Micropitting and Martensite Decay in Gears

by

Chi Lung

A Thesis Submitted for the degree of Doctor of Philosophy

School of Chemical Engineering and Advanced Materials,

Newcastle University

July 2016

Declaration

I declare that the entire content of this thesis is, to the best of my knowledge and belief, original, unless otherwise acknowledged in the text. This thesis records work carried out in the School of Chemical Engineering and Advanced Materials at Newcastle University between September 2011 and September 2015. I have not submitted this material, in whole or part in any other qualification at any other university.

Acknowledgements

The author would like to acknowledge his supervisor, Dr. Adrian Oila, for his guidance, encouragement and support throughout the course of his study. He would also like to thank Prof. Steve Bull for his assistance and support.

The author thanks all the members in the design unit - Gear Technology Centre at the University of Newcastle upon Tyne for their valuable discussions. Special thanks to Dr. Ruth Chantry at the University of Leicester for providing training and assistance for the TEM. Thanks to Dr Jerry Haggon from the School of Chemistry at Newcastle University for providing help with the server to run the simulations.

Finally, the author would like to express his sincere gratitude to his family for their unconditional sacrifices, understanding and love.

Abstract

Micropitting is a type of surface contact fatigue often observed in gears and rolling element bearings operating under mixed or elastohydrodynamic lubrication (EHL). Once initiated, micropitting will lead to the catastrophic failure of the affected components which then requires unplanned industrial shutdowns to allow for their replacement. Micropitting in carburised gears has become a major concern in the wind power industry and other sectors where gears operate at relatively high loads and relatively slow speeds. It occurs most often in parallel axis gears (spur and helical) but it was also reported in other types of gears such as spiral bevel gears.
An important feature that has been observed in bearings damaged by micropitting was the transformation of the initial microstructure. The change in microstructure, known as martensite decay, consists in the development of a new phase known as Dark Etching Region (DER) due to its appearance in reflected light microscopy. This microstructural change which has also been observed in gears leads to changes in the mechanical properties in the affected regions with implications on the initiation and propagation of the cracks leading to the formation of micropits. The fatigue life of gears can be extended by controlling the formation of micropits but this requires an in-depth understanding of the micropitting mechanism.

The aim of this project was to identify and characterise the microstructural changes accompanying micropitting in gears and to develop a micropitting mechanism which describes the formation of the micropits. The microstructure has been investigated by electron microscopy techniques such as Scanning Electron Microscopy (SEM), Electron Backscattered Diffraction (EBSD), and Transmission Electron Microscopy (TEM). Nanoindentation was used to determine the mechanical properties of the affected regions. The proposed micropitting mechanism is based on the results from the above investigations as well as Density Functional Theory (DFT) calculations of material properties combined with Finite Element Analysis (FEA) of the contact region.

Nomenclature

A_c (m^2)	area of contact
a (m)	radius of the contact area
a,b (m)	semi-axes of the contact ellipse
a_i (m/s^2)	acceleration
a_γ (Å)	lattice parameter
b (m)	half width of the contact rectangle
B	bulk modulus
C_γ (Å)	lattice parameter of austenite
c_s	carbon concentration at the surface
c	ellipticity
d(m)	interplanar spacing
D (m^2/s)	diffusion coefficient
D_0 ($cm^2s^{-1)}$	frequency factor
E (N/mm^2)	Young's modulus
E' (N/mm^2)	elastic contact modulus
E_{xc}	exchange correlation density
F (N)	frictional force
g (m/s^2)	acceleration due to gravity
G_{Reuss}	shear modulus
G	reciprocal lattice vector
G_m	materials parameter
$h_{min}(m)$	minimum film thickness
h_p (m)	depth penetaration
H (GPa)	hardness
J_x (mol/m^2s)	diffusion flux
k_c	carbon coefficient
K	symmetry label in the brillouin zone

l (m)	length
M_s (C)	martensite start temperature
m (kg)	mass
N (N)	normal force
N_s	fatigue life
N_v	number of martensite plates per unit volume
n_t	rate of conversion
P (N)	applied load
p_o (Pa)	contact pressure
Q (kJ mol^{-1})	activation energy
R (m)	radius of curvature of the equivalent cylinder
R', R'' (m)	major and minor relative radii of curvature
R_a	roughness average
S	slide-to-roll ratio
$S(N)$	probability of survival
t (s)	time
T (K)	temperature
U_s	speed parameter
U	potential energy
u (m/s)	velocity
V (m^3)	volume
W (N)	load
W_L	load parameter
X	number of atoms
z (m)	depth of maximum shear stress
Z	atomic number
τ (Pa)	shear stress
v	poisson's ratio

μ	friction coefficient
β	weibull dispersion parameter
β_g	geometry correction factor
$\hat{H}$	hamiltonian operator
E	total electronic energy
ψ	molecular wave function
$\hat{T}$	kinetic energy operator
$\hat{V}$	potential energy operator
$\hbar$	Planck constant
λ (m)	wavelength
ϵ_{xc}	exchange correlation energy
ρ (kg/m^3)	density
$\partial c/\partial x$	concentration gradient
η_0 (cp)	viscosity
ρ (kg/m^3)	density
σ (MPa)	stress
θ	Bragg angle

List of Figures

List of Tables

Table of Contents

Chapter 1. Introduction

1.1 Background

Micropitting is a form of fatigue damage most commonly found in gears and bearings. It occurs due to small cracks formed at the tooth surface which grow until small volumes of material break out of the tooth surface. The British Standard Institution (BS 7848: 1996) defines micropitting as "a form of surface fatigue phenomena which consists of degradation of gear tooth working surfaces under lubrication conditions where the film is too thin for the load". Micropitting is a type of surface fatigue damage influenced by load, surface topography, and lubrication conditions. It often causes catastrophic failures such as tooth breakage. The lubricant thickness in rolling/sliding contacts depends on load and speed (Crook, 1957). If the lubricant film is too thin, micropitting can occur on the gear surface.

Contact fatigue was first observed in the early 1800s (Grassie, 2003) when investigators found that railroad tracks were cracking due to repeated loading. As technology developed over the years, the increase use of machines means more and more components are subjected to surface contact fatigue. The material choice for gears depends on the application and operation conditions such as load requirements, speed and torque as well as cost. Some common materials used include steel, cast irons, aluminium alloys, bronzes, and plastics. Steel remains one of the most common materials used in gears due to the high load carrying capacity compared to nonferrous gear materials. Micropitting is observed by industry in all types of gears. It can occur after a short period of operation, less than 10^6 cycles (Akamatsu, 1989; Spikes, 1986).

Micropitting is commonly observed in wind turbine gearboxes after failure analysis (Al-Tubi, 2013). Surface damage and failures in contacting components such as gears and bearings are among the most costly types of failures and the main cause of failure in a gear box. As a result of micropitting, the wind turbine operation costs increase due to part replacement and shutdowns. Micropitting occurs in all models and sizes of the wind turbines, in both gears and bearings. Micropitting occurs in wind turbine gearboxes due to a variety of factors including unsteady winds, contamination of lubricant, high loads, and rotational speed (Al-Tubi, 2013). It should be mentioned that in wind turbines the lubricant film thickness varies considerably depending on the speed of the gears and micropitting is most likely to occur at the addendum and dedendum of the driving gear (Al-Tubi, 2013).

Microstructural transformations were often found below the contact surface of steel bearings and gears subjected to micropitting (Dinner, 2011; Choi et al., 1998; Voskamp 1985). The affected regions have a distinct colour when etched and observed under the light microscope. These regions are referred to as martensite decay.

1.2 Objectives

The aim of this project was to investigate the martensite decay phenomena observed on gear subjected to surface contact fatigue and determine the mechanical properties by various microstructural characterisation techniques. The data is used to compare the mechanical properties of dark etching region with tempered martensite, and use the properties calculated and measured as input in a FEM.

1.3. Surface Contact Fatigue

The definition of contact fatigue contained in ASTM E 1150 states that it is the process of progressive localised structural change which occurs in a material subjected to fluctuating

stresses and strains and that it may culminate with formation of cracks or in complete fracture after a sufficient number of fluctuations (ASTM, 1995). Surface contact fatigue occurs in many applications including ball bearings, gears, and track rails. In gears, plastic deformation can occur in the regions below the contact where the shear stress is maximum.

1.4 Surface Contact Fatigue in Gears

Surface contact fatigue phenomena can be classified into four main categories (BS 7848:1996). These categories are pitting, micropitting, flaking, and case crushing.

Pitting occurs when small volumes of materials are broken out of the surface and leave a cluster of pits (see Figure 1). It is often observed in rolling element bearings and gears. According to the American Gear Manufacturers Association (AGMA), pitting can be further categorized based on its severity starting from initial pitting, destructive pitting and spalling. The initial pits have diameters in the range between 400 μm to 800 μm. (Dudley, 1962). Initial pitting occurs in locations with high contact stresses such as grinding marks but it can initiate in the subsurface regions at defects that intrinsic to the steel.

Figure 1. Pitting damage on gear (www.novexa.com).

Destructive pitting refers to pits that are larger in size compared to initial pitting. As a result, the amount of the material lost from the surface is higher and changes in the gear surface profile will also occur.

Spalling occurs when larger volumes of material break away from the surface (Figure 2) when smaller size pits combine into a larger pit.

Figure 2. Spalling damage on pinion (www.orientalmotor.com).

Micropitting is the damage of metallic surfaces by the formation of asperity scale microcracks and asperity scale micropits (ASM Handbook, 2001). Micropitting occurs when microscale pits (pits with diameter below 100 microns) develop on the surface (see Figure 3). These micropits appear similar to an etched steel surface.

Figure 3. Micropitting damage on gear tooth surface (www.xtek.com).

Flaking occurs when large volumes of material with a shallow depth detach from the surface (see Figure 4). It occurs most often in bearings. Flaking can be further classified into surface originated flaking and subsurface originated flaking.

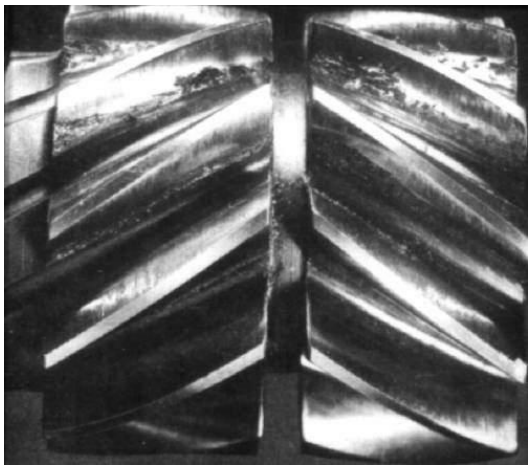

Figure 4. Flaking damage on gears (Dudley, 1962).

Surface originated flaking is generally caused by contaminated lubrication conditions (Furumura,1993). Flaking initiated below the surface is generally initiated at non-metallic inclusions (Furumura,1993).

Case crushing occurs in case hardened gears when cracks propagate along the case core interface (Figure 5).

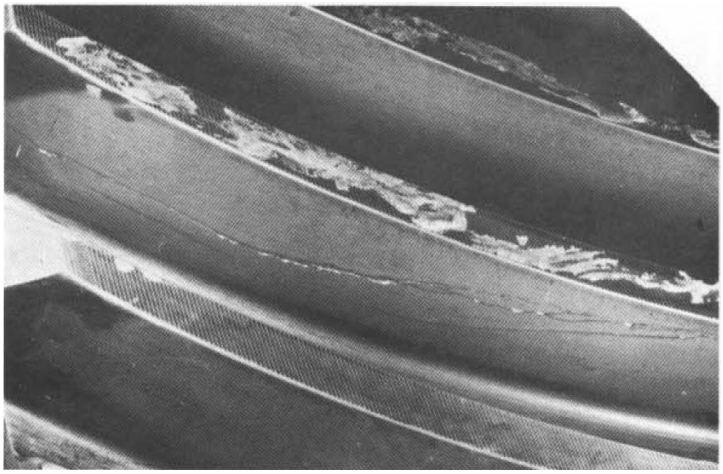

Figure 5. Case crushing in gears (www.orientalmotor.com).

1.5. Phase changes due to surface contact fatigue

The microstructure of case carburised gears consists of tempered martensite in the case and ferrite, pearlite and/ or bainite in the core. Martensite is a metastable phase and diffusion may occur during gear operation due to temperature rise and alternating stresses. Martensite can transform to ferrite and carbides due to an increase in temperature or due to elastic and/or plastic cyclic loading, or a combination of these. As a consequence, certain properties of the material including hardness, elastic modulus and yield strength will also change and the affected regions will no longer support the contact loads. Microstructural transformations associated with surface contact fatigue in general and micropitting in particular in steel bearings have been well documented (Swahn et al., 1976; Osterlund & Vingsbo, 1980; Voskamp, 1985; Bhargava et al., 1990; Gentile et al., 1965) but very few reports have been published on martensite decay in gears (Oila et al,. 2004). Figure 6 shows martensite decay in a gear tooth affected by micropitting after etching the polished cross section with nital.

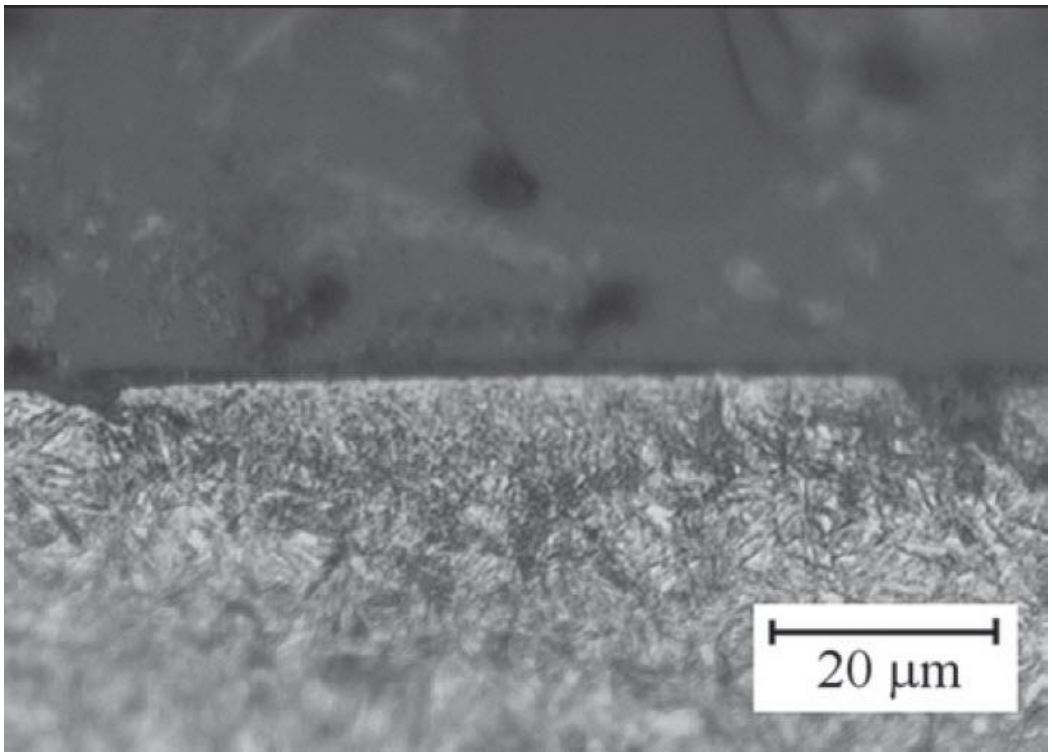

Figure 6. Micrograph showing martensite decay as a semi-circular zone beneath the gear tooth surface (Oila et al., 2004).

1.6 Thesis Structure

The thesis is organised into seven chapters. The introductory chapter provides an overview of the whole thesis. The thesis is grouped into two topics; the micropitting/microstructure characterisation and computational simulation. The microstructure characterisation and its related background are discussed in chapter 2 and 3. Chapter 4 explains the experimental methods. The computational simulations are discussed in chapter 5. Chapter 6 and 7 discuss the results and conclusions.

Chapter 2 includes background information related to gears and steel microstructure. It begins with a brief description of gear manufacturing, heat treatment and microstructure of gear steels. Chapter 2 also includes a section on contact mechanics and a presentation of the microstructural transformations associated with surface contact fatigue along with the characteristics of the micropitting.

Chapter 3 discusses the theory behind the computational experiments. It begins with an introduction to the ab-initio method and the density functional theory.

The experimental method and microstructural analysis are discussed in Chapter 4. It includes the process of the sample preparation. The mechanical properties and the treatment applied are discussed with the preparation required for each analysis.

The computational simulation method is presented in Chapter 5. The chapter describes the steps taken for the simulation of martensite. It includes methods used for geometry optimisation and calculation of the elastic constants.

Chapter 6 presents the assessment of micropitting and its appearance.

Chapter 7 presents the results from the microstructural investigation with light microscopy, scanning electron microscopy, nanoindentation, and transmission electron microscopy.

Chapter 8 presents the modelling of the effect of martensite decay on gears.

Chapter 9 presents the finite element part of the investigation.

The concluding remarks and future extensions of this study are presented in Chapter 10.

Chapter 2. Micropitting and Microstructural Changes

2.1 Background

Two of the most important requirements for metallic gears are high power transfer efficiency and long lasting durability. Both the efficiency and the durability of gears are strongly influenced by their tribological performance. Micropitting is a type of surface contact fatigue often observed in gears and rolling element bearings operating under elastohydrodynamic lubrication (EHL). Once initiated, micropitting leads to the catastrophic failure of the affected components which then requires unplanned industrial shutdowns to allow for their replacement. An important feature that has been observed in bearings damaged by micropitting was the transformation of the initial microstructure (Gentile et al., 1965; O'Brien & King, 1966; Buchwald & Heckel, 1968; Swahn et al., 1976; Becker et al., 1976; Osterlund & Vingsbo, 1980; Voskamp, 1985; Bhargava et al., 1990). This change in microstructure, known as martensite decay has also been observed in fatigued gears (Hoeprich, 2001; Oila, 2005). Martensite decay consists in the development of a new phase known as Dark Etching Region (DER) due to its appearance in reflected light microscopy. Changes in material microstructure result in changes in material properties with adverse effect on the gear lifetime because of the implications on the initiation and propagation of the cracks leading to the formation of micropits. The fatigue life of gears can be extended by controlling the formation of micropits but this require an in-depth understanding of the micropitting mechanism.

Most failure predictions are based on the initial material properties and contact conditions (Bush et al., 1961; Berthe et al., 1980; Yoshida, 2008; Webster, 1995; Wang et al, 2000), but failure predictions based on microstructure evolution during operation (Becker et al, 1990; Rangaswamy, 2000; Polonsky, 1994) are employed to a lesser extent.

The microstructure and the mechanical properties of gear steels depend on manufacturing methods and the heat treatments applied while understanding the tribological behaviour of gears requires knowledge of contact mechanics. The aim of this chapter is to provide the background information on gears manufacturing, gear heat treatment, metallurgy of gear steel, and contact mechanics. Micropitting and martensite decay are treated in the last two sections of this chapter.

2.2 Gear Manufacturing Methods

Gears can be divided into two main categories (Dudley, 1962): parallel axis gears which connect parallel shafts, and non-parallel axis gears which are designed to redirect the axis of rotation. Examples of parallel axis gears are spur gears and helical gears and non-parallel axis gears include bevel gears, worm gears, and hypoid gears. In general, slow speed gears are more prone to micropitting due to thinner lubricant films.

2.2.1 Gear Types

Spur gears have teeth parallel to the axis of the shafts. The shape of the tooth is in an involute form projected radially. Spur gears (see Figure 7) are usually the least expensive of all gear types due to the ease of manufacturing. Micropitting is a common failure mechanism found in spur gears. It is observed frequently by the industry in carburised gears and all types of steel gears are susceptible to micropitting (Hohn & Michaelis, 2004; Brandao et al., 2009; Cardoso et al., 2009; Amarnath et al., 2012).

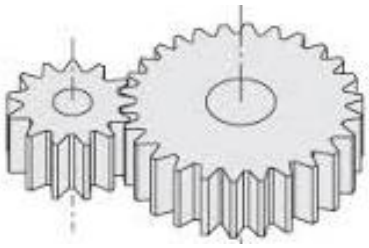

Figure 7. Spur gear (Dudley, 1962).

Helical gears (Figure 8) have teeth that are inclined to the axis of rotation. Because of the spiral shape teeth, the contact is continuous so helical gears are generally quieter than spur gears. The helix angle ranges from a few degrees to 45^0 (Dudley, 1962). The power transfer efficiency of helical gears is generally lower than spur gears. Due to the angle of the teeth, they have a larger surface area of contact, which increases the load capacity. However, this creates higher sliding friction compared to spur gears. Micropitting has been observed in helical gears (Hoeprich, 2001) and it was associated with the dark etching regions caused by high asperity contact and traction stresses (Hoeprich, 2001).

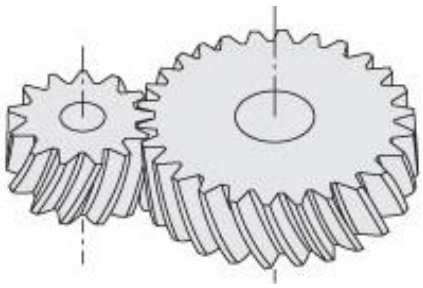

Figure 8. Helical gear (Dudley, 1962).

Non parallel axis gears are characterised by their ability to redirect the power to another direction. Most non parallel gears operate on nonparallel, coplanar axis. Some of the gear types include worm, bevel, and hypoid gears.

Worm gear sets (see Figure 9) consists of a worm gear in contact with a worm element. The axes of the two elements are positioned at a right angle and they have the role to

reduce the rotational speed at high torque transmission. Micropitting was found on worm gears under lubricated conditions (Snidle & Evans, 1997).

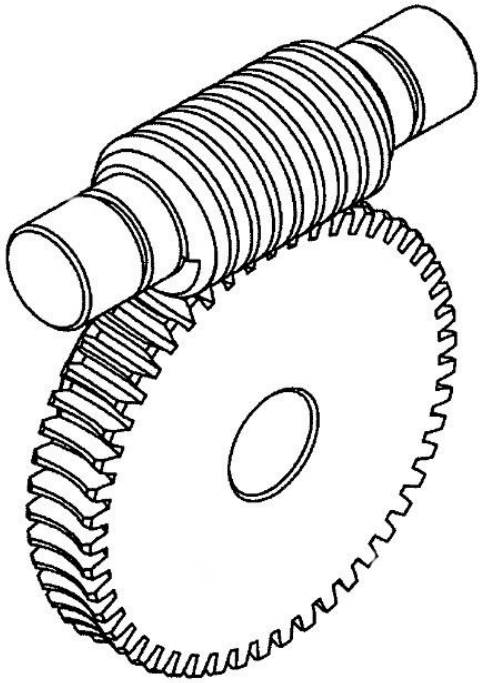

Figure 9. Worm gear (Dudley, 1962).

Bevel gears (Figure 10) are conical and the axes of the two shafts intersect each other. They can be used to change the axis of rotation depending on the application. Bevel gears are used when high efficiency is required and they can achieve 98% or higher efficiency (Davis, 2005). Micropitting has been reported to occur on bevel gears under lubricated condition (Hoehn et al., 2012).

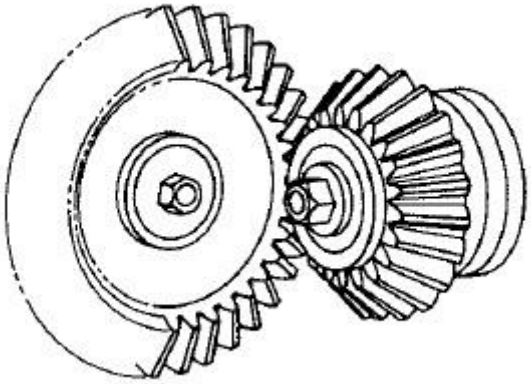

Figure 10. Bevel gear (Davis, 2005).

Hypoid gears (Figure 11) have curved teeth and the pinion axis offset from the gear axis.

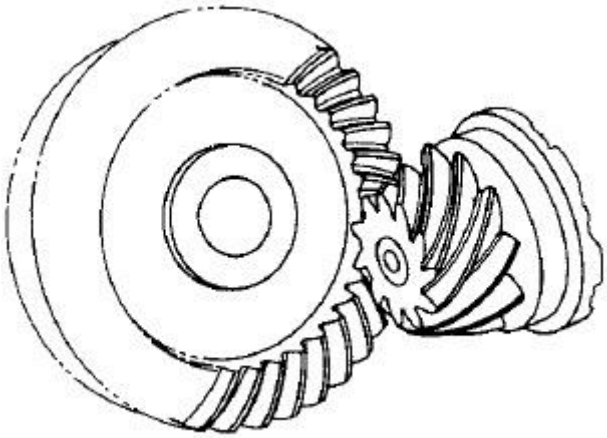

Figure 11. Hypoid gear (Davis, 2005).

2.2.2 Gear Manufacturing

Gears can be manufactured by a wide variety of processes such as casting, forging, and extrusion. The manufacturing method depends on the material used and the intended application. Machining is generally required to achieve the final dimensions and shape.

Gears can be manufactured by one of the following methods: generation, cutting and moulding. The generation principle consists of the synchronised movement of a tool - as an equivalent rack, with the rotation of the gear blank. Hobbing, shaping, planing, shaving, honing and some grinding methods are based on the generation principle.

Cutting involves a tool profile that is shaped as an exact replica of the tooth gap to be machined. Milling, broaching, blanking and some grinding methods are based on cutting.

Moulding is achieved by: injection moulding, die casting, sintering and forging.

The blanking operation is the operation that produces a blank for gear machining and it involves the initial metal cutting in order to remove the excess material and to machine the blank to the required specifications.

Gear milling is a machining process where gear teeth are produced by a form cutter. Milling is used to roughing and finishing spur gear. It can be used to machine any tooth form with different teeth spacing or angle.

Hobbing is a machining process where gear teeth are formed through the use of a hob. A hob is a cutting tool used to cut the teeth into the material. The cylindrical hob rotates and cuts out the desired shape of the teeth in a blank. The hob and the gear blank rotate continuously until all the teeth are formed.

Gear shaping is a process which involves continuous, same-plane rotational cutting. It produces a high quality surface finish which is beneficial for the reduction of wear (Britton, 2000). Shaping is based on the action of a cutter mounted on a cutter-spindle. The cutter makes passes through the gear blank generating a series of cuts with the workpiece rotating 360 degrees.

Shaving is a gear finishing operation that removes small amounts of material and corrects the inaccuracies from gear cutting.

Gear finishing is a generation process that uses a pinion type tool. The cutter rotates and the cutting axis is parallel to the gear axis. The cutter rotates slowly in relationship with the gear blank at the same pitch-cycle velocity. The advantage of this type of process is the high accuracy able to be achieved. Finishing is done to ensure proper surface finish and dimensional accuracy. The conventional finishing process used are shaving, grinding, honing, and lapping.

Grinding is a technique of finishing using an abrasive wheel. It is a process that is used to form fine-pitch teeth or remove heat treatment distortions (Davis, 2005).

Honing is a process that was developed to reduce noise generated during gear operation by removing nicks, improving surface finish, or making correction in tooth related to heat-treat distortions.

Lapping is a polishing operation used to refine the tooth surface. The mating gear and pinion run together under a controlled load while a mixture or abrasive compound is pumped onto the pair.

2.3 Heat Treatments and Microstructure

Heat treatment is an important step in gear manufacturing used to enhance the properties (most often to increase the surface hardness) of the gear steel. The change in microstructure due to the heat treatment enhances the case hardness and the core strength so the gear will withstand the contact stress (case) as well as bending stresses (core).

2.3.1 Fe-Fe_3C Phase Diagram

The phase transformations which occur by diffusion in steels can be studied using the Fe-Fe_3C phase diagram (see Figure 12). A phase diagram is a graphical representation of the relationships among environmental constraints (e.g., temperature and pressure),

composition, and regions of phase stability, typically under conditions of equilibrium (Callister, 2015). A phase is defined as a homogeneous portion of a system that has similar physical and chemical characteristics (Callister, 2015). The microstructure of an alloy is related to the characteristics of its phase diagram, and the microstructure and mechanical properties are closely related to one another. *Microconstituent* is a term used to describe an element of microstructure having identifiable and characteristic structure (Callister, 2015).

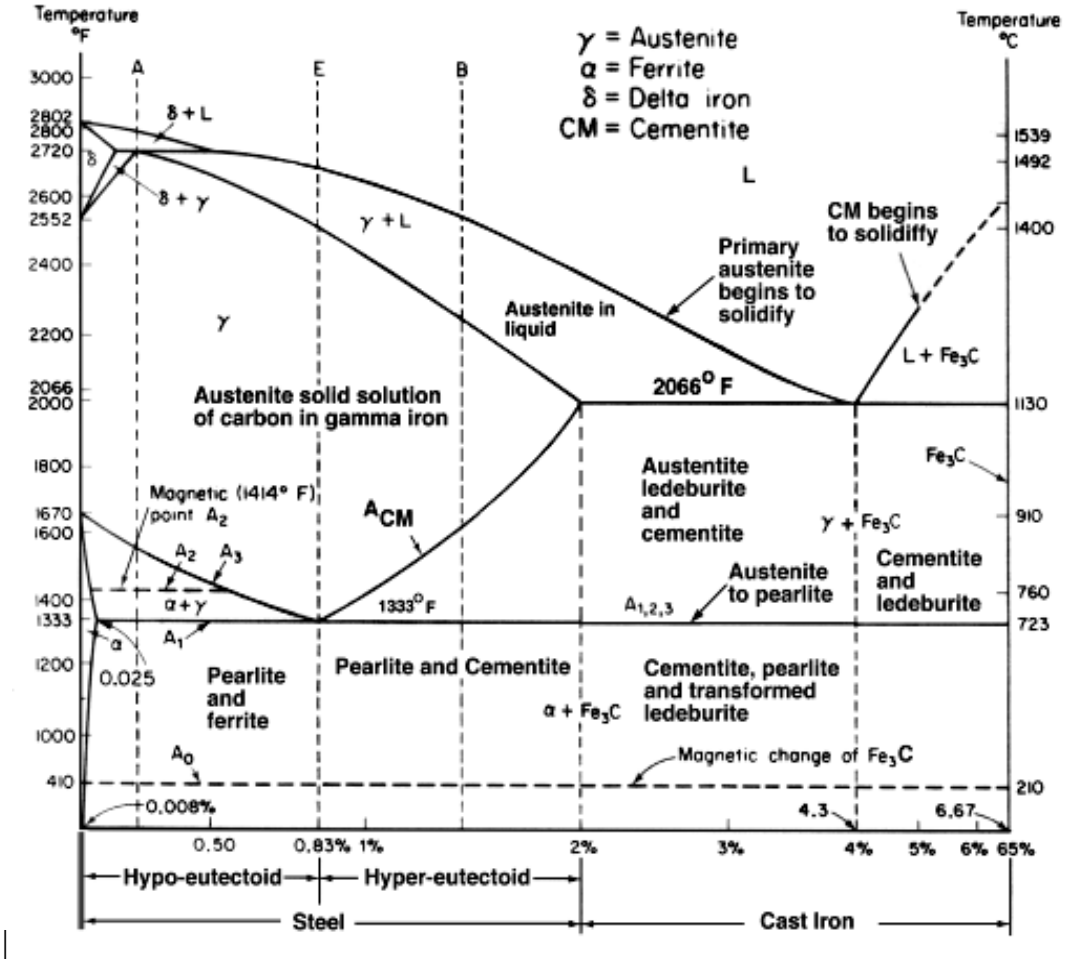

Figure 12. Fe-Fe_3C phase diagram (Pollack, 1988).

2.3.2 The allotropy of Fe

The property of a chemical element to exist in more than one crystalline form is known as *allotropy*. The property is called *polymorphism* when it applies to chemical compounds.

An allotrope is a crystal structure stable at a given temperature and pressure. There are three allotropic forms of iron under atmospheric pressure (see Figure 13): α-iron which has a body centred cubic (BCC) structure and is stable below 912 °C, γ-iron with a face centered cubic (FCC) structure and stable between 912 °C and 1394 °C, and δ-iron with a BCC structure which is stable between 1394 °C and 1530 °C. The α to γ transformation is accompanied by a 1% volume change which can lead to internal stress (Honeycombe, 1995).

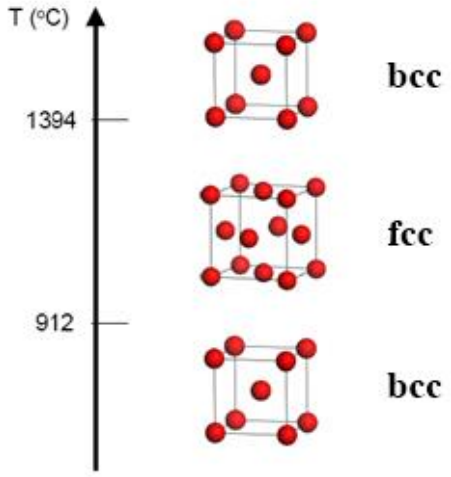

Figure 13. Unit cells of BCC-Fe and FCC-Fe.

The solid solutions of carbon in iron are named: *ferrite* (carbon in α-iron), *austenite* (carbon in γ-iron) and δ-*ferrite* (carbon in δ-iron). The only chemical compound found in the Fe-Fe_3C phase diagram is the iron carbide Fe_3C known as cementite. At 0.8 % wt. C below 723 °C a lamellar constituent called *pearlite* forms in the binary system Fe-Fe_3C.

2.3.3 Ferrite

Ferrite is the BCC solid solution of carbon in iron. The lattice parameter varies with the carbon content but because there is very small amount of carbon in ferrite (maximum 0.025 % wt. at 723 °C) the effect on the lattice parameter is negligible. Ferrite belongs to the space group $Im\bar{3}m$ (no. 229). The Wyckoff positions of the iron atoms are Fe 2*a*

(0,0,0). At 25°C the maximum solubility of carbon in iron is 0.008 *wt*. %, and at 723°C the maximum solubility of carbon is 0.025 *wt*. %. The lattice parameter is 2.868 Å, and the supercell represents the correct amount of carbon content in the ferrite. The carbon content is calculated by dividing the total weight of carbon atom in the unit cell by the total weight of the atoms inside the unit cell.

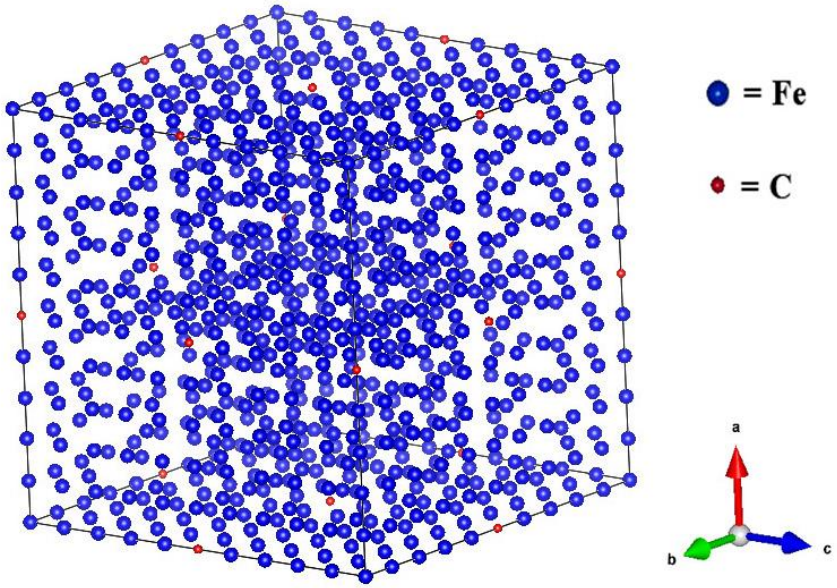

Figure 14. Supercell of ferrite with 0.008 *wt*. % carbon.

2.3.4 Austenite

Austenite is a solid solution of carbon in face centre cubic (FCC)-iron and is stable at temperatures between 723°C and 1492°C as shown in the phase diagram. Austenite belongs to the space group *Fm3m* (no. 225). The carbon content relates to the lattice parameter by the equation

$$a_\gamma = a_0 + k_c C_\gamma \quad (1)$$

where a_0 is the lattice parameter of austenite at 25°C and k_c is the carbon coefficient which has been determined as 0.00528Å by experiments (Roberts, 1953). The Wyckoff position of the iron atoms is Fe 4*a* (0,0,0). The carbon atoms are located at Wyckoff positions 4*b* (0.5, 0.5, 0.5). In austenite the iron atoms are present at the corners and at the

centre of each face of the unit cell. Austenite is more densely packed than the BCC structure of ferrite (APF_{bcc}=0.68, APF_{fcc} =0.74, APF is calculated by multiplying the number of particles in the unitcell by the volume of the particle and divide by the volume of the unit cell). The maximum solubility of carbon in austenite is 2.1 *wt*. % carbon at 732°C. Excess carbon that cannot diffuse into the iron structure can either form as iron carbides or carbides of the alloying elements such as chromium, molybdenum, tungsten, or vanadium. Figure 15 shows the supercell of austenite. A supercell is a repeating unit cell of the crystal that contains many primitive cells. The use of supercell is necessary because it represents the correct amount of carbon content in the austenite.

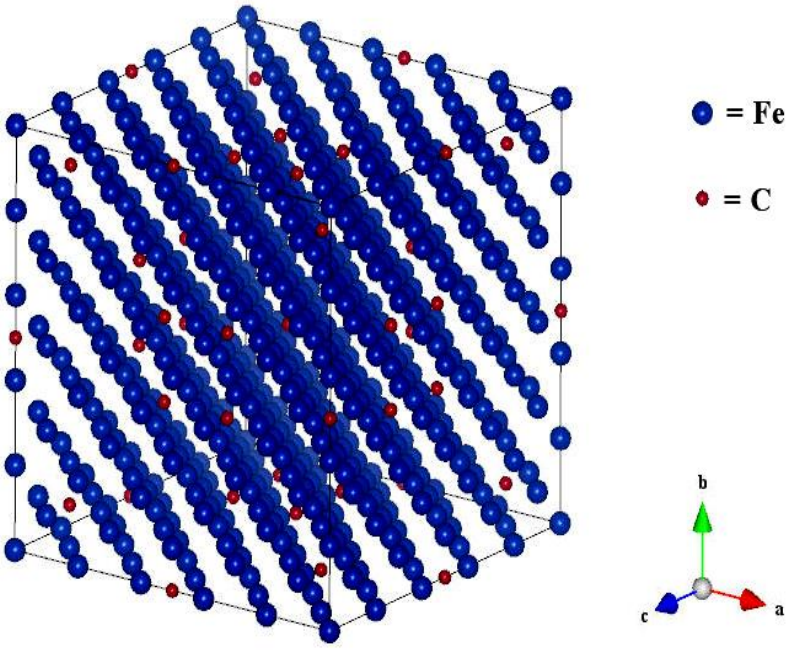

Figure 15. Fe-C austenite 5x5x5 supercell with 0.8 *wt*. % carbon.

2.3.5 Cementite

The main carbides found in steels are cementite, Fe_3C and the transition carbides epsilon (ε) ($Fe_{2.4}C$ with hexagonal lattice) and eta (η) (Fe_2C with monoclinic lattice) (Porter and Easterling, 1992). Cementite has an orthorhombic crystalline structure (space group *Pnma*, No.62) with lattice parameters *a* = 4.51 Å, *b* = 5.08 Å, and *c* = 0.673 Å (Wood et

al., 2004). It contains 6.67 wt. % carbon. There are 12 iron atoms and 4 carbon atoms per unit cell (Fasiska & Jeffrey, 1965; Wood et al., 2004). Each carbon atom is surrounded by eight iron atoms and each iron atom is surrounded by three carbon atoms. The twelve Fe atoms are distributed between lattice sites at Wyckoff position Fe 4*c* and Fe 8*d*. Figure 16 shows the cementite crystal structure. Cementite usually precipitates from martensite after the tempering treatment.

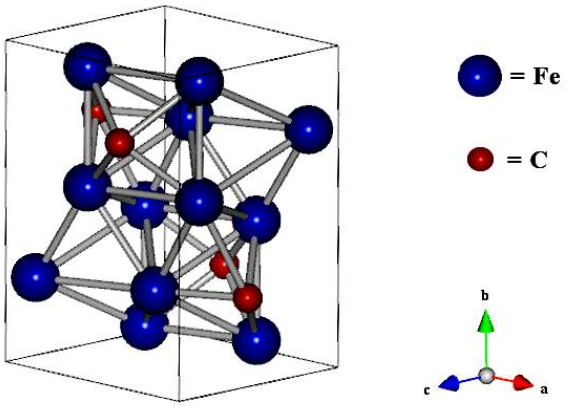

Figure 16. Fe_3C unit cell.

2.3.6 Pearlite

In steels having the composition of 0.83 *wt*. % carbon the constituent stable below 723°C (A1 line on the phase diagram) consists of a mixture of ferrite and cementite called pearlite (see Figure 17).

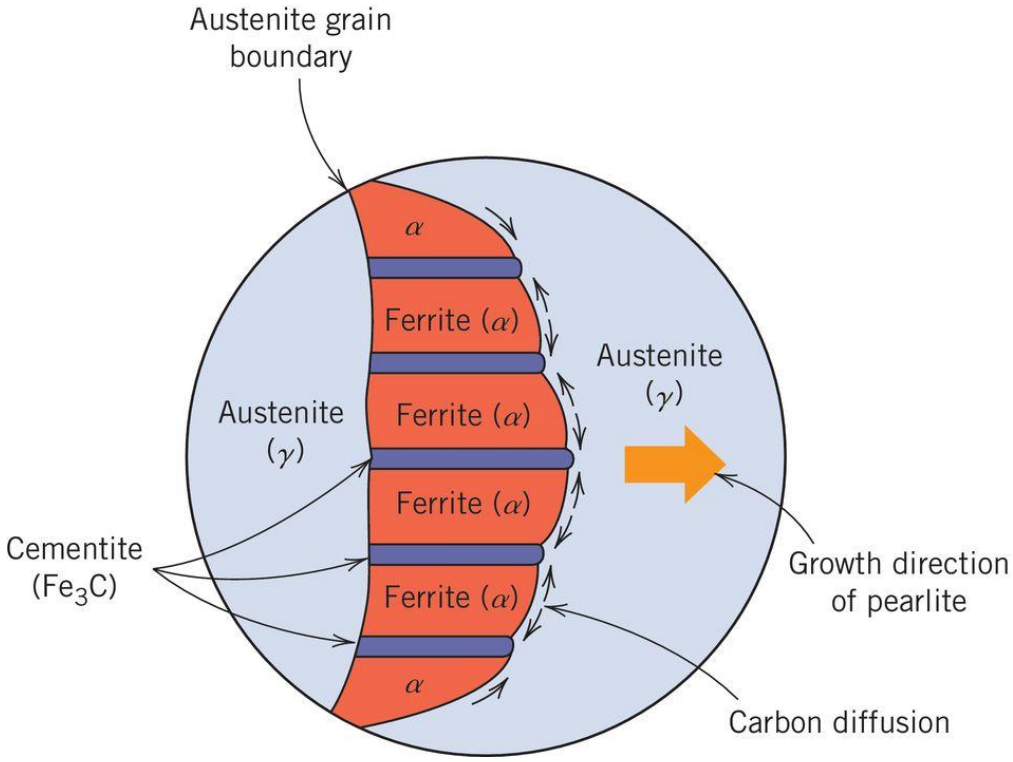

Figure 17. Schematic representation of the formation of pearlite from austenite (Callister, 2015).

Pearlite grains contain a mixture of ferrite and cementite plates arranged in an alternating fashion. The thickness of the plates depends on the volume of ferrite and cementite within the material. The strength of pearlite increases as the interlamellar spacing decreases. Pearlite nucleates from the austenite grain boundary and grows into the austenite grain. The nucleation process is controlled by atomic diffusion. During quenching (heat treatments with high cooling rate) diffusion is suppressed and the austenite transforms to martensite by shearing (see section 2.4.1).

The transformation of austenite to pearlite, illustrated in Figure 17, is called the *eutectoid transformation*. Steels are classified function of the carbon content and the eutectoid point as: *eutectoid* or *pearlitic* steels, *hypoeutectoid* steels with carbon less than that corresponding to the eutectoid composition and *hypereutectoid* steels with carbon above 0.8 *wt*. %.

2.3.7 Eutectoid Steels

The eutectoid point (see Figure 12) represents the temperature and composition where a eutectoid reaction occurs. A eutectoid reaction occurs when one solid phase (austenite) transforms into a mixture of two solid phases (ferrite and cementite). The phase diagram shows that the eutectoid point exists at a composition of 0.83 *wt*. % carbon. The microstructure of steels containing 0.83 *wt*. % carbon at room temperature consists of pearlite (see Figure 18).

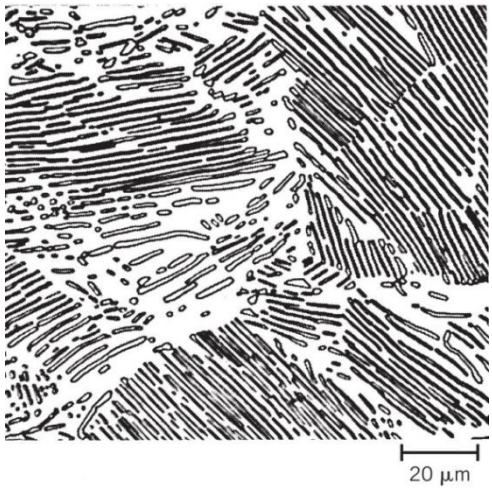

Figure 18. Micrograph of eutectoid steel showing pearlite microstructure consisting of alternating layers of ferrite and cementite (Callister, 2015).

2.3.8 Hypoeutectoid Steels

Hypoeutectoid steels contain less than 0.83 *wt*. % carbon. The microstructure of hypoeutectoid steels at room temperature consists of a mixture of ferrite and pearlite (see Figure 19). As the carbon content is increased, the volume percentage of pearlite increases until it is 100% at the eutectoid composition.

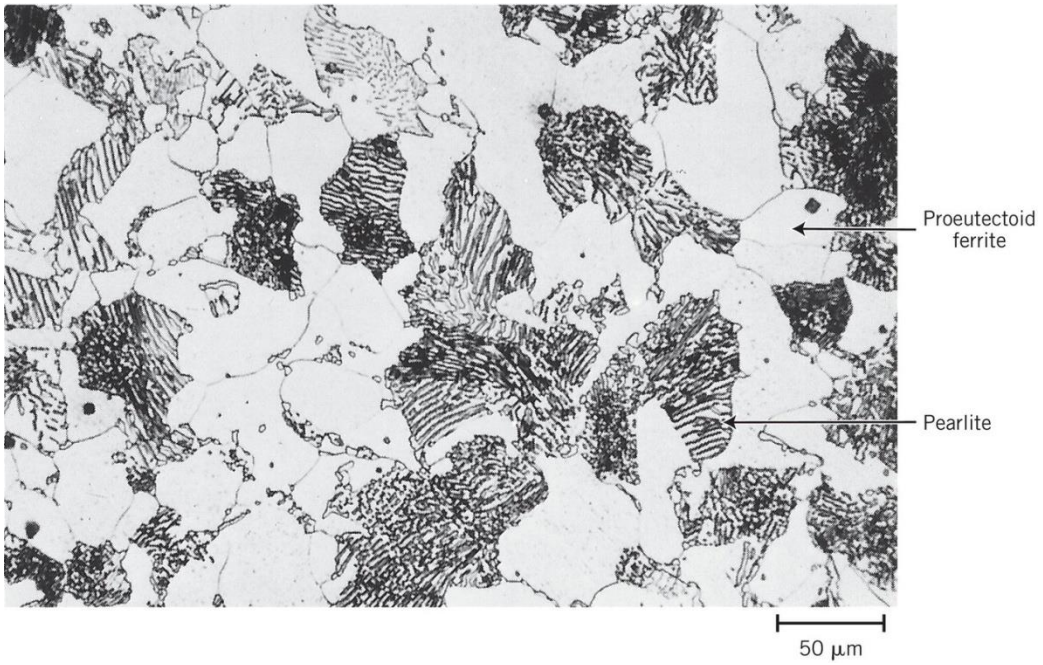

Figure 19. Micrograph of a 0.38 *wt*. % carbon steel microstructure consisting ferrite and pearlite (Callister, 2015).

2.3.9 Hypereutectoid Steels

Hypereutectoid steels contain more than 0.83 *wt*. % carbon. At room temperature their microstructure consists of pearlite and cementite at the grain boundaries (see Figure 20).

