning. The fractographical analysis showed that the cause of the blade damage was high-cycle fatigue-crack growth in the vicinity of the connection part of the blades and the gas-turbine shaft.

The cracks started on the blade surfaces in the upper groove of the connection area (see **Figure 12**) and then propagated with transcrystalline decohesion. It can be assumed that the δ-ferrite particles acted as the initiation points of the fracture. Many progressing lines were detected on the fracture areas of the damaged blades. The presence of these lines indicates that the loading was not constant and this is typical for fatigue fracture. From the number of progressing lines, the number of starting regimes of the gas turbine could be estimated.

From the width of certain growth rings it can be concluded that during stable operating conditions the crack-propagation stopped. From the microfractography it can be further concluded that the fatigue failure propagated with the striation mechanism (see **Figure 13a**).

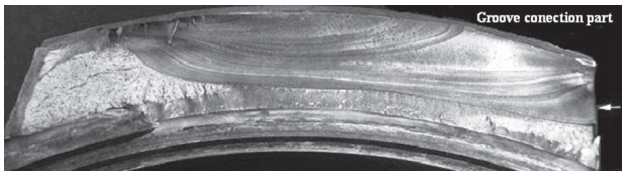


Figure 12: Fracture area of blade No. 152 (the white arrow shows the front line beneath the striking edge)
Slika 12: Prelomna površina lopatice št. 152 (svetla puščica kaže na čelno črto pod udarnim robom)

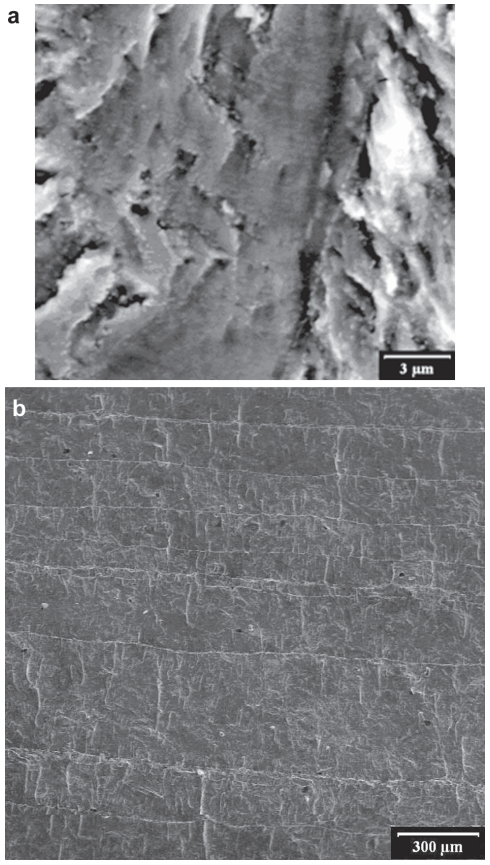


Figure 13: Characteristic microfractographical signs of fatigue failure. (a) Blade No. 152, field of submicron striations, (b) Blade No. 152, propagation lines
Slika 13: Karakteristični mikrofraktografski znaki utrujenostnega preloma. (a) Lopatica št. 152, področje submikrometrskih brazd, (b) lopatica št. 152, brazde propagacije

In addition, the presence of the progressing line determined the location and the shape of the fatigue crack tip at the moment of the overloading cycle. The dependence between these lines and the operating conditions made it possible to reconstruct the whole history of the failure process (see **Figure 13b**).

Estimation of the time-dependent failure propagation (modified 12 % Cr martensitic steel)

On the blades from the 3.TW (modified 12 % Cr martensitic steel AK1 TD. 9) on LP3 extensive maps were prepared (at various magnifications up to 300-times). This made it possible to obtain detailed information about the history of the fatigue-crack propagation.

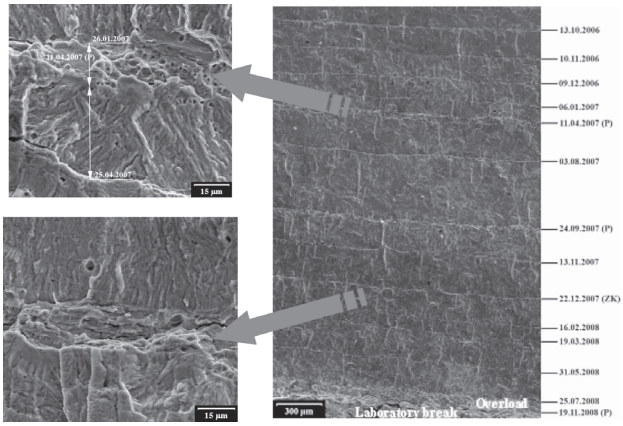


Figure 14: Blade No. 156. Propagation lines on the fracture area and dates of turbine operation