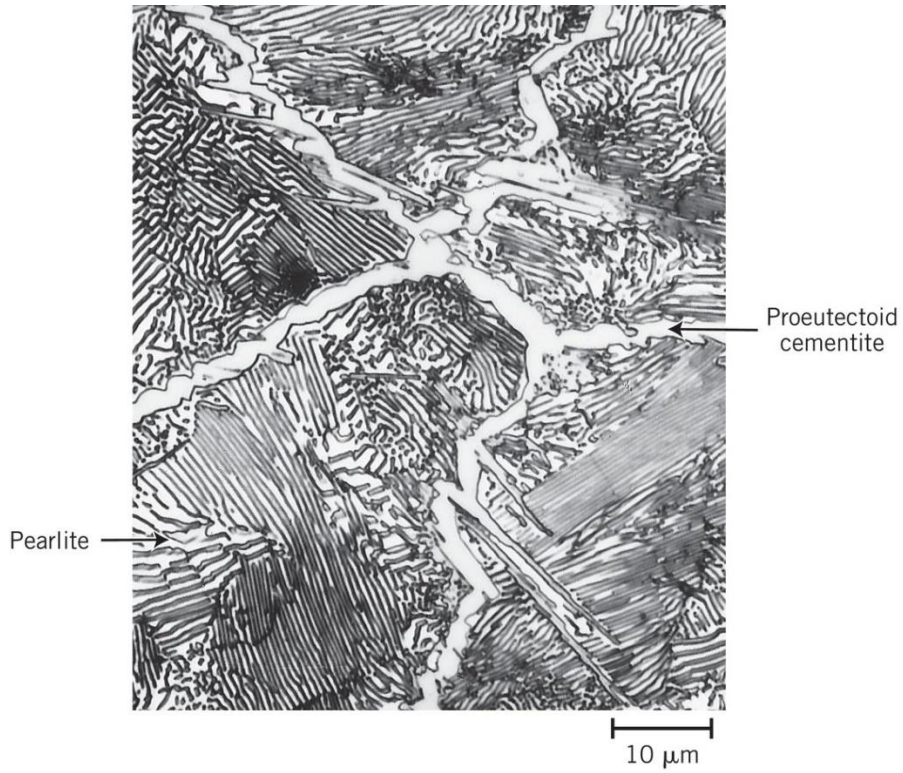

Figure 20. Micrograph of a 1.4 *wt*. % carbon steel microstructure consisting of cementite and pearlite (Callister, 2015).

2.3.10 Alloying elements

The alloying elements added to steel give a high flexibility to adapt to a wide range of applications. Increasing carbon content increases the yield strength of the steel but reduces its ductility. The heat treatments applied can also change the microstructure.

Steels may contain small amounts of elements including copper, phosphorous, manganese, sulphur, nickel, chromium, and many others. Some of these elements are impurities generated during steel making and others are added intentionally. One of the most common elements added to steel is chromium. Chromium increases both the corrosion resistance and the hardenability of steel (ASM Handbook, 2001).

The hardenability is influenced by all alloying elements. The hardenability is defined as the ability of the steel to increase its hardness during quenching (ASM Handbook, 2001). Molybdenum is usually added in order to increases the hardenability of steels. Elements

such as sulphur and phosphorous are undesired because they cause brittleness. In most cases, high quality steel requires that sulphur and phosphorous content to be brought to low amounts (0.05 *wt*. %). Nickel is sometimes added in steel in order to increase the toughness of ferrite and increase the overall impact resistance.

2.3.11 Basics of diffusion

Diffusion is a phenomenon of material transport by atomic motion (Callister, 2015). In order for the atoms in solids to change their initial position there must be sufficient activation energy and an empty site adjacent to it. There are two dominant types of diffusion models proposed for metals. Vacancy diffusion involves the interchange of an atom from its original lattice position to the neighbour vacant lattice site (Callister, 2015). The amount of vacant lattice sites present varies depending on the amount of defects in the material and the temperature. Interstitial diffusion involves the migration of atoms from interstitial position to a vacant neighbouring site. It often occurs when atoms are small enough to fit into the interstitial positions (carbon in iron).

The diffusion processes which transport the atoms within the solid occur at rates which depend exponentially on temperature and the rate of cooling. The diffusivities can be expressed by the equation:

$$D = D_0 \exp\left(-\frac{Q}{RT}\right) \quad (2)$$

where D is the diffusion coefficient of the element, D_0 is the frequency factor [cm^2s^{-1}], Q is the activation energy [kJ mol^{-1}]. The diffusivities of various elements in α-iron and γ-iron are listed in Table 1.

Table 1. Diffusivities of elements in iron adapted from Honeycombe & Bhadeshia (2005).

Solvent	Solute	Activation energy, Q (kJ mol^{-1})	Frequency factor, D_0 (cm^2 s^{-1})	Diffusion coefficient, $D_{910°C}$ (cm^2 s^{-1})	Temperature range (°C)
γ-iron	C	135	0.15	1.5×10^{-7}	900–1050
	Fe	269	0.18	2.2×10^{-13}	1060–1390
	Co	364	3.0×10^{2}	24.0×10^{-12} (at 1050°C)	1050–1250
	Cr	405	1.8×10^{4}	58.0×10^{-12} (at 1050°C)	1050–1250
	Cu	253	3.0	15.0×10^{-11}	800–1200
	Ni	280	0.77	7.7×10^{-13}	930–1050
	P	293	28.3	3.6×10^{-12}	1280–1350
	S	202	1.35	1.5×10^{-9}	1200–1350
	W	376	1.0×10^{3}	12.0×10^{-12} (at 1050°C)	1050–1250
α-iron	C	80	6.2×10^{-3}	1.8×10^{-6}	
	N	76	3.0×10^{-3}	1.3×10^{-6}	
	Fe	240	0.5		700–750
	Co	226	0.2	2.1×10^{-11}	700–790
	Cr	343	3.0×10^{4}		
	Ni	358	9.7	3.7×10^{-11}	700–900
	P	230	2.9	2.0×10^{-10}	860–900
	W	293	3.8×10^{2}		

Diffusion is a time dependent process. The rate of diffusion is referred to as the diffusion flux (J). The time dependent diffusion process can be described with the Fick's first law

$$J_x = -\frac{D\partial c}{\partial x} \tag{3}$$

where J is the flow rate, $\frac{\partial c}{\partial x}$ is the concentration gradient, D is the diffusion coefficient or diffusivity. The Fick's first law relates the diffusion flux to the concentration gradient. Steady state diffusion describes the situation when the diffusion process reaches a state where the diffusion flux no longer changes with time and there is no net accumulation of the diffusing species (Callister, 2015).

In non-steady state diffusion the diffusion flux and the concentration gradient in solid vary with time. There is a net accumulation of the diffusion species in the material. The

Fick's second law describes the time dependent concentration gradient of the diffusion and can be represented by

$$\frac{\partial c_x}{\partial t} = \frac{D\partial^2 c_x}{\partial x^2} \tag{4}$$

where *D* is the diffusion coefficient, $\frac{\partial c}{\partial x}$ is the concentration gradient.

Fick's second law can be used for carburisation calculations where the diffusion concentration of carbon atoms into the steel is given by

$$\frac{c_x - c_0}{c_s - c_0} = 1 - \text{erf}(\frac{x}{2\sqrt{Dt}}) \tag{5}$$

where *D* is the diffusivity, *t* is time, c_s is the carbon concentration at the surface, c_0 the initial carbon concentration of the carbon and erf is the Gaussian error function. The erf value can be found in material handbooks. In carburising under the same conditions, the diffusivity of carbon in Body Centred Cubic (BCC) iron is higher than carbon in Face Centred Cubic (FCC) iron. This is expected since the atomic packing of the bcc structure (APF=0.68) is less than FCC structure (APF=0.74). The temperature also plays an important role in the rate of diffusion.

2.4 Diffusionless Transformations

A diffusionless transformation is a phase transformation occurring without atomic diffusion. Diffusionless transformations produce metastable states that are not represented in the equilibrium phase diagram. The transformation of austenite to martensite during quenching is an example of such transformation.

2.4.1 Martensite

Martensite is a metastable phase formed when diffusion of carbon atoms cannot take place due to the fast cooling rate, and the carbon atoms are trapped in the structure of

iron. Martensite forms by a shear process in the austenite lattice without atomic diffusion or change in chemical composition. The martensite transformation begins at the martensite start temperature (M_s) and finishes at the martensite finish temperature (M_f). Martensite has a body center tetragonal (BCT) structure. Figure 21 presents the supercell of martensite.

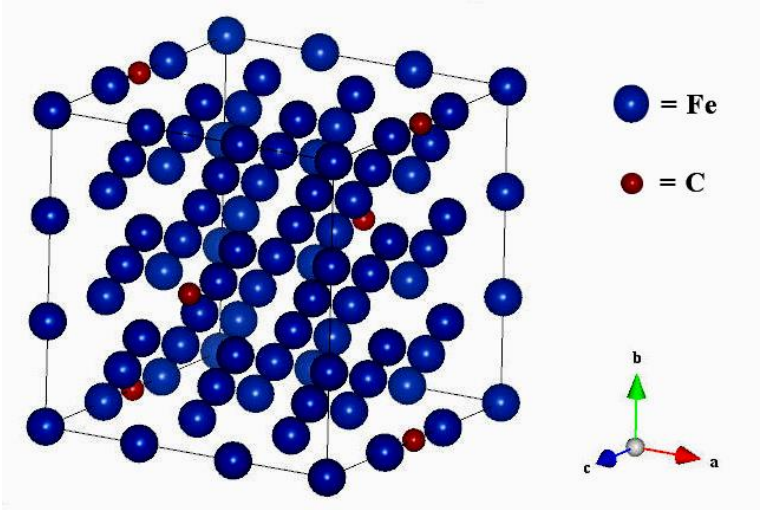

Figure 21. The supercell of martensite with 0.8 *wt* % carbon.

Iron suffers an allotropic transformation from FCC to BCC at 912°C upon slow cooling. The excess of carbon atoms leave the strucure of Fe by diffusion and they form pearlite and/or cementite function of the carbon content (hypo- or hypereutectoid). When austenite is subjected to the quenching process (fast cooling), the carbon is trapped in the iron solid solution because there is not enough time for diffusion to occur and the formation of ferrite is suppressed. The trapped carbon creates a slight stretch in the original BCC lattice and forms the body center tetragonal (BCT) structure called martensite. The result is a fine structure with carbon atoms trapped within the octahedral interstitial sites of the BCT structure. Any remaining austenite is termed retained austenite. Tempering (a heat treatment in which the temperature does not exceed 600°C) the martensite facilitates diffusion and cementite particles precipitates while martensite

tetragonality is reduced. Tempering for long periods of time will completely result in the complete transformation of martensite into ferrite and carbide precipitates.

Bain (Bain & Dunkirk, 1924) suggested that the transformation involves a compression of the *c* axis of the austenite unit cell and an expansion of the *a* axis as shown in Figure 22.

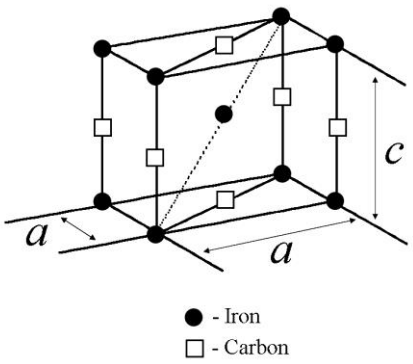

Figure 22. FCC unit cell contracted on the *c*-axis and expanded on the *a*-axis

The microstructure of martensitic steel is obtained by quenching the austenite phase. The lattice parameter of martensite varies with carbon content in a linear fashion. The tetragonality and *c/a* ratio, increases with carbon content, and the effect of this lattice distortion is the increase in hardness. Figure 23 shows the microstructure of martensite.

Figure 23. Photomicrograph of the martensitic microstructure. Dark needle shape grains are martensite phase. White regions are austenite phase (Callister, 2015).

The martensitic transformation occurs because the free energy of martensite is lower than that of austenite. Additional energy is needed for the transformation to take place due to the surface energy and transformation strain energy. Therefore the austenite to martensite transformation occurs when the free energy available from the transformation is enough to provide the strain energy for the shearing operation. When the starting temperature of quenching is high enough, the resulting structure become coarse and the individual crystal of martensite can be seen as plate or needles.

In steels the transformation front moves at the speed of sound in the material. The nucleation rate of martensite can be described by the equation:

$$N_v = n_t v e^{-Q/RT} \quad (6)$$

where N_v is the number of martensite plates per unit volume, n_t is the rate of conversion, v is the rate of attempts per unit time, Q is the activation energy for nucleation, T is the temperature.

The number of active sites is difficult to predict because their number vary and is constantly changing (Magee, 1971).

The aim of tempering is to increase the toughness of martensite without significantly reducing hardness. If tempering occurs at temperatures below 200°C (stage 1) carbon atoms diffuse throughout martensite and fine carbide particles nucleate and grow within martensite. This type of carbide is metastable and has a hexagonal structure and is referred to as epsilon carbide ($Fe_{2.4}C$) (Ohmori & Tamura, 1992). At subsequent stages of tempering eta carbide (Fe_2C) with orthorhombic structure (Hirotsu & Nagakura, 1972) has been observed to precipitate.

2.4.2 Bainite

Bainite was first discovered by Edgar Bain (Bain, 1936). The bainitic transformation consists in the formation of fine aggregates of ferrite plates and cementite particles at temperatures between 250 °C and 550 °C. Upper bainite refers to fine plates of ferrite grains formed at the temperature range between 400~500°C (Honeycombe, 1995) while lower bainite forms in the temperature range between 300~400°C (Honeycombe, 1995). The microstructure of lower bainite is similar to upper bainite as it contains needle like ferrite. The difference is that small carbide particles precipitate inside the ferrite plates for lower bainite. Figure 24 shows the microstructure of upper bainite.

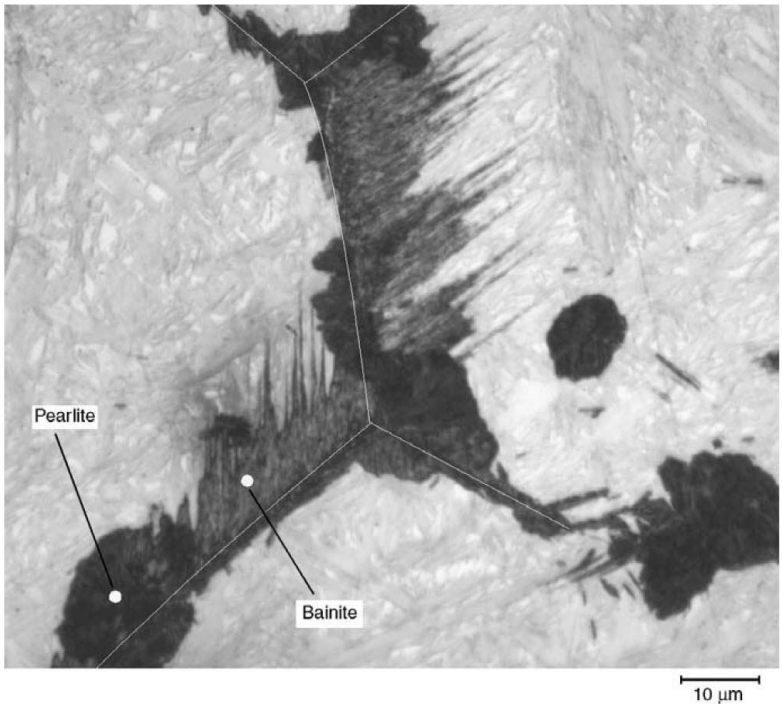

Figure 24. Photomicrograph of a 1.4 wt. % carbon steel microstructure consisting pearlite and bainite (ASM, 1992).

2.4.3 TTT diagram and cooling rates

The rate of cooling from the austenitic region has a strong influence on the microstructural products because cooling rate affects the atomic diffusion. The martensitic transformation begins at the grain boundaries of austenite, therefore prior grain size is an important factor of the transformation.

A TTT (Time-Temperature-Transformation) diagram is specific for each steel grade because the transformation of austenite is affected by the amount of carbon and other alloying elements. Figure 25 presents a TTT diagram for a steel containing 0.89 wt. % carbon.

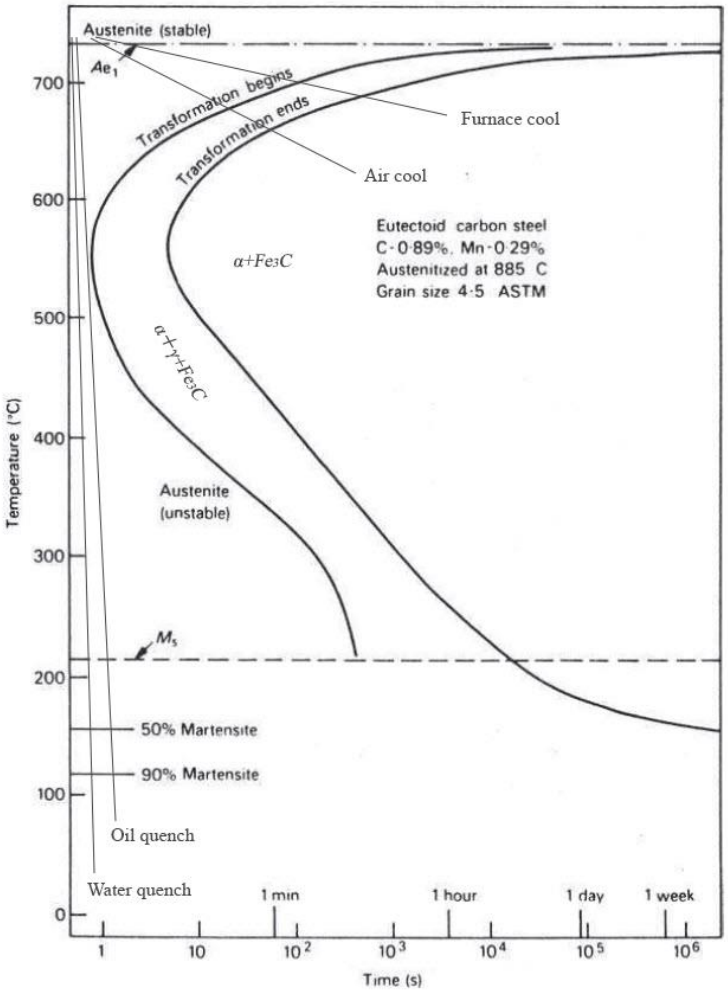

Figure 25. TTT diagram of a 0.89 wt. % carbon steel (Bhadeshia & Honeycombe, 2006).

The TTT diagram in Figure 25 shows a schematic of the cooling routes of some common processes. The diagram represents the rate of transformation of austenite at a constant temperature. Air cooling and furnace cooling are usually not sufficient for martensitic transformation to occur because of the slow cooling rate allowing the formation of pearlite. A fast cooling medium such as water or oil (quenching) is required to prevent the formation of pearlite and bainite.

The martensitic transformation begins when the temperature reaches the martensite start temperature (M_s) and the transformation is complete when the temperature reaches martensite finish temperature (M_f). Steels with high carbon content will cause the M_f

marker to drop below zero degrees Celsius, which means that a high rate of cooling is required to achieve complete austenite transformation. In practice this is rarely the case.

The austenite that does not transform into martensite is called retained austenite. When stresses or heat are applied during operation, the retained austenite may transform into martensite (strain assisted transformation). The optimum microstructure of gears can be achieved by controlling the amount of retained austenite in order to obtain the desired fatigue resistance (Matlock et al., 2005; Senthilkumar & Rajendran, 2014).

2.5 Steel gear heat treatment

The ideal material for gears must have high strength to support the load and it should be easily shaped and machined with acceptable tolerances. It also needs to have a relatively high hardness in order to provide good wear resistance. Ductility is a property generally associated with low carbon steels while hardness and wear resistance is generally associated with high carbon steels. In order to combine both ductility and hardness in gears, low or medium carbon steel is used but the chemical composition at the surface is altered by thermo-chemical treatments followed by case hardening treatments. The vast majority of gears today are made with either plain carbon steel or alloyed steel. These steels are usually heat treated to obtain enhanced strength. Alloyed steels have the advantage of having high hardenability. In most cases, the surface of the gear teeth and the subsurface regions are the most important areas because those are the regions where failures most often originate. A properly controlled treatment gives the part resistance to bending and contact fatigue. Carburising causes distortions of the gear tooth geometry, therefore grinding is often required in some cases. Although grinding is a method to remove the undesired distortions, this adds to the manufacturing cost. Grinding also removes some of the hardened case thus lowers the overall hardness of the surface.

The most common thermochemical treatments include carburising, nitriding, carbonitriding, and nitrocarburising and the most common heat treatments are induction hardening and flame hardening. Generally, a heat treatment for steel involves heating the steel to its austenitising temperature (800~900°C) and then quenching in oil or water followed by tempering. The quenching ensures the formation of martensite. Tempering is done by reheating the steel to increase its ductility.

Increasing the carbon content on the surface of the gear increases wear resistance and resistance to contact fatigue. A level of 15 to 20% retained austenite is sufficient for resisting sliding and contact fatigue (Krauss, 1995).

The case depth is an important factor in terms of contact fatigue resistance. The total case depth is the distance from the surface of a hardened tooth to a point inside the tooth where the difference in chemical and mechanical properties of the case and core can no longer be distinguished. In practice the case depth is specified when the hardness drops to a fixed level, often 550VHN where the microstructure is 50/50 martensite/ferrite and pearlite. The optimal case depth depends on the size of the gear and whether it will be sufficient to carry the load. The strength of the core is also an important factor in determining the bending resistance at the root of the tooth.

2.5.1 Through-Hardening

Through hardening is a process applied to gears that does not result in high surface hardness. In through hardening, gears are heated to the austenitic temperature and quenched in air, gas, or liquid. This process may be used either before or after the gears are cut. The case and core hardness of the product are similar. Common heat treatment methods used for through hardening are quenching and tempering.

2.5.2 Carburising

Carburising is a process in which a metal is in contact with a carbon source and the surface is enriched in carbon. This process makes use of the fact that carbon dissolves interstitially in iron. Since FCC iron can dissolve carbon, the steel is carburised above the austenitic temperature at around 900°C. The process of diffusion occurs due to a concentration gradient of carbon from the metal surface to core. The case depth depends on the carburising temperature, time, and the carbon source. A typical depth of the case ranges between 1 and 2 mm. The case depth depends on the carburising process such as vacuum conditions and gas. A low-carbon steel subjected to the carburising treatment will have high carbon content at the surface (0.8~1%) with low carbon core (0.1~0.3% carbon). Surface carbon is often limited to 0.9% due to the presence of retained austenite. The goal is to obtain a hardened surface and retain the ductility of the core.

The austenite grains become larger with increasing carburising time (Geoffrey, 1999). There is a limit where the case depth becomes too deep where the compressive residual stresses generated by the carburised region increase the possibility of crack initiation (Genel & Demirkol, 2000). Carburising increases the carbon content and the wear resistance increases due to the formation of martensite during quenching (Genel & Demirkol, 1999; Genel & Demirkol, 2000).

There are three general methods of carburising. The source of carbon can be in the form of either gas, liquid or solid.

Gas carburising is a common form of carburising treatment used today. It is carried out in a carbon rich atmosphere. The gears are heated in the gaseous atmosphere with a carbon potential that allows the diffusion of carbon. The most common used medium is an endothermic gas produced by reacting hydrocarbon methane and propane with oxygen. The gaseous atmosphere contains carrier gas such as nitrogen, carbon monoxide, and

hydrogen in order to increase gas circulation and increase case uniformity. The main chemical reactions taking place in carburising are (ASM Handbook, 2001):

$$Fe + 2CO \rightarrow Fe(C) + CO_2$$

$$Fe + CH_4 \rightarrow Fe(C) + 2H_2$$

$$Fe + CO + H_2 \rightarrow Fe(C) + H_2O$$

Fe(C) is the carbon in solution in austenite.

The gears are heated above 900°C for a certain period of time depending on the application (ASM Handbook, 2001). The carbon is adsorbed onto the surface and diffusion takes place from surface to core. The amount of carbon and the diffusion depth depends on the carbon potential of the environment.

Solid carburising or pack-carburising is an old process. It is called packed carburising because the steel surface is in contact with a mixture of charcoal and alkali carbonates packed in a box (ASM Handbook, 2001). The alkali carbonate or sodium carbonate serves as an energizer to increase the rate of carbon diffusion by chemical reaction that causes carbon atoms to be released at the surface of the steel. The regions where carburising is not desired are electroplated with copper with a thickness of 0.1 mm since carbon does not dissolve in copper. The box is slowly heated to the carburising temperature (900°C) and maintained for 5-10 hours depending on the depth of the case required.

Liquid-carburising or cyanide hardening is done in baths of molten salts that contain 20~50% sodium cyanide with 40% sodium carbonate and barium chloride (ASM Handbook, 2001). Carburising takes place due to the decomposition of sodium cyanide at the steel surface. Atoms of both carbon and nitrogen are released and are absorbed by the steel. The cyanide mixture is heated at around 900°C and the steel is immersed in the

mixture. The immersed time is relative to the case depth. One advantage of this process is that quenching can be done directly from the bath.

2.5.3 Nitriding

Nitriding is a case-hardening process used for producing wear resistance surfaces on alloyed steels. The process is carried out between 500°C to 600°C. It is most suitable for gears that are not subjected to high shock loads or are required to work at high temperatures due to the softening of carburised surfaces above 150°C. Nitriding process can be done in either gas or liquid media containing nitrogen such as ammonia. The nitrogen atoms combine with the alloying elements and form nitrides that give high hardness. The case depth of nitrided gears is usually smaller than 1.0 mm. A smaller case depth means that nitrided gears should not be subjected to high loads. The case depth is smaller because the process temperature is below the austenitising temperature so carbon diffusion is slower.

2.5.4 Carbonitriding

Carbonitriding is a process where carbon and alloy steel gears are held at a temperature above the transformation range between 750°C and 900°C in a gaseous atmosphere containing carbon and nitrogen. The case depth is usually less than 0.7 mm. The nitrogen increases the hardness by forming nitrides and the carbonitrided case has better hardenability compared to carburised case.

2.5.5 Nitrocarburising

Nitrocarburising is a case hardening process that diffuses nitrogen and carbon into steel at the temperature range between 500°C to 650°C. The advantage of nitrocarburising is that

the process occurs in the ferritic phase. The result is increased resistance to wear, and fatigue with negligible shape distortion.

2.5.6 Quenching

Quenching is carried out after the carburising process. One advantage of quenching directly from the carburising temperature is the simplification of the process and reduced time and cost. The disadvantage is the risk of increasing thermal stresses and thermal shocks. The thermal expansion in different parts of the gear will generate stresses which may be a cause of crack initiation. In order to avoid distortions the temperature can be slowly reduced to room temperature and then the gears are reheated for quenching. The disadvantage is the amount of energy used for reheating and the cost for additional processing steps (Krauss, 1995). Quenching directly after the carburising process has the advantage of reduce the amount of heating and cooling cycle. The temperature cycle causes phase transformations generating internal stresses and volume changes. Therefore, reducing the amount of processing steps helps increase the metallurgical stability and reduces dimensional distortions. In the case of gears, careful control of carbon content is required in order to reduce the precipitation of carbides because carbide formation at the grain boundary is usually detrimental to gear life.

2.5.7 Cryogenic treatment

Cryogenic treatment has been found to have the effect of improving wear and fatigue behaviour of steel gears (Baldissera & Delprete, 2008; Baldissera, 2009; Bensely, 2006; Bensely, 2011; Manoj, 2004; Paulin, 1993). There are two types of cryogenic treatment, deep cryogenic treatment and shallow cryogenic treatment. Deep cryogenic treatment takes place in the temperature range between -120°C to -196°C, which is near the temperature of liquid nitrogen. Shallow cryogenic treatment takes place from -80°C to -

120 °C. A “wet” cryogenic process describes the process where the material is immersed directly in liquid nitrogen. In a “dry” cryogenic process the material is placed in a chamber which is cooled with liquid nitrogen. The dry process gives more temperature control and reduces the thermal stresses caused by rapid cooling of the wet process. However, due to the separation between the liquid nitrogen and the material, the cryogenic treatment temperature is often much higher than the temperature of the liquid nitrogen, usually the lowest achievable temperature is around -120°C to -140℃. Figure 26 presents a plotted example of temperature versus time for a cryogenic process.

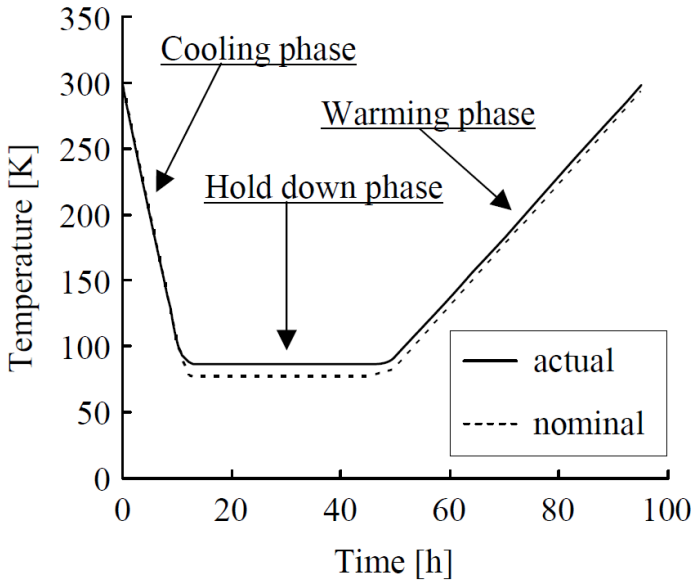

Figure 26. Plot of temperature versus time for a cryogenic process (Bensely, 2006).

The controlled cryogenic process starts with the loading of an insulated treatment chamber with the material. Liquid nitrogen is used as a cooling medium to decrease the temperature to treatment temperatures. The temperature is lowered at a slow rate with very precise control since rapid change in temperature can induce stresses caused by thermal shock. The process usually takes 10~20 hours. Once the desired temperature is reached, the process enters the soak phase and maintains the temperature for 20~40 hours depending on the size of the component. The soak phase describes the phase where the

steel must be maintained at a particular temperature for a complete austenitic reaction. The soak phase ensures the entire cross section of the material is completely treated. After the soak phase, the temperature is gradually raised to room temperature. Addition tempering may stabilize the martensite depending on the material and application. The effectiveness of the cryogenic treatment greatly depends on the details of the process such as hold down phase, cooling and warming rate.

2.5.8 Tempering

Tempering is a process in which the quenched steel is heated below the eutectoid temperature to reduce internal stresses generated due to the martensitic transformation. A hardened steel is brittle therefore tempering is required to reduce internal stresses. After the tempering process the brittleness and hardness are reduced while ductility increases. Heating the gears around 120°C~200°C activates diffusion and residual stresses are reduced. Some carbon precipitates as transition carbides initially but after longer periods as cementite.

In general, low temperature tempering (around 200°C) is used when hardness is the most important factor. Higher tempering temperature (around 400°C) is used when strength and toughness are more important.

Case microstructure

The dominant microstructural constituent in a case carburised steel is tempered martensite (see Figure 23). The carbon content decreases with increasing depth from the surface. Other microstructural constituents include small amounts of retained austenite, carbides, bainite, and pearlite (Krauss, 1989; Parrish, 1980). The first martensite plate forms at the martensite start temperature (M_s). Martensite has a needle-like structure under the microscope. Massive carbides form isothermally in carbon-saturated austenite at the austenite grain boundaries. Tempering introduces a high density of dislocations. The

carbides and the dislocations makes deformation difficult, therefore result in increased hardness and strength.

During quenching some of the austenite will not transform into martensite. This is known as the retained austenite. The presence of retained austenite can be the source of dimensional instability, residual stress, and may contribute to cracking in gears. Retained austenite is generally unavoidable in the high-carbon case of carburised steels. Cryogenic treatments can be used to minimize the amount of retained austenite.

Retained austenite under cyclic stress may transform into martensite. Stress assisted transformation occurs when low amounts of retained austenite are present and the transformation does not require new nucleation sites. Strain assisted transformation occurs when large amounts of retained austenite are present which require large amount of strains to generate nucleation sites. When stresses or heat are applied such as during gear operation, some retained austenite may transform into martensite. Having a small amount of retained austenite is not always a disadvantage. For example, research shows that 10~20% of retained austenite in steel gives higher fatigue life (Baldissera, 2009). According to literature, cryogenic treatment increased the surface hardness from 62.4 to 68 HRC (Brown, 1995). Cryogenic treatment can reduce residual stresses and convert some of the retained austenite to martensite (Brown, 1995).

Core microstructure

The core microstructure of case carburised gears include ferrite, lath martensite, and small amounts of pearlite (Krauss, 1989; Parrish, 1980; Stormvinter et al., 2011). Lath martensite occurs in steels containing less than 0.6 wt% C (Krauss, 1995; Stormvinter et al., 2011). If the steel has low hardenability, the core may transform to ferrite and pearlite. If the steel has high hardenability, the core may transform to martensite.

2.6 Contact Mechanics

Contact mechanics is the study of stresses and deformations which arise when the surfaces of two solid bodies are brought into contact. It provides information needed for tribology applications and stress analysis. Contacts can cause high stresses between surfaces and cause failure in the form of surface contact fatigue and wear.

2.6.1 Surfaces

Conforming contacts occur when two surfaces fit exactly or closely together without deformation (Hertz, 1881). Non-conforming contact refers to surfaces which deform plastically under the applied load. A perfectly smooth surface is referred to as an ideal surface. A nominal surface is the ideal surface designed as intended by the drawing or other technical documents (BS EN ISO 4287). Real surface deviates from the nominal surface as a result of the manufacturing processes. The apparent area of contact depends on the geometry of the two bodies and the real area of contact depends on the roughness.

2.6.2 Hertzian Theory

The origins of contact mechanics are linked with Heinrich Hertz who developed a theory for solving the problem of elastic deformation between contacting surfaces in 1881 (Hertz, 1881). When two curved bodies are brought into contact they touch at either a point or along a line. At very small loads they deform elastically giving rise to contact areas. The Hertzian theory includes some assumptions called the Hertz assumptions (Hertz, 1881). These assumptions are that: surfaces are continuous, smooth, nonconforming, and frictionless; the contact area is small compare to the size of the bodies; each solid can be considered to behave as an elastic half-space in the vicinity of the contact zone; the contacting bodies are homogeneous and isotropic; the surface shape does not change in time.

2.6.3 Circular contact

Circular contact occurs between two spheres or between a sphere and flat (see Figure 27)

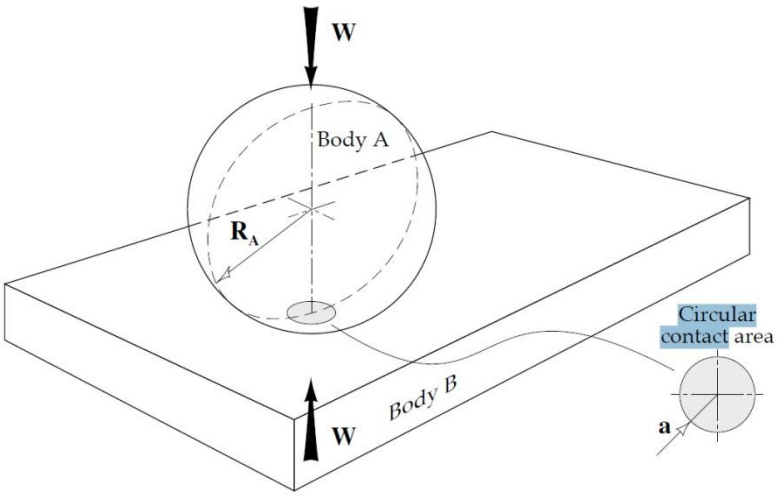

Figure 27. Contact between a sphere and a flat surface (Stachowiak, 2014).

The area of contact between two spheres is described by the equation 6:

$$a = \left(\frac{3WR'}{4E'}\right)^{1/3} \tag{7}$$

where a is the radius of the contact area [m], W is the applied load [N], R' is the reduced radius of curvature, E' is the reduced elastic modulus [Pa].

The reduced Young's modulus is described by the equation 8:

$$\frac{1}{E'} = \frac{1 - v_A^2}{E_A} + \frac{1 - v_B^2}{E_B} \tag{8}$$

$$\frac{1}{R'} = \frac{1}{R_A} + \frac{1}{R_B} \tag{9}$$

where v are the Poisson's ratios of the contacting bodies A and B, E are the elastic moduli of the contacting bodies A and B, E' is the reduced elastic modulus.

The maximum contact pressure is:

$$P_{\max} = \frac{3W}{2\pi a^2} \tag{10}$$

The average contact pressure is:

$$P_{\text{average}} = \frac{W}{\pi a^2} \tag{11}$$

The maximum shear stress can be described by the equation:

$$\tau_{\max} = \frac{1}{3P_{\max}} \text{ at a depth of } z = 0.638a \tag{12}$$

where a is the radius of the contact area [m], W is the load [N], P is the contact pressure [Pa]. τ is the shear stress [Pa], z is the depth where the maximum shear stress occurs, E' is the reduced elastic modulus [Pa], R' is the reduced radius of curvature.

2.6.4 Line contact

Line contact occurs between two parallel cylinders (see Figure 28).

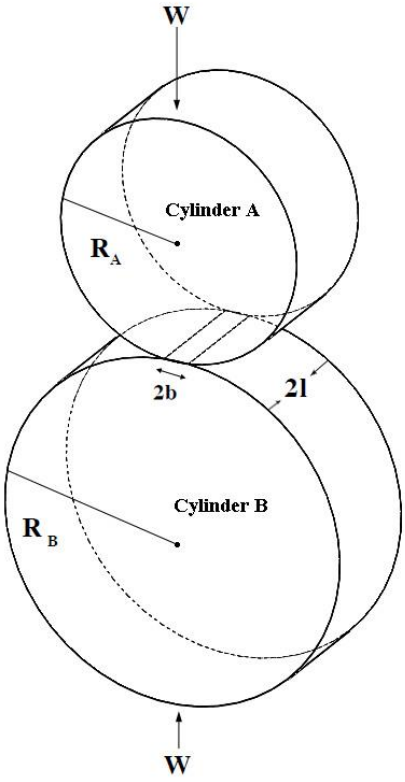

Figure 28. Contact between two parallel cylinders.

The Hertzian contact theory states that the maximum shear stress τ_{max} occurs at a depth z below the contact. For the two disks in contact shown in Figure 28, the half width of the contact rectangle b is given by

$$b = \sqrt{\left(\frac{4WR'}{\pi l E'}\right)} \tag{13}$$

Where W is the load, l is the half length of the contact rectangle, E' is the effective contact modulus, R' is the effective radius of curvature of the two disks. The equation to obtain the effective contact modulus and the effective radius of curvature are given in equations 8 and 9.

Where R_A and R_B are the radii of the two cylinders respectively, ν_1 and ν_2 are the Poisson's ratios. The maximum contact pressure P_{max} and average contact pressure P_{avg} are given in equation 14 and 15.

$$P_{\max} = \frac{W}{\pi b l} \tag{14}$$

$$P_{\text{avg}} = \frac{W}{4 b l} \tag{15}$$

The maximum shear stress can then be calculated by using the maximum contact pressure with equation 16.

$$\tau_{max} = 0.304 P_{max} \tag{16}$$

The depth where this maximum shear stress occurs below the contact surface can be estimated from equation 17.

$$z = 0.786 b \tag{17}$$

The depth of maximum shear stress gives indication of the location where plastic deformation is most likely to occur. It has been shown that martensite decay in ball bearings occurs at depths which correspond to the maximum shear stress as predicted by Hertzian theory. In gears the martensite decay occurs at a distance from the surface which coincides with the depth of maximum shear stress caused by contacting grinding marks (Oila, 2004).

2.6.5 Ellipital contact

The gear tooth contact is best described by an elliptical shape (see Figure **29**).

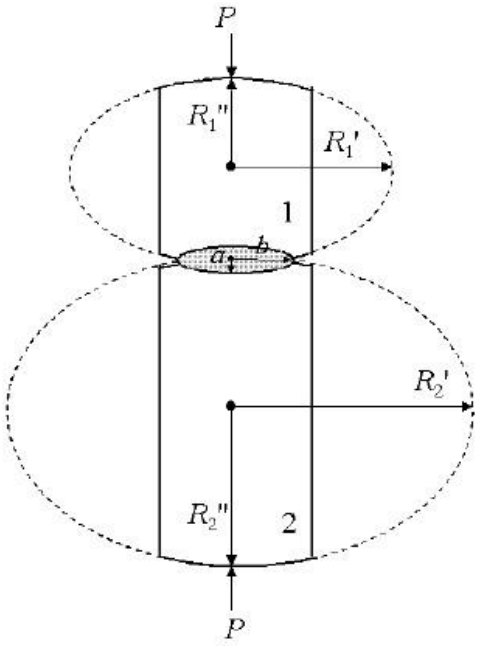

Figure 29. Elliptical contact (Oila, 2003).

In Figure 29, *P* is the load, *a* and *b* represents the major and minor semi-axis of the contact ellipse, *R′* is the maximum radius of curvature, and *R″* represents the minimum radius of curvature. The equivalent radius of curvature is defined by equation 18.

$$R_e = \sqrt{R' \cdot R''} \tag{18}$$

The equation for the ellipticity *c* is

$$c = \sqrt{ab} = \left(\frac{3PR_e}{4E'}\right)^{\frac{1}{3}} \tag{19}$$

The maximum contact pressure is

$$p_0 = \left(\frac{6PE'^2}{\pi^3 R_e^2}\right)^{\frac{1}{3}} \tag{20}$$

The samples investigated in this work were full scale gears and crowned disks used to simulate the gear tooth contact. The crowned disks simulates the contact between the gear teeth at point *P* as shown in Figure 30.

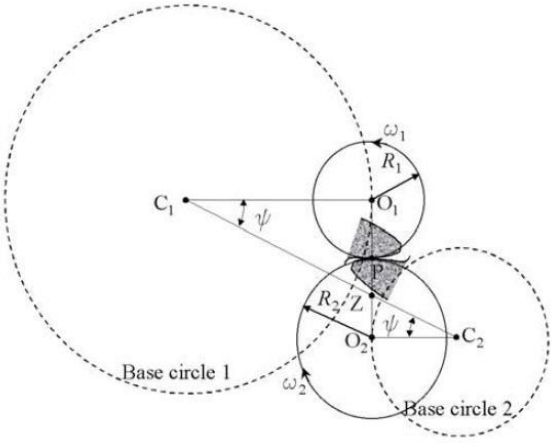

Figure 30. Contact between two gear teeth at point P can be described as the contact between the disks with their centres in O_1 and O_2 and their radii R_1 and R_2 respectively (Oila, 2003).

2.6.6 Surface roughness

The surface of a solid contains many defects and imperfections. An ideal surface is defined as a perfectly smooth surface and referred as the nominal surface. The real surface is not ideal due to manufacturing processes. The real surface contains peaks and valleys. A peak is defined as a point higher than its two adjacent points. A valley is a point lower than its two adjacent points. The irregular spacing forms a pattern or texture on the surface. The surface finish process can affect the surface topography depending on the tools used.

The topography of a surface contains certain components. Surface roughness represents the high frequency irregularities caused by the interaction of the material and the cutting tools (ASM Handbook, 2001). Roughness is characterised by asperities. The real area of contact is the combined area of these asperities. Surface waviness is defined as the medium frequency irregularities on the surface on which the surface roughness is superimposed. A Lay refers to a pattern that is directional such as parallel, circular, or radial.

The most commonly used parameter is the amplitude parameters roughness average (R_a).The roughness average is defined as the mean height of the roughness profile. R_a can be described by the equation:

$$R_a = \frac{1}{l}\int_0^l |y(x)|dx \tag{21}$$

where *y(x)* is the surface profile function with the mean value subtracted. *l* is the sample length. Surface roughness measurements are made according to a set of parameters. Some common parameters include amplitude parameters, spacing parameters, and hybrid parameters. The amplitude parameters measure the vertical characteristics of the surface. The spacing parameters measure the spacing of the irregularities on the surface. The hybrid parameters account the combination of amplitude and spacing irregularities.

2.6.7 Lubrication of gears

Lubrication is used to prevent direct metal-to-metal contact, to remove contaminants, and to remove heat from the gear contact. Inadequate lubrication will often cause early failure in gears. A gear lubricant consists of a base fluid with a combination of chemicals. For gears and bearings, the use of liquid lubricant is the most common. When choosing the right lubricant factors such as the film thickness, gear speed, chemical stability and, oxidation resistance must be taken into account. The high contact pressure may cause plastic deformation on the surface of the gear tooth, which can change the lubrication conditions. High temperature may change the chemistry of the lubricant and the friction coefficient of the lubricant. This may cause degradation in the lubricant chemistry or change the film thickness with consequences in reduced cooling and metal-to-metal contact which leads to failure. In terms of load, high asperity contact pressure leads to high friction and high lubrication temperature. In general, high rolling speed will reduce the effect of asperity contacts, but high sliding speed will increase the effect of the

asperity contact. In regards to micropitting, clean lubrication without contaminant is best for preventing micropitting. Studies done by Graham and Olver show that micropitting is a result of inadequate lubrication (Graham & Olver, 1980).

The lubrication mechanism can be divided into fluid pressure lubrication and surface film lubrication. Fluid pressure lubrication is defined as the mechanism where the fluid pressure is used to prevent surface to surface contact. Surface film lubrication is defined as the mechanisms where a protective film is used to separate the contact surfaces. The different regimes with the corresponding values for lubrication film thickness, *h*, and friction coefficient are presented in table.

Table 2. Lubrication regimes and lubrication film properties (Holmberg & Matthews, 1994)

Lubrication Regime	Film thichness (μm)	Specific film thickness (λ)	Friction coefficient (μ)
Boundary	0.005-0.1	<1	0.03-0.2
Mixed	0.01-1.0	1-4	0.02-0.015
Elastohydrodynamic	0.001-10.0	3-10	0.01-0.1
Hydrodynamic	1-100	10-100	0.001-0.01

Dowson and Higginson developed the minimum film thickness formula for lubricated line contacts (Dowson & Higginson, 1959) for predicting the film thickness and found that the film thickness is described by the equation

$$h_{min} = 2.65 G_m^{0.54} {U_s}^{0.7} W_L^{-0.13} R \tag{22}$$

where h_{min} is the minimum film thickness, G_m is the materials parameter, U_s is the speed parameter

$$U = \frac{\eta_0 u}{E' R_e} \tag{23}$$

η_0 is the viscosity, *u* is the velocity, *R* is the equivalent radius of curvature, *E'* is the elastic modulus, *W* is the load parameter described by

$$W_L = \frac{w}{LE'R_e} \tag{24}$$

The amount of friction developed between two contact surfaces depends on the pressure, heat generated and the surface conditions. High contact pressures can increase the local temperature and induce plastic deformation and local welding. The friction energy is measured by the friction coefficient

$$\mu = \frac{F}{N} \tag{25}$$

where *F* is the frictional force, *N* is the normal force.

Under dry conditions the friction coefficient due to sliding contact is between 0.1-0.4 while the friction coefficient due to rolling contact of the gear teeth is between 0.002 and 0.004 (Oila, 2004). This shows that sliding friction is much higher than rolling friction. In order to reduce the friction coefficient, gears need to be lubricated.

Viscosity defines a measure of internal friction of a fluid. It is the degree to which a fluid resists flow and it is the most important factor in determining a suitable lubricant. Viscosity can be described as being dynamic or kinematic. Dynamic viscosity also called absolute viscosity describes the internal resistance to shear flow.

Viscosity can be described by the equation:

$$\frac{F}{A} = \frac{\eta dv}{dx} \tag{26}$$

The *dv/dx* is a measure of change in fluid layer velocity and the *F/A* is the shear stress. Viscosity can be determined by dividing shear stress by the change in fluid layer velocity.

Kinematic viscosity also called static viscosity is the ratio of the dynamic viscosity to the density of the fluid. The standard unit of dynamic viscosity is poise (P). The standard unit for kinematic viscosity is stoke (St). Equation 22 gives the relationship between dynamic viscosity and kinematic viscosity:

$$\nu = \frac{\eta}{\rho} \tag{27}$$

Where ν is the kinematic viscosity of a fluid, η is the dynamic viscosity ensity, ρ is the density.

Hydrodynamic lubrication occurs when the pressure in the lubricant is large enough to prevent metal to metal contact. Hydrodynamic lubrication was first presented by Osborne Reynolds (Reynolds, 1886). He found that when the velocity of the contact surface increases, the liquid flow velocity also increases and produces liquid with a pressure sufficient to keep the two surfaces separated. The most important factors governing hydrodynamic lubrication are viscosity and speed. The higher the viscosity, the higher is the friction coefficient. High friction generates heat, which will decrease viscosity. Additives could be added to manipulate the viscosity temperature dependence if needed. In hydrodynamic lubrication high speed gives better lubrication and low speed may cause lubrication failure.

In elastohydrodynamic lubrication (EHL), the viscosity of lubricant is high enough so that it behaves similar to a solid and the contacting surfaces are subjected to elastic deformation.

2. 7 Surface contact fatigue

Contact fatigue during sliding can cause severe plastic deformation of the material at the surface and close to the surface. The direction of the deformation corresponds to the

sliding direction and higher friction coefficient will result in more severe deformation. The strain level caused by shearing in sliding is schematically illustrated in Figure 31.

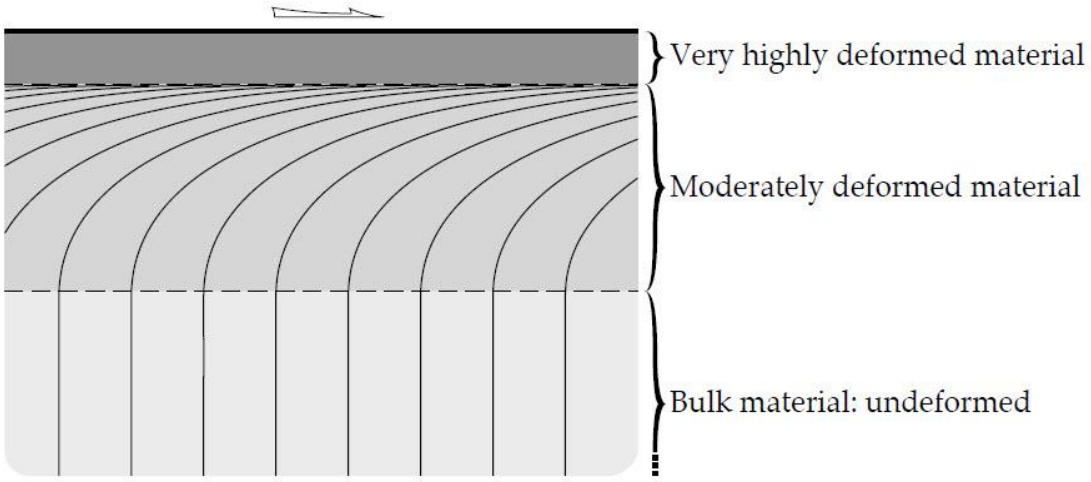

Figure 31. Strain level in a deformed surface (Stachowiak, 2014).

The friction coefficient contributes to the severity of deformation. A low coefficient of friction due to lubrication will result in less deformation. The high friction during sliding is mainly due to adhesion between the contacting surfaces. The high strains result in dislocations forming in the original grain structure.

The crack initiates at the surface and propagate downward and then the crack may go deeper into the material or continue back to the surface and create a pit and wear debris. Figure 32 shows a schematic of a surface crack initiation and propagation mechanism.

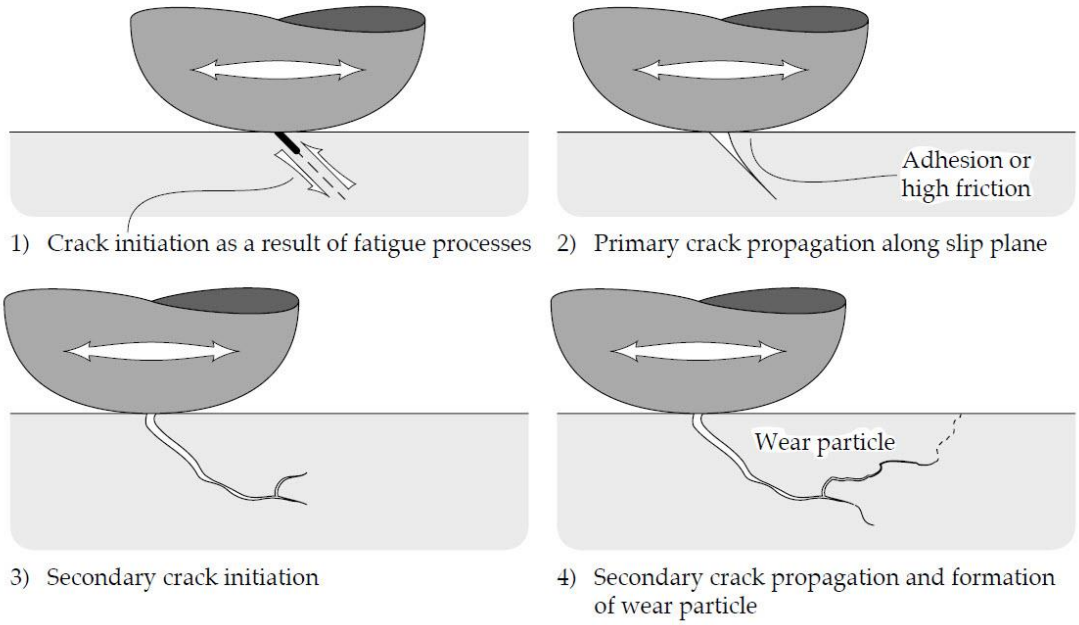

Figure 32. Schematic of the process of surface crack initiation and propagation (Stachowiak, 2014).

The fatigue wear during sliding contact is a result of surface initiated fatigue cracks in the plastically deformed layer. Subsurface crack can also initiate under the surface of the contacting bodies. Subsurface cracks originate at imperfections such as inclusions, internal defects, areas with weak grain boundaries, or zones of high residual stress. The crack nucleation is controlled by the cyclic plastic deformation and the cracks nucleate at the areas where cyclic plastic deformation is high.

In pure rolling conditions lubrication provides an EHL film which prevents metal to metal contact. Figure 33 shows the mechanism of wear during lubricated rolling contact between two surfaces.

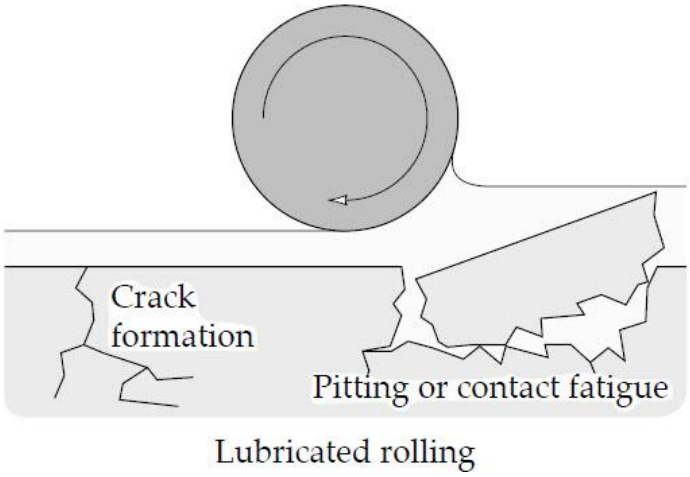

Figure 33. Schematic of wear occuring during lubricated rolling (Stachowiak, 2014).

The rolling contact is defined by the rolling of one body over another body. The repeated application of stress results in the initiation of subsurface or surface cracks. The rolling contact wear is most often associated with cyclic stress so it is commonly referred to as rolling contact fatigue. In gears, pure rolling occurs only at the pith diameter while in addendum and dedendum a sliding component is also present.

The film thickness may reduce due to roughness, high loads or reduced viscosity. Figure 34 shows the scenario of imperfect EHL and wear debris trapped by the asperities of the two contacting bodies.

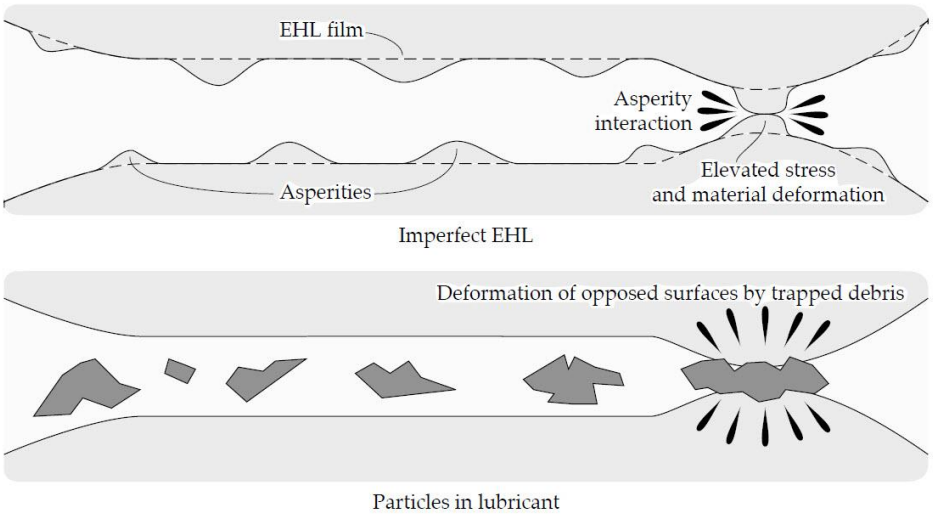

Figure 34. Contact between asperities as a cause of contact fatigue (Stachowiak, 2014).

When the dimension of the wear debris exceeds the EHL film thickness, the debris can be trapped between asperities as shown in Figure 34. Wear debris in the lubricant can result in accelerated wear or contact fatigue, therefore it is important to avoid wear debris in order to increase the lifetime of gears.

In rails plastic deformation occurs as a result of the cyclic contact caused by the train wheel. The cracks propagate along the deformed grain structure. Figure 35 shows the crack development in steel rail subjected to surface contact fatigue.

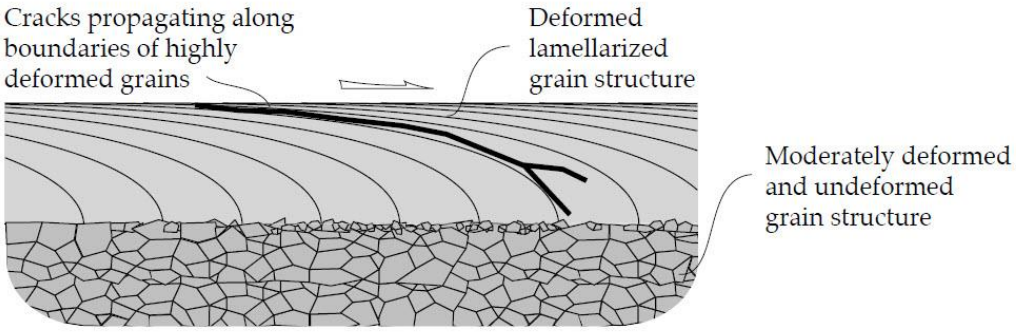

Figure 35. Cracks development caused by plastic deformation during rolling contact (Stachowiak, 2014).

A crack present on the surface can propagate due to the mechanism known as the hydraulic pressure crack propagation. This mechanism suggests that a crack opening under lateral tensile stress is filled with a lubricant due to high pressure. The presence of lubricant inside the crack has been demonstrated before (Bower, 1988). The crack then propagates due to the compressed lubricant in the cracks. The hydraulic pressure theory does not explain the direction of crack propagation. A schematic of the process of hydraulic pressure crack propagation is shown in Figure 36.

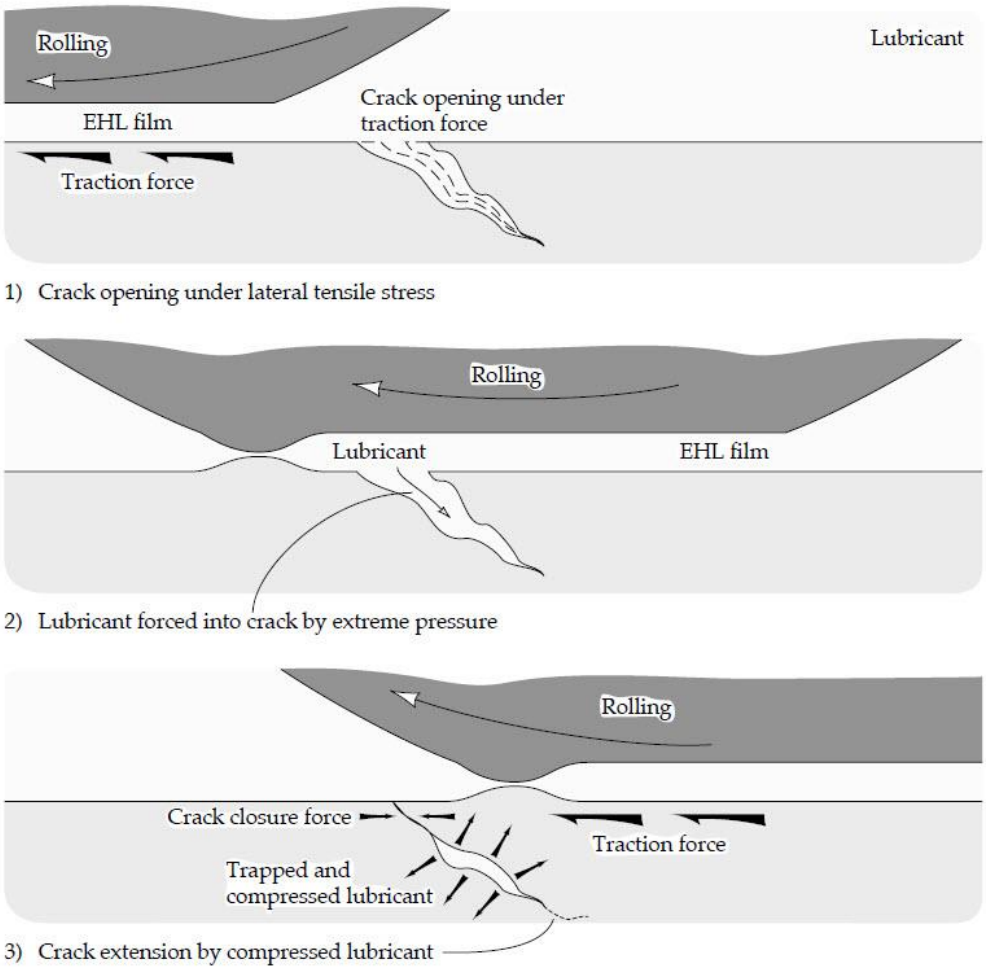

Figure 36. Schematic of the process of hydraulic pressure crack propagation (Stachowiak, 2014).

2.8 Micropitting

Micropitting is a form of surface contact fatigue observed in elastohydrodynamic lubrication conditions (Yoshida & Taki, 2008). The micrometer size pits can develop on the surface of the gear teeth in a short period of time and introduce noise and vibrations due to the change in the surface roughness. If left unattended, micropitting may lead to tooth breakage and other catastrophic failures. Previous research has shown that surface contact fatigue is associated with microstructural transformations in steel (Swahn et al.,1976; Voskamp et al., 1980; Voskamp et al., 1985; Voskamp et al., 1997; Bush et al., 1961; Bhargava et al., 1990; Becker et al., 1976; Buchwald & Heckel, 1968; Gentile et al., 1965; Oila et al, 2005; Osturlund & Vingsbo, 1980; O'Brien & King, 1966). This section provides the background information related to micropitting and the phase change phenomenon (martensite decay).

2.8.1 Micropitting in Gears

All types of gears with various heat treatments are susceptible to micropitting (Errichello, 2002). Micropitting can be identified as a collection of micrometer size holes on the surface of the gear teeth. The asperity contact and sliding between the two gears creates plastic deformation which is followed by crack initiation and crack growth. Heavy loading and high temperature are contributing factors in micropitting. Micropitting starts with a microcrack at or near the surface that propagates at around 30° into the material. The slip line field theory gives prediction of a region of tensile residual stress and the maximum tensile stress is at an angle around 30°, which corresponds to the crack propagation direction. The crack returns to the surface and small volumes of material detach from the surface leading to a small cavity forming a micropit. The lubricant trapped in the crack as the gears are moving increases the pressure inside the crack

leading to accelerated crack propagation (Yoshida & Taki, 2008). Figure 37 shows a schematic of the mechanism of micropitting.

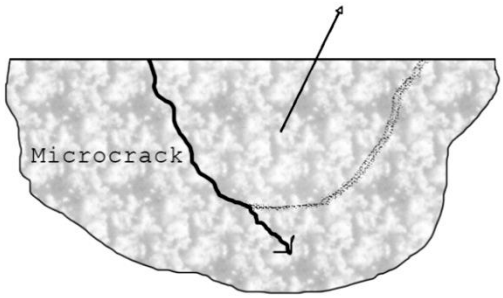

Figure 37. Schematic representation of micropitting.

Micropitting may occur in the addendum and dedendum regions on the gear teeth, and it most often forms in regions of high rolling/sliding contact stresses (Oila, 2003). Figure 38 shows a schematic of a gear tooth contact and the locations where micropitting is most often observed. Rolling and sliding directions are also indicated.

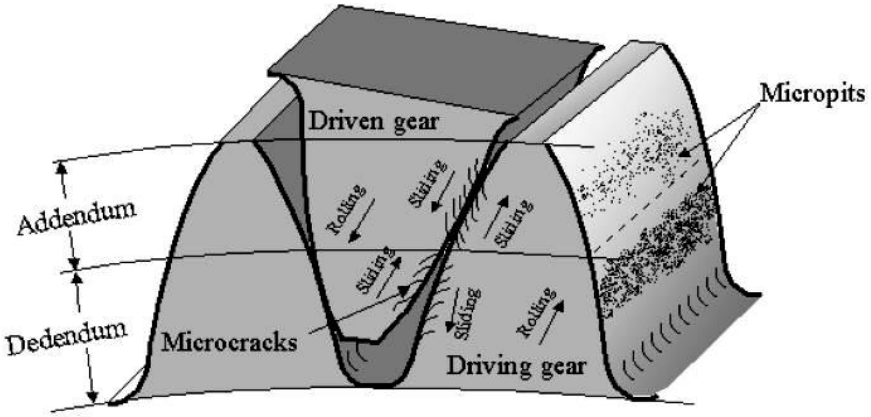

Figure 38. Rolling and sliding contact between gear teeth and the regions where micropitting occurs (Oila, 2003).

The complexity of micropitting is due to the large number of influencing factors. Operating conditions such as load and speed, material microstructure, alloying elements, amount of retained austenite and tempered martensite are all important factors. The condition of the surface is important since rough surfaces tend to create stress

concentration at asperities (Oila, 2003). The contact pressure is also a major factor in micropitting initiation (Oila & Bull, 2004).

The phenomenon of micropitting in gears is related to the high stress concentration areas developed under the contact. Cracks may initiate at or near the surface and they propagate for a few microns below the contact. The crack will change its propagation path towards the surface and when it reaches the surface small volumes of material will detach leading to the formation of a micropit. The damage can initiate both, at the surface or below the surface depending on the conditions. To the naked eye this region appears etched compared to the surrounding area. Under the microscope, small pits are observed in these regions (Winter, 1980). Micropitting can occur with all heat treatments and lubrications after around 10^6 cycles in gears (Winter, 1980). The mechanism of this form of surface contact fatigue is not well understood in gears and previous research (Oila, 2004, Hoeprich, 2000) suggests that micropitting is related to the phase changes observed (martensite decay).

2.8.2 Martensite Decay

Martensite decay is the term used to describe peculiar microstructural features observed in fatigued components such as bearings and gears. The products of martensite decay are the Dark Etching Regions (DER) and the White Etching Bands (WEB) (see Figure 39).

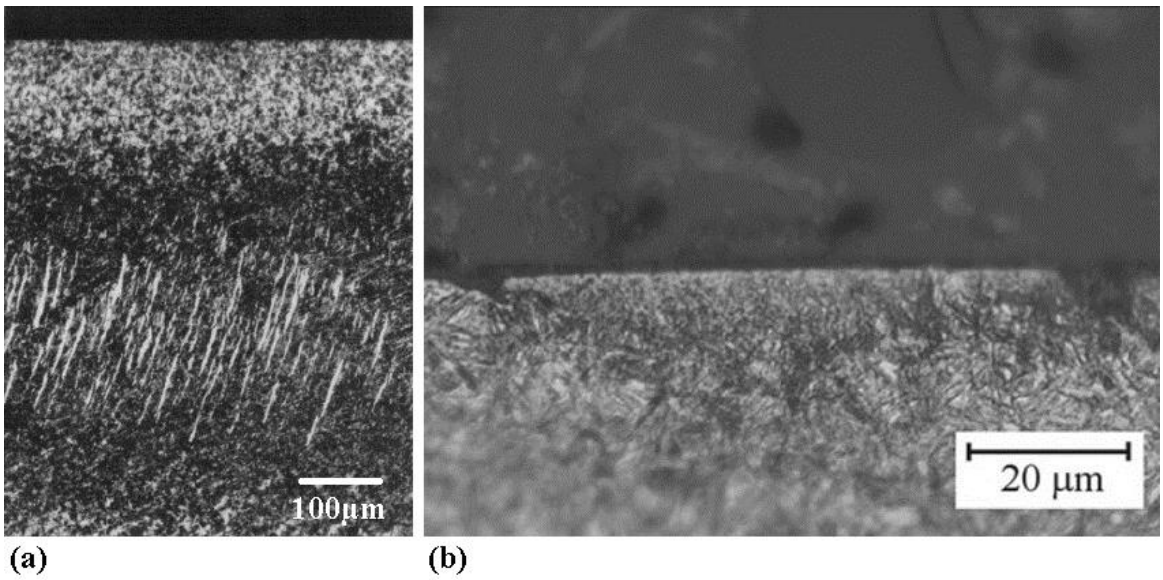

Figure 39. (a) DER observed in bearings (Swahn et al., 1976) (b) DER observed in gears (Oila et al., 2004)

In rolling element bearings DER were first observed in 1947 by A. B. Jones (Jones, 1947). WEB were also observed in the subsurface regions (Tunca & Laufer, 1986). The DER mainly consists of overtempered martensite (Tunca & Laufer, 1986). The WEB was found to have a hardness of 11.3 GPa and DER with a hardness of 6.5 GPa while tempered martensite has a hardness of 7.4 GPa (Tunca & Laufer, 1986). The DER developed around the depth of maximum shear stress and consisted of regions of high density of dislocations. The migration of carbon atoms from martensite to the heavily dislocated region was suspected to be the cause of the phenomenon. It was also found that the material softens after 10^6~10^7 cycles of stress (Furumura et al., 1996).

According to Swahn et al. (1976) the DER mixture contains martensite, carbides, and ferrite-like phases. It was suggested (Swahn et al. 1976) that the transformation

mechanism of DER occurs by the successive recrystallisation in a preferred crystallographic direction and that martensite decay is linked to the oversaturation of carbon atoms at grain boundaries during cyclic loading.

In gears, DER was associated with high asperities (Hoeprich et al. 2001). This point made by the author leads to the idea that the DER can be prevented by surface engineering or treatment via either coating or lubrication by decreasing the amount of asperity contact. It was shown that at high loads and for poorly lubricated contacts density of micropitting increases (Hoeprich et al. 2001).

Transmission Electron Microscopy (TEM) studies found that the initial DER matrix is made up of homogeneously distributed spherical carbides that gradually decay to a new structure along the maximum shear stress beneath the surface (Swahn et al. 1976). It was also found that the WEB is made up of ferrite free of carbides (Swahn et al. 1976). Eta-carbides resulting from tempering appeared as plates in the investigated structure (Swahn et al. 1976).

Martensite is a supersaturated interstitial solid solution of carbon in α-iron. All carbon atoms occupy one of the third possible locations of the octahedral sites. The interstitial carbon atom creates a strain field (Zener & Wert, 1950) which can be understood by using the elasticity theory.

Because martensite is constrained in an austenite matrix, any redistribution of interstitial atoms between different sub-lattices of octahedral interstitial sites leads to changes in the lattice parameter of the martensite. An increase in dislocation density by plastic deformation of the martensite at low temperature leads to an increase in the release of heat. Comparing the internal friction and heat release rate suggest that the migration of carbon atoms to dislocations is the cause of carbon atom redistribution in martensite (Bronfin et al., 1986).

Rolling contact fatigue has been related to the decomposition of the retained austenite and microstructural changes (Voskamp et al. 1985). The relationship between cyclic contact and microstructural changes is influenced by the transformation of retained austenite and other factors related to volume change (Hahn et al. 1987).

A study by Buchwald (Buchwald & Heckel, 1968) showed that lubricant must be present if micropitting is to occur, but if the lubricants have a viscosity above a critical value depending on the load, then micropitting can be prevented (Buchwald & Heckel, 1968; Way, 1995). It was shown that the change of microstructure is a result of cyclic stress associated with the formation of carbide and plastic deformation (Buchwald & Heckel, 1968). The carbon atoms diffuse into the undeformed regions adjacent to the WEB and reprecipitates as carbides. Voskamp (Voskamp et al. 1985) discussed the DER's association with the carbon movement to the heavily dislocated regions and the association with the martensite rejection of excess carbon (Voskamp et al. 1985). He also showed that the WEB develop later than DER and the bands have lower hardness.

The WEB was found to be the result of stress induced carbon diffusion that breaks up primary carbides (Swahn et al. 1976). Since ferrite is soft, the deformation during the rolling contact will concentrate on the ferrite, resulting in increased dislocation density. The increased dislocation density may lead to the microstructural change known as the DER.

Previous studies (Swahn et al. 1976) on ball bearings suggest that the martensite decay due to rolling contact fatigue depends largely on the stress applied. Although there are very few studies on gears there are correlations that can be extracted. The transformation mechanism proposed by Swahn (Swahn et al. 1976) assumes that carbon redistribute from martensite and carbides. The microstructure of DER shows a strongly deformed ferrite structure, which suggests that the mechanism requires carbon diffusion from martensite

matrix to carbides. Martin et al. (1966) found that the WEB has a fine cell structure that might be deformed ferrite and the phase transformation may generate residual stress.

In conclusion, the dark etching regions are regions with strain-induced microstructural changes which occur as a result of cyclic strain. The change in microstructure contributes to the crack nucleation and growth at the DER boundaries (Swahn et al. 1976). Surface contact fatigue produces a microstructural change which increases the strain around the DER and ultimately leads to nucleation of cracks.

The DER have been observed by others (Obrien & Kings, 1966;Swahn et al., 1976; Gentile et al., 1965; Voskamp, 1998; Marten et al, 1966). Bush (Bush, 1961) found that carbides formed as a result of diffusion.

A phase transformation is caused by the free energy difference between two phases. The free energy is influenced by pressure and temperature. The martensitic transformation is also influenced by external stress. The martensitic transformation takes place by lattice deformation of the parent crystal, and the lattice deformation is induced by the shear deformation, so the phase transformation is favoured by applying shear stresses. The martensite start temperature is raised when stress is applied to the specimen and transformation will occur at that particular temperature as shown by Kawakami.

The existence of dark etching regions (DER) was observed in the carburised gear shown in Figure 40.

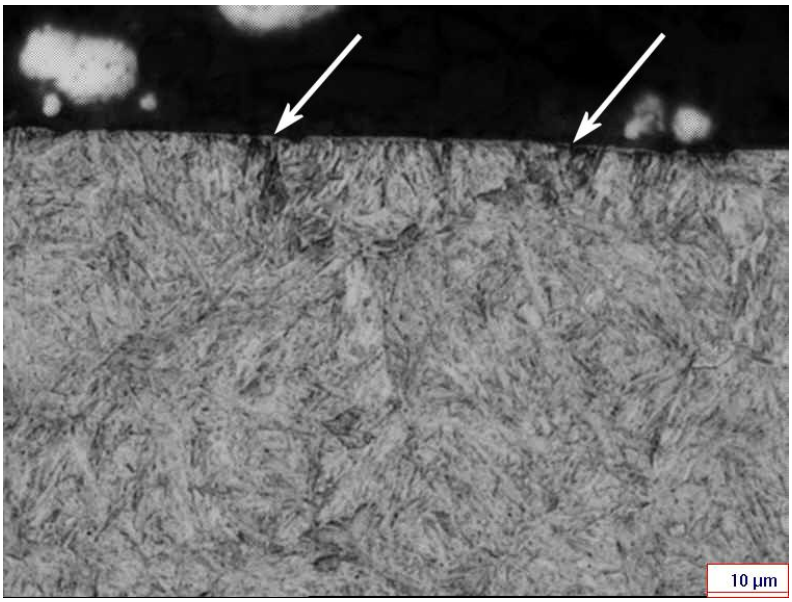

Figure 40. Dark etching region observed in the caburised gear.

2.9 Surface Contact Fatigue Models

2.9.1 Background

In this section the most widely used models developed for surface contact fatigue are presented.

2.9.2 The Lundberg-Palmgren Model

In 1947 Lundberg and Palmgren (Lundberg & Palmgren, 1947) proposed the first mathematical model to predict the bearing fatigue life. The Hertzian theory is valid when the contact area is small compared to the dimensions of the bodies and the frictional forces in the contact area are neglected. Their theory describes the probability of survival S expressed as a function of the shear stress and the depth of the maximum shear stress under the surface. The equation is given by

$$\ln\frac{1}{S(N)} \propto \frac{N^{\mathrm{e}}\tau_0^{\mathrm{c}}V}{z_0^{\mathrm{h}}} \tag{28}$$

where τ_0 is the maximum shear stress, z_0 is the depth of maximum shear stress, V is the stress volume, e,c, and h are constants. The stress volume is given by

$$V = alz_0 \tag{29}$$

where a is the semi-major axis of the hertzian contact ellipse, l is the length of the running track of the race.

The assumptions of this model are:

1. The crack initiation period is much greater than the crack propagation period.

2. Fatigue failure is always initiated at the depth of maximum shear stress.

3. The subsurface stress distribution is a function of contact pressure.

The model does not include friction, material imperfections such as inclusions, residual stress, surface roughness, or lubricant properties.

2.9.3 The Tallian and Colleagues Bearing Fatigue Model

Tallian (Tallian, 1992) attempted to improve the Lundberg-Palmgren model by incorporating the crack propagation as a result of extensive cyclic load. The model takes into account the surface and subsurface defects and the elastic properties of the material. Over the years the Tallian model has been improved several times with varying modifying factors including lubricant film thickness, surface roughness, friction, and asperity interaction. The model assumes that the fatigue life is the combination of surface and subsurface initiated fatigue cracking.

$$N^{-\beta} = N_s^{-\beta} + N_{ss}^{-\beta} \tag{30}$$

Where N represents fatigue life related to cracks, β is the Weibull dispersion parameter, N_s and N_{ss} are fatigue lives related to surface and subsurface initiated cracks respectively. This model assumes that the crack initiation is dominant. Tallian improved his model by

including several parameters. The following equation gives the local probability of survival

$$\mathrm{p}(N) = \exp(-m\Delta V \mathrm{F}(N)) \tag{31}$$

where m is the defect volume density, F(N) is the defect distribution, V is the stressed volume. The probability of survival is:

$$S(N) = \exp(-CN^{\beta}V) \tag{32}$$

where C is the parameter incorporating factors such as load, defect density, and fatigue resistance, V is the volume, and β is the Weibull dispersion parameter.

This model includes the following assumptions.

1. The fatigue crack initiation time is small compared to the crack propagation time.

2. The crack propagation equation used is similar to Paris law which describes the crack growth under fatigue stress. The Paris law relates the stress intensity factor to the crack growth.

$$\frac{da}{dN} = C\Delta K^{m} \tag{33}$$

where a is the crack length, ΔK is the range of the stress intensity factor, C and m are material constants, N is the number of load cycles.

3. The fatigue crack propagation follows the mode 2 shear stress fracture mechanics.

4. The crack initiates from surface defects and subsurface defects are neglected.

5. The defect severity distribution follows a power law.

6. The defect density per contact surface element follows a Poisson distribution.

2.9.4 Models for Fatigue Life of Gears

The fatigue life model by Lundberg and Palmgren applies to rolling element bearings. In order for the model to be applicable to gears modifications to account for gear geometry are required. Antoine and Besson (Antoine & Besson, 2002) developed a model for micropitting for gears. It was found that the equivalent dynamic stress could be approximated by

$$\sigma_{eq} = \frac{\sigma_{max}/2}{1 - [(\sigma_{max}/2)/R_{0.2}]^2} \tag{34}$$

σ_{max} is maximum stress (MPa), $R_{0.2}$ is yield strength (MPa). A model for gears was also developed by Blake in 1991 (Blake, 1991). He developed a pitting life model based on fracture mechanics to measure the service live of spur gears. The contact fatigue model includes the combination of assumptions and approximations using fundamental physics. This model narrows down the number of factors affecting the surface contact fatigue and reaches a prediction by using the probability of failure. The most important factors are the magnitude of stress, the size of the contact area, the distribution of stress and depth, and the material properties. These factors will affect the crack initiation and propagation mechanism.

Chapter 3. Computational Modelling and Simulation

3.1 Background

Computational Modelling is the term used to describe the use of mathematical models to predict certain material properties and material behaviour under specified conditions. The development of a mathematical model can help us further the understanding of the structural mechanics of material. First principle methods can be employed to calculate atomic configurations and elastic properties of individual atomic clusters or phases without relying on empirical parameters as in empirical methods.

In the computational atomistic simulation, quantum mechanics is used as a tool to predict certain properties of material. Using these calculated material properties, it is possible to use finite element analysis (FEA) to simulate material behaviour under different conditions at much larger scales.

3.2 Atomistic Simulation

In material science there are various simulation methods in terms of length scale and time scale. Macroscopic simulation applies to length scales larger than 10^{-3}m and time scales larger than 10^{-6}s. Mesoscopic simulations are appropriate for length scales between 10^{-6}m and 10^{-3}m with time scale between 10^{-10}s and 10^{-6}s. Nanoscopic simulations are used to investigate materials at length scale between 10^{-9} and 10^{-6}m and time scale between 10^{-13} and 10^{-10} s. Quantum simulations are used at length scales between 10^{-11}m and 10^{-8}m with a time scale between 10^{-16} and 10^{-12}s. Figure 41 shows the diagram of timescales and length scales.

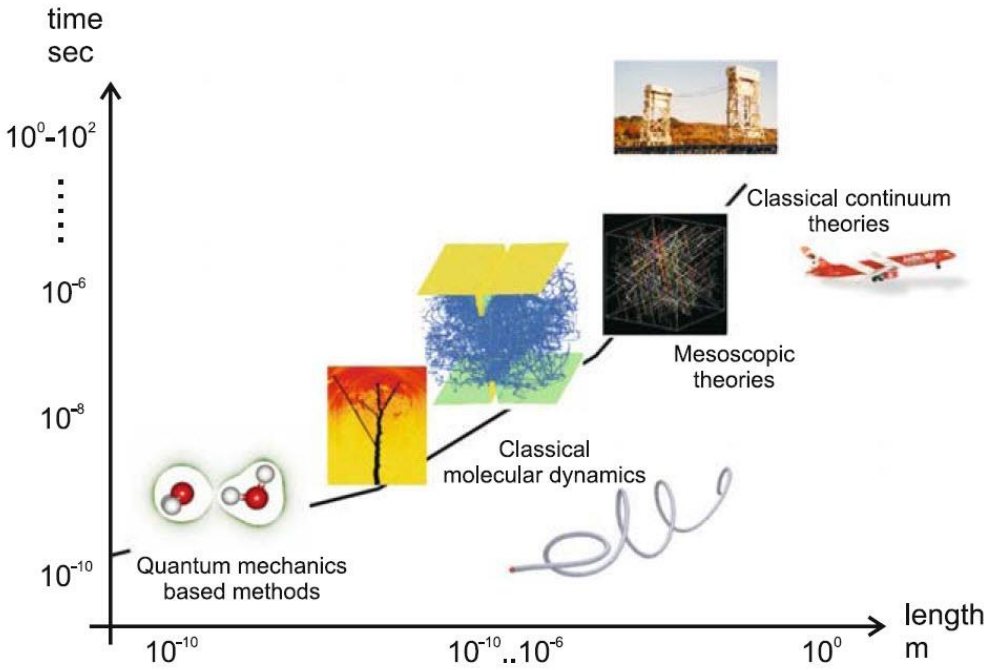

Figure 41. Diagram of different simulation tools associated with lengthscales and timescales (Buehler, 2008)

The goal of quantum mechanics methods is to solve the time-dependent Schrodinger equation and determine the electronic structures of atoms and molecules. In molecular dynamics the atom movements follow the Newton's law of motion. The Newton's equation of motion is

$$F_i = m_i a_i \tag{35}$$

where *m* is mass and *a* is acceleration for each particle *i* in the system. The dynamics of the system is described by

$$m_i \frac{d^2 r_i}{dt^2} = -\frac{dU(r)}{dr_i} \tag{36}$$

where *U* represents the potential energy, $\frac{d^2 r_i}{dt^2}$ is the acceleration and *m* is the mass.

The first principle quantum mechanical simulation of molecular properties involves the wave equation

$$H\Psi = E\Psi \tag{37}$$

Where *H* is the hamiltonian operator, *E* is the total electronic energy, Ψ is the molecular wave function. Depending on the operators, the molecular wave function can give the eigenvalues of all molecular properties. An operator is a symbol for an instruction to perform an action. The Hamiltonian operator corresponds to the total energy of the system denoted by *H*. It is the sum of the kinetic energy of all the particles plus the potential energy of the particles associated within the system and is represented by the equation:

$$\hat{H} = \hat{T} + \hat{V} \tag{38}$$

where *T* is the kinetic operator and *V* is the potential energy operator.

The Schrödinger equation is a probabilistic statement about the position and impulses of the particles. It is based on the idea that the position and impulses cannot be measured arbitrarily accurately at the same time according to the Heisenberg's uncertainty principle. All statements that can be made about a quantum system can be derived from a wave function or state function.

$$\Psi = \Psi(R_1, \dots R_N, r_{1,\dots} r_{k,}\, t) \tag{39}$$

$$3(N + K) = \text{coordinate of particles in vector space}$$

R represents the position of nuclei in space, *r* represents the position of electron in space and *t* is the time. By solving this equation it is possible to describe the system at time *t* in the volume at coordinate (*R*,*r*). Integrating over a volume gives the probability to find the system in a certain state. Ψ represent the wavefunction of the system and contain all the properties of the system.

The electrostatic potential is also called the Coulomb potential. The Coulomb charge is given by $e/4\pi r\varepsilon_0$ where ε_0 is the dielectric constant, *r* is the distance from the position of

the charged particle, $1/4\pi\varepsilon_0$ is also known as the coulomb constant. An electron moving in this potential has the potential energy of $-e^2/4\pi r\varepsilon_0$.

The Hamilton operator associated to a system consisting of a nucleus and electrons is given as the sum over the operators for the kinetic energy and coulomb potentials where

$$H = T_e + T_k + V_{ee} + V_{ek} + V_{kk} \tag{40}$$

where T_e and T_k represents the operators of kinetic energy of electron and nuclei, V_{ee}, V_{ek}, and V_{kk} represents the operators of the potential energy of the interactions between electrons, between electron and nuclei, and between only nuclei respectively. The state function is now given as the solution of the Schrödinger equation

$$i\hbar\frac{\partial\Psi(R,r,t)}{\partial t} = \widehat{H}\Psi(R,r,t) \tag{41}$$

The Hamiltonian contains terms describing the motion of electrons around the nucleus and the motion of nucleus. The Schrodinger equation is too difficult to solve even by using computers. Therefore, the problem is simplified by assuming that the mass of the electron is very small comparing to the mass of the nucleus. The movement of nucleus is small due to the large difference in mass, which makes the movement of nucleus negligible. This is referred to as the Born-Oppenheimer approximation. Because the forces acting on electrons and ions are similar, their momentum is also comparable. Because the ions are much larger than the electrons, the kinetic energy of nuclei is much smaller than that of the electrons. The electrons are assumed to respond instantaneously to nucleus motion. It is assumed that electrons are in the instantaneous ground state. The Schrodinger equation is used in each time step of Newton's equation to directly compute the potential energy for the nucleus, which is the basic idea of ab-initio (first principles) method. The electronic Schrodinger equation is solved in order to determine the effective potential energy of the nucleus and the movement of the electrons according to the Newton's equation of motion. There are two main approaches to solving the Schrodinger

equation, the ab initio method and the semiempirical method. In this thesis we will only focus on the ab initio method.

3.3 Ab-initio method

Ab-initio is a Latin term translated as "from beginning." In computational simulations it is used to describe the most fundamental properties of a material. The calculation is “ab-initio” when it is based on the laws of nature without additional assumptions. For example, the calculation of a physical property of steel starts with the properties of Fe and C and the laws of electrostatics and quantum mechanics. From these basics, the properties can be derived by computations of the interactions of atoms until the bulk properties of steel have been determined. Ab-initio calculations are performed by computing the forces acting on the nucleus from electronic structure calculations as the molecular dynamic trajectory is generated. A model must be chosen for the electronic wave function and the Schrödinger equation is solved by using the atomic numbers of the nuclei involved.

In order to find the single particle solutions to the Kohn Sham equation the single particle orbitals can be represented in a basis set. A basis set contains the linear combination of atomic orbitals, a real space grid, and plane waves. The plane waves represent the orbitals in the Fourier space.

For a periodic system following the Bloch’s theorem

$$\psi_{\boldsymbol{k}}(r) = \exp(i\mathbf{k}\cdot\mathbf{r})\, u_{\mathrm{k}}(\mathbf{r}) \tag{42}$$

$$\psi_{\boldsymbol{k}}(r) = \sum_{\boldsymbol{G}} C_{\boldsymbol{G}+\boldsymbol{K}} e^{i(\boldsymbol{G}+\boldsymbol{K})\cdot\boldsymbol{r}} \tag{43}$$

where G is the reciprocal lattice vector and K is the symmetry label in the 1st Brillouin zone. Bloch wave is a wavefunction of a particle placed in a periodic potential. For solids there is a fundamental symmetry with a periodically extended system represented with a unit cell. Each k value represent a region within k space. The energy eigenfunction for

such a system may be written as the product of a plane wave function and a periodic function which has the same periodicity as the potential.

The cut-off energy limit the number of plane wave components to those such that

$$\frac{(G+K)^2}{2} \leq E_{cut} \tag{44}$$

This defines a length scale where

$$\lambda = \frac{\pi}{\sqrt{E_{cut}}} \tag{45}$$

Choosing a cutoff energy requires choosing the minimum length, and it depends on the elements in the system. According to the variational principle the energy decreases to the ground state as the cut off energy increases.

3.4 Self consistent fields

The self-consistent field (SCF) method was originally introduced by Douglas Hartree (Hartree, 1928) and later improved by Fock and Slater (Fock, 1930). The improved method is called the Hartree Fock method. The Hartree Fock method assumes that the exact wave function can be calculated by a single Slater determinant. This simplifies the problem significantly. The Schrodinger equation is converted to a set of Hartree-Fock equations and is integrated for that particular electron and the effective potential in which it moves. The total charge is defined by the position of the electron and the distance of the electron from the nucleus. The wavefunction of other electrons is assumed to be already known. The Hartree-Fock equation for the electron is solved and the procedure repeated for all the electrons in the atom. A set of improved wavefunctions for all the electrons are generated and used to calculate the new effective potential. The computation continues until the improved set of wavefunction does not change significantly from the previous

result. At this point the wavefunction is called "self-consistent" and is accepted to be a good approximation to the many electron atoms.

3.4.1 Iterations to self-consistency

The following steps are necessary for a self-consistency calculation:

1. Initial guess at the orbital.

2. Construction of all operators and external potential. Specify positions of atoms and pseudopotential.

3. Assign a trial electronic density

4. Construct Hartree operators and exchange correlation.

5. Solve Kohn-Sham equation for $\widehat{H}$ by a specified diagonalization (solving the eigenstate equation). There are various methods for this: Davidson, Lancsoz, or conjugate gradient method.

6. Generate solution of the single-particle Schrodinger equation and calculate the charge density.

7. With the new set of orbitals, construct the Hartree operators again.

8. Iterate the procedure until the solution converges or if the Hartree operator does not change anymore.

Due to the nature of the approximation, when the Hartree equation is solved some information on what happens instantaneously to all other electrons due to the correlation principle is lost.

The complexity of the many body problem can be represented by an arbitrary atom. If the atom has an x number of electrons, the electron wave function has $3x$ the number of

variables and a full sampling would require 10^{3x} calculations because each electron would interact with every other electron in order to know what every other electron is doing. It is because of this reason that the Hartree-Fock provide the solution of not considering what each electron instantly does, but treating it as what one electron would do in a field on average of what all the other electrons would be doing. This method allows one to only require the solution of one electron with the average charge distribution of what other electron does. This is called the "mean-field effective potential". The Hartree solution leads to a Schrodinger equation in which we are trying to solve the problem of a single electron at a time but that electron feels the average electrostatic charge distribution of all the other electrons.

3.5 Pseudopotentials

Pseudopotentials are used to approximate complex systems. They can be used to describe the core electron behaviour and the motion of nuclei. This allows the Schrodinger equation containing the modified potential instead of the Coulomb potential. Pseudopotentials replace the potentials that describe the electrons in an atom such that all the core states are eliminated. The density functional theory (DFT) in practice uses the total energy Pseudopotentials approach which is the approach to study most solids. This approach removes the tightly bound core electrons. The reason is that in solids describing the core electron is exceedingly complex. A large number of electrons are unaffected by the chemical environment because they are so tightly bound to the nucleus. Also the spatial variation of core electron will be sharp and their contribution to the valence bound really does not change much. These are often the 1*s*, 2*s* electrons. The core electrons are assumed to be one with the core forming a non-polarisable ion core. This is also called the frozen core approximation (Barth & Gelatt, 1980). This step simplifies the properties of the element and significantly improves the computational speed of the simulation.

The ultrasoft pseudopotentials (Vanderbelt, 1990) are a modern form of pseudopotentials. They have the characteristics of adding more freedom for the 2*p* and 3*d* states. They also add augmentation charges inside the cutoff sphere to correct the charge.

3.6 Density Functional Theory

The ab initio method requires wavefunctions for the calculations but these wavefunctions do not physically exist or can be measured, so density functional theory (DFT) is used. Density Functional Theory (DFT) is a quantum mechanical method used to investigate the electronic structure of atoms and molecules at the ground state (the lowest energy state or vacuum state). It is a method of obtaining an approximate solution to the Schrodinger equation of a many body system in order to investigate the structural, magnetic, and electronic properties of molecules and materials. It was developed by L.H.Thomas and E. Fermi in 1927 (Fermi, 1927) and is a powerful tool for investigating the complex behaviour of phases in material. The advantage of using DFT is that it uses the electron density for an electron and does not involve inter-electron distances. The single particle approach is simpler to solve than the Hartree-Fock equations. This makes it possible to consider more complex systems by using the Kohn-Sham equation.

The dynamic behaviour of atoms is simulated using Newton's law and atom-atom interactions. The atomic interaction is simulated using the density functional theory which utilises the Schrodinger's equation to obtain a solution for the model. The results of the atomistic simulation provide information on the material behaviour and the most fundamental description of a material. Density functional theory uses an assumption of the reciprocal relationship between potential energy and electron density in a molecule in order to simplify the computations. The electronic energy *E* is treated as a functional of the electron density at a coordinate in space. The concept of DFT is that the many body problems of interacting electrons and ions can be mapped to a series of Kohn-Sham

equations (one electron equation). This means that all ground state molecular properties can be calculated from the ground state electron density. The main purpose of this is to simplify the molecular electronic Hamiltonian

$$H = -\sum_{i=1} \frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{2} \sum_i \sum_{j \neq i} \frac{e^2}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_i \sum_I \frac{e^2}{4\pi\varepsilon_0} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \tag{46}$$

where *i* and *j* represent electron. *I* represents nuclei. *Z* is the atomic number. The energy eigenvalues of this Hamiltonian occur as solutions of the Schrödinger equation.

$$\left[-\frac{1}{2}\nabla^2 + \phi\right] \psi_i(r) = \epsilon_i \psi_i(r) \tag{47}$$

$$\phi = v(R) + \int \frac{\rho(r)}{|R - r|} dr + E_{xc}[\rho] \tag{48}$$

$v(R)$ is the nuclear potential, $\int \frac{\rho(r)}{|R-r|} dr$ represents the coulomb potential, $E_{xc}[\rho]$ represents the exchange correlation potential written as a function of density.

Density functional theory roots in the Thomas Fermi model and Hohenberg-Kohn theorems. In density functional theory the functional is the electron density. The electron density is used as a fundamental property. Using electron density significantly increases the calculation speed. Hartree Fock theorem suggests that the density of any system determines all ground state properties of the system. In this case the total ground state energy of the system is a functional of the electron density. So if we know the electron density functional, we obtain the total energy of the system. Therefore it is possible to write the total energy of the system in terms of all the functionals of the charge density, which are

- Ion-electron potential energy
- Ion-Ion potential energy

- electron-electron energy
- kinetic energy
- exchange correlation energy

In density functional theory, the external potential and the number of electrons completely define the problem, and the energy is a functional of the external potential and the number of electron. In general a basis set is chosen that fit and the accuracy depend on the number of plane waves used. In solids, the periodic functions such as sin and cosine are used.

3.6.1 Exchange correlation

The challenge of DFT is that the exchange correlation functional is unknown. The functional consists of the exchange energy and correlation energy. The exchange correlation energy functional $E_{xc}[\rho]$ can have different forms. Many schemes have been developed for obtaining the approximate forms for the functions for the exchange correlation energy such as the Hartree schemes that define the exchange energy (Hartree, 1929). The main source of errors usually comes from the exchange correlation term and its approximation.

3.6.2 Local Density Approximation

Local density approximation (LDA) describes a non-homogeneous charge density in space and determines the quantum kinetic energy. The LDA functional is the functional of only the electron density.

$$E_{xc}[\rho] = \int \rho(r)\, \epsilon_{xc}(\rho) dr \tag{49}$$

where ρ is the electron density. ϵ_{xc} is the exchange correlation energy per particle of the electron charge density ρ. LDA is important in that it allows the construction of more

sophisticated approximations of the exchange correlation energy such as generalised gradient approximation.

3.6.3 General Gradient Approximation

General Gradient Approximation (GGA) functional is the functional correcting the LDA functional with the density gradient. It is a functional which depends on the charge density and its gradient. There are cases where LDA approach does not lead to sufficiently accurate results. In the case of iron it was found that the GGA correctly predicts the ferromagnetic ground state while the LDA incorrectly predicts its ground state to be non-magnetic (Jiang et. al, 2008). This often happens in cases that require very accurate estimations. Some common examples are when the total energy differences when calculating structural properties and binding energies are in the order of 0.1ev~10eV per bond whereas the total energy are in the order of keV. GGA improves the LDA by including the effects of electron density variation, which means that it is in essence a non-local approximation (Perdew et al., 1998). The GGA exchange correlation functional is

$$E_{xc}[\rho] = \int f\big(\rho(r)\big)\,\nabla\rho(r)dr \tag{50}$$

where *f* is a chosen function. The equation shows that the exchange correlation functional depends on both electron density and its gradient. The first principles calculations are performed within DFT by using GGA method with Perdew-Burke-Ernzerhof (PBE) scheme approximation (Perdew et al., 1998) for the exchange correlation potential. There are many other schemes such as the von Barth-Hedin scheme (Barth & Hedin, 1972) and Lee-Yang-Parr (Lee et al., 1988).