Slika 14: Lopatica št. 156. Linije propagacije razpoke in datumi rasti razpoke

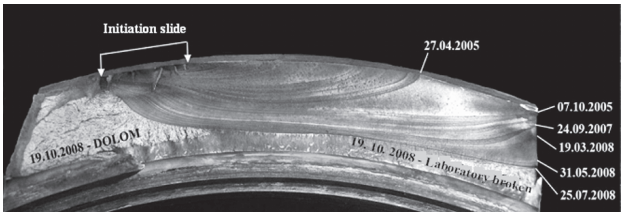


Figure 15: Blade No. 156. Dates, when the crack front reached the line of final blade failure

Slika 15: Lopatica št. 156. Na makrografiji je označen datum, ko je čelo razpoke doseglo mejo končnega preloma lopatice

The maps were closely analyzed with respect to the operation history of the gas turbine.

Figure 14 shows the propagation lines identified by these observations. These lines correspond to the data on the left-hand side of the micrograph. As mentioned above, the lines are related to the table operation of the turbine, and from the known data of the specific factory test the propagation regime on the fracture can be derived. The main results of the fatigue fraction history of the blades from 3.TW on LP3 in the fracture area can be seen in **Figure 15**.

4 CONCLUSION

The mechanical properties confirm that the requirements for the mechanical properties of most of the blades from the modified 12% Cr martensitic steel are fulfilled. However, the following exceptions were identified:

1. The blade No. 344 (X2CrNiMo13-4 steel) does not fulfil the requirements for ultimate strength. The lower value of 1.5 % is negligible, but it also positively affects the increase in the ductility parameters A_5 and Z.
2. The Charpy impact tests of blades No. 152, 156 and 106 revealed that measured values do not fulfil the requirement for impact toughness.
3. In contrast to the impact toughness of blade No. 344 (X2CrNiMo13-4 steel) was 2.5 times higher than required.
4. The hardness values fell within the required interval that allows higher hardness values within the blade No 344 (X2CrNiMo13-4 steel) in comparison with the modified 12 % Cr martensitic steel.

The low resistance to fatigue failures (mainly the resistance to fatigue-crack propagation) of the used materials in the working conditions can be stated as the main reason for the fatigue cracks' initiation.

The initiation of the fatigue cracks was a result of the presence of δ-ferrite (modified 12 % Cr martensitic steel). The fractographical reconstruction revealed that the initiation of the fatigue cracks in the blades started quite early after the initialization of the working process.

From the microstructural analysis it was concluded that the presence of δ-ferrite (the shape and the particles) negatively influences the resistance of the modified 12 % Cr martensitic steel to crack initiation. The subsequent propagation of fatigue cracks is controlled by the main operating load. The new material, 1.4939, without tungsten, which eliminates the presence of δ-ferrite to a large extent, is proposed for the manufacture of 3.TW blades.

5 REFERENCES

[1] J. Janovec, J. Siegl, CTU in Prague, 12/2008, Research report No. 10-08 Příčiny porušení lopatek 3. oběžných kol NT3 a NT1 TG1 ETE na 1. HVB (část 1) – (Causes of failure of 3th low pressure turbine wheel in LP3 and LP1 of TG1 in 1. HVB (part 1))

[2] J. Siegl, CTU in Prague, FJFI V-KMAT-746/08 – Research report Fraktografická analýza porušených lopatek NT dílů TG1 ETE (část 2) – (Fractographical analysis of failure blades LP parts of TG1 ETE (part 2))

[3] J. Siegl, CTU in Prague, FJFI V-KMAT/2009 Presentation Fraktografická analýza porušených loptek NT2 dílu TG2 ETE (Fractographical analysis of failure blades in LP2 parts of TG2 ETE))

[4] J. Janovec, J. Siegl, CTU in Prague, 12/2008 Presentation Příčiny porušení lopatek 3. oběžných kol NT3 a NT1 TG1 ETE na 1. HVB (Causes of failure 3rd in LP3 and LP1 of TG1 ETE)

[5] J. Zenkl, O. Novák, ČEZ, a.s. ETE, ČEZ, a.s. EDU, Analýza příčin poškození lopatek na NT dílech TG 1000 MW (Analysis of the causes of damage to the blades on LP parts TG 1000 MW)