3.6.4 Hohemberg-Kohn Theorems

Hohemberg-Kohn Theorem is the fundamental theorem of quantum mechanics based on the electron density. The development of DFT came with the derivation of a set of one-electron equations where the electron density can be obtained. The theorem was introduced in order to produce the actual solvable set of equations. The Kohn-Sham method was later developed based on the theorem for electronic state calculations. The theorem shows that it is possible to express the ground state energy of a molecule as a function of the ground state electron density

$$E[\rho] = T[\rho] + V_{ee}[\rho] + \int \rho(r)v(r)dr + E_{xc}[\rho] \tag{51}$$

$E_{xc}[\rho]$ is the exchange correlation energy. The Kohn-Sham theory states that the ground state energy and all other ground state electronic properties are uniquely determined by the electron density.

The external potentials corresponding to the nuclear-electron potential are determined by the electron density. The energy variation principle is always established for any electron density.

3.7 Elements of Solid State Physiscs

3.7.1 Reciprocal Lattice

A reciprocal lattice (Bravais lattice) is the lattice in which the Fourier transform of the spatial wavefunction of the original lattice is represented. This space is known as the momentum space or the *K*-space. This is due to the relationship between the Pontryagin dual momentum and position. The reciprocal lattice of a reciprocal lattice is the original lattice (see Figure 42).

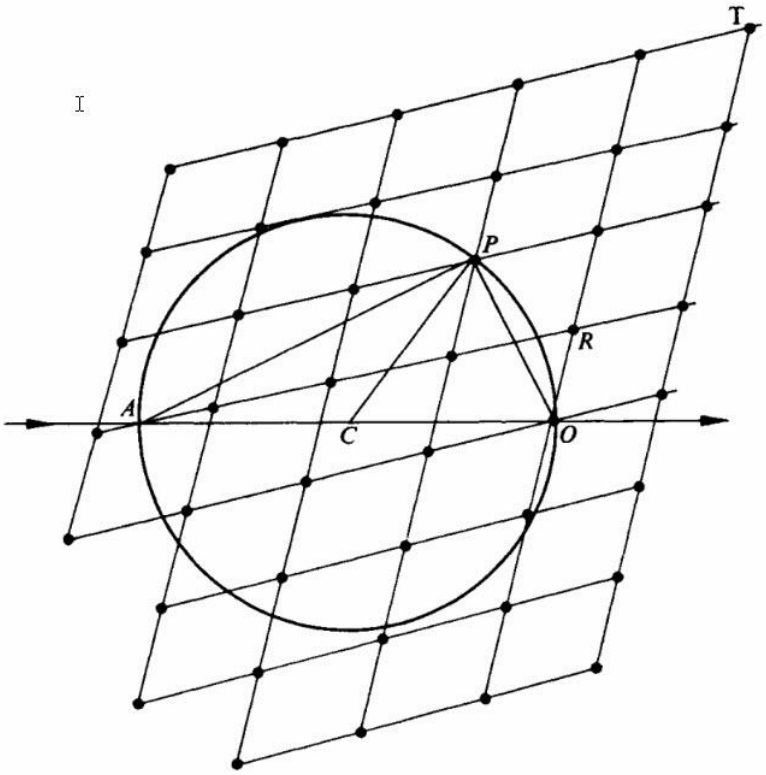

Figure 42. Reciprocal lattice centered at O on a sphere of radius 1/λ (Sands, 1993).

In physics there are two vector spaces, the real space and the momentum space. The real space is the position space or the coordinate space. It is the set of all position vectors $\vec{\boldsymbol{r}}$ of an object in a three-dimensional space. It defines a point in space where $\vec{\boldsymbol{r}}$ has the dimension of length. The momentum space or *K*-space is a set of all wavevectors K associated with particles. The momentum and the wavevectors are given by the De Broglie relation

$$P = \hbar k \tag{52}$$

where *P* is the momentum, $\hbar$ is the Planck constant. The equation suggests that the momentum and the wavevector are proportional. *K*-vector has dimensions of reciprocal length, so *k* is the frequency. In *K* space the dimension is of 1/length.

$$K = \frac{2\pi}{\lambda} \tag{53}$$

Working in *K* space makes it easier to relate *K* vectors to the crystal themselves.

3.7.2 Brillouin Zone Integrations

The first Brillouin zone is a uniquely defined primitive cell in the *K*-space. The boundaries of this cell are given by planes related to points on the reciprocal lattice. The same method applies as for the Wiegner-Seitz cell in the Bravais lattice. The first Brillouin zone is mapped by a continuous set of points throughout *K*-space. The occupied state of each *K*-point contributes to the electronic potential of the bulk material. The importance of the Brillouin zone stems from the Bloch wave description of waves in a periodic medium in which the solutions can be completely characterised by their behaviour in a single Brillouin zone. The first Brillouin zone is the volume including the same distance from one element of the lattice and its neighbours. Figure 43 shows a diagram of the Brillouin zone.

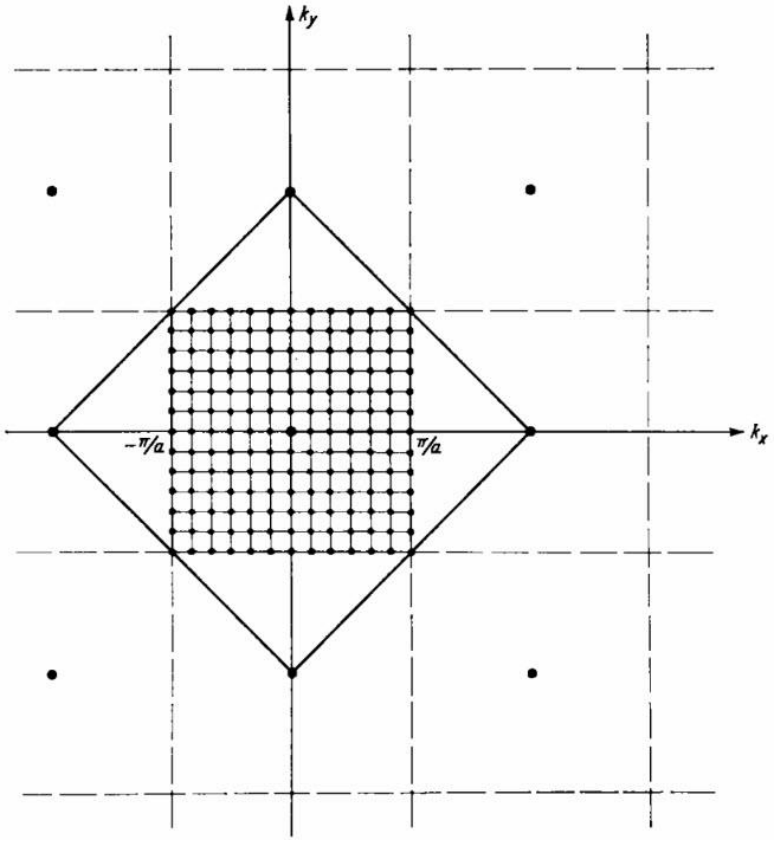

Figure 43. Schematic diagram of first and second Brillouin zone: The set of *K*-points is indicated inside the first Brillouin zone (Balkanski & Wallis, 2000).

A set of points in *K*-space can be reached from the origin without crossing the Bragg plane. One problem arises when an infinite number of electrons are used to express the infinite number of *K*-space vectors within the first Brillouin zone. In this case there are an infinite number of *K*-points in the Brillouin zone at which the wavefunction must be calculated. In order to solve this problem, an assumption is made that the electronic wavefunction for *K*-points grouped very close together are almost the same. Therefore it is possible to represent a small region of *K*-space with a single point. This procedure allows only a set of the *K*-points to be calculated and determine the total energy. Each *K*-point has an energy eigenvalue associated with it represented in the form of a band structure. The average energy eigenvalue is calculated at each *K*-point. This gives the energy associated with the electrons that occupy that region of *K*-space.

When there are large differences in energy in the Brillouin zone, it is necessary to converge the total energy of the structure to within a certain energy value. Methods have been developed for the purpose of obtaining accurate approximation to the electron potential. The two most common methods are "Chadi and Cohen"(Chadi & Cohen, 1973) and "Monkhorst and Pack." (Monkhorst & Pack, 1976) The more common of the two methods of defining the *K*-point grid is the Monkhorst-pack method.

The electronic wavefunction at each *K*-point is expressed in terms of a plane wave basis sets. The Fourier series is infinite, so in order to reduce the basis set to a finite set, a plane wave energy cutoff must be used. Each plane wave has a kinetic energy associated with it. The plane waves with the smaller kinetic energy are more important than the plane waves with a high kinetic energy because a wave function is expressed by an infinite number of K vectors in the first Brillouin zone, and only the plane waves with lower kinetic energies within the cutoff are considered.

3.7.3 Hellmann-Feynman Theorem

In quantum physics, a particle is described by a quantum state. The quantum state can be represented as a combination of basis sets. The Hellmann-Feynman Theorem is based on the idea that the exact wavefunction for the system is a solution of the Schrodinger equation with its energy depending on the internuclear distance parameter. If the parameter varies, the energy of the system also varies. This is represented by the equation

$$\frac{dE}{dP} = \langle \partial H / \partial P \rangle \tag{54}$$

The theorem relates the derivative of the total energy with respect to a parameter. A system may be characterised by a Hamiltonian which depends on this parameter such as the strength of the electric field or the interatomic distance. The exact wave function for the system is a solution of the Schrodinger equation. The Hellmann Feynman theorem ultimately describes how the energy of the system varies as the parameters are varied.

3.8 Ab Initio Calculations on Steels

The electronic configuration and the elastic properties of Fe-C martensite have been thoroughly investigated by material characterisation techniques but, to date there are no ab initio calculations reported in the literature.

The limitations in spatial resolution of experimental techniques do not allow analysis on an event in solids such as solid state diffusion and phase transition. The energy and thermodynamic properties of the martensite are difficult to measure, so theoretical method is a viable option.

The kinetic energy barrier for the diffusion of point defects in BCC iron supersaturated with carbon has been investigated by Kabir (Kabir, 2010) using ab initio calculations. They showed that a reduced self-diffusivity in BCC iron is a result of increasing carbon in

the solid solution. Potentials for semi-empirical methods have been determined for iron (Simonelli et al. 1992; Calder et al. 1993; Ackland et al. 1997).

The description of pure Fe can predict the configuration of self interstitials and the correct dislocation core structure (Ackland et al., 1997). For martensite, carbon is important due to the fact that it is trapped within the iron atoms. Many phenomena correlate with the interaction of carbon atoms and their behaviour (Tapasa et al., 2007). Different configurations of the Fe-C system were investigated with increased accuracy of the interatomic potentials (Jiang et al. 2003; Slane et al. 2004).

Jiang (Jiang et al. 2003) performed DFT calculations of carbon interstitial diffusion using GGA. It was found that the supercell of 128 Fe atoms and one C atom is sufficient for describing dilute concentrations of carbon in BCC Fe (Jiang et al. 2003). 32 Fe atoms and one C atom were found to be sufficient for the cell size (Jiang et al. 2003).

The interaction between dislocations and vacancies with atoms plays an important role in the decay of martensite due to surface contact fatigue. Molecular dynamics may be used to study the interaction of dislocations with each other. Kabir (Kabir, 2010) used Monte Carlo simulations to study the dislocation climb in heavily deformed BBC iron with a supersaturation of vacancies (Kabir, 2010). The simulation of heavily deformed BCC Fe was used to predict the dislocation climb from interactions between point and line defects (Kabir, 2010). Miyazaki (Miyazaki, 2004) used molecular dynamics simulation to study the ferrite behaviour around the crack tip under cyclic loading.

Chapter 4. Experimental Methods

4.1 Background

Several material characterisation methods have been used to investigate micropitting and the associated martensite decay. The micropitting damage on the surface of the gear was observed with reflected light microscopy. The microstructure of the disks and gears has been investigated by electron microscopy techniques such as Scanning Electron Microscopy (SEM), Electron Backscattered Diffraction (EBSD), and Transmission Electron Microscopy (TEM). Nanoindentation was used to determine the mechanical properties of the DER regions.

4.2 Metallographic samples

Material characterisation was carried out on specimens obtained from disks used to simulate the gear tooth contact as well as specimens cut from full size gears subjected to micropitting test. The micropitting tests have been carried out prior to this research (Oila, 2004) and for convenience details are given below.

4.2.1 Disk Samples

The material used to manufacture the disks was BS S156 steel. The steel specifications can be found in (BS S 156:1976). S156 is commonly used for heavy duty gears such as those used in gearboxes for helicopters and aircrafts. Table 3 shows the chemical composition of the S156 steel (wt %).

Table 3. Chemical composition of BS S156 steel in wt % (BS S156, 1976).

	Min	Max
C	0.14	0.18
Si	0.10	0.35
Mn	0.25	0.55
P	-	0.015
S	-	0.012
Cr	1.00	1.40
Mo	0.20	0.30
Ni	3.80	4.30

Table 4 shows the mechanical properties of the steel before carburising.

Table 4. Mechanical properties of BS S156 (BS S156, 1976).

Mechanical Properties		
	Minimum	Maximum
UTS, MPa	1320	1520
Elongation, %	1030	-
Reduction of Area, %	40	-
Hardness, MPa	-	939

The disk samples have been subjected to contact fatigue experiments in a rolling/sliding machine (Oila, 2004). In order to introduce a sliding component, the experiments were carried out with pair of disks with different diameters. This way the slide-to-roll ratio is controlled according to the equation:

$$S = \frac{2(u_1 - u_2)}{u_1 + u_2} = \frac{2(R_1 - R_2)}{R_1 + R_2} \tag{55}$$

where u_1 and u_2 are the linear velocities of the two surfaces, R_1 and R_2 are the radii of the two disks. The resulting slide to roll ratio after applying the radius of the two disks is 0.3.

The dimensions of the disks are shown in Figure 44.

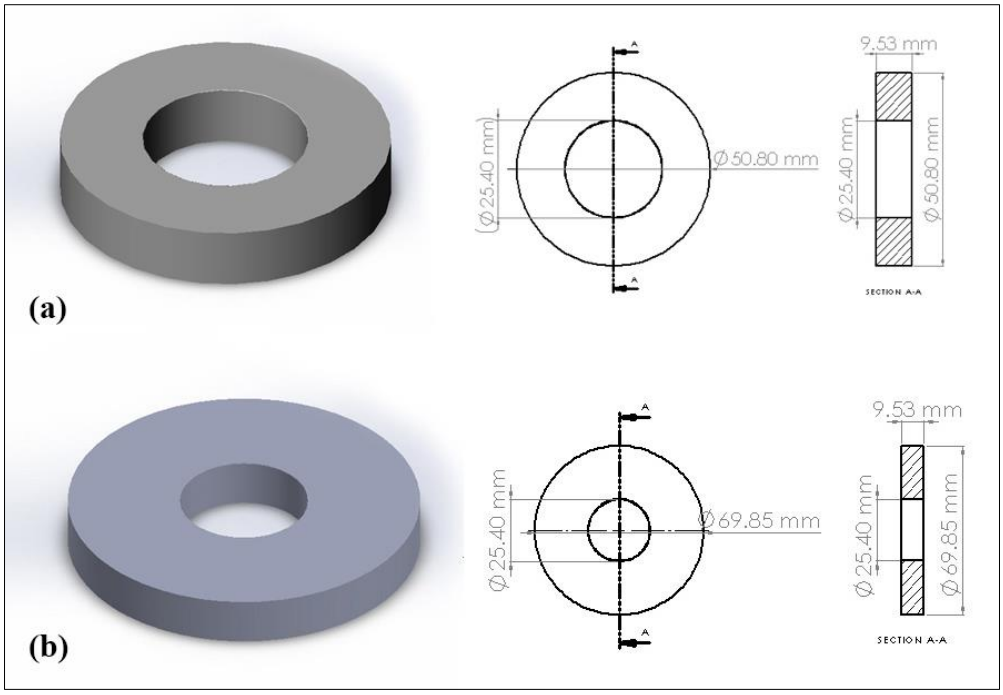

Figure 44. The geometry of the two disks subjected to the wear test.

After carburising followed by quenching and tempering, the disks were axially ground to achieve a surface lay perpendicular to the direction of motion as in gears (Figure 45).

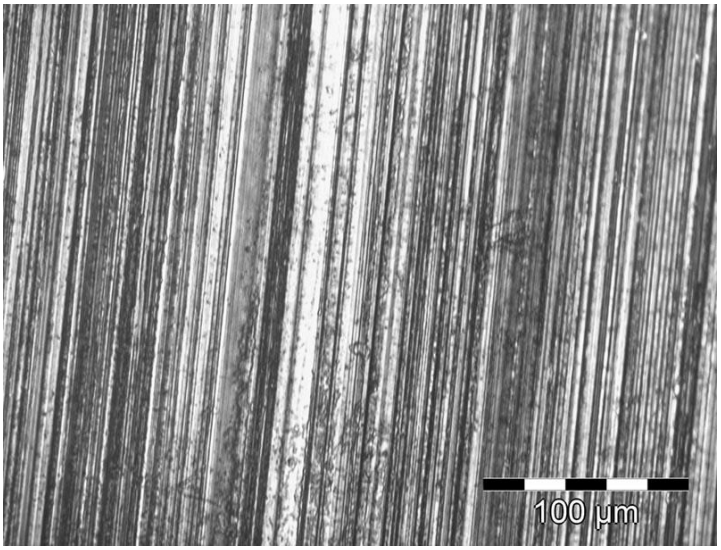

Figure 45. Grinding marks perpendicular to the direction of motion on the disk surface.

Grinding was carried out with a Cincinnaiti Tool and Cutter Grinder model No. 2 with C-BN grinding disk.

4.2.2 Gear Samples

The case hardened gears were from made from 16CrMn5 steel. After carburising at 900°C and quenching, they were tempered at 200°C. Table 5 shows the chemical composition of the steel.

Table 5. Chemical composition of 16CrMn5 steel (BS EN 1090-1).

	Min	Max
C	0.14	0.19
Si	-	0.40
Mn	1.00	1.30
P	-	0.035
S (16MnCr5)	-	0.035
S (16MnCrS5)	0.020	0.040
Cr	0.80	1.10

The two samples investigated were identified as B29 and AC9P (Oila, 2010)

Gear Code	Temperature	Load, MPa	Speed, rpm
B29	50	1550	3000
AC9P	90	1550	3000

4.3 Surface treatments, coatings and lubrication

All tests were performed under elastohydrodynamic lubricating conditions. Two types of lubricants were used during the wear test: SAE 85W-140 and FRL-46 (Formost, 2005).

SAE 85W-140 is a heavy duty mineral based multigrade gear oil designed to reduce thermal stress under high loads and high stress in order to improve service life. The lubricant properties are given in Table 6.

Table 6. SAE 85W-140 oil and FRL-46 oil properties (SubsTech).

Property	SAE 85W-140	FRL-46
Density	0.901 g/cm^3 at 15.6°C	0.884 g/cm^3 at 15.6°C
Dynamic viscosity at 60 °C	138 cP	54 cP
Flash point	300°C	225°C
Pour point	-12°C	-55°C

Both lubricants have been used as received and also after adding WS_2 nanoparticles (Foremost, 2005). 5%WS_2 were added to 85W140 and 3% WS_2 to FRL46. Several pairs of disks which were lubricated with 85W140 oil were subjected to cryogenic treatments (Frozen Solid, UK) and coated with Balinit (Oerlikon Balzers, UK) or CrN before the contact fatigue test. The chart describing the various treatments is presented in Figure 46.

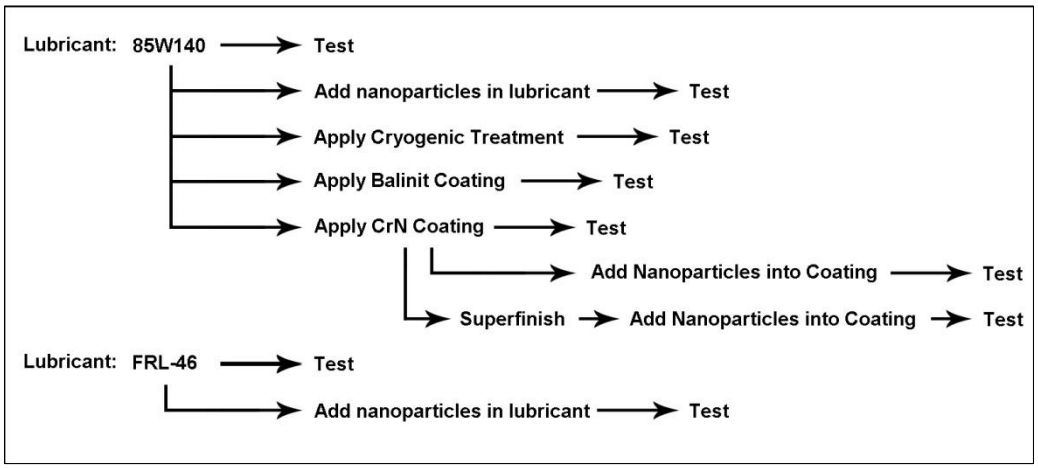

Figure 46. Chart showing the lubricants used and the specific treatment/coating applied in each case. 5%WS_2 nanoparticles added to 85W140 and 3% WS_2 nanoparticles added to FRL-46.

The Balinit C coating consist of WC/C (a-C:H:W) with a hardness of 1500 HV. It was deposited by Oerlikon Balzers, UK. The technique used to deposit the coating was PVD (Oerilikon, UK). The deposition temperature is around 160 °C. The coating thickness is around 2-5 µm.

CrN coatings were deposited by IonBond UK based in Consett, Durham. The technique used to deposit the coating was arc evaporative PVD. The deposition temperature is 250 °C. The film thickness is around 2-5 µm.

The cryogenic treatments (Oila et al., 2014) were performed by Frozen Solid, UK**.** The treatment was done in a stainless steel chamber at -185 °C using liquid nitrogen. The samples were slowly cooled until their core reached -185 °C, then 'soaked' for 2 hours and then heated slowly to room temperature. After the sample reached the room temperature, they were tempered for 2 hours at 180 °C.

4.4 Micropitting test

A complete description of the micropitting tests can be found in literature (Oila et al., 2004, Oila, 2005). The wear test operating conditions are shown in Table 7.The two disks have different outer diameters. This difference is necessary to achieve a slide to roll ratio of 0.3. Load was applied once the disk rotation reached the desired speed. The tests were stopped after 500,000 cycles (Oila et al., 2004, Oila, 2005).

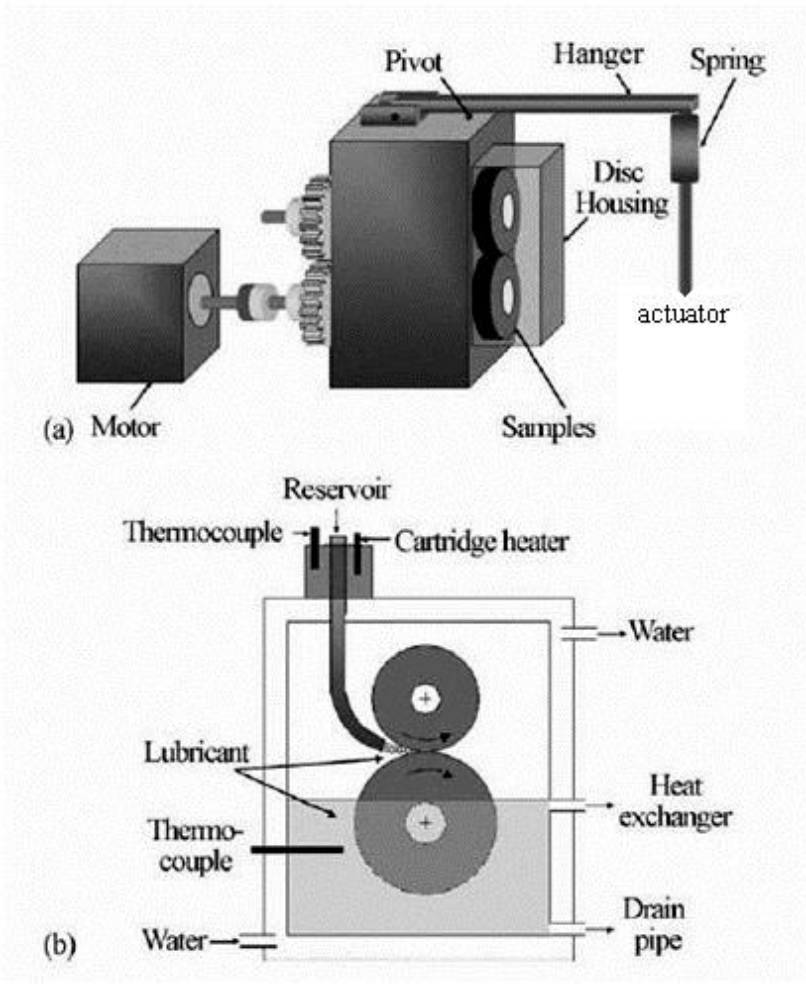

Figure 47. The disk machine and the configuration of the disk housing (Oila, 2004).

The wear test operating conditions are shown in Table 7. Wear test conditions.

Table 7. Wear test conditions (Formost, 2005).

Contact pressure (GPa)	1.5
Temperature (˚C)	60
Speed (rpm)	1200
Slide-to-roll ratio	0.3
Total number of cycles	500,000

4.5 Sample preparation

Surface investigations and microstructural analyses were carried out after sample preparation. The surface of the samples was investigated using reflected light microscopy.

Figure 48 shows a schematic of how the samples were cut from the disk. Because martensite decay occurs in regions near the surface of the disks, the area of interest is below the curved surface. The samples were cut into small sections for detailed analysis.

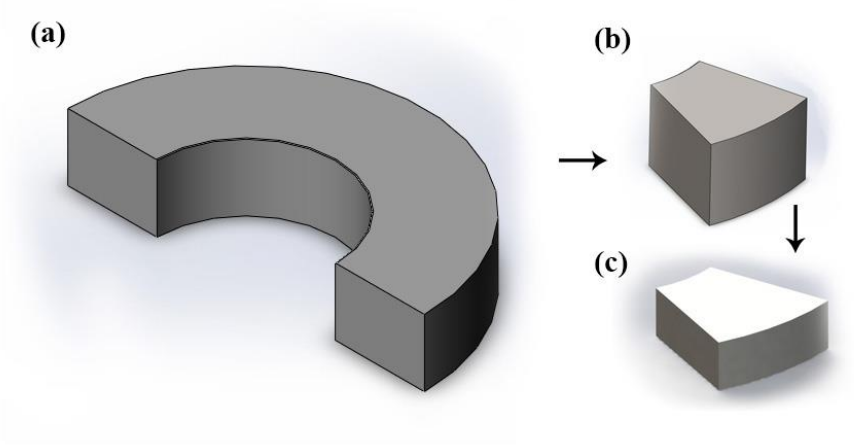

Figure 48. Location of metallographic investigations with the sectioning steps. (a)The disk was first cut in half. (b) A small volume of the half disk was cut out. (c) The small section is further cut in order to expose the microstructure under the surface for investigation.

The specimens were then mounted in resin using a hot mounting press Struers Prontopress. The pressure was set at 25 kPa with 6 minutes heating and 3 minutes cooling cycles. The resin used was a Buehler ProbeMet conductive copper moulding compound. The conductive copper moulding compound ensures that the charging effects during SEM investigation are eliminated.

The metallographic samples were ground using a Buehler Metaserv Motopol 8 grinder. The average wheel speed was set at 250 rpm. Grinding was carried out using a resin bonded diamond grinding disk MD Piano^{TM} 220 from Struers (comparable to the SiC

paper with 220 grit size). During grinding the samples were lubricated by streams of cold water.

The samples were than polished using abrasive disks. The first stage of polishing was performed with the Struers MD Allegro polishing disk with the sample lubricated by streams of cold water. The continuous flow of water is necessary to ensure cooling and removal of debris during polishing. The specimens were then polished using satin woven acetate cloths impregnated with diamond suspension of 9 µm for 3 minutes. After the 9 µm polish the samples were polished using 3 µm diamond suspension for 3 minutes and then 1 µm diamond suspension for 3 minutes. The specimens intended to be used for electron backscatter diffraction (EBSD) investigation were further polished using colloidal silica with particle size 500nm. The colloidal silica polish was performed manually for 10 minutes with minimum pressure applied.

In order to reveal the micro-structure and to produce optical contrast between microstructural constituents the samples were etched with 2% nital solution. Nital is a solution of nitric acid in methanol.

4.6 Surface Damage Assessment

Surface damage assessment was carried out by estimating the average micropitting area on the specimen surface from images acquired using reflected light microscopy. The micropitting area was estimated using Scion image software. Firstly, the image was processed in order to enhance the contrast. The image was then made binary with the "threshold" function to remove unnecessary pixels. Figure 49 shows an example of the image used for the calculation.

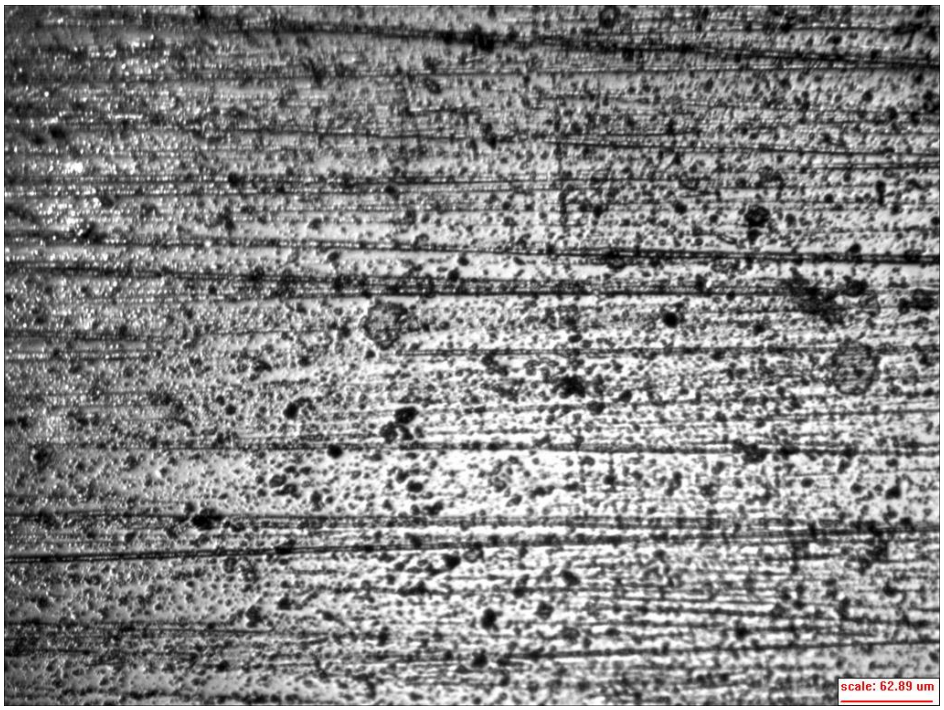

Figure 49. Surface image captured with light microscopy.

Figure 50. Analysed image with micropits counted. Number of micropits=220, mean area = 66.5 μm^2, minimum area = 0.4 μm^2, maximum area = 493 μm^2, total micropitting area = 5%.

Using this image the number of particles (micropits) can be counted and the total area of micropitting damage can be computed.

4.7 Microstrutural Analysis

The microstructure was observed in the cross section of the disk just below the surface where martensite decay has been reported to occur. The microstructure was investigated using reflected light microscopy (RLM), scanning electron microscopy (SEM), electron backscatter diffraction (EBSD), and transmission electron microscopy (TEM). Nanoindentation tests were also performed in order to assess the mechanical properties in these regions.

4.7.1 Reflected Light Microscopy

Figure 51 shows the Olympus BH2-UMA microscope that was used in the light microscopy investigation. The light microscope is equipped with the 10x eyepiece and 5x, 10x, 20x, 50x, and 100x objective lenses. The magnification is respectively 50x, 100x, 200x, 500x, and 1000x. The optical microscope has an illumination system that is necessary for obtaining the best resolution and contrast of the specimen. An adjustable objective lens turret allows for changing the magnification. The fine and coarse knob adjust the sample distance relative to the objective lens. Images were taken with a camera attached to the microscope and a JVC video monitor model TM-210PS connected to the computer via Analysis software.

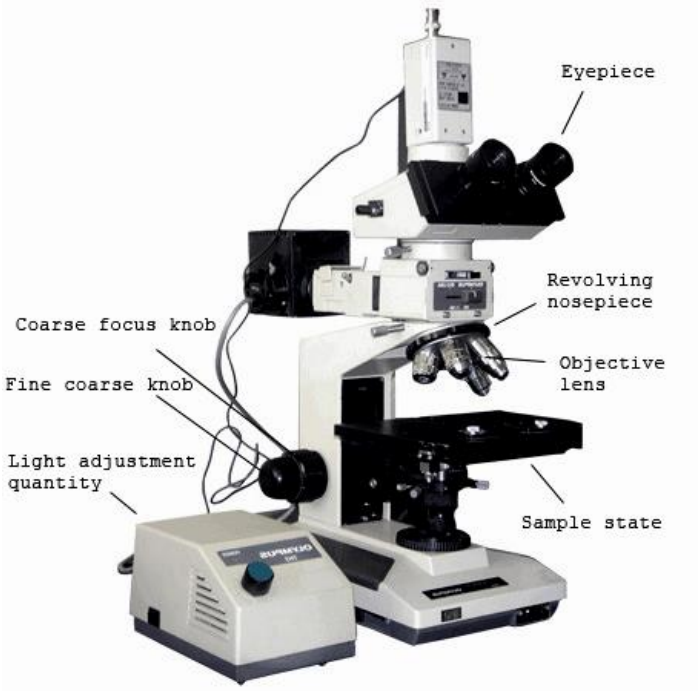

Figure 51. Olympus BH2-UMA microscope.

4.7.2 Scanning Electron Microscopy

Scanning electron microscopy (SEM) uses an electron beam to interact with the sample. It provides a direct display image of a specimen surface at a magnification ranging from 20 to 20,000 times. The basis for electron microscopy can be understood by looking at the de Broglie Equation

$$\lambda = \frac{h}{\sqrt{2m_0 eU}} \tag{56}$$

where e is the charge of an electron, U is the accelerating voltage, m_o is the mass of electron, h is the Planck's constant. If the accelerating voltage is 20keV, the wavelength can be found to be around 0.008 nm. Comparing this value with the visible light range of 380~780 nm, there is a significant difference between the wavelength of electron and visible light.

The type of SEM used in this project is Philips XL30-ESEM-FEG. It can operate in both, low and high vacuum modes. The resolution ranges from 2nm at 30kV and 5nm at 1kV. It is capable of imaging wet samples. The stage consists of a 4-axis motorised unit with a manual tilting system capable to tilt the stage at angles between -15 to 75 degree. The hkl EBSD system is attached on the SEM.

The main components of an SEM include a vacuum pump, a sample chamber, an electron gun, and detectors. There are also several user computer interfaces required to control the microscope. The source of electrons is an electron gun, which consists of a cathode assembly and an anode. The cathode assembly is composed of a tungsten filament enclosed within a cylindrical shield (cathode cap) bearing a high negative potential with respect to the filament. The electron from the tip of the filament is accelerated to 50 to 100 KV before it reaches the anode which is grounded. After passing through the anode, much of the beam is lost against walls and apertures except for a narrow pencil which passes through a condenser aperture. The magnetic lens acts as a lens on an electron beam which travels along its axis. The lens must be axially symmetric both magnetically and geometrically. Focusing the image is done by changing the lens currents.

Once the electron beam hits the sample, different signals are emitted in different directions and the various detectors can pick up the signals and send the information to the computer. Figure 52 shows the schematic of a scanning electron microscope.

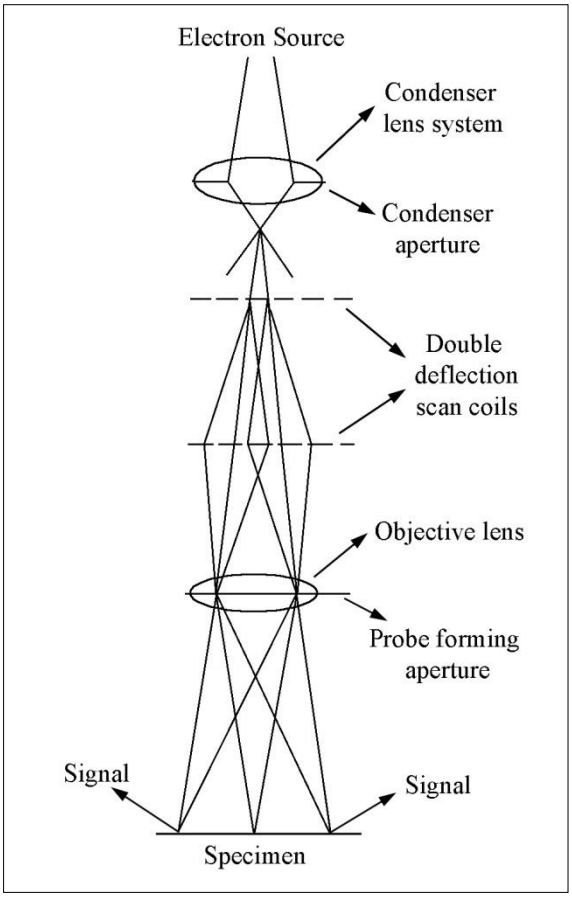

Figure 52. Schematic showing components of a SEM.

The ultra high vacuum keeps the electron beam from scattering due to the atmosphere so that the resolution can be increased. There are several types of cathodes used in SEM as shown in Table 8.

Table 8. SEM cathode comparison

SEM Cathode Comparison				
	Tungston FIlament	**LaB6**	**Schottky**	**Field Emission**
Apparent Source Size	100 micrometer	5 micrometer	<100 Angstroms	<100 Angstroms
Brightness	1 A/cm^2 steradian	20-50 A/cm^2 steradian	100-500 A/cm^2 steradian	100-1000 A/cm^2 steradian
Vacuum Required	10^{-5} Torr	10^{-6} Torr	10^{-8} Torr	10^{-9} Torr

The wavelength of an electron is much smaller compared to the wavelength of light, so the resolution in SEM is much better than light microscopy. For the electron microscope, the interaction volume determines the image resolution. The interaction volume depends

on the electron beam diameter and the type of electron beam gun. Figure 53 shows a diagram of interaction volume of an electron beam hitting the surface of a specimen.

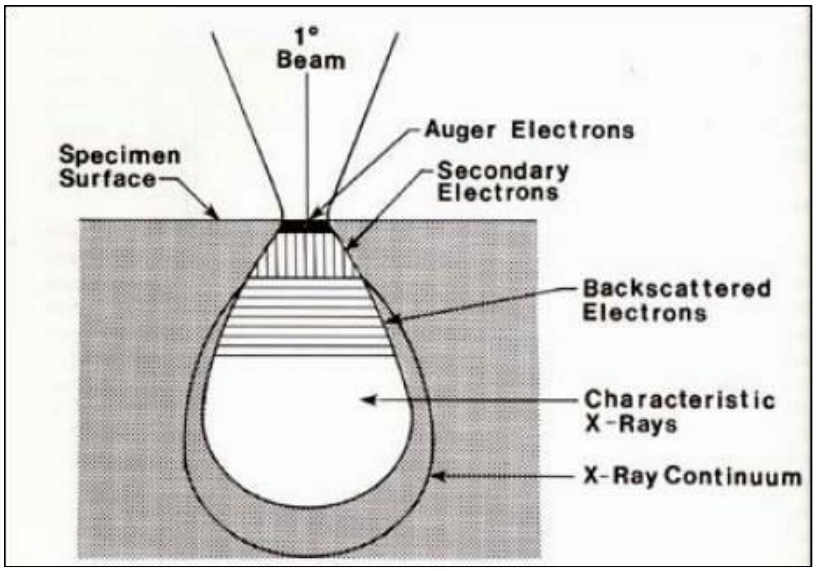

Figure 53. Interaction volume of an electron beam hitting a specimen surface (Balonek, 2013).

The interaction volume depends on the diameter of primary beam, the energy associated with the primary beam and atomic weight of the specimen. Different signals can be collected with the SEM. The most common is the secondary electrons signal generated from the sample surface. The low energy secondary electrons usually have energy less than 50eV. Most of the electrons emitted from the sample will be secondary electrons and they are often used to obtain surface topographic information. The back scattered electrons originate from the collision between the electron beam and the atoms from the deeper part of the interaction volume. When the electrons penetrate into the material, different signals can be detected depending on the depth. High energy electrons can escape from deeper areas in the sample. The back scattering electrons have higher energy than both the secondary electrons and Auger electrons.

The Auger electrons originate from the electron beam knocking out an electron from the inner shell of an atom, which causes the electron from the outer shell to drop down to the energy level of the vacancy. During the process, it kicks out another electron from the

outer shell. The electron at the outer shell has a higher energy, therefore when it drops down to the inner shell it release a particular energy.

The X-ray emission works similar to the Auger electron emission. When an electron is removed from the inner shell, a vacancy is created. The electron from the higher energy level drops down to the vacancy for keeping stability. During the process, the electron releases an X-ray which can be measured by an X-ray detector. Energy Dispersive X-ray microanalysis (EDX) is a method that uses the X-rays emitted by the specimen in order to obtain elemental information. Cathodoluminescence is the light emission after the sample is excited by the eletron beam. It occurs when an electron and a hole recombine, and photons are released.

4.7.3 Electron Backscatter Diffraction (EBSD)

Electron Backscatter Diffraction (EBSD) is also known as backscatter Kikuchi diffraction. Kikuchi patterns were first observed in the transmission electron microscope (TEM). Recently the technique was also implemented in SEM.

In 1982, Kikuchi (Kikuchi & Nishikawa, 1928) discovered that the electron diffraction through a mica crystal can create a diffraction pattern, so called “Kikuchi lines”. He found that the spreading of the incident beam correlates to the source of the specimen. The source of the electron scattering is the lattice plane. Due to the development of computer technology today, the indexing of EBSD patterns is automated in the SEM. One application of EBSD is phase identification. Phase identification can be used to discriminate between different phases in a sample and is useful for separating different phases in the specimen.

In EBSD the sample is tilted at 70° in order to increase the contrast and the diffraction pattern collected. Once the electron beam hits the sample, the diffracted electrons hit a phosphor screen. A camera is positioned to capture the diffraction pattern. Figure 54 shows the origins of the Kikuchi lines in EBSD.

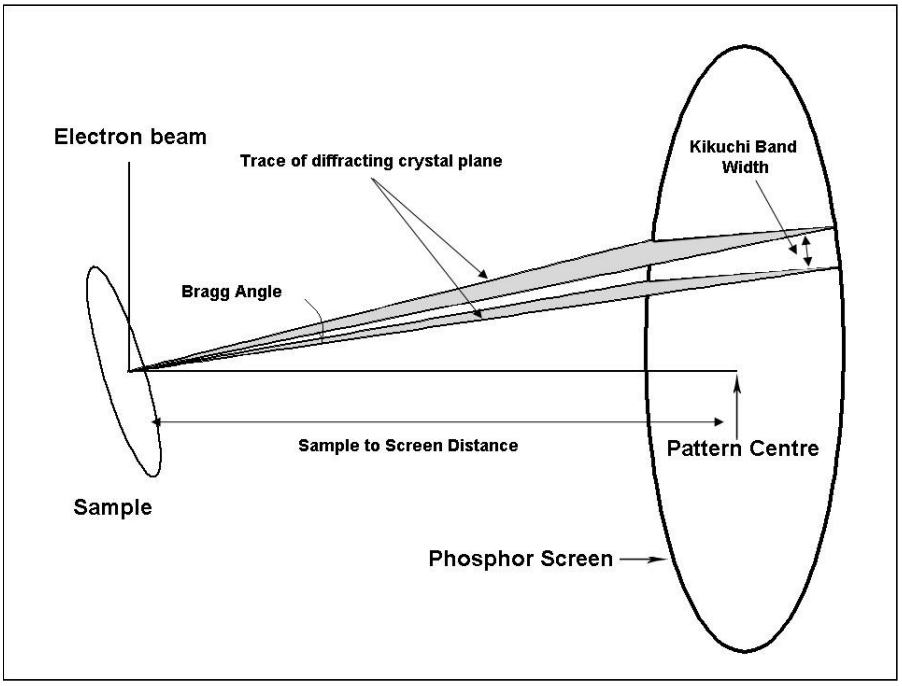

Figure 54. EBSD geometry with Kikuchi lines.

The equation used to describe the reflected beam is the Bragg's Law

$$n\lambda = 2d\,sin\,\theta \tag{57}$$

where λ is the wavelength, n is an integer, d is the interplanar spacing of scattering atoms, θ is the Bragg angle. When an electron beam enters the sample, it is scattered in all directions. The inelastic scattering between Bragg reflections is influenced by the crystal periodicity and results in Kikuchi lines and bands. The electrons are diffracted to form a large angle cone corresponding to each lattice plane of the material. The cone forms the Kikuchi lines on the phosphor screen. A pair of Kikuchi lines forms a Kikuchi band. The thickness of a Kikuchi band corresponds to a specific plane. The diffraction pattern is related to the crystal structure of the sample, so the Kikuchi bands can be used to

calculate the orientation of the crystal. The intersection of the Kikuchi bands corresponds to the zone axis which gives information about the angular relationship within the crystal. An example of a Kikuchi pattern is shown in Figure 55.

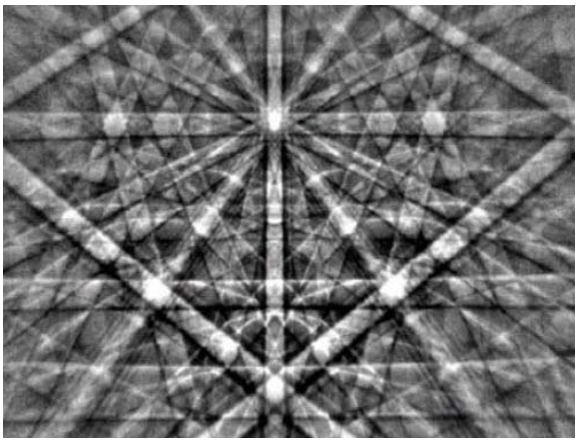

Figure 55. EBSD Kikuchi pattern on silicon. The crystal direction corresponds to the Kikuchi band intersection (Winkelmann & Nolze, 2010).

4.7.4 Nanoindentation

Nanoindentation is a test used to measure the mechanical properties of a small volume of material. The martensite decay observed in gears occurs in regions that are usually less than 30x30 µm area, therefore only nanoindentation can be used to accurately measure its mechanical properties. For this research, the goal of the nanoindentation testing was to obtain information on the elastic modulus and hardness of the phases present in the microstructure. The nanoindenter used was the Hysitron Ti-900 Triboindenter. Figure 56 shows a photograph of the indenter.

Figure 56. Hysitron Ti-900 Triboindenter (www.coherent.com.au).

The Triboindenter provides automated testing and imaging before and after the tests. The software controlling the Triboindenter is Triboscan 5.0. During imaging, the force between the indenter tip and sample is measured by the capacitance sensor of the transducer. The piezo can be adjusted by the voltage and the surface topography is monitored.

Micro-indentation is useful for investigating the hardness of the material. However, it is not suitable for investigation on specific regions of microstructure because of the large indenter size. A large indenter size covers a larger area and also results in deeper penetration depths. The microindenter obtains the data from multiple grains, whereas a nanoindenter has the capability to obtain information from a single grain.

The type of indenter used in this work was the pyramidal Berkovich tip (Figure 57). The Berkovich indenter has the advantage of having three sides joined to a single area

compared to the four sided Vickers indenter. The Berkovich indenter has a face angle of 65.3°.

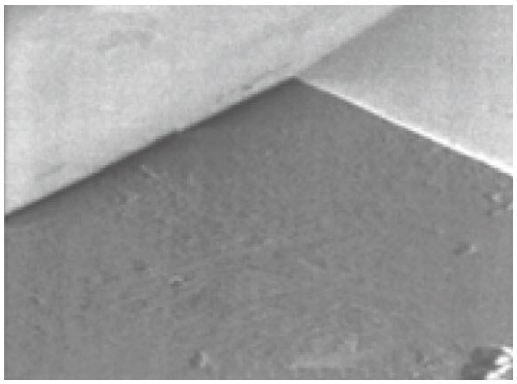

Figure 57. Berkovich indenter tip (Cripps, 2011).

Oliver and Pharr (Oliver & Pharr, 1992) introduced a method for measuring hardness and elastic modulus from nanoindentation data. The relationship between hardness and maximum indentation load is given by equation 58.

$$H = \frac{P_{\max}}{A} \tag{58}$$

H is the hardness, *P* is the maximum indentation load, *A* is the tip contact area.

The contact area for a Berkovich indenter is described by the equation

$$A = 3\sqrt{3}h_p^2 \tan^2 65.3 = 24.5h_p^2 \tag{59}$$

Where h_p is the depth penetration.

For a Berkovich indenter, the contact area is evaluated using a face angle of 65.27 [31]. The hardness equation becomes

$$H = \frac{P_{\max}}{24.5h_p^2} \tag{60}$$

P_{max} is the maximum indentation load. The equation for reduced elastic modulus is given in equation 3.4.

$$E = \left(\frac{dP}{dh}\right)\left(\frac{1}{2h_p\beta_g}\right)\sqrt{\frac{\pi}{24.5}} \quad (61)$$

dP/dh is the stiffness, β_g is the geometry correction factor (1.034 for Berkovich indenter).

The load-displacement curve gives information on the load profile and the displacement of the indenter. An example of the load displacement curve is shown in Figure 58.

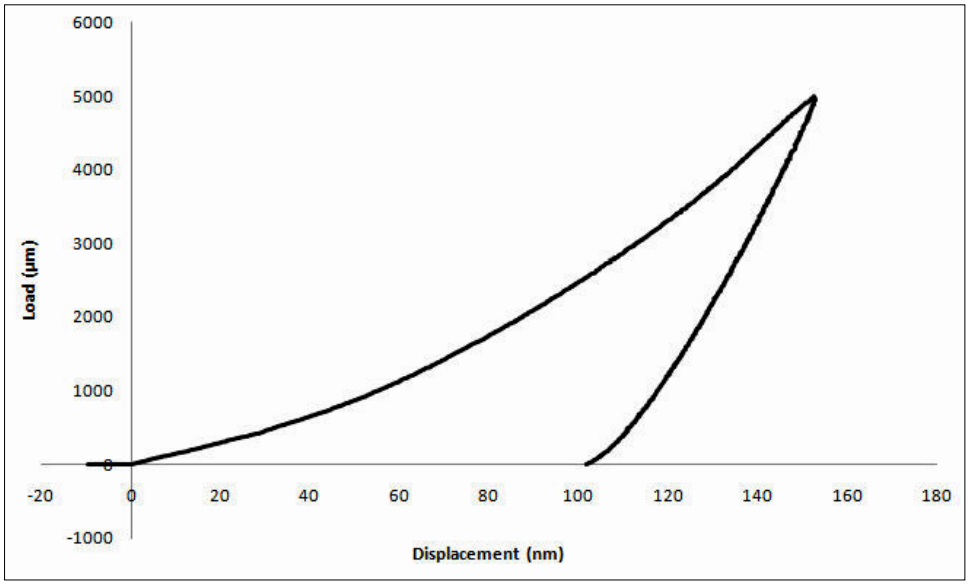

Figure 58. Load-displacement curve for the 85W140 lubricated sample.

The nanoindenter offers in-situ Atomic Force Microscopy (AFM). The AFM measures the voltage needed to move the piezoelectric sensor up and down and a topographic image of the surface can be obtained.

The separation between each indentation was set to 5 µm so the plastic deformation from each indent does not interfere with other neighbouring indentation. Maximum indentation load was set at 5 µN. The scan area around the indentation after each indentation was 3 × 3 µm. The loading/unloading cycle is represented in Figure 59.

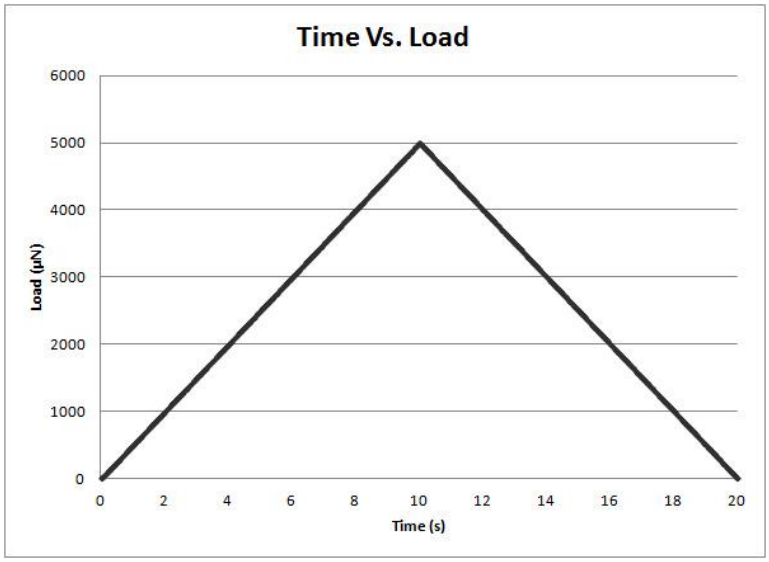

Figure 59. Loading/Unloading Cycle versus Time.

4.7.5 Transmission Electron Microscopy

Transmission electron microscopy (TEM) uses a beam of electrons which passes through a very thin specimen. The transmitted beam can be analysed to obtain structural information about the specimen. The TEM was done using the JEM-2100 Plus Transmission Electron Microscope. It contains three condenser lenses with accelerated voltage up to 200 kV with spot size up to 20 nm . The electron beam is generated by an electron gun which consists of a cathode assembly composed of a tungsten filament similar to SEM. The acceleration voltage can range between 100-500 kV. A series of lenses focuses the electron beam on a thin electron transparent sample. The final image can be viewed on a fluorescence screen and recorded by allowing the electrons to strike the photographic plate.

The electrons passing through a material may lose energy as a result of absorption and may scatter with a change in direction by electron fields within the atoms of the material. The scattering of the electrons is a function of atomic number. The beam leaving the specimen is spread out over a broad angle depending on the specimen thickness. If the

diffracted beam falls outside of the aperture of the objective lens, they do not contribute to the final image.

The electrons interact with atoms by elastic and inelastic scattering so the sample must be very thin, generally 100 nm. Special preparation techniques are needed such as electropolishing (PIP) and ion beam etching (FIB). Because the sample is extremely thin, it is prone to bending and breaking, so careful consideration needs to be taken when handling the specimen.

TEM sample preparation

To prepare a TEM sample from a tooth of carburised gear, the following procedure was used:

- A representative sample containing Dark Etching Region was cut from the gear tooth.
- Grinding and polishing the sample to a thickness of 50-80 µm.
- Making disks with 3 mm diameter as standard size to fit into the TEM sample holder.
- Further thinning of the sample using the PIP method or FIB method.
- Checking the thin edges of the prepared thin foil on a TEM for sufficiently large areas of good electron transparency.

The steel was cut into thin slices with thickness in the order of 250 to 300 µm. A low speed diamond saw was used for the cutting. The thinning process was performed by grinding using silicon carbide grinding paper grade P240 (average particle size 58.5 µm) until the sample thickness was in the order of 120 µm to 150 µm. The surface deformation layer produced during cutting is removed by this step. At this stage there are some scratches on the surface, which are minimised by polishing on finer grit. In the next step the sample was polished using silicon carbide papers P600 (average particle size 25.8

μm) for 60 seconds, P800 (average particle size 21.8 μm) for 30 seconds, and P1200 (average particle size 15.3 μm) for 30 seconds. This step is repeated until the sample has a thickness between 50 to 80 μm.

Once the sample is between 50 to 80 μm, diamond lapping film is used for further polishing. The diamond abrasive film used was from South Bay Technology with 8" diameter. The polishing starts with the film containing 15 μm size particles for 2 minutes. Then 8 μm, 3 μm, 1 μm, and 0.5 μm film for 2 minutes on each film.

The sample is then put in acetone in an ultrasonic bath for 3 minutes. The sample slice is punched using a disk punch that matches the specifications of the TEM specimen holder (3 mm diameter). The sample is then put into a capsule for storage.

The electropolishing technique

The electropolishing technique was applied using a Gatan Model 691 precision ion polishing system. A rotation is applied to the disk at 3 rpm and ion gun beam energy was set at 3 keV so the total beam energy was 6 keV.

Figure 60. The Gatan Model 691 precision ion polishing system.

TEM specimens preparation by focused ion beam milling

The focused ion beam (FIB) instrument consists of a vacuum system, a liquid metal ion source, a sample chamber, detectors, gas injection system (GIS), and a computer. The vacuum used was around $1x10^{-8}$ torr in order to eliminate contamination. The GIS was used for ion beam assisted vapour deposition with the injection capillary positioned 100 microns above the sample surface. The ion beam milling takes place from the physical sputtering of the target area. Sputtering occurs as a series of elastic collisions from the momentum energy of the ion to the region. The surface atom is ejected as a sputtered particle when the ion energy exceeds the surface binding energy of the material.

Focused ion beam is an ideal technique due to the amount of control it offers. The FIB uses a gallium ion beam originated from a liquid metal ion source. The beam is accelerated over the specimen surface so that site specific milling can be performed.

The steps used for sample preparation using FIB were as follows:

- Identification of the region of interest.
- Platinum deposition using GIS with dimension: length 15 µm, breadth 2 µm, and thickness 0.3~1 µm. The thickness is controlled by deposition time.
- Ion milling to make a lamella by milling adjacent to the platinum coated region.
- Lift the lamella using the omniprobe.
- Transferring the lamella onto the TEM copper grid.
- Weld with carbon or platinum metal.
- Clean specimen surface with a low energy gallium beam at 5 kV or lower. Tilt the specimen by 1 degree with respect to the beam so the final thickness is in the order of 100 nm.

Chapter 5. Computational Experiments

5.1 Background

In this project, the structural and elastic properties of martensite were calculated from first principles including the lattice parameters and the single crystal elastic constants. To date there is no published work for this type of calculations for martensite.

The calculations were performed using the Quantum Espresso package (Giannozzi et al., 2009). Quantum Espresso is an integrated suite of open source codes developed for performing electronic structure calculation and materials modelling. The structure of martensite was created using Wyckoff positions of C and Fe atoms with iron atoms at Fe 4*a* (0,0,0) and carbon atoms located at Wyckoff positions C 2*b* (0, 0, 0.5). Martensite has a total of 27 Fe atoms and one carbon atom in the conventional cell, which corresponds to a carbon content of 0.8 wt % carbon in the 3x3x3 super cell.

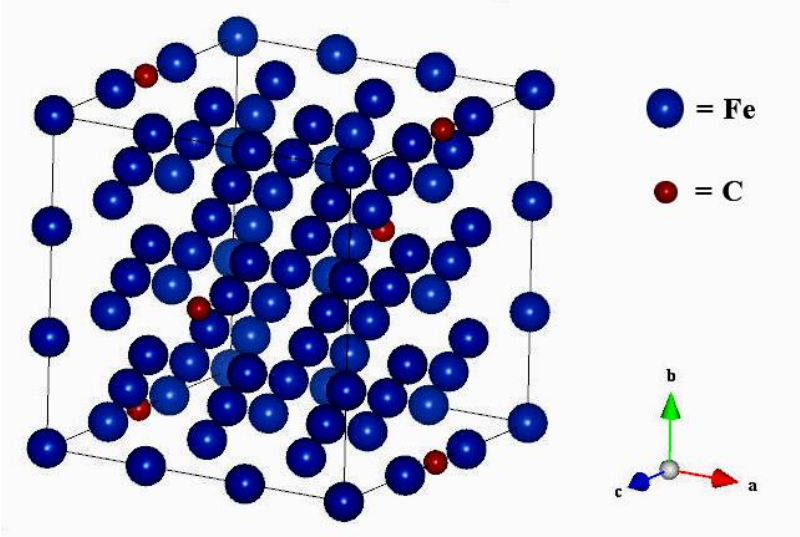

Figure 61. Supercell of martensite with 0.8 wt % cabon.

This carbon content corresponds to that normally found in the tempered martensite present in the case of carburised gears. The weight percentages can be calculated from the atomic percentages and atomic weight through the equation

$$wt\ \%\ of\ A = X_A \times \frac{m_A}{X_A \times m_A + X_B \times m_B} \times 100 \quad (62)$$

Where *m* is atomic weight. *X* is the number of atoms. A supercell is a repeating unit cell of the crystal that contains many primitive cells. The use of supercell is necessary because it represents the correct amount of carbon content in the martensite.

5.2 The K-Points Grid

Once the atomic positions have been designated, the *K*-points mesh was constructed in the Brillouin zone following the Monkhorst Pack scheme (Monkhorst & Pack, 1976). The tetragonal crystal structure of martensite belongs to the space group I4/MMM. The initial martensite lattice parameter are set to a=8.5692 Å, b=8.5692 Å, and c=8.863 Å (Wechsler et al., 1954). The Brillouin zone integrations were performed using the Marzari-Vanderbilt method (Marzari & Vanderbilt, 1997). The kinetic energy cutoff and charge density cutoff were set to arbitrary values in order to find the minimum number of *K*-points required for the simulation. Increasing the number of *K*-points increases the computational time significantly, therefore it is important to find a suitable number of *K*-points.

The initial kinetic energy cutoff was set to 50 Ry (680 eV). The initial charge density cutoff was set to 400 Ry (5442 eV). A total number of 10 simulations were run in this step with *K*-points meshes from 1x1x1 to 10x10x10, which proved sufficient to find a suitable *K*-points mesh with an accuracy of within 0.5 meV.

Table 9. *K*-points meshes and the corresponding number of *K*-points.

K-Points mesh	Number of *K*-Points
1x1x1	2
2x2x2	8
3x3x3	12
4x4x4	26
5x5x5	36
6x6x6	60
7x7x7	80
8x8x8	118
9x9x9	150
10x10x10	204

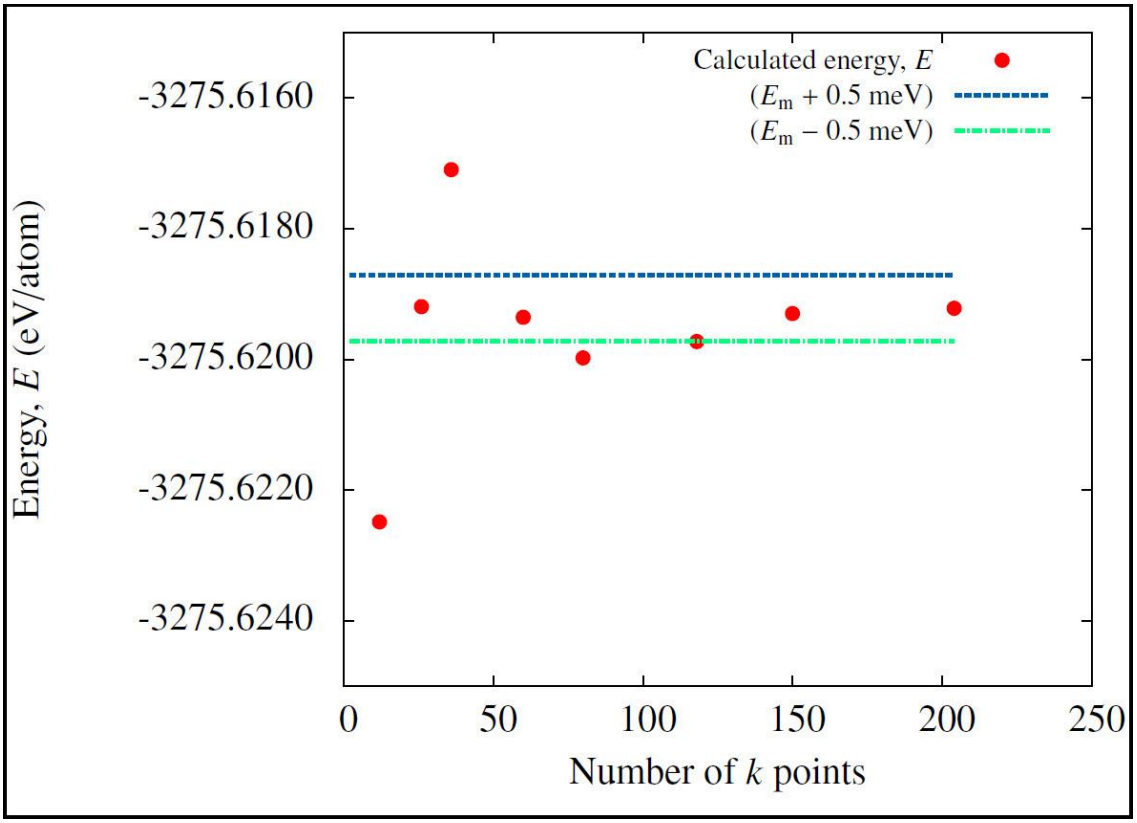

Figure 62. Total energy as a function of the number of *K*-points.

The graph in Figure 62 shows that a *K*-points mesh 8x8x8 which gives 118 *K*-points in the Brillouin zone is sufficient for an accuracy of 0.5 meV and was selected for the consequent calculations. An input file for this simulation is located in Appendix 1.

5.3 Estimation of the Kinetic Energy Cutoff

Once the suitable *K*-points mesh has been found, a total of nine simulations were performed in order to determine the plane wave kinetic energy cutoff and the charge density cutoff. The optimum cut-off energy was found by starting at 40 Ry (544 eV) and incrementally increasing the cutoff energy by 5 Ry (68 eV) until the change in consecutive energy per atom was less than 5 meV/atom. The initial charge density cutoff was set to 320 Ry (4354 eV).

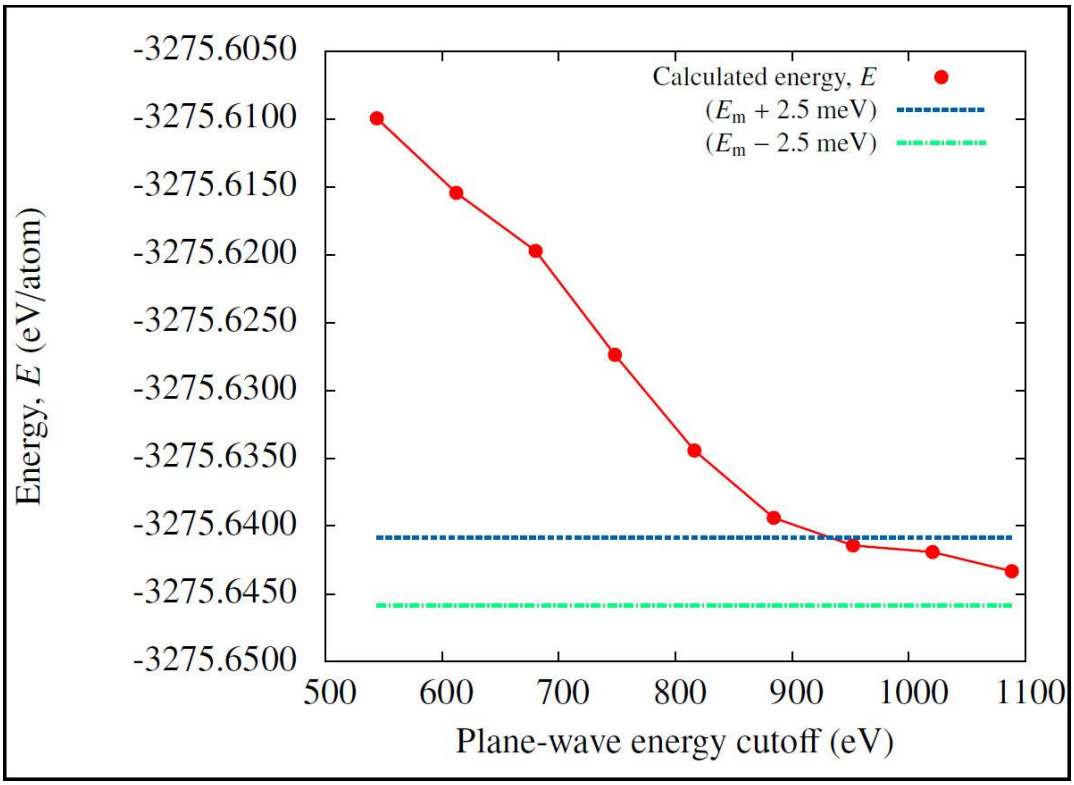

Figure 63. Energy convergence test for martensite.

After the convergence test, the plane wave kinetic energy cutoff of 70 Ry (952 eV) and charge density cutoff of 560 Ry (7619 eV) were found to be sufficient to converge the total energy to less than 5 meV/atom (Figure 63).

5.4 Geometry optimisation

Geometry optimisation was performed in order to find the optimised atomic configuration and the optimum lattice parameters by allowing the atomic coordinates to relax (for input file see Appendix 2). The total energy was computed as a function of the unit cell volume from the lattice parameter ratio b/a and c/a. Since the total energy varies significantly with small changes in the structure and the arrangement of atoms geometry optimisation is the step required to calculate the position of the atoms corresponding to the minimum total energy.

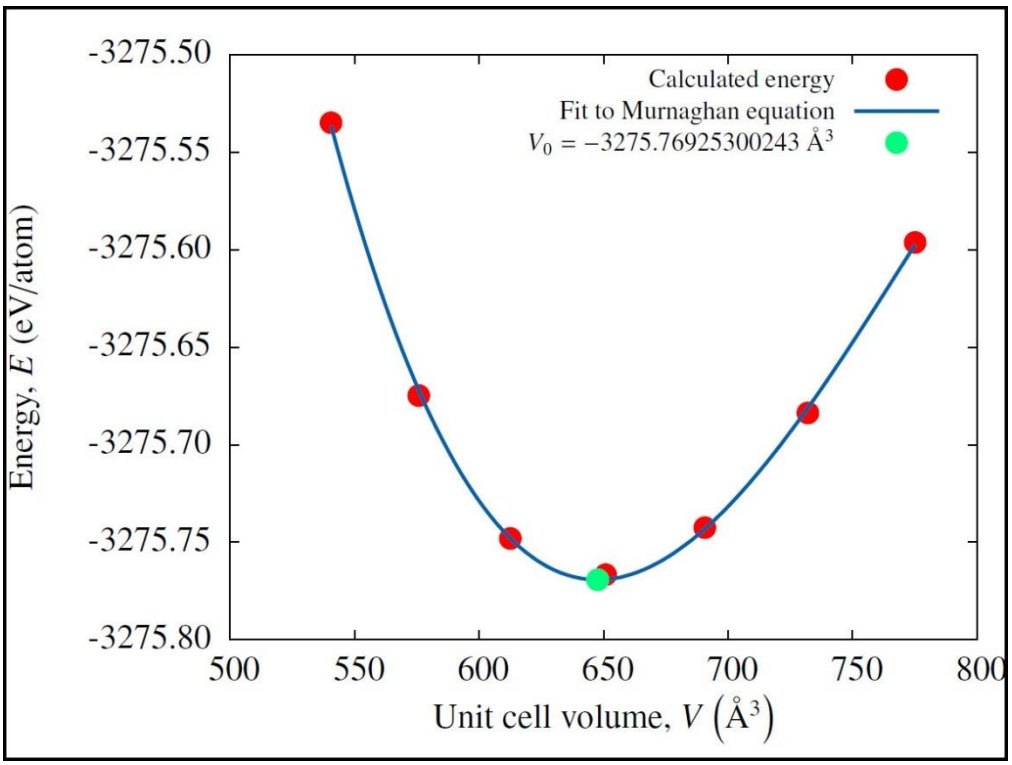

Figure 64. Calculated energy as a function of the martenstie unit cell volume.

In the first step of the geometry optimisation the martensite lattice was deformed by applying small strains of -0.06, -0.04, -0.02, 0.02, 0.04, and 0.06. A total of seven simulations were then performed with these new values for the lattice parameters in order to determine the minimum total energy. The calculated total energy was plotted versus the unit cell volume (Figure 64). The optimised unit cell volume corresponding to the minimum energy was found to be 647.605 $Å^3$.

Table 10. Optimised cell volume of martensite.

Cell dimension (c/a)	1.034285581		
Cell volume ($Å^3$)	647.60568962492		
a,b,c (Å)	8.555066607 Å	8.555066607 Å	8.848382035 Å

.

In the second step of the geometry optimisation the optimised unit cell volume was used to determine the optimised lattice parameters by using the relationship

$Volume = a * b * c$. The optimised lattice parameters were determined by curve fitting (see Figure 65). A total of eleven simulations were performed in this step.

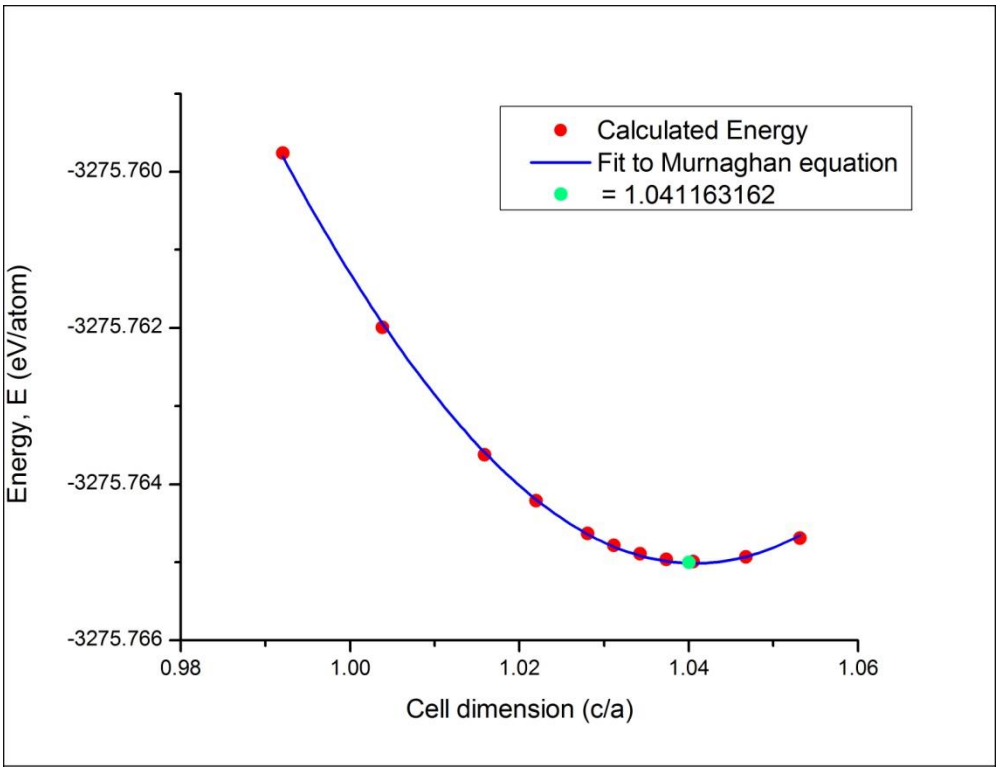

Figure 65. Calculated energy as a function of c/a.

The optimised cell dimensions were used to determine the atomic positions as shown in Table number 11.

Table 11. Optimised cell parameters for martensite.

	Optimised parameter
a (Å)	8.536187658
b (Å)	8.536187658
c (Å)	8.887564135
(c/a)	1.041163162

Table 12. Optimised atomic position for martensite.

Fe	0.0000000000	0.0000000000	0.0000000000
Fe	0.0000000000	-0.0100226040	0.3317941840
Fe	0.0000000000	0.0100226040	-0.3317941840
Fe	0.0000000000	0.3317941840	-0.0100226040
Fe	0.0000000000	-0.3317941840	0.0100226040
Fe	0.3418167880	-0.0100226040	-0.0100226040
Fe	-0.3418167880	0.0100226040	0.0100226040
Fe	0.0000000000	0.2971672640	0.2971672640
Fe	0.0000000000	-0.3378200030	0.3378200030
Fe	-0.3363412810	0.0000000000	0.3363412810
Fe	0.3330427380	-0.0014896130	0.3315531260
Fe	0.0000000000	0.3378200030	-0.3378200030
Fe	0.0000000000	-0.2971672640	-0.2971672640
Fe	0.3363412810	0.0000000000	-0.3363412810
Fe	-0.3330427380	0.0014896130	-0.3315531260
Fe	0.3330427380	0.3315531260	-0.0014896130
Fe	-0.3330427380	-0.3315531260	0.0014896130
Fe	-0.3363412810	0.3363412810	0.0000000000
Fe	0.3363412810	-0.3363412810	0.0000000000
Fe	0.3243589930	0.3378200030	0.3378200030
Fe	-0.3330427380	-0.3354031360	0.3315531260
Fe	-0.3418167880	0.3317941840	0.3317941840
Fe	0.3330427380	-0.3315531260	0.3354031360
Fe	-0.3330427380	0.3315531260	-0.3354031360
Fe	0.3418167880	-0.3317941840	-0.3317941840
Fe	0.3330427380	0.3354031360	-0.3315531260
Fe	-0.3243589930	-0.3378200030	-0.3378200030
C	0.0000000000	0.5000000000	0.5000000000

5.5 Calculation of elastic constants

The ElaStic software (Andris et al., 2014) was used to calculate the elastic constants of martensite. ElaStic uses Quantum Espresso for energy and stress calculation. The elastic constants C_{ij} were calculated by applying small strains (-0.001; -0.0005; 0.00001; 0.0005; 0.001) to the optimised unit cell (see Appendix 3 for input file). The total energy obtained from these calculations was used by ElaStic to calculate the elastic constants.

Chapter 6. Assessment of micropitting appearance

As indicated in the previous chapters micropitting is influenced by load, surface roughness, speed, temperature and, lubrication. Light microscopy examination shows that not all disks subjected to wear testing suffered micropitting damage but this depends on the lubricant, the presence of coatings, whether or not the samples were subjected to additional treatments such as cryogenic treatment.

The surface investigation was performed by quantifying the damage introduced by the wear test. As expected, the light microscopy observation revealed that micropitting tends to initiate on the highest ridges of the grinding marks, as shown in Figure 66.

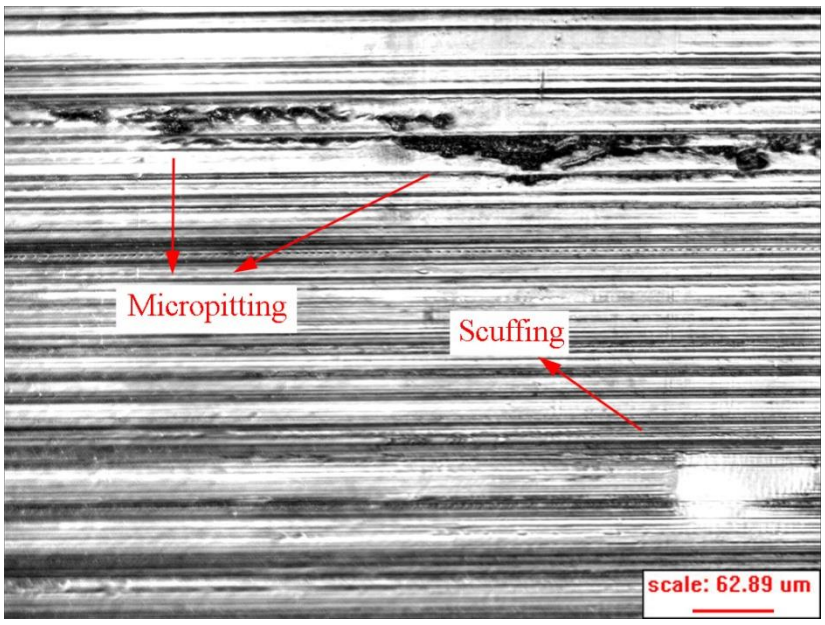

Figure 66. Micropitting damage on the sample tested with 85W140 lubricant.

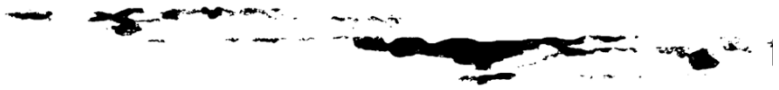

Figure 67. Processed image of the light microscopy image.

Figure 66 shows the damage introduced by the surface contact fatigue on the surface of the disks. The average micropitting damage on each specimen was calculated using Scion Image software. Five images were taken from the surface of each sample and converted to black and white. Figure 67 shows the image converted to only black and white using the binary function in Scion Image. The binary image is necessary in order to calculate the amount of black (representing the micropitting damage) with the software. Scion Image counts the number of micropits and computes the total area of the micropits. The grinding line was removed using the eraser tool leaving only the micropitting damage as shown in Figure 67. The damage initiated from a cluster of micropitting, which increased the damage area over time during testing. In some areas scuffing occurred as indicated in Figure 66 due to the sliding between the disks.

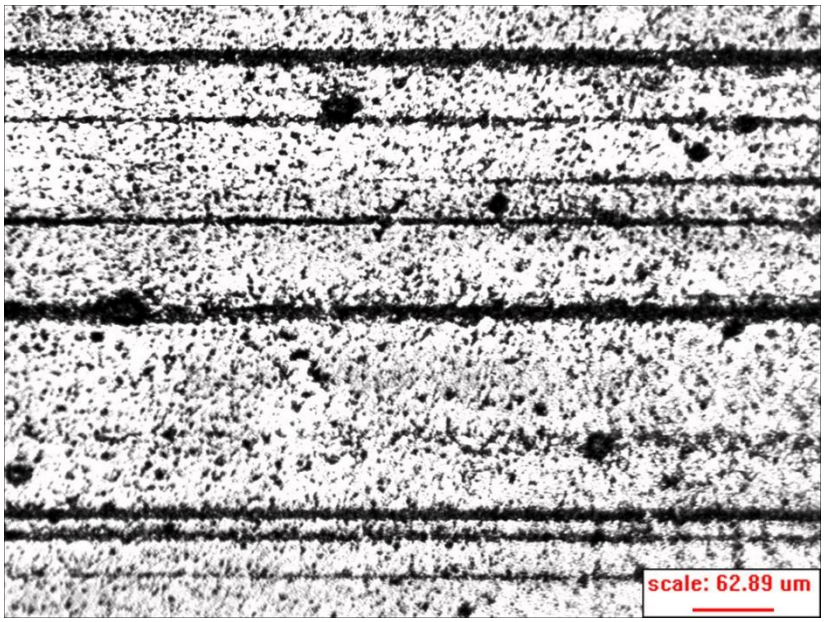

Figure 68. Surface of the sample coated with Balinit C coating.

Figure 68 shows the surface of a sample coated with Balinit C coating. The coating consists of tungsten carbide and carbon. Minimum damage was found on the surface confirming that these types of coatings provide good micropitting protection.

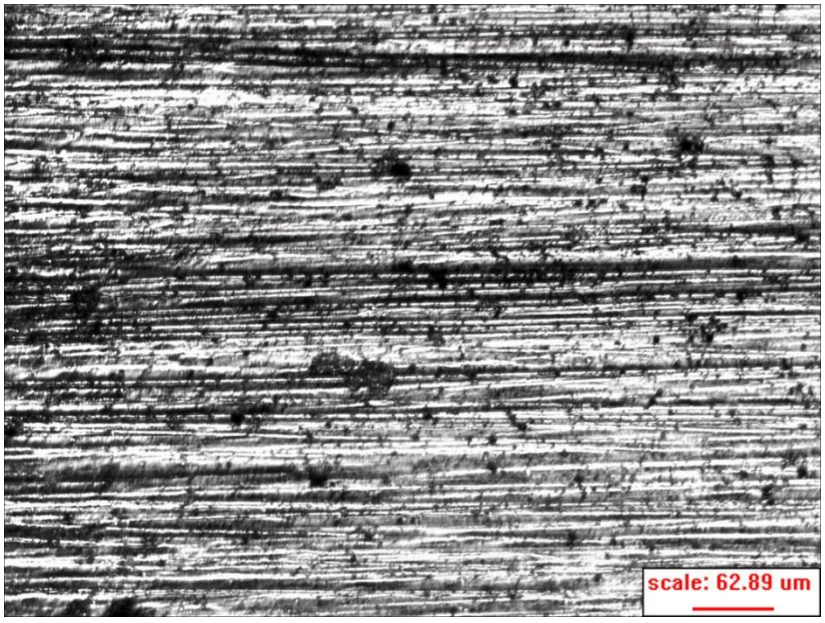

Figure 69. Surface of the sample coated with CrN.

Figure 69 shows the surface of a disk coated with CrN after the wear test. The image shows minimum damage.

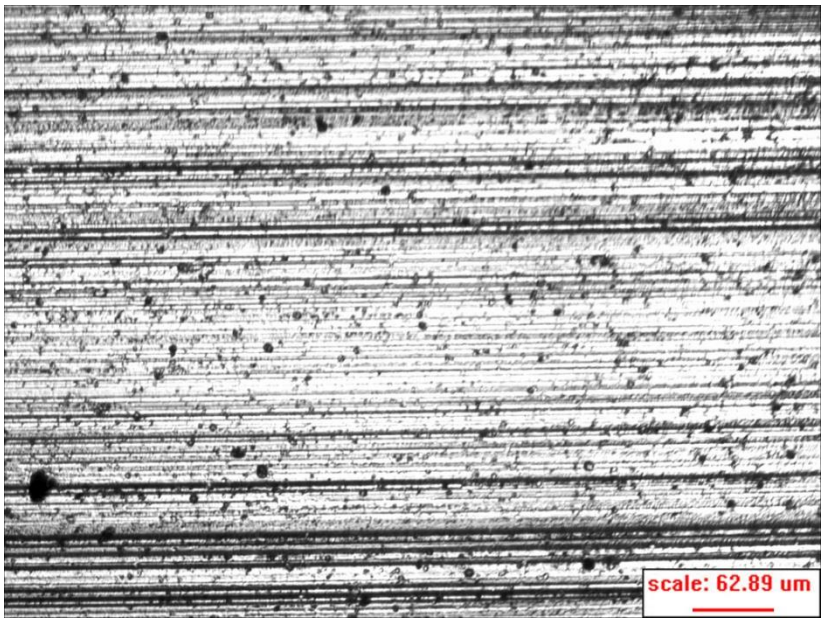

Figure 70. Surface of the sample coated with CrN containing WS_2 nanoparticles.

Figure 70 shows the surface of the sample coated with CrN containing WS_2 nanoparticles. The WS_2 nanoparticles were added into the coating during deposition in order to assess the influence of the nanoparticles on the fatigue performance of the coatings. The light microscopy image shows that there is no wear damage. However, there are issues with the adhesion of these coatings. Figure 71 shows a region where adhesion failure of the coating occurred.

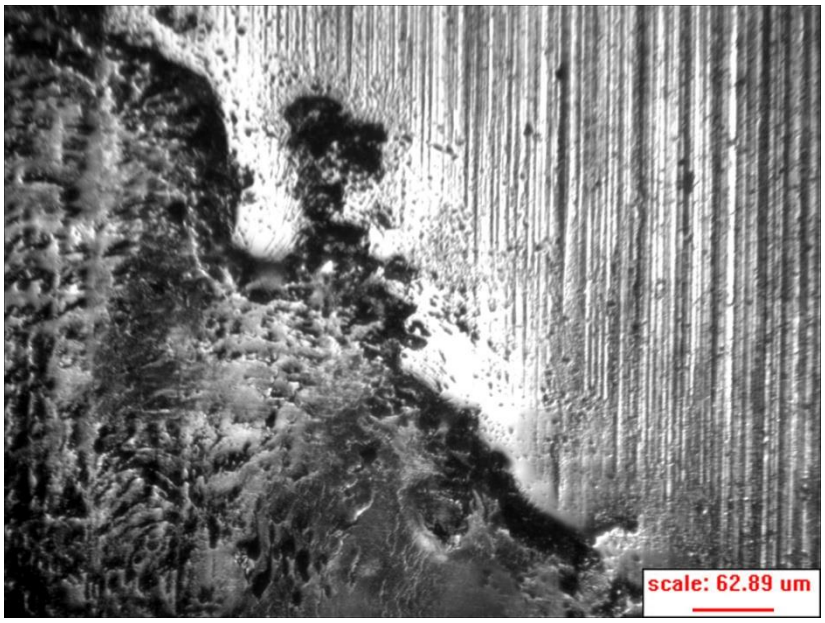

Figure 71. CrN coating containing nanoparticles showing coating failure after the wear test.

A large portion of the coating was peeled off from the surface of the disk.

Figure 72 shows the surface of a sample subjected to deep cryogenic treatment before the wear test.

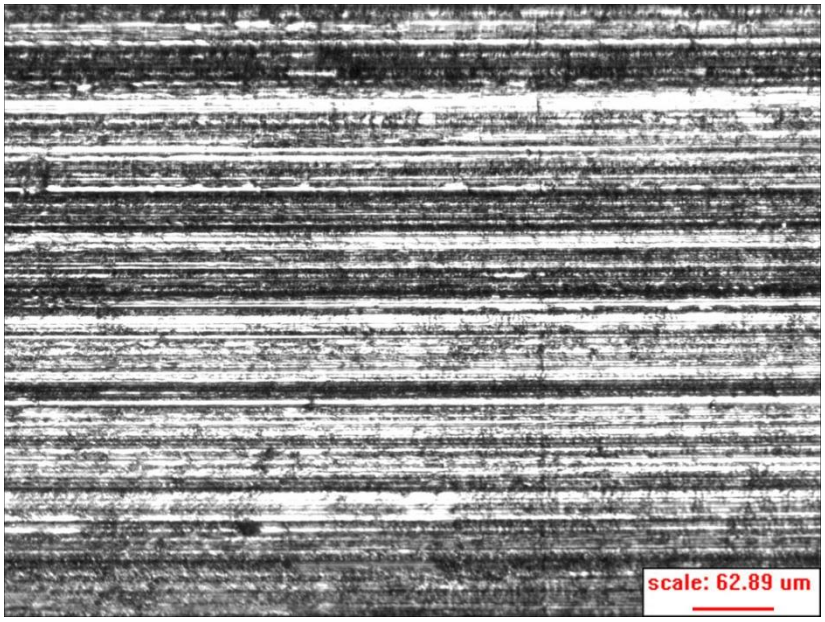

Figure 72. Sample subjected to deep cryogenic treatment without wear test.

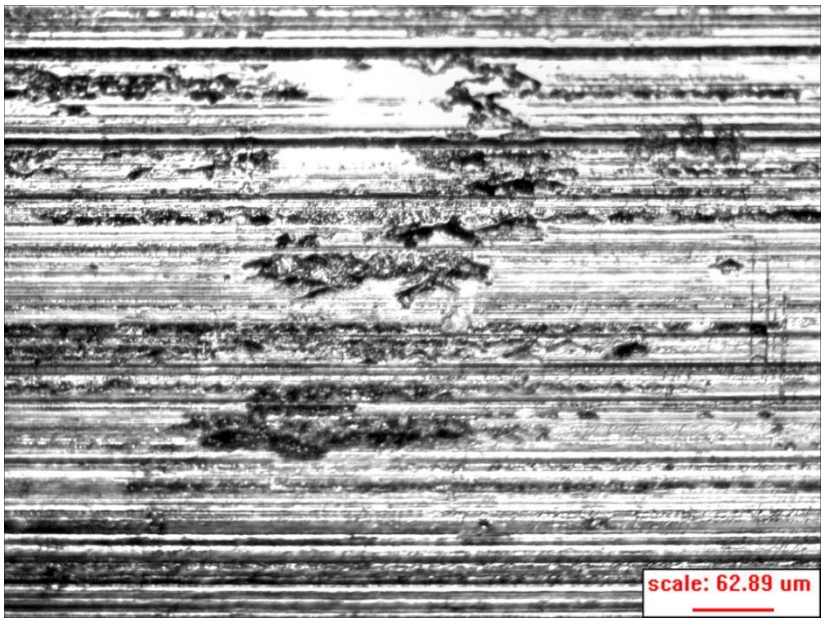

Figure 73. Micropitting damage on sample subjected to deep cryogenic treatment.

Micropitting damage was found on the disk surface of the cryogenically treated disk after wear test as seen from Figure 73.

Figure 74 shows the surface of a sample tested using the FRL-46 lubricant containing WS_2 nanoparticles. All other testing parameters were kept the same.

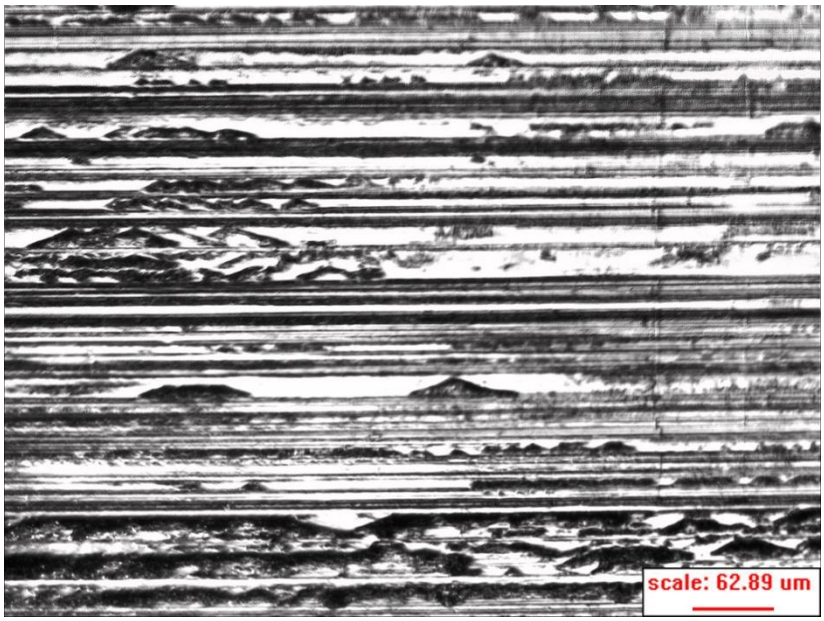

Figure 74. Micropitting damage on the disk lubricated with FRL46 lubricant containing nanoparticles.

Although the nanoparticles were added in FRL-46, the results show that the amount of damage is very high. This is possibly due to the difference in viscosity of the lubricant.

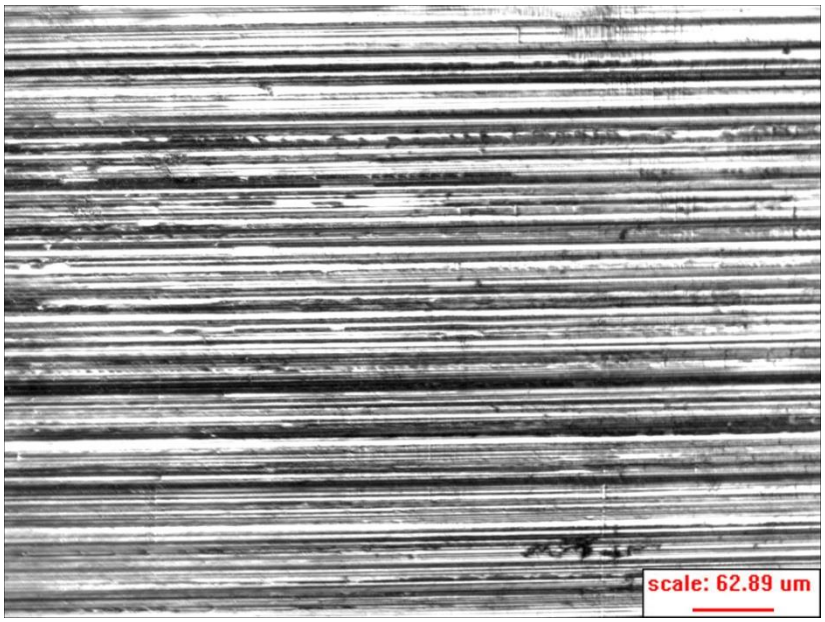

Figure 75. Micropitting damage on the disk lubricated with 85W140 lubricant containing nanoparticles.

Figure 75 shows the surface of the sample lubricated with 85W140 lubricant containing nanoparticles. The light microscopy reveals very little damage compared to the sample tested with the 85W140 lubricant without IFLM added (Figure 66).

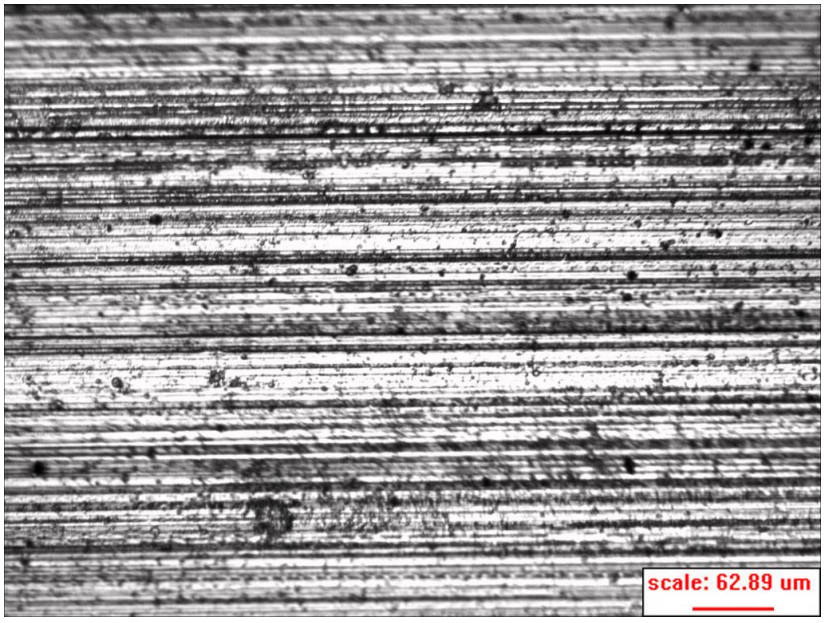

Figure 76. The surface of the disk superfinished and coated with CrN containing nanoparticles.

Figure 76 shows the surface of the disk that was superfinished and was coated with CrN containing nanoparticles. The superfinishing reduces the roughness of the disks and increases the adhesion of the CrN coating. This seems to be effective since there are no adhesion issues for this sample. The amount of damage is minimum and no large cluster of wear damage was observed.

Figure 77 shows the compiled data for average micropitting damage for the disks.

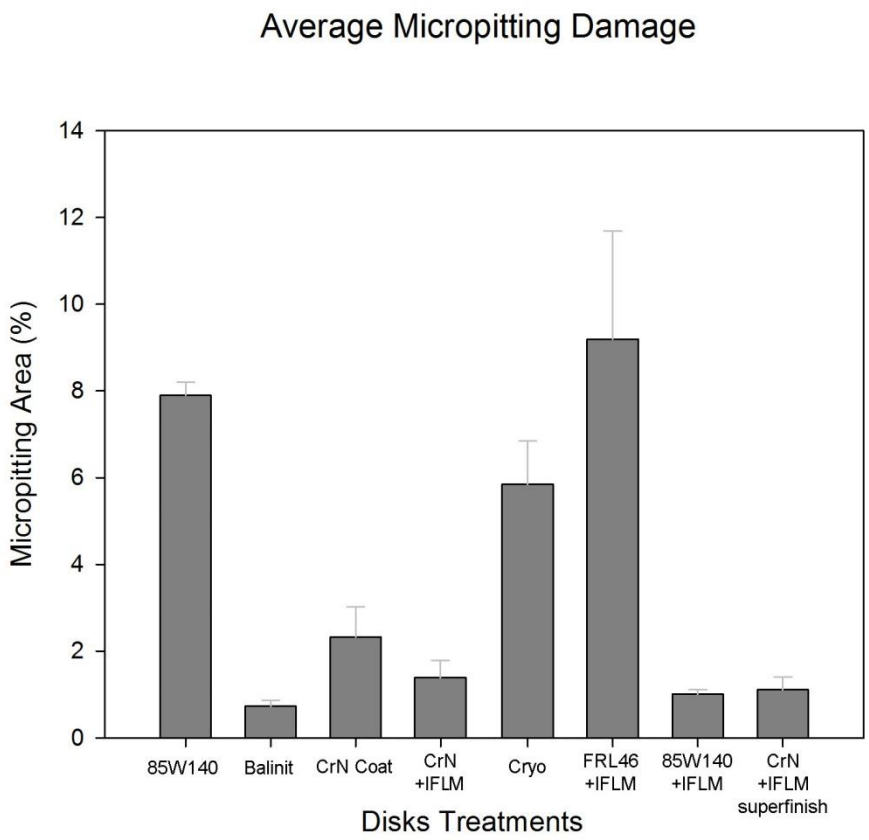

Figure 77. Average surface micropitting damage for each treatment.

For 85W140 lubricated disk, the average micropitting percentage is 7.80±0.30%. The Balinit C coated sample has an average of 0.72±0.13%. The CrN coated sample has an average of 2.32±0.8%. The CrN coated disk with nanoparticles IFLM (Inorganic fullerene-like material) additive has an average of 1.38±0.50%. The Cryogenic treated

disk has an average of 5.84±1.02%. The disk lubricated with FRL46 with IFLM additive has an average of 9.20±3.10%. The 85W140 lubricated disk with IFLM additive has an average of 1.00±0.15%. The CrN coating with IFLM added with superfinish has an average of 1.10±0.3%.

Although the underlying core material is the same, the surface treatment and the lubrication play a significant role in altering the wear behaviour of each disk. From the initial wear damage assessment, the results show that not all treatments will improve the wear resistance. For example, the deep cryogenic treatment appears to increase the amount of micropitting on the surface. By switching the lubricant from 85W140 to FRL46, the amount of damage on the disk surface also increased. The main role of lubrication is to decrease the metal to metal contact, and the viscosity proved to play an important role. FRL46 lubricant has a lower dynamic viscosity (54 cP) compared to the 85W140 lubricant (138 cP), therefore it is reasonable to assume that the reason for the decrease in wear resistance is due to the lowered viscosity.

Chapter 7. Microstructural changes associated with micropitting

7.1 Light Microscopy Investigation

Figure 78 shows the micrograph of a cross section of the sample lubricated with 85W140 and etched with 2% Nital. The grains under the disk surface shows primarily tempered martensite with the needle like structure. Dark Etching Regions are visible in the subsurface regions.

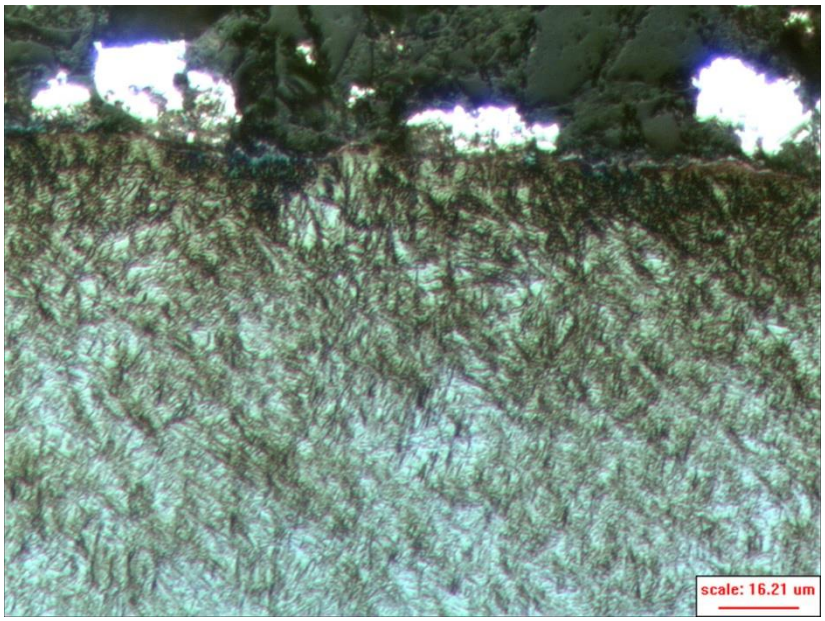

Figure 78. Micrograph showing a cross section of the disk lubricated with 85W140 lubricant. Etch: Nital 2%.

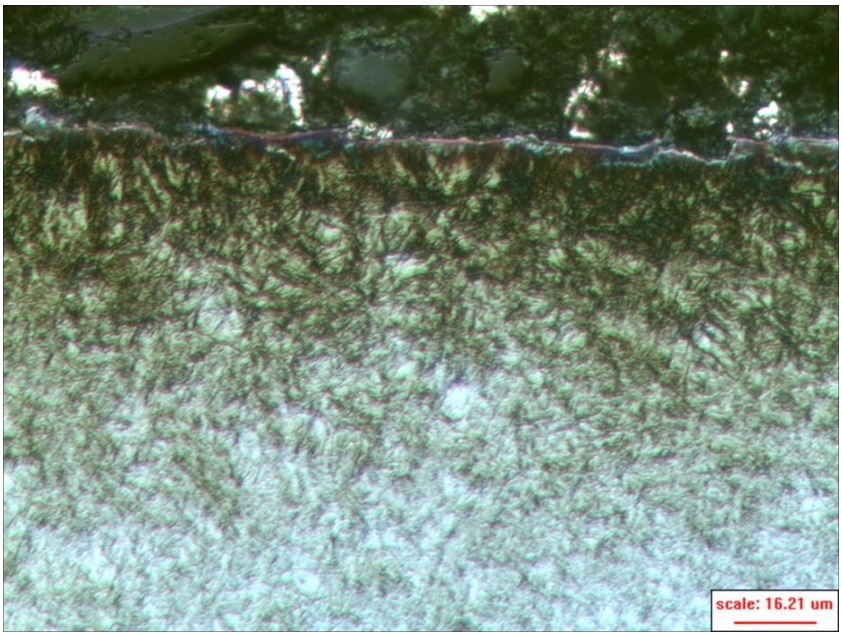

Figure 79. Micrograph showing a cross section of the disk lubricated with 85W140 lubricant containing nanoparticles. Etch: Nital 2%.

Figure 79 shows the microstructure of the disk lubricated with 85W140 with IFLM additive. Dark Etching Regions were observed in this sample which correlates well with the fact that micropitting on the surface was present.

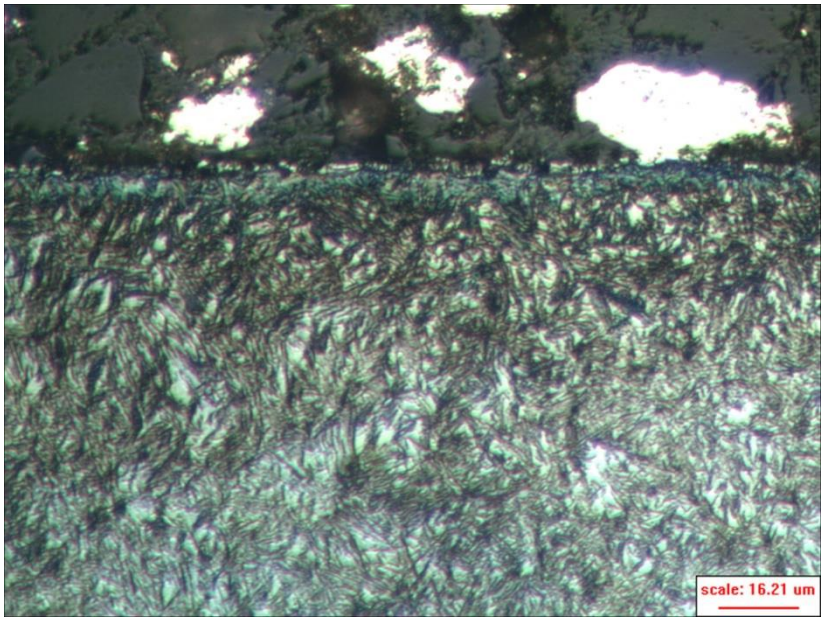

Figure 80. Micrograph showing a cross section of the disk coated with Balinit C. Etch: Nital 2%.

Figure 80 shows the cross section of the disk coated with Balinit C. Balinit C coating has a thickness of 3~5 micron. The needle structure underneath the surface shows tempered martensite with retained austenite (light areas). No discernible Dark Etching Regions were observed but this was expected since no micropitting damage occurred on the surface.

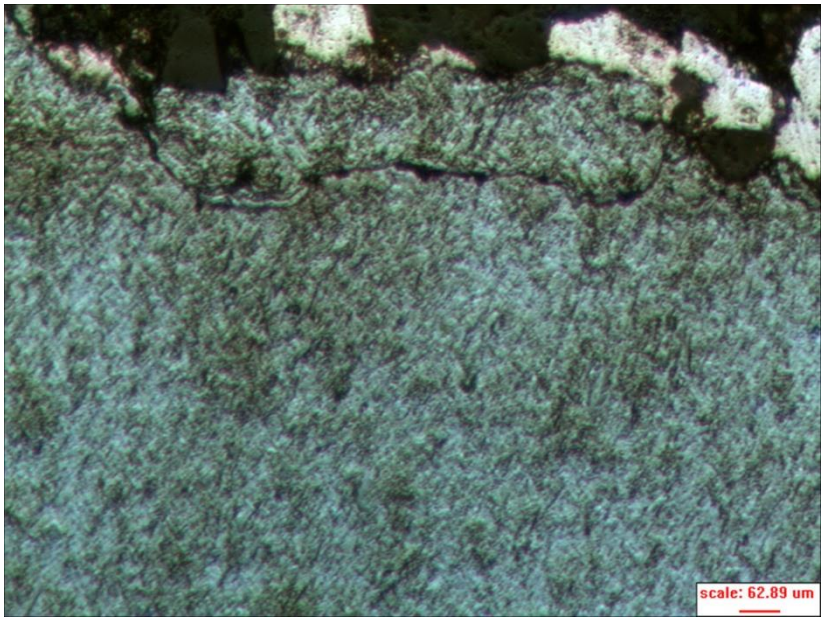

Figure 81. Micrograph showing a cross section of the cryogenically treated disk. Etch: Nital 2%.

Figure 81 shows the area underneath the surface of the cryogenically treated sample. Although martensite decay cannot be observed a relatively large crack is easily observable.

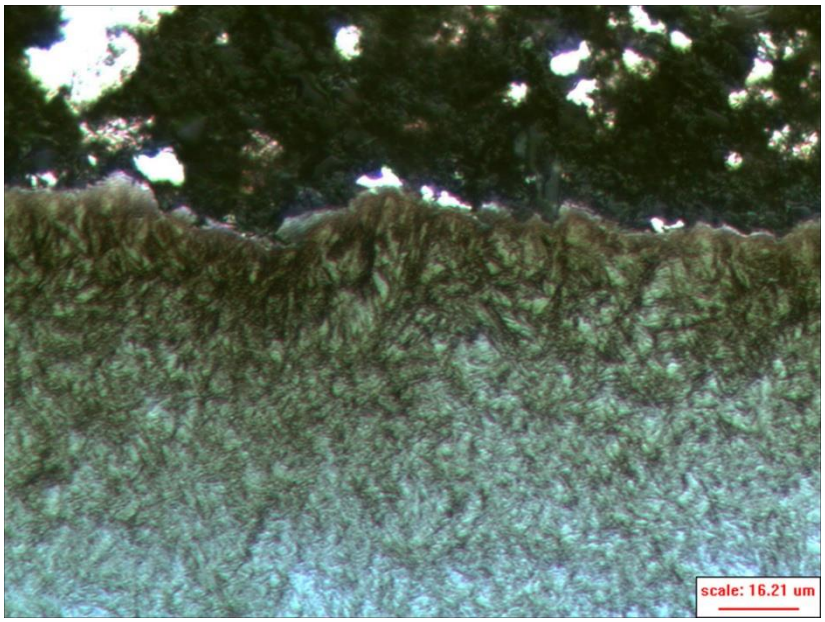

Figure 82. Micrograph showing a cross section of the disk lubricated with FRL46 lubricant. Etch: Nital 2%.

Figure 82 shows the cross section of disk lubricated with FRL-46 lubricant. The FRL-46 lubricated disk shows tempered martensite structure underneath the surface and some dark etching areas.

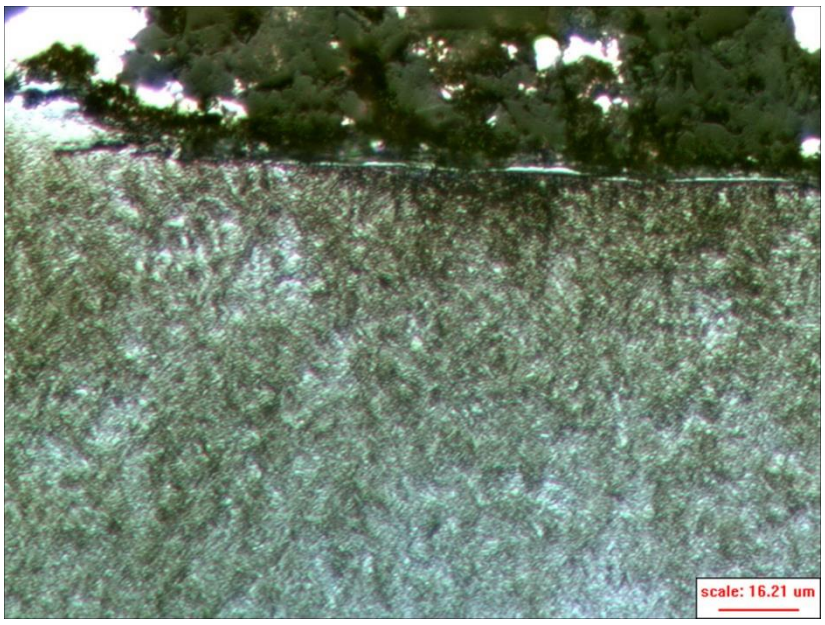

Figure 83. Micrograph showing a cross section of the disk coated with CrN. Etch: Nital 2%.

Figure 83 shows the cross section of the disk coated with CrN. The microstructure consists of very fine grains, which is the result of tempering from the coating process. The micrograph also shows dark etching region.

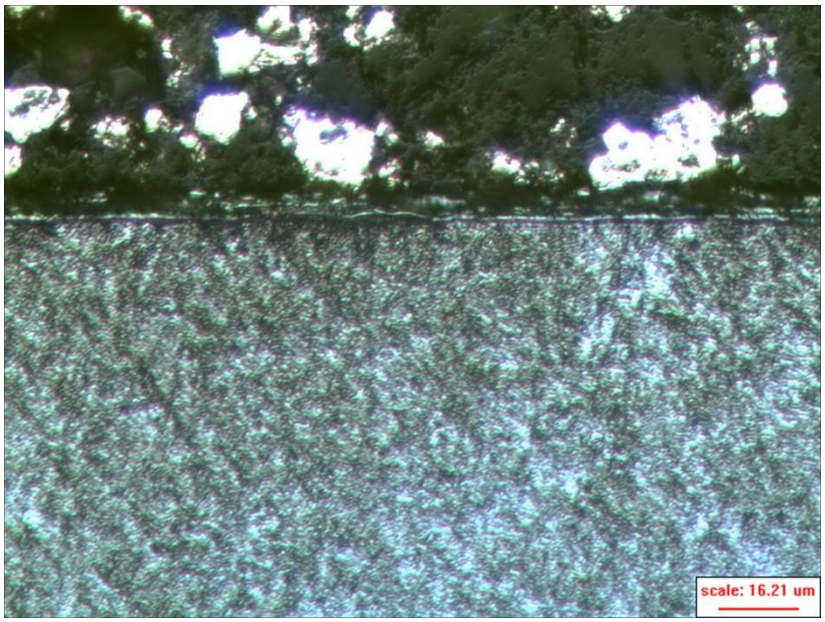

Figure 84. Micrograph showing a cross section of the disk coated with CrN containing IFLM. Etch: Nital 2%.

Figure 84 shows the cross section of the disk coated with CrN containing IFLM. The microstructure shows very fine grains and dark etching region.

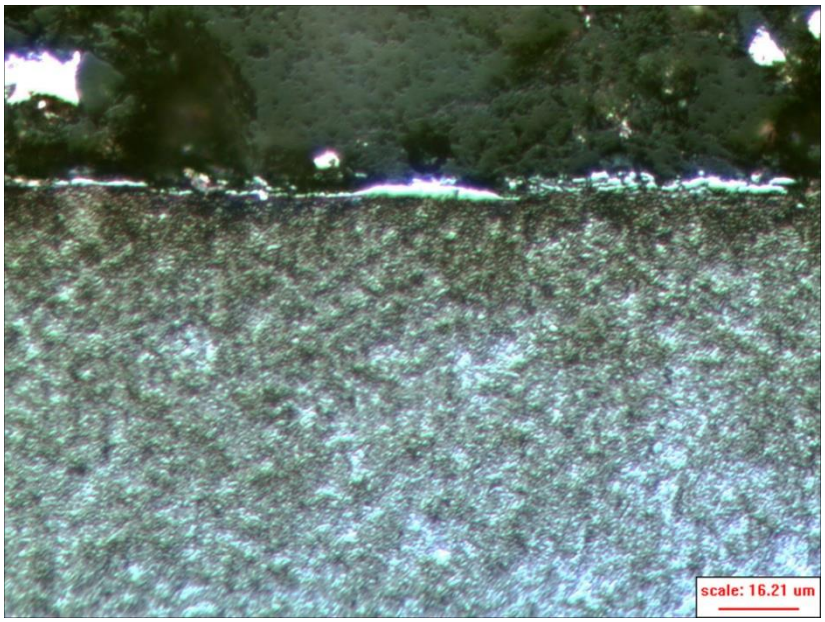

Figure 85. Micrograph showing a cross section of the superfinished disk coated with CrN containing IFLM. Etch: Nital 2%.

Figure 85 shows the cross section of the superfinished disk coated with CrN containing IFLM showing fine grains and dark etching region. The typical tempering temperature for the S156 steel is 180°C, and the CrN coating was deposited at 230°C. The CrN coating process causes the further tempering of the steel.

7.2 Nanoindentation testing

Nanoindentation was performed in order to determine the hardness and elastic modulus of martensite and Dark Etching Regions, respectively. Multiple patterns were used depending on each sample and the area chosen for the test. The load was set to 5 mN with 5 µm separating each indent. Figure 86 shows a nanoindentation pattern with 45 indentations.

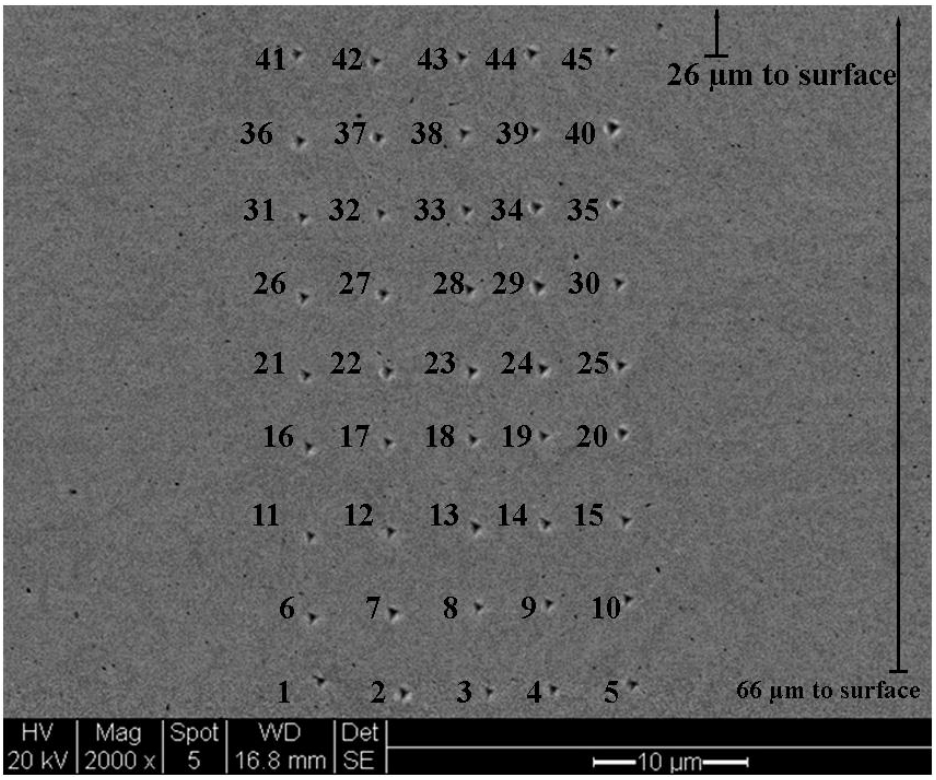

Figure 86. Nanoindentation pattern used on the cryogenically treated sample.

The results of the nanoindentation test are shown in Table 13.

Table 13. Nanoindentation data of cryogenically treated sample (corrected).

Indentation	Hardness (GPa)	Elastic Modulus (GPa)
1	11.6	185.68
2	11.4	193.38
3	10.8	194.15
4	10.6	190.19
5	11.7	193.38
6	10.6	198.99
7	12.5	194.92
8	12.3	196.46
9	12.1	198.99
10	11.7	202.95
11	11.2	194.92
12	11.3	198.44
13	11.9	189.53
14	10.8	195.03
15	10.4	189.42
16	11.9	197.12
17	10.8	192.72
18	11.7	193.82
19	10.8	194.37
20	11.3	197.56
21	11.6	199.76
22	10.9	193.93
23	11.0	201.63
24	11.1	199.21
25	11.8	197.23
26	11.4	199.21
27	11.5	195.69
28	11.3	196.79
29	11.7	200.53
30	11.1	194.81
31	10.7	198.11
32	12.3	198.99
33	11.2	194.26
34	11.4	198.66
35	11.3	193.82
36	10.6	212.85
37	10.6	203.06
38	10.8	202.84
39	12.1	205.15
40	11.7	199.1
41	10.9	202.84
42	11.0	197.67
43	12.1	193.6
44	11.6	195.69
45	12.0	203.72

The average elastic modulus is 198 GPa. The average hardness is 11.4 GPa.

The “pile up” phenomenon occurs during nanoindentation due to the material being pushed towards the side of the indentation. The increase in height often results in

inaccurate calculations using the Oliver-Pharr formula and causes the overestimation of hardness and elastic modulus. The correction factors K_H =0.82 for hardness and K_E =0.75 for elastic modulus were used to correct the elastic and hardness values (Oila & Bull, 2004). Figure 87 shows an example of pile up from an indent on the cryogenically treated sample.

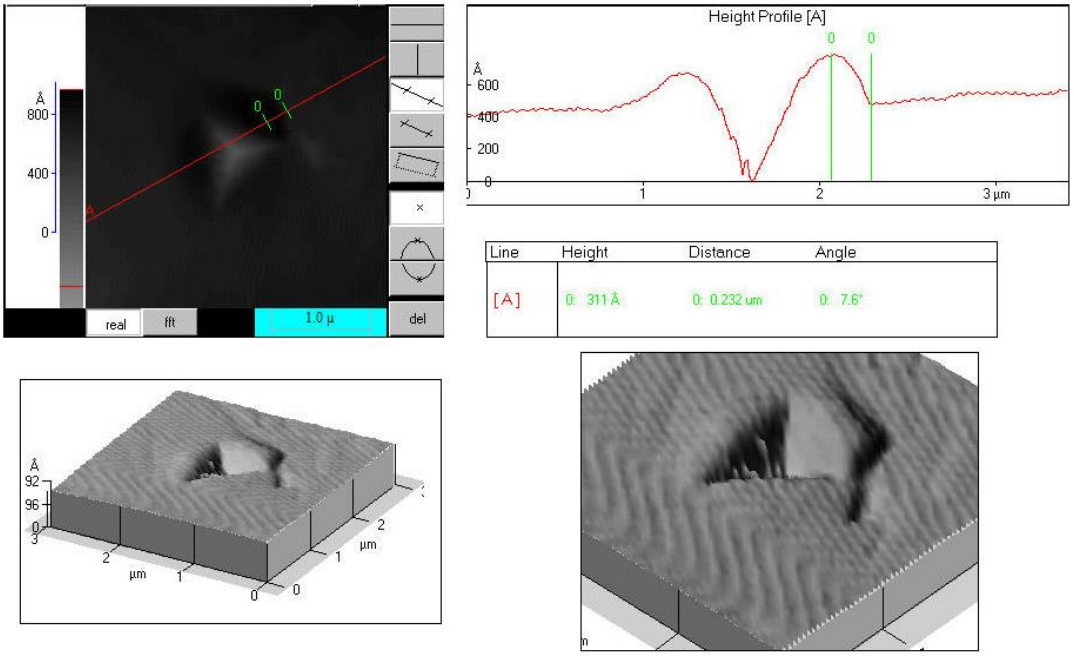

Figure 87. AFM images showing the pile-up effect in the cryogenically treated sample. Top left shows the line analysis of a single indentation with a red line. Top right shows the surface topography along the line and the height of the pile up is measured between the green markers. Bottom shows the 3D representation of the same indentation.

7.3 Scanning Electron Microscopy and EBSD

SEM was used to examine the cross sections of the samples subjected to the micropitting test. EBSD was used on the Balinit C coated sample and gear samples. Figure 88 shows a crack initiated from the surface.

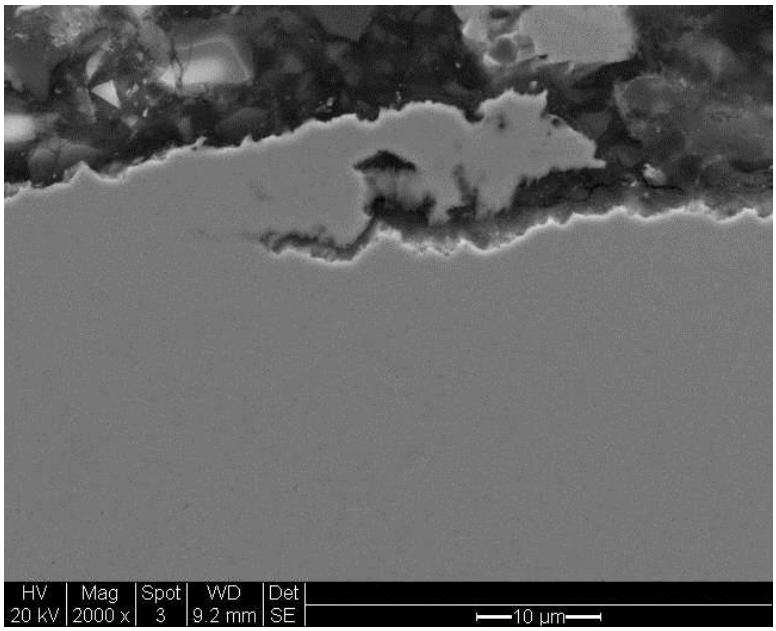

Figure 88. SEM image showing a crack in the sample lubricated with 85W140.

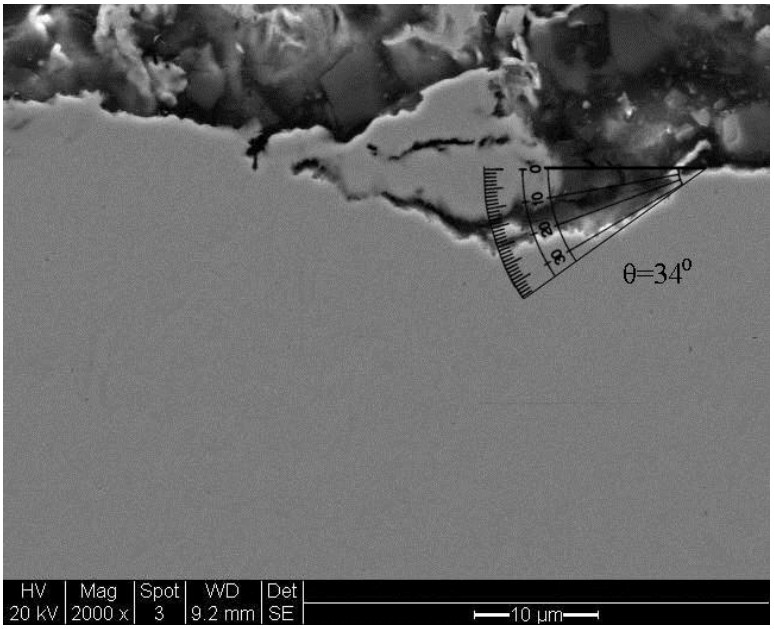

Figure 89. SEM image showing a cross section from sample lubricated with 85W140 showing a crack.

Figure 89 shows an SEM image with crack initiated from the surface with an angle of around 34°.

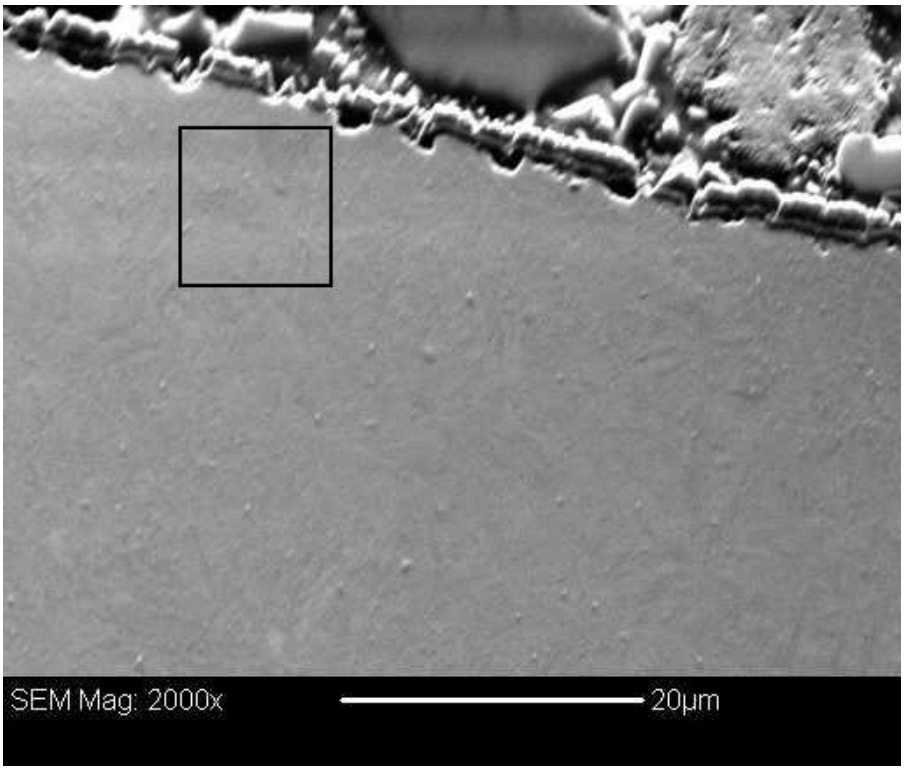

Figure 90. The location marked for EBSD analysis on the Balinit C coated sample.

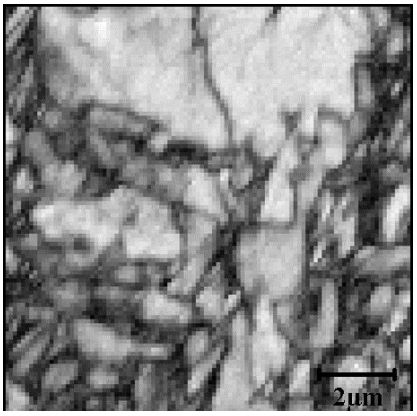

Figure 91. Band contrast image obtained from the Balinit C coated sample.

Figure 90 shows the area chosen for the band contrast image. Figure 91 shows the band contrast image which is a quality factor that describes the average intensity of the Kikuchi bands within an area. It shows the grains and grain boundaries. Two large grains were seen on top of the figure and finer grains on the bottom of the figure. The lighter grains are grains of retained austenite. From the Kikuchi bands, the automated indexing capability of the software automatically locates the individual Kikuchi bands and compares the data with the internal library as shown in Figure 92.

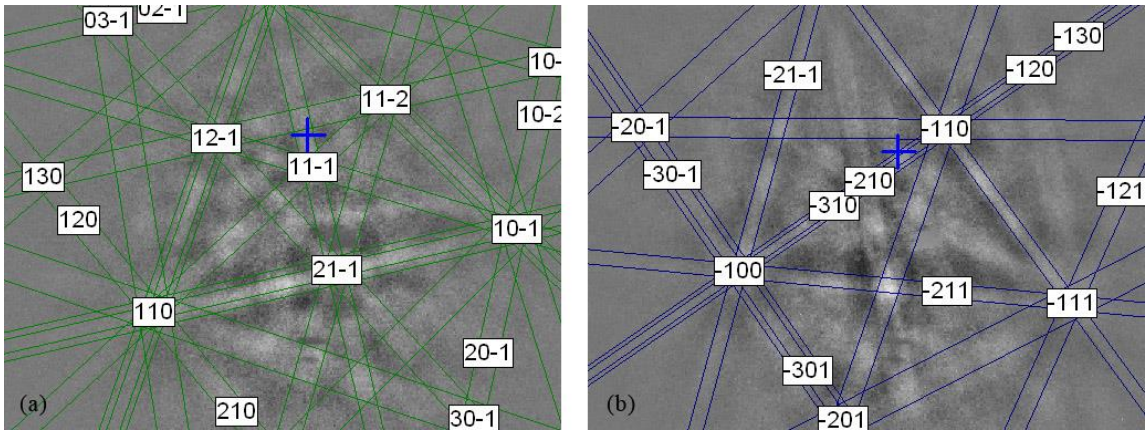

Figure 92. Kikuchi pattern of (a) austenite, and (b) martensite with Miller indices.

7.4. Investigation of Gear Samples

Figure 93 shows the needle shape grains of the martensite phase on gear sample subjected to 1.5 GPa load. The prior austenite grain boundaries are clearly visible.

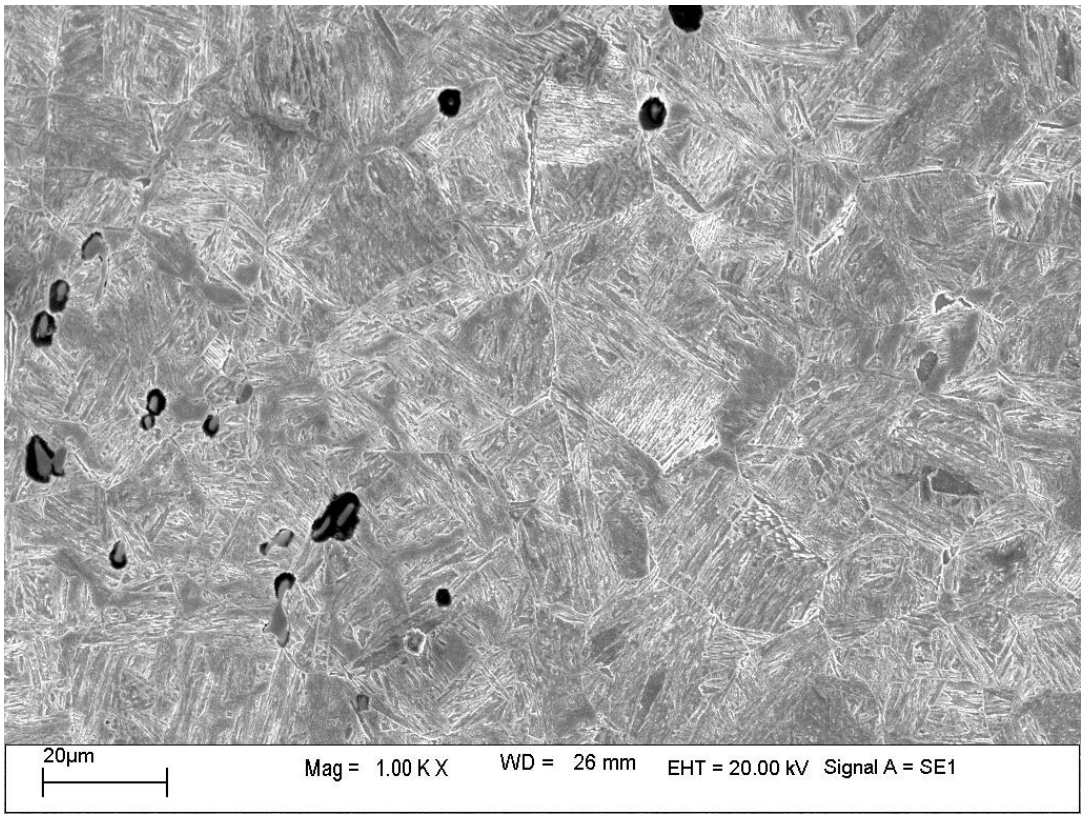

Figure 93. SEM image of the martensite grain structure from the gear sample B29.

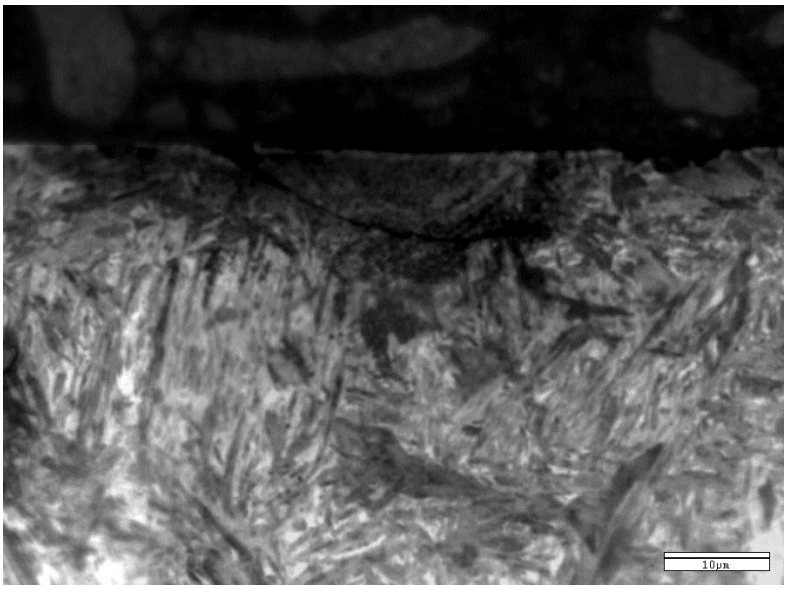

Figure 94. Light microscopy of the dark etching region in sample B29.

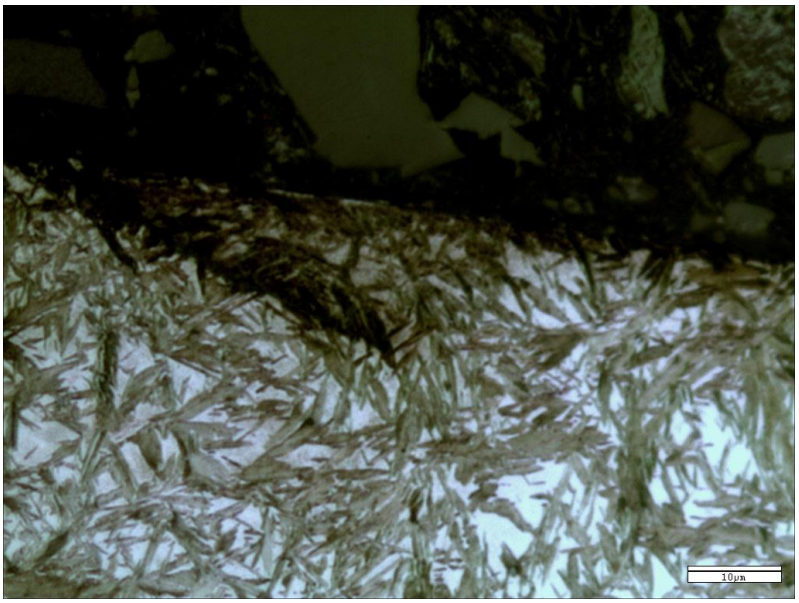

Figure 95. Light microscopy of the dark etching region in sample AC9P.

The dark etching regions (DER) have been identified on gear samples B29 and AC9P by light microscopy and SEM. Figure 94 and Figure 95 show light microscopy images of the DER near the contact surface. The etching reveals a much darker color compared to the surrounding region. The region is semi-circular in shape. Multiple DER's were found near the surface and they have an average size between 20 x20 µm and 50 x 50 µm. The metallographic examination indicates that the boundaries of the DER are prone to crack

initiation and propagation. If the crack continues to propagate, it will generate a wear debris particle which will detach from the surface. The size of the DER corresponds to the width of the grinding marks, which suggest that the DER develop beneath the grinding marks. The region below the grinding mark is exposed to high contact stresses and becomes an area prone to crack initiation.

7.5 Electron Backscatter Diffraction of DER

Dark etching region have been identified on sample B29 (Figure 96) and AC9P (Figure 97) by light microscopy

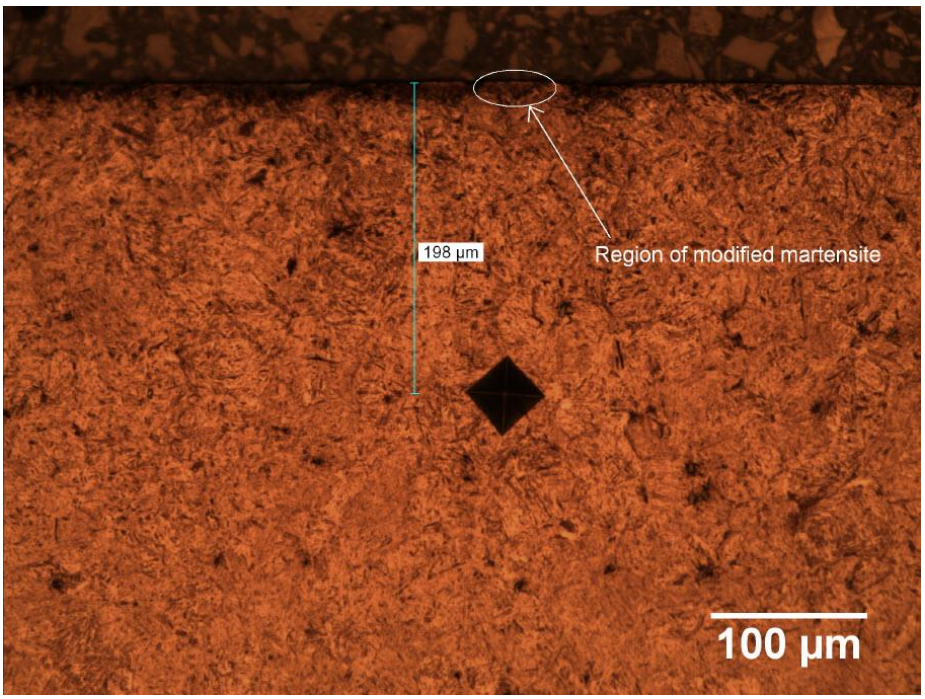

Figure 96. Dark etching region - sample B29. Indent marked at 198 µm below the surface (Oila, 2010).

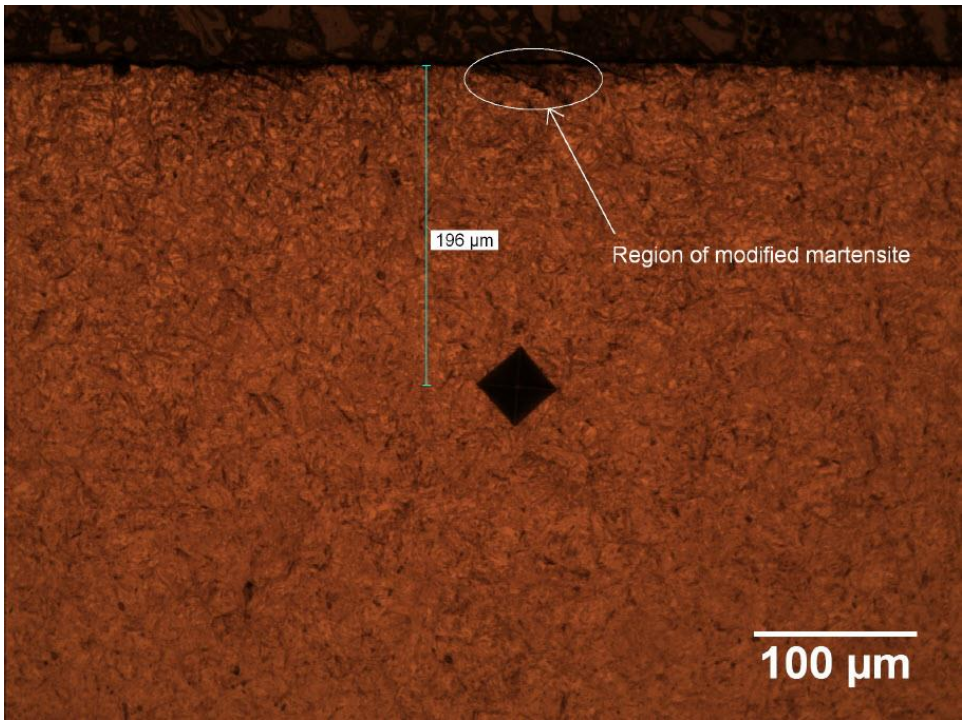

Figure 97. Dark etching region - sample AC9P. Indent marked at 196 µm below the surface (Oila, 2010).

Figure 98 shows the forescatter image of an area selected for EBSD without fatigue test. The slight roughness on the surface is the result of the colloidal silica polishing which also etches the surface to some extent.

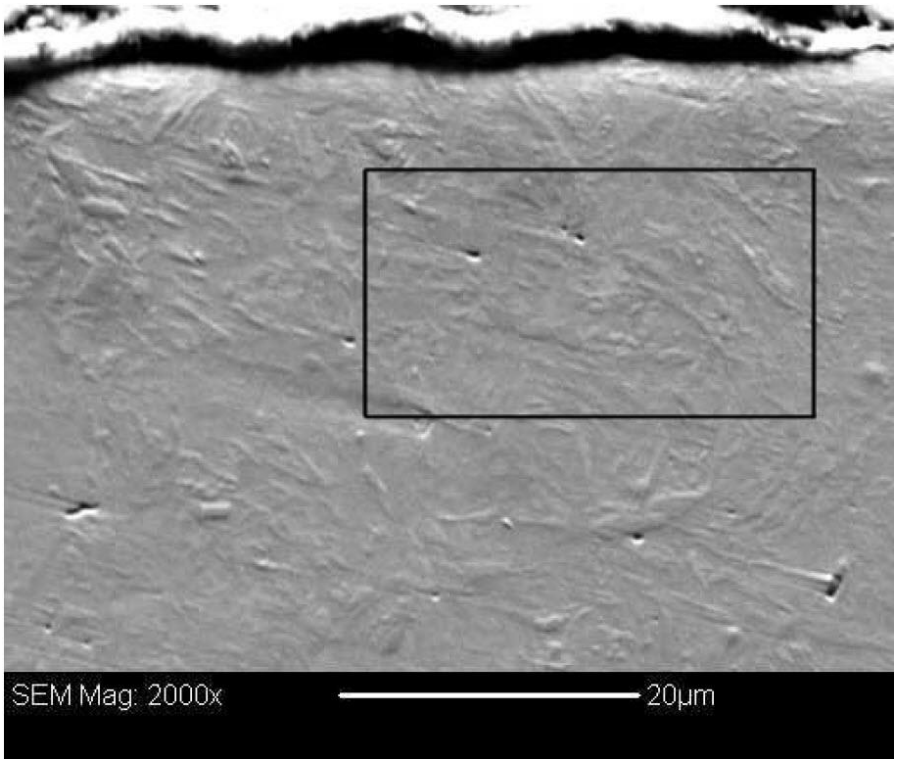

Figure 98. Forescatter image showing the area selected for EBSD (Oila, 2010).

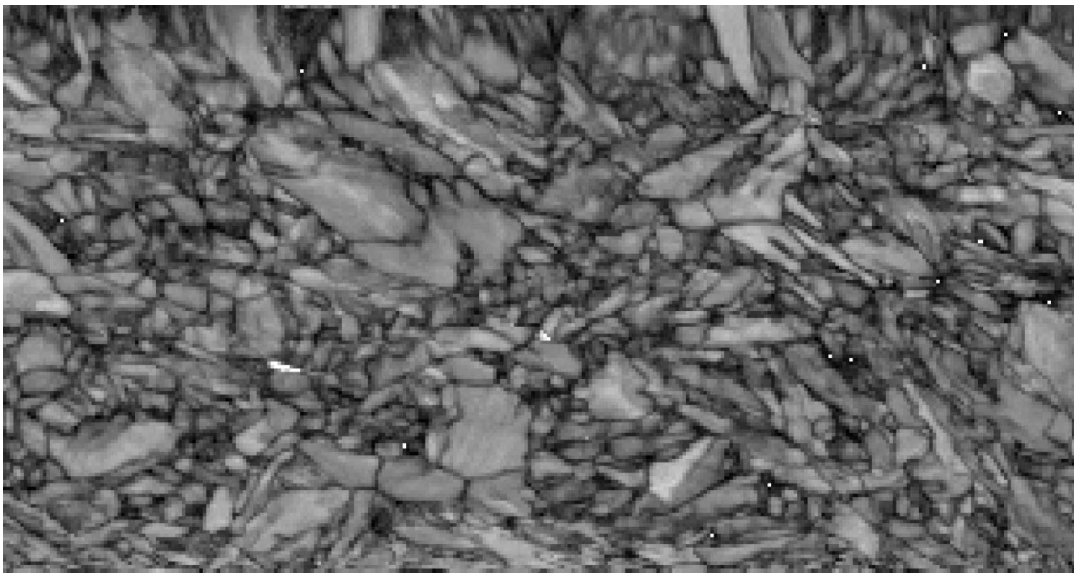

Figure 99. EBSP quality map for untested sample (Oila, 2010).

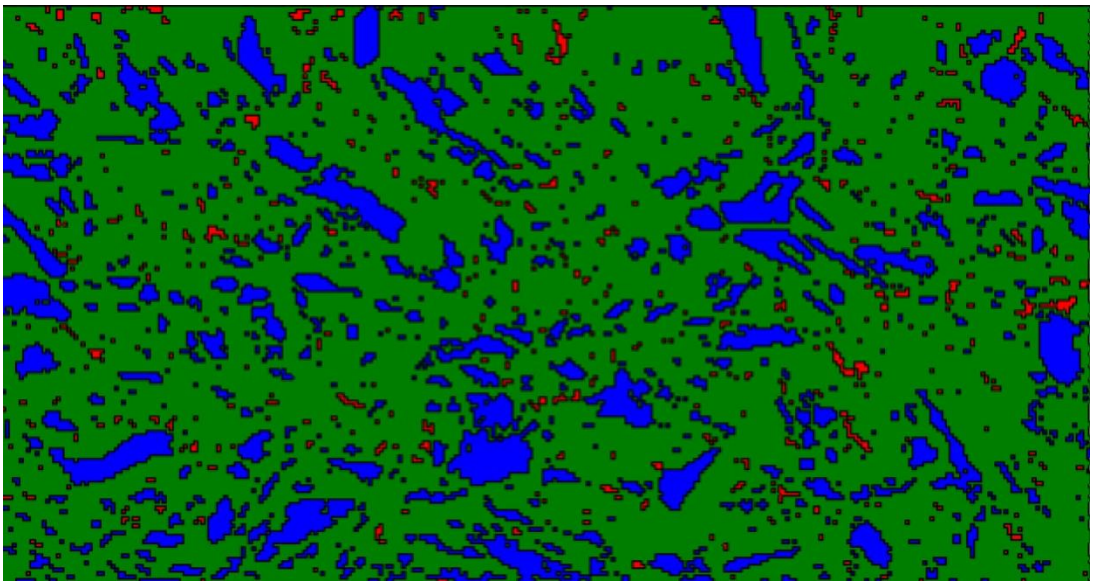

Figure 100. Untested sample phase map (Red-cementite, blue retained austenite, green-martensite, white-not indexed) (Oila, 2010).

From the EBSP quality map and the phase map shown in Figure 99 and Figure 100 it is apparent that the untested sample contains mostly martensite, some retained austenite, and a small amount of cementite.

The rectangle from Figure 101 shows the EBSD mapped area of AC9P. The phase map is shown in Figure 103.

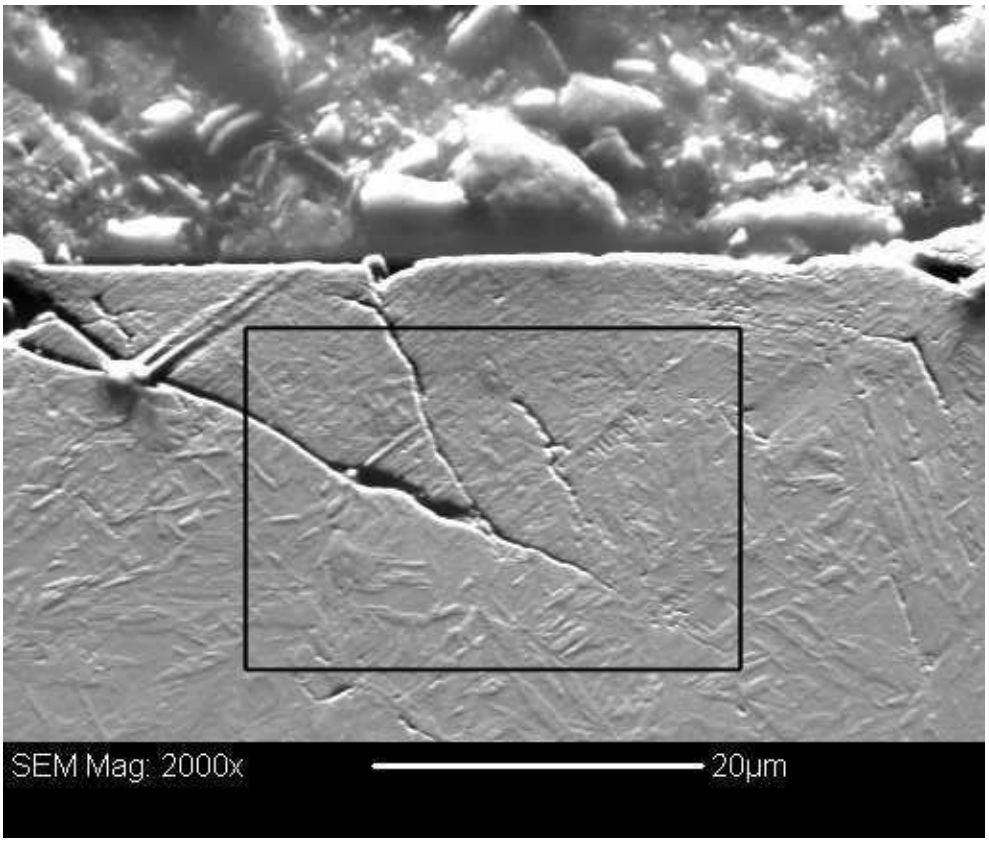

Figure 101. Forescatter image of the area selected for EBSD on AC9P with area marked for EBSD analysis (Oila, 2010).

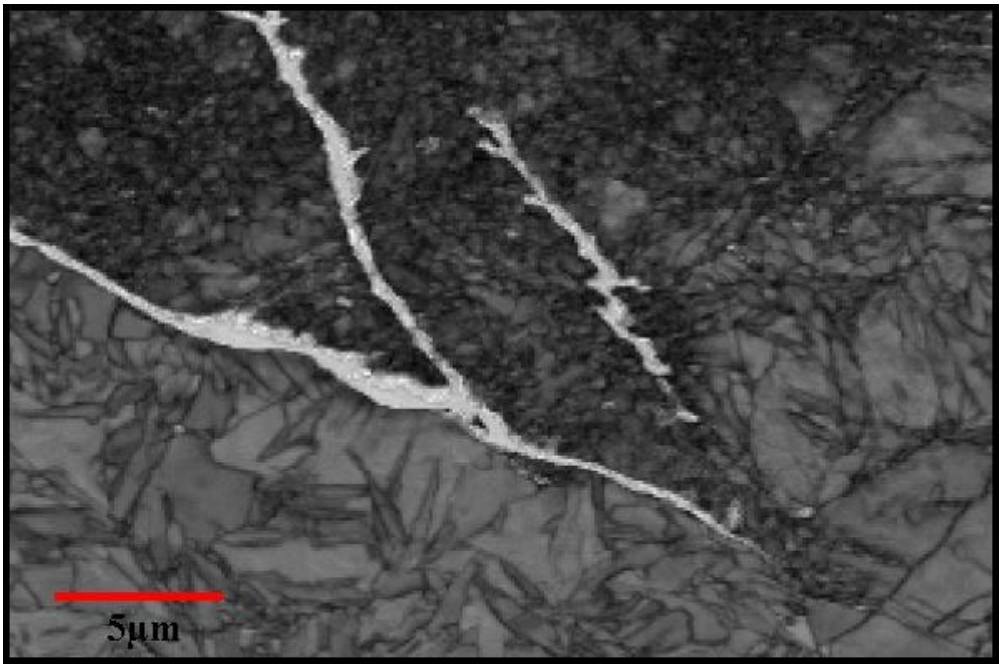

Figure 102. EBSP quality map for fatigue tested sample at the DER (Oila, 2010).

The EBSP (Electron backscatter diffraction pattern) quality map of DER shown in Figure 102 suggests that the crack is associated with DER. It shows that the martensite grains breaks down and form a very fine microstructure of the DER. The surface initiated crack propagates along the boundary of DER which is severely deformed.

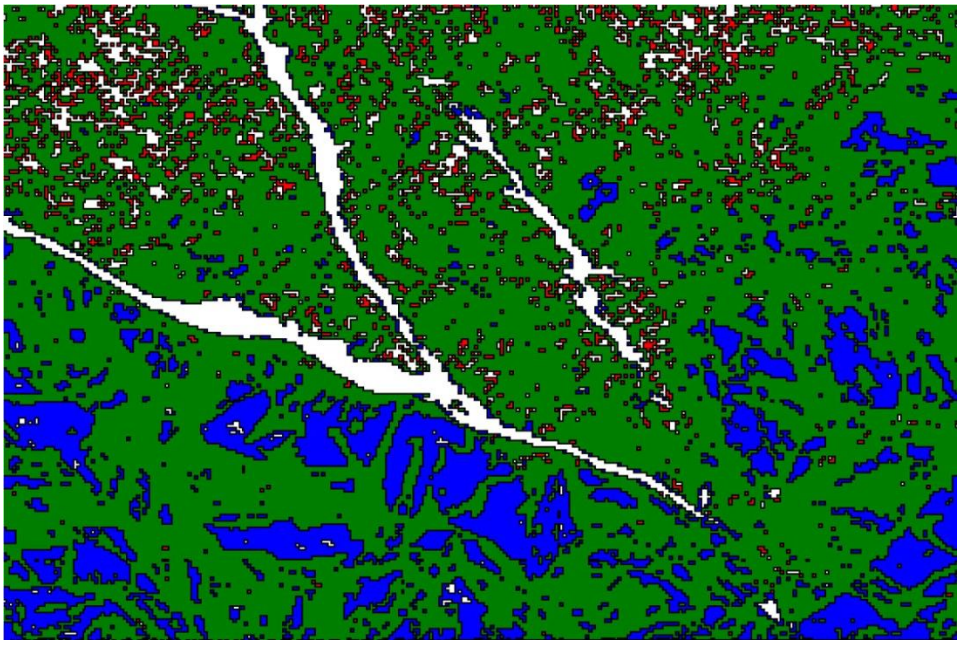

Figure 103. Sample AC9P phase map (Red-cementite, blue retained austenite, green-martensite, white-not indexed) (Oila, 2010).

The phase map in Figure 103 shows that inside the DER the amount of retained austenite is reduced, and the presence of cementite is increased. Outside the DER region consists of mainly martensite and retained austenite.

7.6 Nanoindentation of DER

Nanoindentation was performed in order to determine the hardness and elastic modulus of martensite and Dark Etching Regions. The load was set to 10 mN with a distance of 5 μm separating each indent.

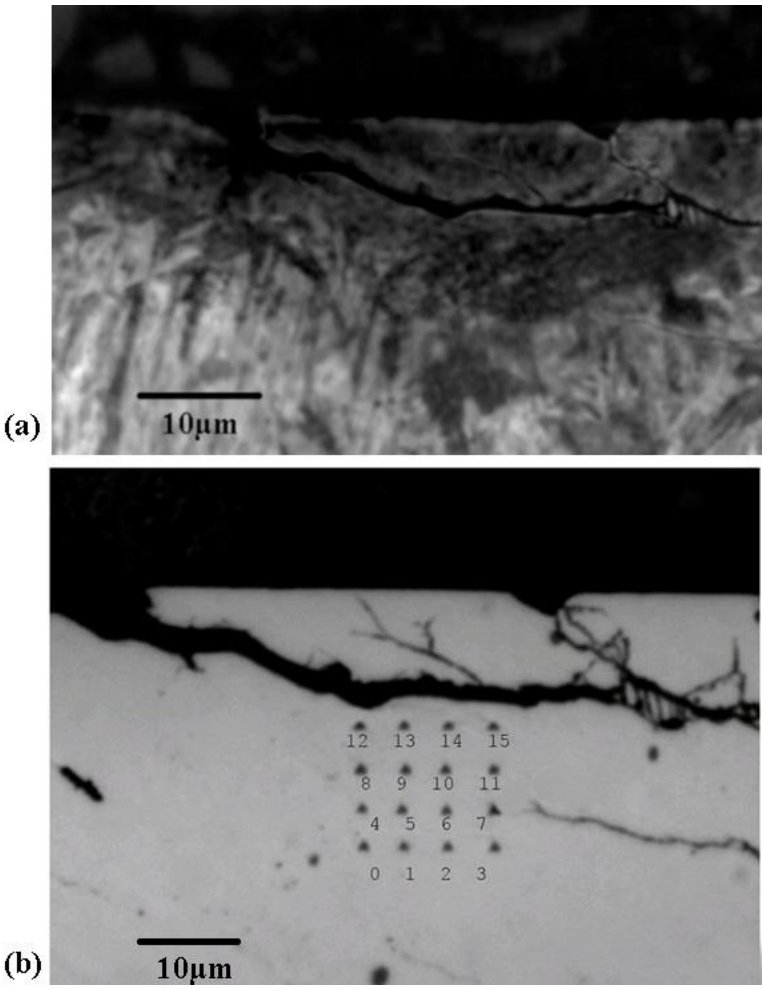

Figure 104. (a) Light microscopy showing DER on sample B29. (b) Nanoindentation pattern inside DER.

The hardness and elastic modulus data are shown in Table 14.

Table 14. Nanoindentation data corresponding to the indentations shown in Figure 104 (corrected).

Indentation	Hardness(Gpa)	E(GPa)
0	11.70	171.0
1	11.77	173.9
2	11.25	174.4
3	11.57	174.5
4	12.35	178.9
5	11.90	176.8
6	12.22	178.4
7	11.31	175.7
8	12.03	181.8
9	12.09	177.8
10	11.57	173.6
11	11.05	174.7
12	11.83	175.1
13	10.99	177.4
14	11.64	176.9
15	10.99	176.8

The results show that the DER has an average hardness of 11.64 GPa and an average elastic modulus is 176.1 GPa.

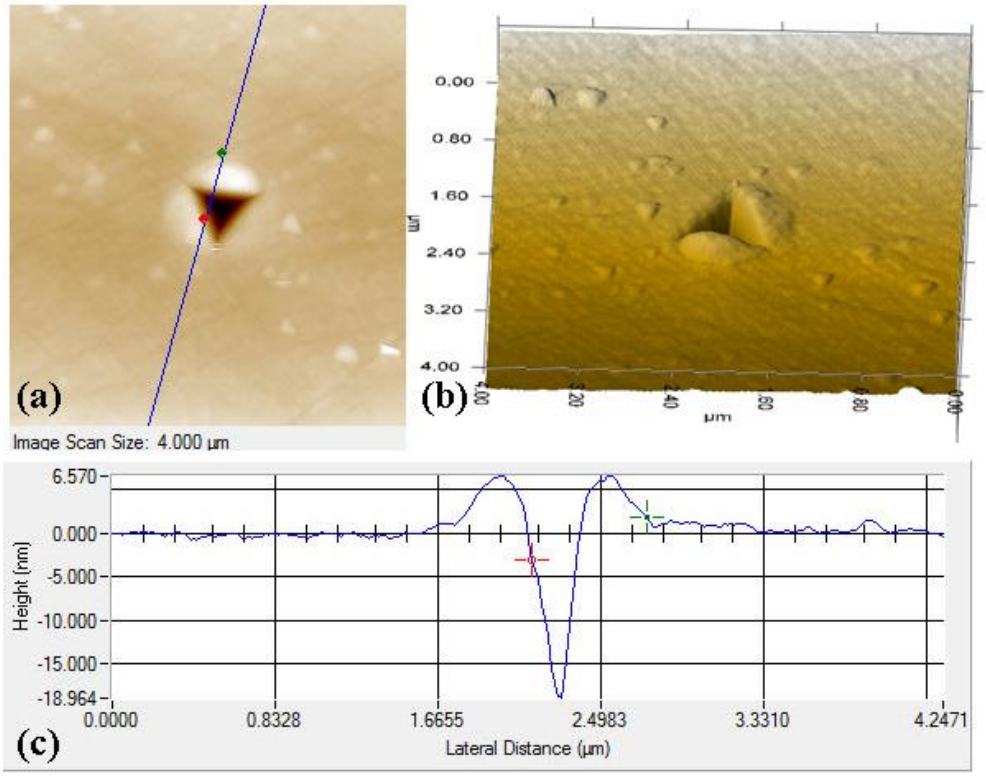

Figure 105. Indent 10 - DER - (a) AFM image from Figure 104. (b) 3D image of indent (c) height (nm) vs lateral distance (µm) indicating the height of pile up=6.57 nm.

Pile-up can be observed from Figure 105 near the edge of the indent, therefore the use of correction factor is justified.

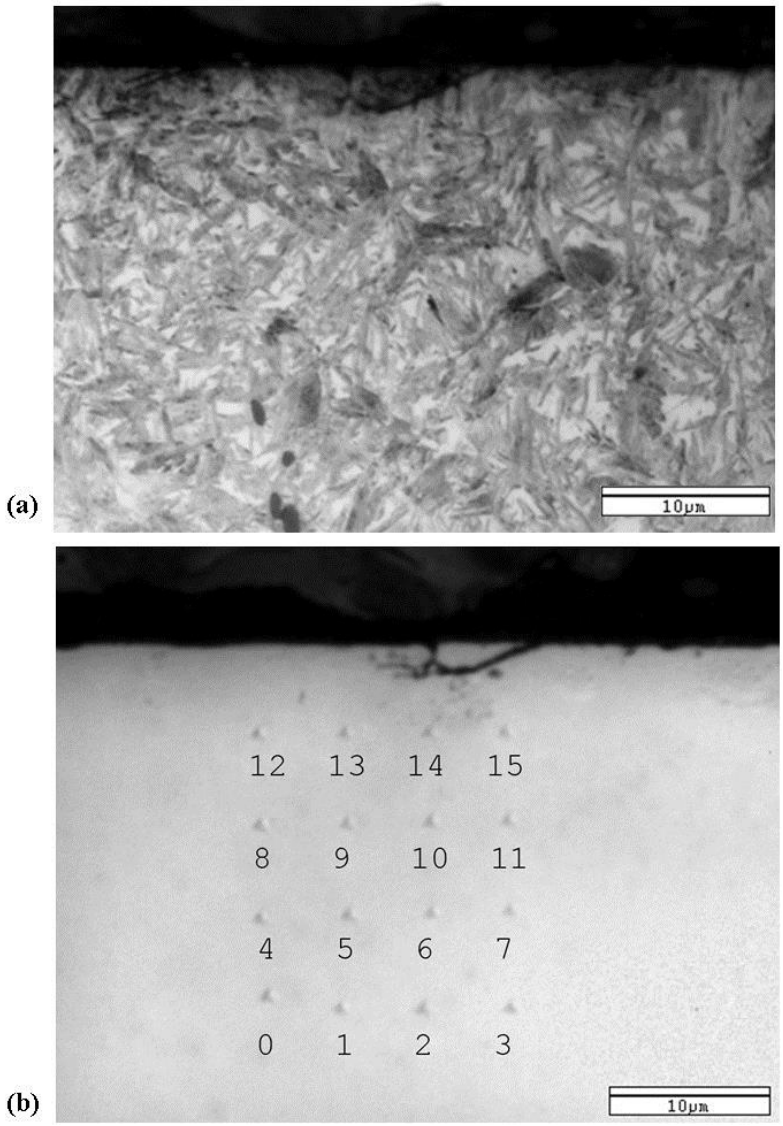

Figure 106. (a) DER on sample B29. (b) Location of nanoindentation outside DER.

Table 15. Nanoindentation data corresponding to the indentation shown in Figure 106 (corrected).

Indentation	Hardness(GPa)	E(GPa)
0	12.09	158.4
1	12.81	159.2
2	12.35	160.3
3	11.83	161.1
4	12.42	163.0
5	11.70	162.6
6	11.96	164.2
7	10.40	160.0
8	10.47	157.3
9	11.57	162.7
10	11.90	164.8
11	10.92	153.8
12	12.09	156.8
13	11.77	158.7
14	11.31	160.5
15	12.09	159.1

The results show that martensite has an average hardness of 11.73 GPa and an average elastic modulus is 163.1 GPa.

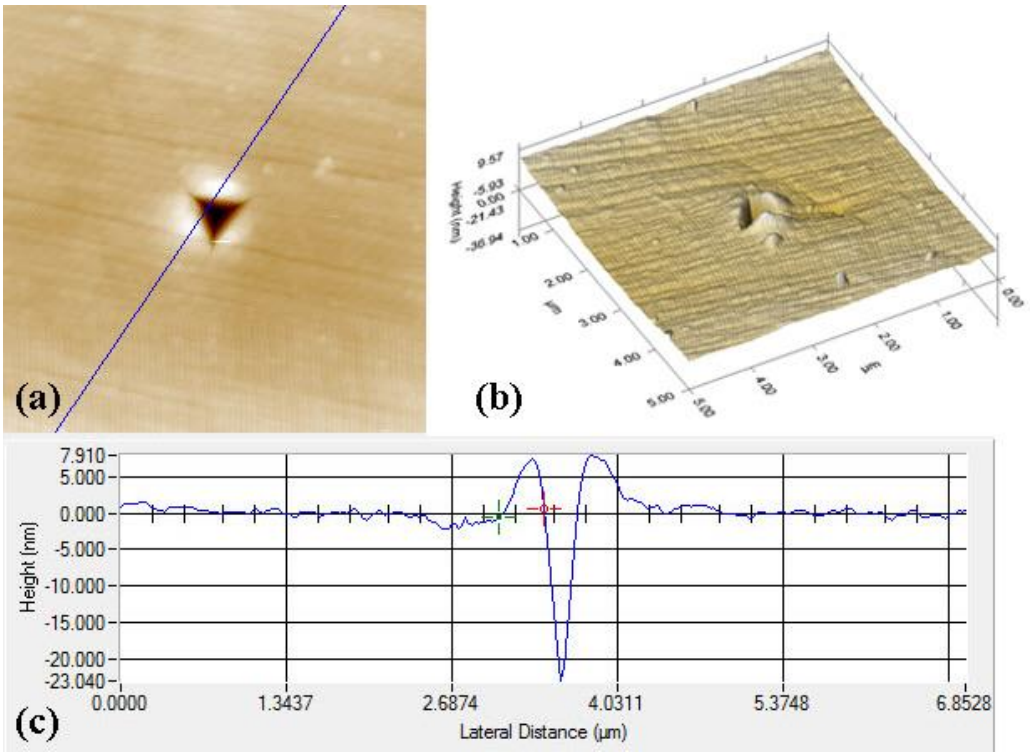

Figure 107. Indent 12 - (a) AFM image from Figure 106. (b) 3D image of indent (c) height (nm) vs lateral distance (µm) indicating the height of pile up up to=7.91 nm.

The pile-up can be observed from the color difference near the edge of the indent mark, therefore the use of correction factor is justified.

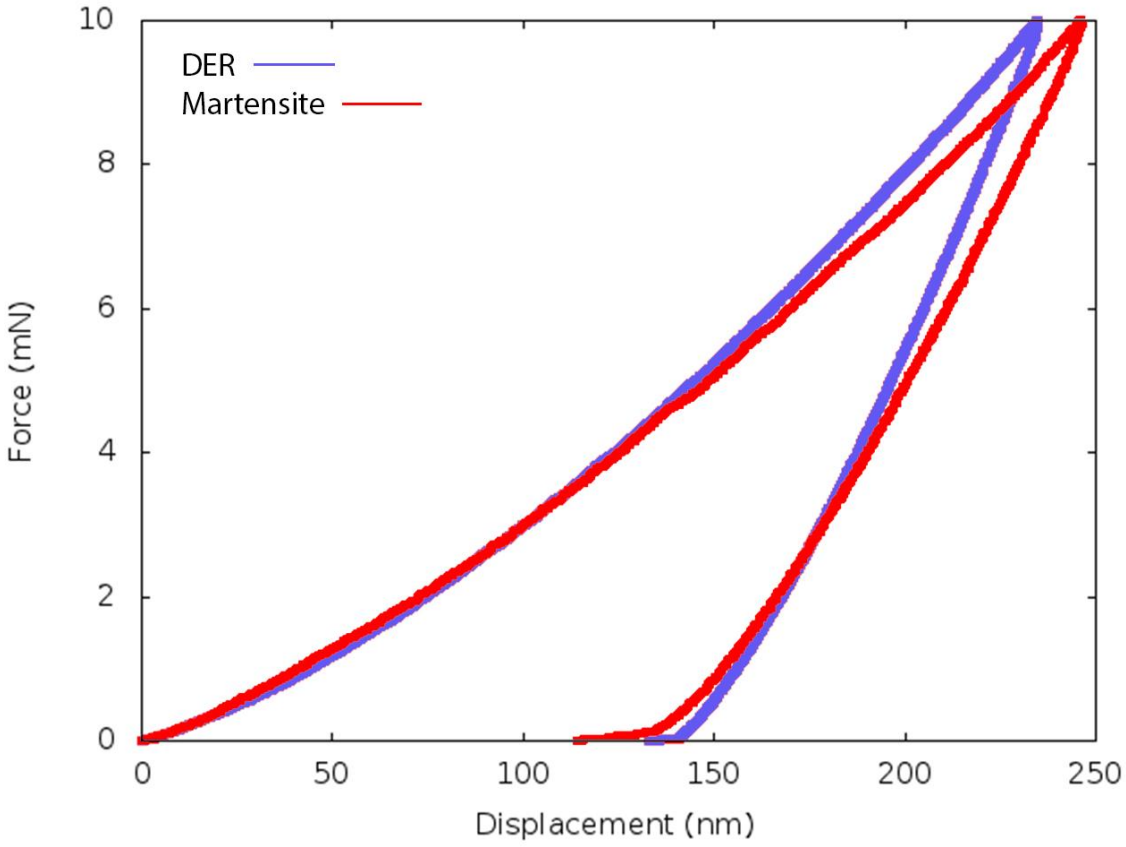

Figure 108. Load displacement curve of DER and martensite.

The load displacement curves from Figure 108 show the behaviour of the two phases during loading and unloading (when the tip of the indenter is removed). The displacement during loading is larger for martensite. When the indenter is pulled away, the recovery portion of the graph indicates that the stiffness of DER is higher compared to that of the martensite.

7.7 Transmission Electron Microscopy (TEM)

TEM investigations were carried out on the DER region marked by the red rectangle in Figure 109.

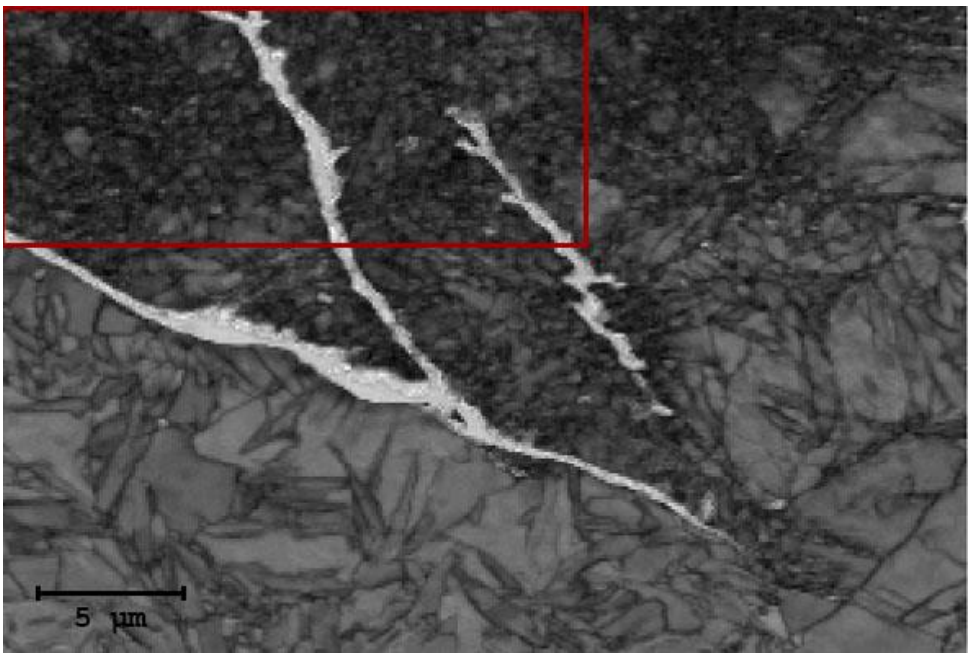

Figure 109. EBSP map with region marked for TEM investigation.

Focused Ion Beam milling (FIB) was used in order to extract the specific area highlighted in Figure 109 as described below.

The FIB Quanta 200 3D system was used to extract a sample wafer (10 µm × 25 µm) and attach to a TEM grid. The sample was placed in the sample chamber and pump until the vacuum is less than 0.001 Pa. The ion beam was set to 30 kV with 10 pA. The lower beam current are used for good resolution imaging (10 pA). The high beam current was used for milling (0.3 nA). The stage is tilt to 52° to centred it since other angle will require beam shift for both electron and ion beam.

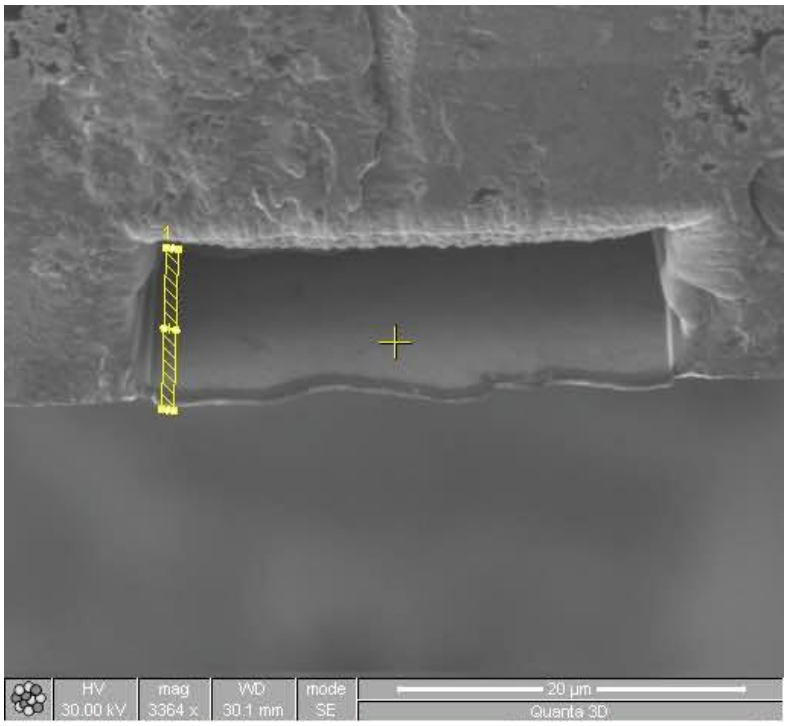

Figure 110. Ion beam was used to mill through the sample to extract the sample with current set to 0.3 nA. The yellow area is the location selected to mill through.

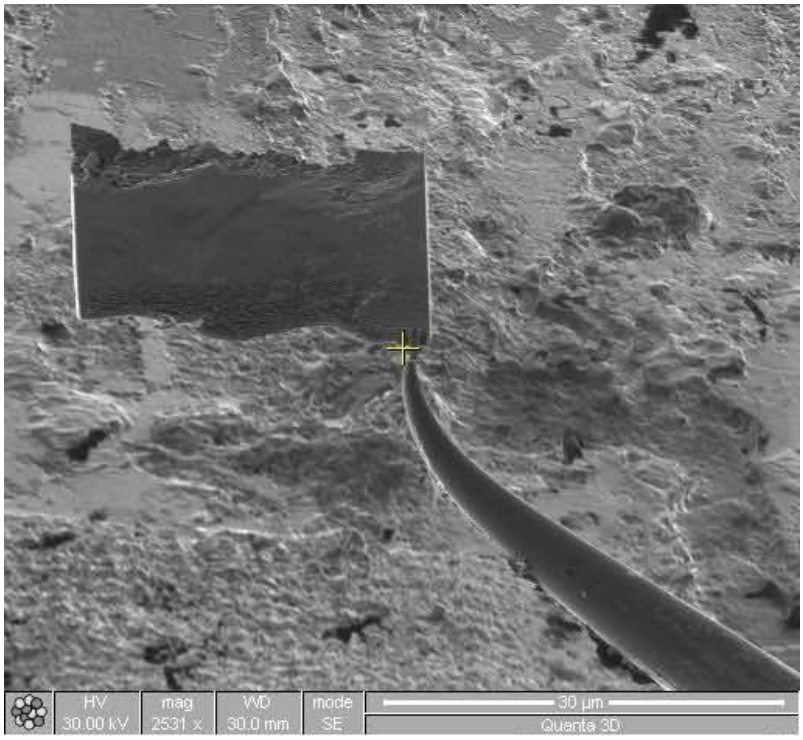

Figure 111. Omniprobe tungsten needle used to attach the sample for extraction. Carbon deposition was used to attach the sample to the probe (parameters: 30 kV, 0.30 nA).

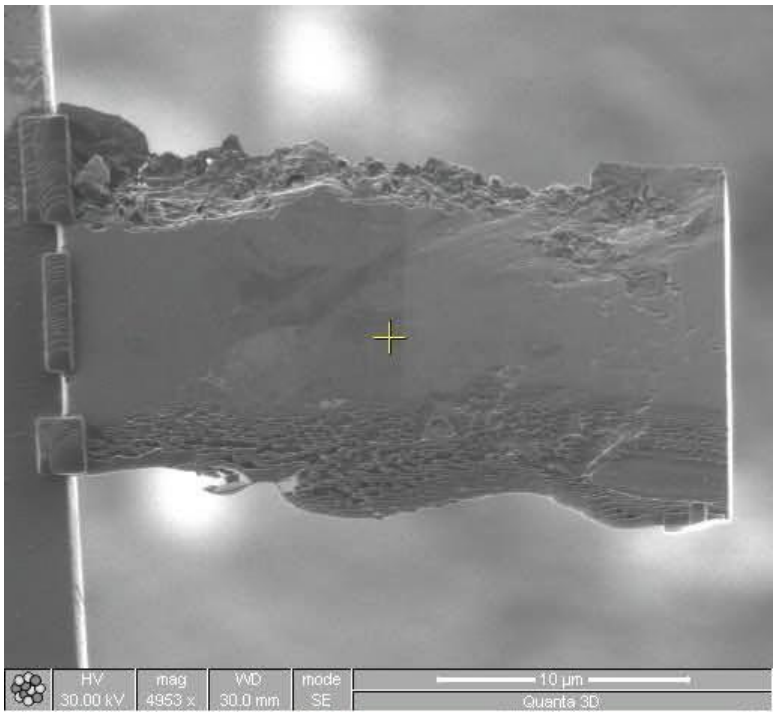

Figure 112. Three locations were chosen for carbon deposition for attaching the sample onto the copper grid (on the left).

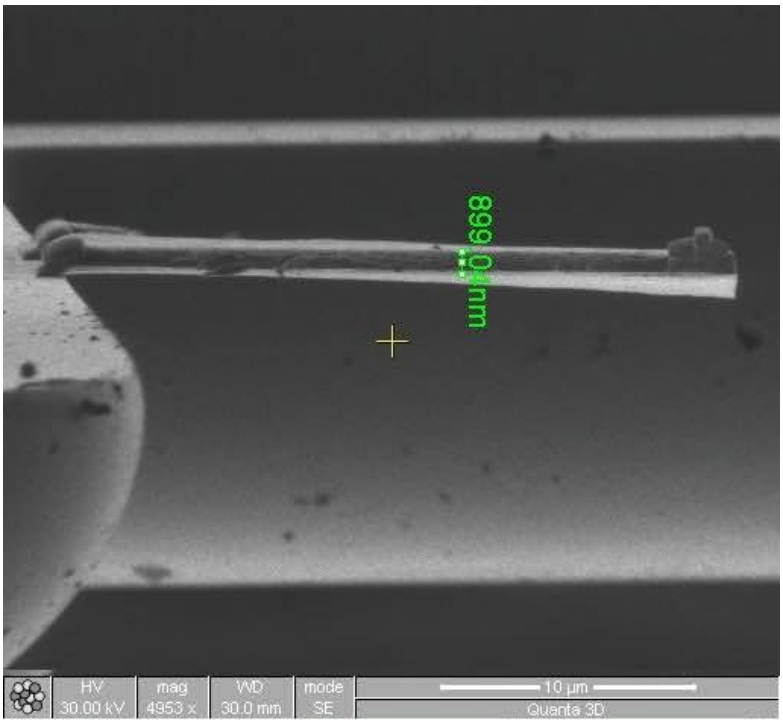

Figure 113. Side view of the sample attached to the copper grid.

Figure 113 shows that the initial thickness of the sample is around 900 nm. In order for the sample to be electron transparent, the thickness needs to be less than 100 nm for a successful TEM analysis.

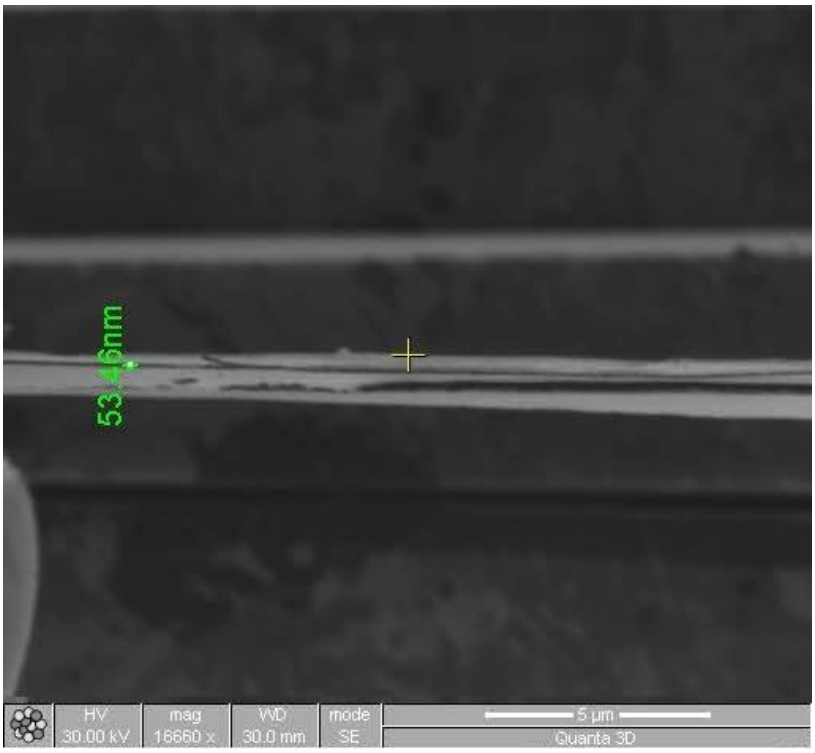

Figure 114. Final thickness of the sample after further milling.

The final thickness was around 53.46 nm, which is less than the recommended 100 nm thickness. Thinning was done by tilting the sample within ±1° and very carefully mill away the sample without cutting the sample in half with current set to 0.1 nA.

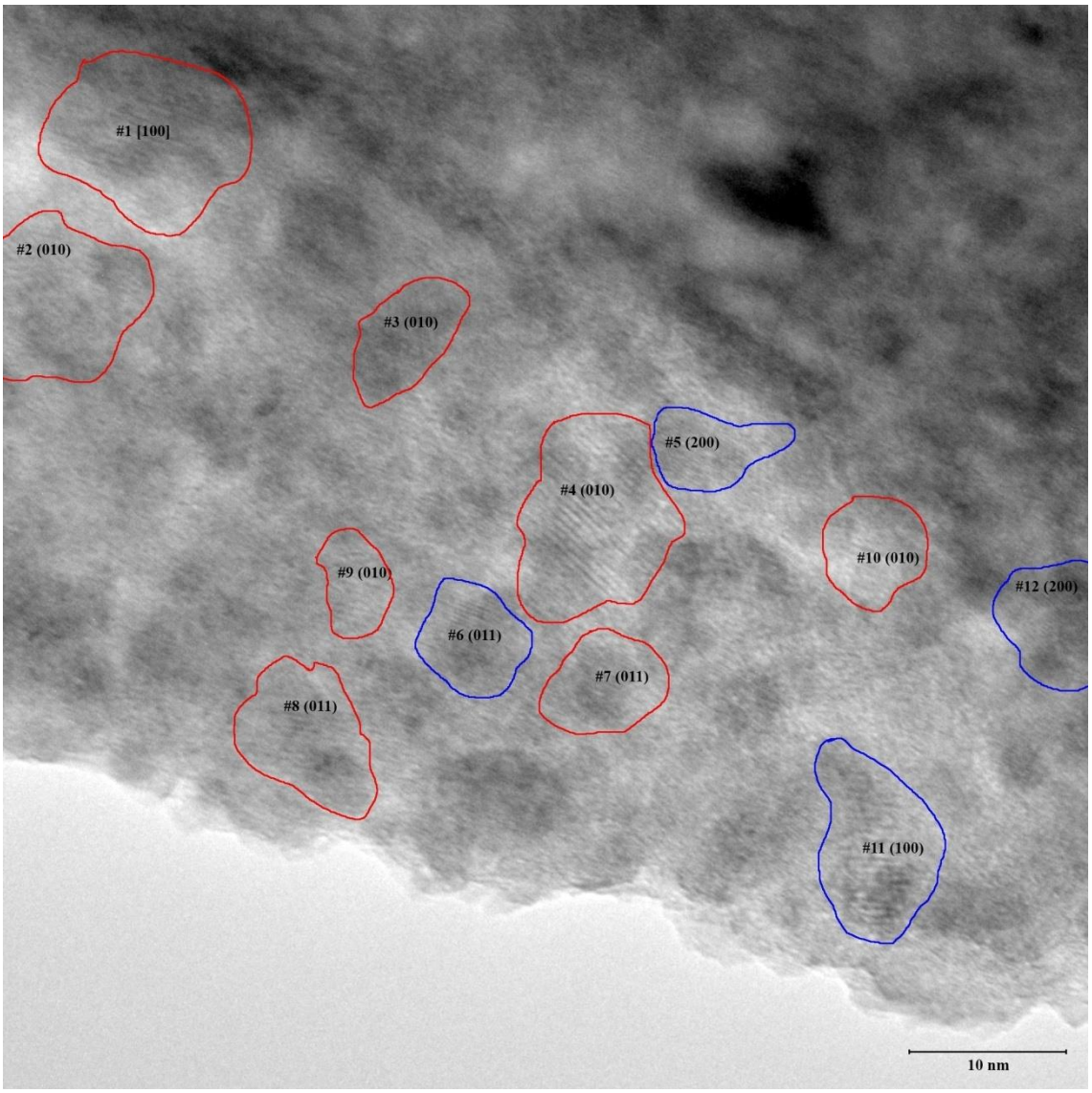

Figure 115. HRTEM of sample AC9P at location A (Red-cementite, blue-η-Fe_2C).

Figure 115 shows an image from the HRTEM of the sample. The lattice fringes (interatomic spacing) were measured using the software Image J. The results were compared with the theoretical interplanar distances for martesite, austenite, cementite and the transition carbides epsilon and eta. In order to perform this calculations a small FORTRAN program was written and compiled. The FORTRAN code is presented in Appendix 4 and Appendix 5. The matching criterion used was that the difference between the measured value and the theoretical (d_{hkl}') one must be less than 0.3. From these data, various grains were identified as η-Fe_2C and cementite in Table 16.

Table 16. The difference between the measured value and the theoretical interatomic spacing at location A.

Grain Number	Grain	d_{hkl} theoretical (Å)	d_{hkl} experimental (Å)	d_{hkl}' (Å)	Percent difference (%)
1	Fe_3C	$d_{(010)}$ =4.530000	4.380000	0.1500	3.3
2	Fe_3C	$d_{(010)}$ =4.530000	4.253000	0.2770	6.1
3	Fe_3C	$d_{(010)}$ =4.530000	4.240000	0.2900	6.4
4	Fe_3C	$d_{(010)}$=4.530000	4.520000	0.0100	0.2
5	η-Fe_2C	$d_{(003)}$=2.256667	2.256000	0.0006	0.02
6	η-Fe_2C	$d_{(011)}$=2.424975	2.410000	0.0140	0.6
7	Fe_3C	$d_{(011)}$=3.764905	3.810000	0.0450	1.2
8	Fe_3C	$d_{(011)}$=3.764905	3.659000	0.1050	2.8
9	Fe_3C	$d_{(010)}$=4.530000	4.215000	0.3000	6.6
10	Fe_3C	$d_{(010)}$=4.530000	4.243000	0.2800	6.2
11	η-Fe_2C	$d_{(100)}$=4.704000	5.037000	0.3300	6.6
12	η-Fe_2C	$d_{(100)}$=2.352 000	2.352000	0.0000	0

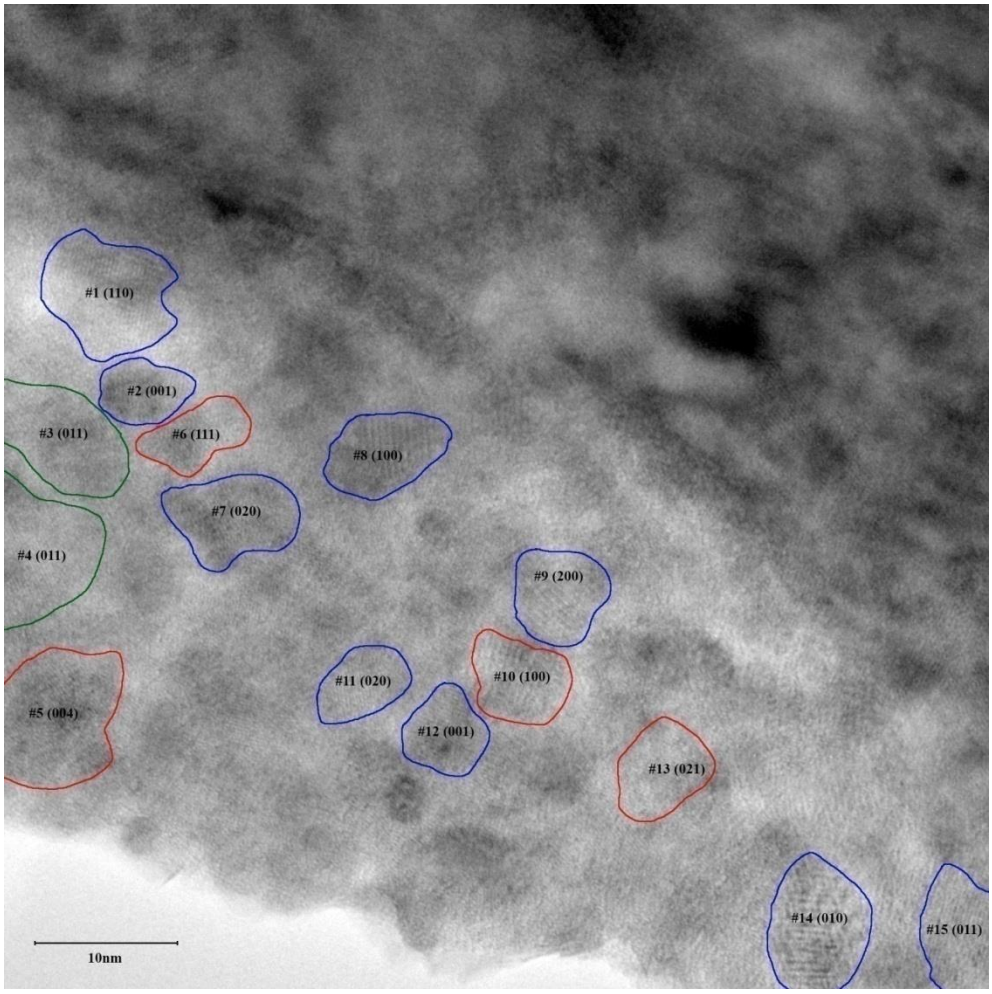

Figure 116. HRTEM of sample AC9P at location B (Red-cementite, Blue-η-Fe_2C, Green-martensite).

Figure 116 is another location chosen for the HRTEM analysis.

Table 17. The difference between the measured value and the theoretical (d_{hkl}') at location B.

Grain Number	Grain	d_{hkl} theoretical (Å)	d_{hkl} experimental (Å)	d_{hkl}' (Å)	Percent change(%)
1	η-Fe_2C	$d_{(110)}$ =2.715856	2.621000	0.0040	0.1
2	η-Fe_2C	$d_{(001)}$ =2.830000	2.827000	0.0030	0.1
3	Martensite	$d_{(011)}$ =2.053269	2.038000	0.0150	0.7
4	Martensite	$d_{(011)}$=2.053269	2.282000	0.0017	0.07
5	Fe_3C	$d_{(004)}$=1.692500	1.687000	0.0055	0.3
6	Fe_3C	$d_{(111)}$=2.439672	2.472000	0.0300	1.2
7	η-Fe_2C	$d_{(020)}$=2.352000	2.367000	0.0150	0.6
8	η-Fe_2C	$d_{(100)}$=4.704000	4.739000	0.0350	0.7
9	η-Fe_2C	$d_{(200)}$=2.352000	2.351000	0.0010	0.04
10	Fe_3C	$d_{(100)}$=4.530000	4.568000	0.0280	0.6
11	η-Fe_2C	$d_{(020)}$=2.352000	2.383000	0.0310	1.3
12	η-Fe_2C	$d_{(001)}$=2.830000	2.827000	0.0030	0.1
13	Fe_3C	$d_{(021)}$=2.147973	2.156000	0.0080	0.4
14	η-Fe_2C	$d_{(010)}$=4.704000	4.821000	0.0040	0.08
15	η-Fe_2C	$d_{(011)}$=2.424975	2.458000	0.0030	0.1

The results show the presence of martensite, cementite, and η-Fe_2C. η-Fe_2C is a transition compound found in the microstructure of quenched steels. Details about the role of η-Fe_2C in surface contact fatigue have been previously published (Oila et al., 2013). In the cited work it was found that the presence of η-Fe_2C will result in an increase in elastic modulus with no change in hardness of the steel. The change in elastic modulus recorded in this work could be attributed to precipitation of η-Fe_2C according to the HRTEM analysis.

Chapter 8. Results of Ab Initio Calculations

The results of the simulations have been compared to data published in the literature. The calculated lattice parameters are in agreement with those reported by other authors. The single crystal elastic constants of η-Fe_2C are presented in Table 18.

Table 18. Single crystal elastic constants of martensite [Units: GPa].

C_{ij}	This Study (0.9% carbon)	Experimental (Kim & Johnson, 2007), SAE 1050 carbon steel. (0.5% carbon)
C_{11}	237.9	268.1
C_{12}	143.3	111.2
C_{33}	227.3	267.2
C_{44}	97.9	79.06
C_{55}	97.9	78.72
C_{66}	84	78.85
C_{12}	143.3	111.2
C_{13}	143.7	110.2
C_{23}	143.0	111.0

The bulk *B* and shear modulus *G* for Reuss procedure (for uniform stress) can be calculated by equation 63 and 64.

$$B_{Reuss} = \frac{1}{s_{11} + s_{22} + s_{33} + 2(s_{12} + s_{23} + s_{13})} \tag{63}$$

$$G_{Reuss} = \frac{15}{4(s_{11} + s_{22} + s_{33} - s_{12} - s_{23} - s_{13}) + 3(s_{44} + s_{55} + s_{66})} \tag{64}$$

The bulk and shear modulus in the Voigt approach (for uniform strain) are

$$B_{Voigt} = \frac{c_{11} + c_{22} + c_{33} + 2(c_{12} + c_{23} + c_{13})}{9} \tag{65}$$

$$G_{Voigt} = \frac{c_{11} + c_{22} + c_{33} - c_{12} - c_{23} - c_{13}}{15} + \frac{c_{44} + c_{55} + c_{66}}{5} \quad (66)$$

The Hill's average can be determined from Voigt and Reuss moduli.

$$B = \frac{B_{Reuss} + B_{Voigt}}{2} \quad (67)$$

$$G = \frac{G_{Reuss} + G_{Voigt}}{2} \quad (68)$$

The elastic modulus *E* and Poisson's ratio *v* can be calculated by the connection with the bulk modulus B and the shear modulus G.

$$E = \frac{9BG}{3B + G} \quad (69)$$

$$\nu = \frac{(3B) - G}{2(3B + G)} \quad (70)$$

Table 19 contain the data for shear modulus, bulk modulus, Young's modulus, and Poisson's ratio (*v*).

Table 19. Shear modulus, Bulk modulus, Young's modulus, and Poisson ratio comparison.

Source	This study (simulation)	Experimental (Kim & Johnson, 2007)	This study (Nanoindentation)
B (GPa)	173.8	163.1	-
G (GPa)	65.36	78.74	-
E (GPa)	174.2	203.5	163.1
v	0.31	0.29	-

The elastic modulus predicted from ab-initio calculations is close to the experimental value obtained from nanoindentation. The lattice parameter of martensite obtained from the simulation are compared with various experimental data in Table 20.

Table 20. Lattice parameters of martensite

Source	*a* (Å)	*b* (Å)	*c* (Å)
This study	8.536	8.536	8.887
Experimental (Wechsler, 1953)	8.569	8.569	8.863
Experimental (Adeline et al., 2014)	8.609	8.609	8.754
Experimental (Bhattacharya et al., 2011)	8.577	8.577	8.811

Chapter 9. Finite Element Analysis

Finite Element is a method to solve systems of equations on a grid so that it can be used to determine the spatial variation of properties given certain boundary conditions. The finite element technique allows the simulation of material behaviour with complex geometric shape or containing different phases. ANSYS is an engineering simulation software developed by Ansys Inc.

In structural simulations the problem consists of solid bodies with some environmental conditions such as loads, and the response can be described by displacements, stresses, or strains. Meshes are created in combination with material properties.

Elastic stress analyses in rolling contact problems using FEM were conducted by various research groups (Bhargava et al., 1985; Jiang et al., 2002; Dahlberg et al., 2009). Recent studies on the finite element modelling of rolling contact provides information about the cyclic plastic deformation under rolling contact for linear kinematic-hardening material behaviour (Tunca & Laufer, 1987). The finite element model for this application describes a cylinder rolling over a surface used to study the stress field and crack growth behaviour observed in a homogeneous structure. Tunca & Laufer (1987) used a layered structure starting with white etching region follow by DER and then the unaffected region. The thickness of the layers is determined from experimental observations. The material properties used in the simulation were determined experimentally from hardness measurements. The results show that sliding can lead to the development of a layered structure and the hard layer structure on the surface increases the wear resistance. The DER serves as a plastic work absorber and is capable of accommodating plastic deformation without fracture and so crack initiation occurs around the DER. The crack growth occurs primarily by shear (Tunca & Laufer, 1987). Only in the tensile zone the crack propagation occur by a combination of tension and shear and during shear crack is

propagated in the direction of the DER structure (Tunca & Laufer, 1987). Under combined rolling and sliding, the required stress to cause failure decreases.

The two-dimensional rolling contact can be simulated by applying a Hertzian pressure over a finite element model surface. The model consists of meshes and skeletons. Most of the studies to date include ball on flat, ellipse on flat, cylinder on flat, or cylinder on cylinder surfaces. Some of these studies use FEA to study the deformation and residual stresses (Ismail, 2010). Increasing the load results in larger deformation of asperities and the residual stress introduces deformations (Ismail, 2010).

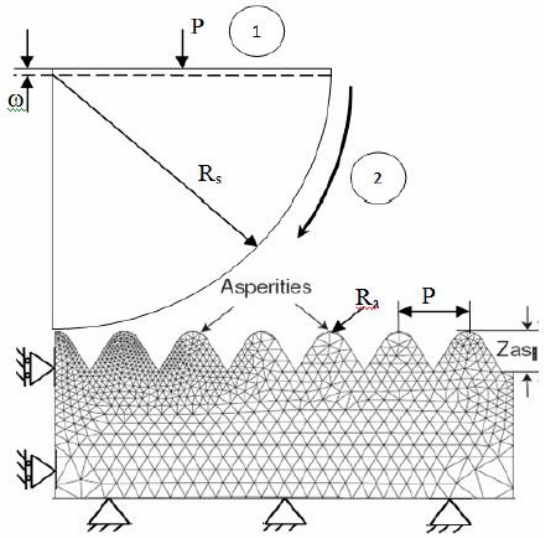

Figure 117. Rolling contact model (Rifky & Ismail, 2010).

According to the results by Rifky Ismail (Ismail, 2010) the elastic relaxation of the boundaries of the model reduces the amount of plastic deformation under the contact. The effects of boundary conditions and residual stress show that rigid boundaries yield 60% larger residual displacements and stresses. It was also found that the adequate mesh refinement can improve results and computational time.

The material is simulated in such a way that each pixel contains the assigned material properties by generating pixel groups and a material map. The properties are assigned to

the skeleton that profiles the mesh geometry. Refining the meshes will increase the number of elements and increase the calculation time.

Finite element analysis was performed with the finite element program Ansys. A mesh was created with material properties assigned to the elements. A 3D material model was created from a sketch. In Ansys a body is made of one kind of material and is the basic building block for the model. A part is a collection of the same type of bodies. All bodies in a part are assumed to bond together. Once the 3D model was created, a mesh was generated considering that the finer the mesh, the more accurate the solution.

The finite element mesh was 500 x500 μm and it contained quadrilateral plane strain elements in order to simulate the contact area. The grinding marks are considered as 50 microns diameter half spheres which is the average size and shape of the grinding marks observed on the gear tooth surface. The applied load was set to an average of 1500 MPa corresponding to the actual contact pressure during the gear tooth contact.

Finite element method was used to model gear behaviour of gear tooth subjected to surface contact fatigue. The contact area is set to 1000 μm by 500 μm with the contact stress of 1.5 GPa simulating the contacting condition of the gear tooth. Pressure was applied at an angle of 45°. The grinding marks are represented with a semicircle with a diameter of 50 μm.

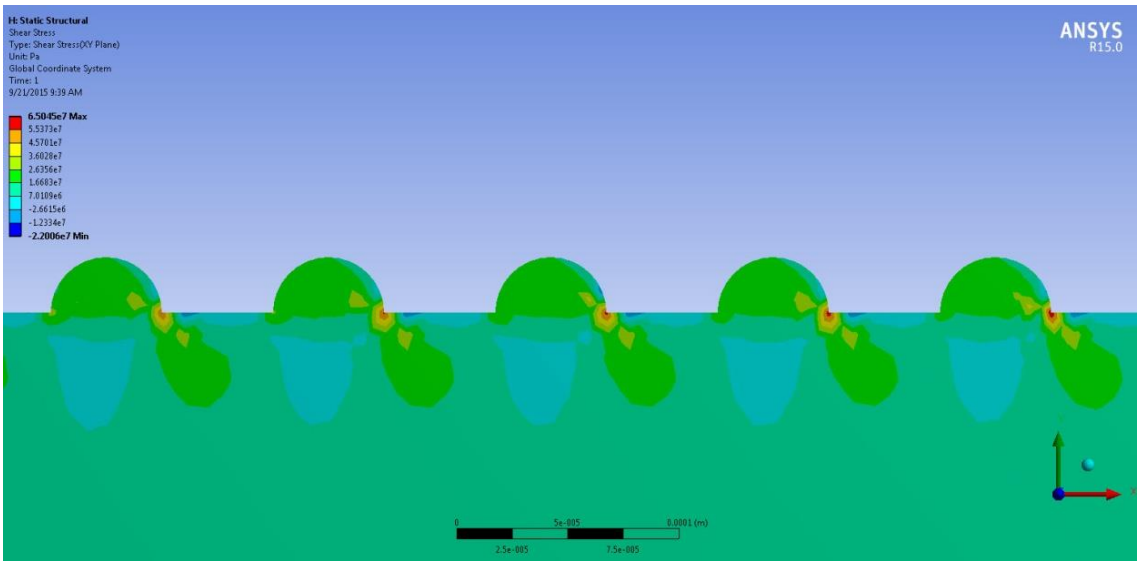

Figure 118. Model of gear teeth surface with the distribution of shear stress (Pa).

High shear stress occur at $6.5045*10^6$ Pa at the edge of the grinding marks. The increase number of cyclic stress generates dislocations in the high stress region. The region below the grinding mark is exposed to contact stresses of $2.824*10^3$ Pa. The area of the shear stress underneath each grinding mark are around 40µm ×40µm.

Once the grinding mark on the gear teeth wear out and the DER phase appears below the surface. Figure118 is a simplified model used to observe the stress distribution when the DER is present as observed in the microscopy data. The exact same condition was applied with the exception of the removal of grinding marks and the addition of the DER phase as shown in figure 119. The mechanical properties of martensite were obtained from the computational simulation, and the properties of DER obtained from nanoindentation.

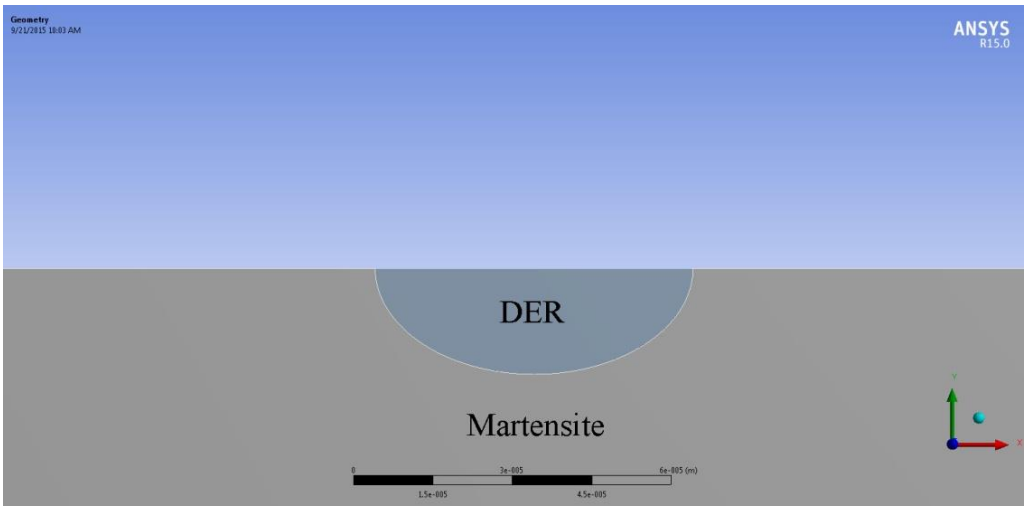

Figure 119. Model of gear teeth surface with the presence of DER

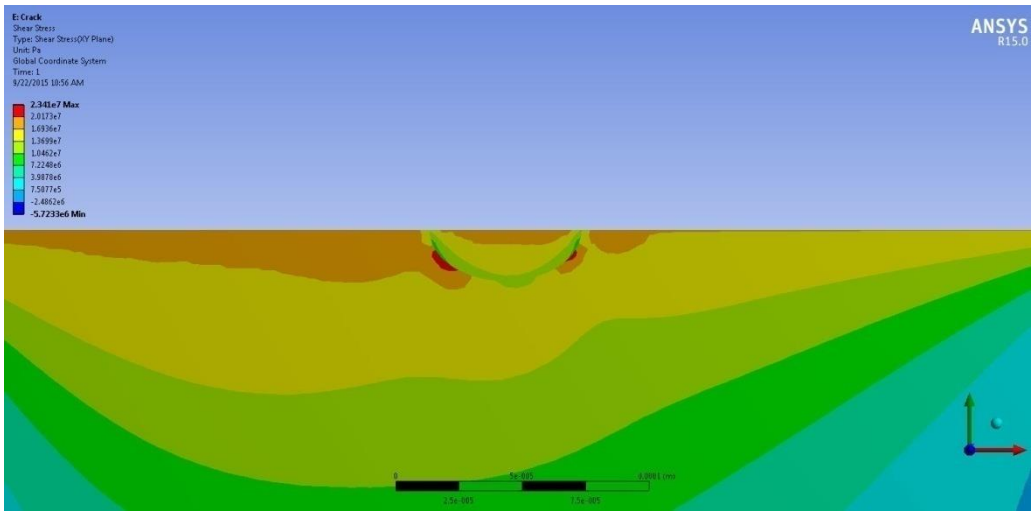

Figure 120. Model of gear teeth surface with the presence of DER with the distribution of shear stress (Pa).

The result shows that the life of the gear is $2.44*10^5$ cycles. The failure criteria occurs when the shear stress reaches the shear strength of the gear at the red region in the figure at the boundaries between DER and martensite.

Chapter 10. Conclusions and future work

In this study, the effect of lubricant viscosity, the effect of cryogenic treatment, the effect of coatings and the effect of nanoparticles added to lubricants on micropitting have been investigated.

The martensite decay associated with micropitting has been investigated by material characterization techniques and the stresses causing micropitting have been modelled.

The electronic structure, lattice parameters and the elastic constants of Fe-C martensite have been calculated using Density Functional Theory and the results used as input in the Finite Element model.

The microstructure of the micropitted samples was observed by reflective light microscopy in cross sections.

The samples showing the least amount of micropitting damage and also the absence of martensite decay are the samples coated with Balinit C and CrN respectively as well as the samples tested in 85W140 containing nanoparticles.

The sample showing the highest amount of micropitting damage was the disk lubricated with FRL-46 containing nanoparticles and the disk lubricated with 85W140 lubricant.

The Dark Etching Regions observed in gears have similar characteristics with those reported in ball bearings but, due to the difference in the contact stress patterns, DER in gears occur much closer to the surface and they form beneath the grinding marks.

The grinding marks deform plastically and martensite decay occurs due to the high strains in these regions. The Electron Backscatter Diffraction study and the Transmission Electron Microscopy investigation of these transformed regions revealed that the initial

tempered martensite ‘decays’. The grains in the Dark Etching Regions are of nanometric scale, cementite but also transition carbides precipitate from the initial tempered martensite. These transformations result in regions with mechanical properties different than those of the initial tempered martensite as the nanoindentation data confirm. The mismatch in mechanical properties (strength and stiffness) favours the initiation and propagation of fatigue cracks at the boundary of the Dark Etching Region as our Finite Element model suggests.

The change in stiffness between DER and martensite provides a preferential path for fatigue crack propagation along the DER boundary. The crack initiates from the surface of the gear and propagates along the DER boundary, it then follows the boundary of DER and propagates back to the surface. The crack continue to propagate until a micropit is formed.

For future research, focused metallurgical study on the DER and crack behaviour may also useful. New coatings and surface treatments that reduce the development of DER during operation are also worthy of further research.

References

Ackland, G. J., Bacon, D. J., Calder, A. F., & Harry, T. (1997) Computer simulation of point defect properties in dilute Fe-Cu alloy using a many-body interatomic potential. *Philosophical Magazine A* 75(3):713-732.

Adam J. Schwartz, Mukul Kumar, and Brent L. Adams. (2000) *Electron Backscatter Diffraction in Material Science*. New York: Kluwer Academic.

Akamatsu, Y. (1989) Peeling damage due to rolling contact fatigue (No. 891909). SAE Technical paper.

Al-Tubi, I. S., Long, H., Zhang, J., & Shaw, B. (2015) Experimental and analytical study of gear micropitting initiation and propagation under varying loading conditions. *Wear*: 328-329.

Al-Tubi, I. S., Long, H. (2013) Prediction of wind turbine gear micropitting under variable load and speed conditions using ISO/TR 15144-1: 2010. Proceedings of the Institution of Mechanical Engineers, Part C: *Journal of Mechanical Engineering Science* 227: 1898-1914.

Amarnath, M., Chandramohan, S., Seetharaman, S. (2012) Experimental investigations of surface wear assessment of spur gear teeth. *Journal of Vibration and Control* 18: 1009-1024.

Andris Gulans, Stefan Kontur, Christian Meisenbichler, Dmitrii Nabok, Pasquale Pavone, Santiago Rigamonti, Stephan Sagmeister, Ute Werner, & Claudia Draxl. (2014) Exciting - a full-potential all-electron package implementing density-functional theory and many body perturbation theory. *Journal of Physics: Condense Matter* 26, 363

Antoine, F., Besson, M. (2002) Simplified modellization of gear micropitting. Proceedings of the Institution of Mechanical Engineers, Part G: *Journal of Aerospace Engineering* 216(6): 291-302.

ASTM E 1150-1987. (1995) *Standard Definitions of Fatigue*.

ASTM. (1995) *Annual Book of Standards.* p 753-762

ASM. (2001) *Annual Book of Standards.*

Bain, E. C., Dunkirk, N. Y. (1924) The nature of martensite. *Transaction AIME 70*(1): 25-47.

Bain, E. C. (1936) On naming the aggregate constituents in steel. *Transaction ASM* 24: 225-261.

Baldissera P., Delprete C. (2008) Deep cryogenic treatment: A bibliographic review. *Open Mechanical Engineering Journal* 2:1–11

Baldissera P. (2009) Fatigue scatter reduction through deep cryogenic treatment on the 18NiCrMo5 carburized steel. *Materials* 30(9):3636–3642.

Baldissera P., Delprete C. (2009) Effects of deep cryogenic treatment on static mechanical properties. *Material Design* 30(5):1435–1440.

Balkanski M., Wallis, R. F. (2000) *Semiconductor Physics and Applications*. Oxford University Press.

Balonek, Gregory. (2013) Coating a grating structure using various deposition techniques. University of Rochester. Institute of optics.

Barth U. V., Hedin L. (1972) A local exchange-correlation potential for the spin polarized case: I. *Journal of Physics C: Solid State Physics* 5:1629

Becker, P.C., Swahn, H., Vingsbo, O. (1990) Modifications structurales provaqueses par la fatigue de contact dans des roulements a billes. *Mecanique Materiaux electricite*: 8-14.

Bensely A., Prabhakaran A., Mohan Lal D., Nagarajan G. (2006) Enhancing the wear resistance of case carburized steel (En 353) by cryogenic treatment. *Cryogenics* 45(12):747–754.

Bensely A., Senthilkumar D., Harish S., Mohan Lal D., Nagarajan G., Rajadurai A., Paulin P. (2011). Cryogenic treatment of gear steel. *Gear Solutions* 37–51.

Bensely A., Senthilkumar D., Mohan Lal D., Nagarajan G., Rajadurai A. (2007) Effect of cryogenic treatment on tensile behavior of case carburized steel-815M17. *Material Characterisation* 58(5): 485–491.

Bensely A., Shyamala L., Harish S., Mohan Lal D., Nagarajan G., Junik K., Rajadurai A. (2009) Fatigue behaviour and fracture mechanism of cryogenically treated EN 353 steel. *Material Design* 30(8):2955–2962.

Bensely A., Venkatesh S., Mohan Lal D., Nagarajan G., Rajadurai A., Junik K. (2008) Effect of cryogenic treatment on distribution of residual stress in case carburized En 353 steel. *Material Science Engineering* A 479(1–2):229–235.

Berthe, D., Flamand, L., Foucher, D., Godet, M. (1980) Micropitting in Hertzian contacts. *Journal of Lubrication Technology* 102: 478-489.

Bhadeshia, H.K.D.H. (1983) Carbon content of retained austenite in quenched steels. *Metal Science* 17: 151-153.

Bhadeshia, H.K.D.H. (2003) Advances in the kinetic theory of carbide precipitation. *Materials Science Forum* 426: 35-42.

Bhadeshia, H. K. D. H., & Honeycombe, R. W. K. (2006). *Steels*. Butterworth-Heinemann, Oxford, third edition edition, 50009-1.

Bhargava, V., Hahn G.T., Rubin, C.A. (1990) Rolling contact deformation, etching effects, and failure of high-strength bearing steel. *Metallurgical Transactions* 21A: 1990-1929.

Bhattacharya, S., Dinda, G. P., Dasgupta, A. K., Mazumder, J. (2011) Microstructural evolution of AISI 4340 steel during Direct Metal Deposition process. *Materials Science and Engineering* A 528 2309–2318.

Brandao, J. A., Seabra, J. O., Castro, M. J. (2009) The prediction of micropitting initiation on gear teeth: a numerical model and its correlation with experimental data. *Mecanica Experimental* 17: 1-12.

Britton, R. D., Elcoate, C. D., Alanou, M. P., Evans, H. P., Snidle, R. W. (2000) Effect of surface finish on gear tooth friction. *Journal of Tribology* 112(1): 354-360.

Bronfin, B. M., Ovchinnikov, V. V., Shveikin, V. P., Amigod, G. G., & Emel'yanov, A. (1986) The Redistribution of Atoms During Strain Ageing of Two-Phase Ferrite-Martensite Steel. *Fiz. Metellurgy 61*(2): 354-360.

Brown, J. (1995) Big chill to extend gear life. *Power transmission Design*: 59-61.

BS 7848: 1996 *Gears - Wear and damage to gear teeth - terminology*. London: British Standard Institution.

Buchwald, J., Heckel, R.W. (1968) An analysis of microstructural changes in 52100 steel bearings during cyclic stressing. *Transactions of the ASM* 61: 750-756.

Buehler, M. J. (2008). *Atomistic modeling of materials failure*. Springer Science & Business Media.

Bush, J. J. G., Robinson, G. H. (1961) Microstructural and Residual Stress Changes in Hardened Steel Due to Rolling Contact. *Transactions of the ASM* 54: 390-412.

Calder, A. F., Bacon, D. J. (1993) A molecular dynamics study of displacement cascades in α-iron. *Journal of Nuclear Materials* 207: 25-45.

Callister, W. D. (2015) *Materials Science and Engineering*. John Wiley & Sons.

Cardoso, N. F. R., Martins, R. C., Seabra, J. H. O., Igartua, A., Rodriguez, J. C., Luther, R. (2009) Micropitting performance of nitrided steel gears lubricated with mineral and ester oils. *Tribology International* 42: 77-87.

Chadi, D.J. & Cohen, M.L. (1973) Special Points in the Brillouin Zone. *Physical Review* B(8): 5747-5753.

Cheng W., Cheng H. S., Mura T., Keer L. M. (1994) Micromechanics modelling of crack initiation under contact fatigue. *Transactions of the ASME* 116: 2-8.

Choi, B. Y., Bahng, G. W. (1998) Characterisation of microstructure and its effect on rolling contact fatigue of induction hardened medium carbon bearing steels. *Materials Science and Technology* 14(8): 816-821.

Choi, B.Y., Shin, J.H., Bahng, G.W., Yoon, K.B. (1996) Metallographic study on rolling contact fatigue of 0.44%C-1.71%Mn induction-hardened bearing steels. *Wear* 192: 1-5.

Chou, C.C., Lin, J.F. (1997) Tribological effects of roughness and running-in on oil-lubricated line contacts. *Pro. Institute of Mechanical Engineers* 211J: 209-221.

Cripps, F. (2011) *Nanoindentation*. Londonderry Drive. New South Wales.

Crook, A. W. (1957) Simulated Gear-Tooth Contacts: Some experiments upon their lubrication and subsurface deformations. *Prceedings of the Institute of Mechanical Engineers* 171:187-124.

Dahlberg, J., Alfredsson, B. (2009) Transient rolling of cylindrical contacts with constant and linearly increasing applied slip. *Wear* 266(1): 316-326.

Datsyshyn, O.P., Panasyuk, V.V. (2001) Pitting of the rolling bodies contact surface. *Wear* 251: 1347-1355.

Davis, J. R. (2005) *Gear Materials, Properties, and Manufacture*. ASM International.

D'errico, F. (2011) Micropitting Damage Mechanism on Hardened and Tempered, Nitrided, and Carburizing Steels. *Materials and Manufacturing Processes* 26(1): 7-13.

Dinner, H. (2011) Micropitting in Wind Turbine Gearboxes: Calculation of the Safety Factor and Optimization of the Gear Geometry". *Applied Mechanics and Materials* 86: 898--903.

Dirand M., Afqir L. (1983) Precise structural identification of carbides precipitated in steels at different stages of tempering. *Journal of Material Science* 49:2383–2390

Draheim, K.J., Gottstein, G. (1996) Simulation of grain boundary motion during high temperature cyclic deformation. *Computational Material Science* 7: 208-212.

Dudley, D.W. (1962) *Gear handbook: the design, manufacture, and application of gears*. McGraw-Hill.

Errichello, R. (2002). How to analyze gear failures. *Journal of Failure Analysis and Prevention 2*(6): 8-16.

Farhat, Z.N., Ding, Y., Alpas, A.T., Northwood, D.O. (1997) The processing and testing of new and advanced materials for wear resistant surface coatings. *Journal of Materials Processing Technology* 63: 859-864.

Fermi, E. (1927). Statistical method to determine some properties of atoms. *Accademia dei Lincei* 6: 602-607.

Fernandes, P.J.L. (1997) Contact fatigue in rolling-element bearings. *Engineering Fatigue Analysis* 4(2): 155.

Fernandes, P. J. L., McDuling, C. (1997) *Surface Contact Fatigue Failures in Gears*. Pergamon: Elsevier.

Formost Report. (2005) Fullerene-based opportunities for robust engineering: making optimised surfaces for tribology. European funded FOREMOST project (FP6-NMP3-CT-2005-515840).

Frozen Solid, Inc. (1998) World Wide Web: http://www.frozensolid.co.uk.

Furumura, K., Murakami Y., Abe T. (1996) Development of long life bearing steel for full film lubrication and for poor and contaminated lubrication. *Motion and Control* No.1: 30-36.

Gavallaro, G.P., Wilks, T.P., Subramanian, C., Strafford, K.N., French, P., Allison, J.E. (1995) Bending fatigue and contact fatigue characteristics of carburized gears. *Surface and Coating Technology* 71: 182-192.

Genel, K., Demirkol, M. (1999) Effect of case depth on fatigue performance of AISI 8620 carburized steel. *International Journal of Fatigue* 21: 207.

Genel, K., Demirkol, M. (2000) Effect of ion nitriding on fatigue behaviour of AISI 4140 steel. *Material Science Engineering* A: 279.

Gentile, A.J., Jordan, E.F., Martin, A.D. (1965) Phase transformations in high carbon, high hardness steels under contact loads. *Metallurgical Society of AIME* 233. 1085-1093.

Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G.L., Cococcioni, M., Dabo, I., Dal, Corso. A., De, Gironcoli, S., Fabris, S., Fratesi, G., Gebauer, R., Gerstmann, U., Gougoussis, C., Kokalj, A., Lazzer,i M., Martin-Samos, L., Marzari, N., Mauri, F., Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A,P., Smogunov, A., Umari, P., Wentzcovitch, R.M. (2009) QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of Physics Condense Matter* 21(39):395502-1–395502-19

Gourguess, A.F. (2002) Electron backscatter diffraction and cracking. *Material Science and Technology* 18: 119-131.

Grabulov, A., Petrov, R., Zandbergen, H.W. (2010) EBSD investigation of the crack initiation and TEM/FIB analysis of the microstructural changes around the cracks formed under rolling contact fatigue. *International Journal of Fatigue*: 576-583.

Graham, R. C., Olver, A. V., Macpherson, P. B. (1980) An investigation into the mechanisms of pitting in high-hardness carburised steels. *Transaction mechanical engineers*: 80-83.

Graham, R.C., Olver, A. V., Macpherson, P.B. (1981) An investigation into the mechanisms of pitting in high hardness carburized steels. *Design Engineering Division of the ASME*: 2-6.

Grassie, S. (2003) Rolling contact fatigue on the British railway system: treatment. *Wear* 258: 1310-1318.

Hahn, G. T., Bhargava, V., Rubin, C. A., Chen, Q., & Kim, K. (1987) Analysis of the rolling contact residual stresses and cyclic plastic deformation of SAE 52100 steel ball bearings. *Journal of tribology*, *109*(4): 618-626.

Hayakawa, M., Matsuoka, S., Tsuzaki, K. (2002) Microstructural analysis of grain boundary carbides of tempered martensite in medium-carbon steel by atomic force microscopy. *Japan Institute of Metals* 65: 734-741.

Hertz, H. (1881). On the contact of elastic solids. *Journal of reine angew Math* 92(110): 156-171.

Hirotsu, Y., Nagajura, S. (1972) Crystal structure and morphology of the carbide precipitated from martensitic high carbon steel during the first stage of tempering. *Acta Metallurgica* 20(4): 645-655.

Hisakado, T., Suda, H. (1999) Effects of asperity shape and summit height distributions on friction and wear characteristics. *Wear* 225-229: 450-457.

Hoehn, B. R., Stahl, K., Michaelis, K. (2012) Lubricant influence on slow speed wear in gears. *Goriva i Maziva* 51(1): 17-28.

Hoeprich, M. R. (2001) Analysis of micropitting of prototype surface fatigue test gears. *TriboTest* 7(4): 333-347.

Hogmark, S., Jacobson, S., Larsson, M. (2000) Design and evaluation of tribological coatings. *Wear* 246: 20-33.

Hohn, B. R., Michaelis, K. (2004) Influence of oil temperature on gear failures. *Tribology International* 37: 103-109.

Hohn, B. R., Oster, P. y Schrade, U. (2005) Studies on the micropitting resistance of case carburised gears Industrial application of the new calculation method. *Vdi Berichte*: 1287-1307.

Holmberg, K., Matthews, A. (2009) *Coatings Tribology: Properties, mechanisms, Techniques and Applications in Surface Engineering*. Amsterdam: Elsevier.

Homnberg, K., Ronkainen, H., Mattews, A. (2000) Tribology of thin coatings. *Ceramics international* 26: 787-795.

Honda, K., Nakano, K., Numakura, H., Yokoi, T., Maeda, D., Yoshinaga, N., & Ushioda, K. (2012). Quantitative analysis of carbon in ferrite in dual-phase steels by mechanical loss measurements. *Solid State Phenomena* 184: 87-91.

Honeycombe, R. W. K. (1995) *Steels: Microstructure and Properties*. London: Edward Arnold.

Iliuc, I. (2006) Wear and micropitting of steel ball sliding against TiN coated steel plate in dry and lubricated conditions. *Tribology International* 39(7): 607-615.

Ismail, R., Tauviqirrahman, M., & Schipper, D. J. (2010). Two-Dimensional Finite Element Analysis on Running-in of Elastic-Plastic Rolling Contact. *4TH ASIAN PHYSICS SYMPOSIUM AN INTERNATIONAL SYMPOSIUM* Vol. 1325(1): 190-193.

Izciler, M., Tabor, M. (2006) Abrasive wear behaviour of different case depth gas carburized AISI 8620 gear steel. *Wear* 260: 90-98.

Jiang, C., Srinivasan, S., Caro, A., Maloy, S. (2008) Structural, elastic, and electronic properties of Fe_3C from first principles. *Journal of Applied Physics* 103(4): 043502-1-043502-8.

Jiang, L., Wang, G. C., Guan, N. J., Wu, Y., Cai, Z. S., Pan, Y. M., Zhong, B. (2003) DFT Studies of CO Adsorption and Activation on Some Transition Metal Surfaces. *ACTA PHYSICOCHIMICA SINICA* 19(5): 393-397.

Jiang, Y., Xu, B. (2002) Elastic-plastic finite element analysis of partial slip rolling contact. *Journal of tribology* 124(1): 20-26.

Johnson, K.L. (1995) Contact mechanics and the wear of metals. *Wear* 190: 162-170.

Kabir, M., Lau, T. T., Rodney, D., Yip, S., & Van Vliet, K. J. (2010) Predicting dislocation climb and creep from explicit atomistic details. *Physical review letters* 105(9): 095501.

Khlestov, V.M., Konopleva, E.V., Mcqueen, H.J. (1998) Kinetics of austenite transformation during thermomechanical processes. *Canadian Metallurgical Quarterly* 37: 75-89.

Kim, S.A., Johnson, W.L. (2007) Elastic constants and internal friction of martensitic steel, ferritic-pearlitic, and α-iron. *Material Science and Engineering* A: 633-639.

Kim, T.H., Olver, A.V. (1999) Stress history in rolling-sliding contact of rough surfaces. *Tribology International* 31(12): 727-736.

Komvopoulos, K., Cho, S.S. (1997) Finite element analysis of subsurface crack propagation in a half-space due to a moving asperity contact. *Wear* 209: 57-68.

Krauss, G. (1995) Heat treated martensitic steels: microstructural systems for advanced manufacture. *ISIJ international* 35(4): 349-359.

Krauss, G. (1995) Microstructure and performance of carburized steel, part III: Austenite and fatigue. *Advanced Materials & Processes* 9.

Laine, E., Olver, A.V., Beveridge, T.A. (2008) Effect of lubricants on micropitting and wear. *Tribology International* 41(11): 1049-1055.

Lee, C., Yang, W., Parr, R. G., (1988) *Physics. Review* B: 785.

Lindahl, E.,Osterlund, R. (1982) 2D transmission electron microscopy applied to phase transformatiosn in ball bearings. *Ultramicroscopy* 9: 355-364.

Littmann, W.E., Widner, R.L. (1966) Propagation of contact fatigue from surface and subsurface origins. *Journal of Basic Engineering*: 624-637.

Long H., Al-Tubi, Issa S. (2013) Effects of operational conditions on wind turbine gear micropitting. *European Wind Energy Association*: 1923-1929.

Lundberg G., Palmgren A. (1947) Dynamic capacity of rolling bearings. *Acta Polytechnica Mechanical Engineering Series* 1(3): 196.

Magee, C. L., Davies, R. G. (1971) Influence of austenite and martensite strength on martensite morphology. *Metallurgical Transactions* 2(7): 1939-1947.

Manoj V, Gopinath K, Muthuveerappan G (2004) Rolling contact fatigue studies on case carburized and cryogenic treated EN353 gear material. *International symposium of research students on material science and engineering*:1–11.

Martin, J.A., Borgese, S.F., Eberhardt, A.D. (1966) Microstructural alterations of rolling bearing steel undergoing cyclic stressing. *Journal of Basic Engineering*: 555-567.

Marzari, N. & Vanderbilt, D. (1997) Maximally localized generalized Wannier functions for composite energy bands. *Physics Review* B 56: 12847-12865

Matlock, D. K., Wagar, R. M., Speer, J. G., Mendez, P. F. (2005) Examination of pitting fatigue in carburized steels with controlled retained austenite fractions. Colorado School of Mines paper (2006-01).

Melander, A. (1997) A finite element study of short cracks with different inclusion types under rolling contact fatigue load. *International Journal of Fatigue* 19(1): 13-24.

Miller, G.R., Keer, L.M., Cheng, H.S. (1985) On the mechanics of fatigue crack growth due to contact loading. *Proc. R. society, London* A397: 197-209.

Monkhorst, H. J., & Pack, J. D. (1976). Special points for Brillouin-zone integrations. *Physical Review B* 13(12):5188.

Murakami, Y., Nomoto, T., ueda, T. (1999) Factors influencing the mechanisms of superlong fatigue failure in steels. *Fatigue Fracture Engineering Structure* 22: 581-590.

Muro, H., Tsushima, T., Nagafuchi, M. (1975) Initiation and propagation of surface cracks in rolling fatigue of high hardness steel. *Wear* 35: 261-282.

Muroga, A., Saka, H. (1995) Analysis of rolling contact fatigue microctructure using focused ion beam sputtering and transmission electron microscopy observation. *Scripta Metallurgica et Materialia* 33: 151-156.

Nishikawa, S & Kikuchi, S. (1928) The Diffraction of Cathode Rays by Calcite. *Proc. Imperial Academy of Japan* 4: 475-477.

O'Brien, J.L., King, A.H. (1966) Electron microscopy of stress induced structural alterations near inclusions in bearing steels. *Journal of Basic Engineering*: 568-589.

Ohmori, Y., Tamura, I. (1992) Epsilon carbide precipitation during tempering of plain carbon martensite. *Metallurgical Transactions* A: 2737-2751.

Ohno, T., Fujinami, Y., Kobessho, M. (2013) The influence of boundary lubrication film on the micropitting of gears. *Politecnico di Torino (DIMEAS)*: 804-805.

Oila, A. & Bull, S. J. (2005) Phase transformations associated with micropitting in rolling/sliding contacts. *Journal of Material Science* 40(18: 4767-4774).

Oila, A. (2003) *Micropitting and Related Phenomena in Case Carburised Gears*. PhD Thesis, University of Newcastle-upon-Tyne, U.K.

Oila, A. & Bull, S. J. (2004) Assessment of the Factors Influencing Micropitting in Rolling/Sliding Contacts. *Wear*: 1510-1524.

Oila, A. (2010) *Martensite decay in gears*. Internal report.

Oila, A., Lung, C., Bull, S. J. (2014) Elastic properties of eta carbide from ab initio calculations: application to cryogenically treated gear steel. *Journal of Material Science* 49: 2383-2390.

Olver, A. V. (1984) Micropitting and asperity deformation. p. 319-323.

Olver, A. V., Dini, D., Laine, E., Beveridge, T. A., Hua, D. Y. (2007) Roughness and lubricant chemistry effects in micropitting. p. 151-160.

Osterlund, R. & Vingsbo, O. (1980) Phase changes in fatigued ball bearings. *Metallurgical Transactions* A (11A): 701-707.

Parrish, G. (1999) *Carburizing: microstructures and properties*. ASM International.

Perdew J. P., Burke K., Ernzerhof M., (1998) *Physics Review Letter* 80, 891-891.

Persson, B.N., Bucher, F., Chiaia, B. (2002) Elastic contact between randomly rough surfacesL comparison of theory with numerical results. *Physical Review* B. Vol 65: 184106.

Philip, C.C. (2001) The calculation of optimum surface carbon content for carburized case hardened gears. *Gear Technology*: 53-57.

Pietikainen, J. (1999) Consideration about tempered martensite embrittlement. *Materials Science and Engineering* A273: 466-470.

Pollack W. (1988) *Material Science and Metallurgy, 4th ed*. Prentice Hall.

Polonsky, I.A. (1994) On white etching band formation in rolling bearings. *Journal of mechanical Physics Solids* 43: 637-669.

Porter, D.A., Easterling, K.E. (1992) *Phase transformations in metals and alloys*. London-New York: Chapman & Hall.

Preciado M, Bravo P, Alegre J (2006). Effect of low temperature tempering prior cryogenic treatment on carburized steels. *Journal of Material Process Technology* 176(1–3):41–44

Rangaswamy, P., Scherer, C.P., Bourke, M.A.M. (2000) Experimental measurements and numerical simulation of stress and microstructure in carburized 5120 steel disks. *Material Science and Engieering* A298: 158-165.

Reynolds, O. (1886) On the theory of lubrication and its application to Mr. Beauchamp Tower's experiments, including an experimental determination of the viscosity of olive oil. *Philosophical Transactions of The Royal Society* 117: 157-234.

Ringsberg, J.W., Loo-Morrey, M., Josefson, B.L., Kapoor, A., Beynon, J.H. (2000) Prediction of fatigue crack initiation for rolling contact fatigue. *International Journal of Fatigue* 22: 205-215.

Roberts, C. S. (1953). Effect of carbon on the volume fractions and lattice parameters of retained austenite and martensite. *Trans. AIME,* 197(2): 203-204.

Rosado, L., Jain, V.K., Trivedi, H.K. (1997) The effect of diamond like carbon coatings on the rolling fatigue and wear of M50 steel. *Wear* 212: 1-6.

Sands, D. (1993) *Introduction to crystallography*. New York : Dover.

Sasada, T.,Nakabayashi, H. (1998) Does Hertzian contact area act as an effective zone generating the friction resistance. *IEICE Trans. Electronics* E81-C No.3: 326-329.

Senthikumar, D., Rajendran, I. (2014) A research review on deep cryogenic treatment of steels. *International Journal of Materials and Structural Integrity* 8(1): 169-183.

Shackelford, J. (2005) *Material Science for Engineers*. 6th edition. New Jersey: Pearson Education.

Shikata, H. (2003) Technological transition of oil-impregnated sintered bearings. *Hitachi Powdered Metals Technical Report* 2: 3-11.

Shotter, B.A. (1981) *Micropitting: Its characteristics and implications on the test requirements of gear oils*. The institute of Petroleum, London: 91-103.

Simonelli, G., Pasianot, R., & Savino, E. J. (1992) Embedded-atom-method interatomic potentials for bcc-iron. *MRS Proceedings* 291: 567

Slane, J. A., Wolverton, C., Gibala, R. (2004) Experimental and theoretical evidence for carbon-vacancy binding in austenite. *Metallurgical and Materials Transactions A* 35(8): 2239-2245.

Song, J. H., Noh, H. G., Yu, H. S., Kang, H. Y., & Yang, S. M. (2004). Estimation of fatigue life by lethargy coefficient using molecular dynamic simulation. *International Journal of Automotive Technology* 5(3): 189-194.

Spikes, H. A., Olver, A. V., Macpherson, P. B. (1986) Wear in rolling contacts. *Wear* 112(2): 121-144.

Stachowiak, G. W. (2014) *Engineering Tribology*. Butterworth-Heinemann.

Stratton P, Graf M (2009). The effect of deep cold induced nanocarbides on the wear of case hardened components. *Cryogenics* 49(7):346–349.

Suchentrunk, R., Fuesser, H.J., Staudigl, G. (1999) Plasma surface engineering – Innovative processes and coating systems for high quality products. *Surface and Coatings Technology* 112: 351-357.

Swahn, H., Becker, P.C., Vingsbo, O. (1976) Electron-microscope studies of carbide decay during contact fatigue in ball bearings. *Metal Science*:35-40.

Swahn, H., Becker, P. C., Vingsbo, O. (1976) Martensite decay during rolling contact fatigue in ball bearings. *Metallurgical Transactions* A: 1099-1109.

Swahn, H., Hellner, L., Andersson, C.G., Zetterlund, B., Vingsbo, O. (1975) Phase transformations and residual stress changes in steel rolls during contact fatigue. *Scandanavian Journal of Metallurgy* 4: 263-266.

Tallian, T. E. (1992) Simplified contact fatigue life prediction model-Part II: new model. *Journal of tribology* 114(2): 214-220.

Tapasa, K., Barashev, A. V., Bacon, D. J., & Osetsky, Y. N. (2007) Computer simulation of the interaction of carbon atoms with self-interstitial clusters in α-iron. *Journal of nuclear materials* 361(1): 52-61.

Tribe, F.J. (1983) Seawater lubricated mechanical seals and bearings associated materials problems. *Lubrication engineering*: 292-298.

Tunca, N., Laufer, E.E. (1987) Wear mechanisms and finite element crack propagation analysis of speed roller bearings. *Wear* 118: 77-97.

Vanderbelt, D., Laasonen, K., Car, R., Lee, C. (1991) Implementation of ultrasoft pseudopotentials in ab initio molecular dynamics. *Physical Review B*, 43(8): 6796.

Voskamp, A.P. (1998) Fatigue and material response in rolling contact. *Bearing Steel* 152-166.

Voskamp, A.P., Osterlund. R., Becker, P.C., Vingsbo, O. (1980) Gradual changes in residual stress and microstructure during contact fatigue in ball bearings. *Metals Technology*: 14-21.

Voskamp, A.P. (1985) Material response to rolling contact loading. *Journal of Tribology* 107: 359-366.

Voskamp, A.P., Mittemeijer, E.J. (1997) The effect of the changing microstructure on the fatigue behaviour during cyclic rolling contact loading. *Z. Metallkd* 88(4): 310-320.

Voskamp, A.P., Mittemeijer, E.J. (1996) Crystallographic preferred orientation induced by cyclic rolling contact loading. *Metallurgical and Materials transactions* 27A: 3445-3465.

Wang, L., Snidle, R.W., Gu, L. (2000) Rolling contact silicon nitride bearing technology: a review of recent research. *Wear* 246: 159-173.

Way, S. (1995) Pitting due to rolling contact. *Journal of Applied Mechanics* A: 49-58.

Webster, M.N., Norbart C.J.J. (1995) An experimental investigation of micropitting using

a roller disk machine. *Tribology transactions* 38(4): 883-893.

Wechsler, M. S., Lieberman, D. S., and Read, T. A. (1953) On the Theory of the Formation of Martensite. *AIME* Trans. J. Metals .

Williamson, D. L., Nakazawa K., Krauss, G. (1979) A study of the early stages of tempering in an Fe-1.2 Pct alloy. *Metallurgical Transactions* 10A: 1351-1363.

Wilson, A.W., Madison, J.D., Spanos, G. (2001) Determining phase volume fraction in steels by electron backscattered diffraction. *Scripta materialia* 45: 1335-1340.

Wilson, A.W., Spanos, G. (2001) Application of orientation imaging microscopy to study phase transformations in steels. *Materials Characterization* 46: 407-418.

Winkelmann, A., Nolze, G. (2010). Analysis of kikuchi band contrast reversal in electron backscatter diffraction patterns of silicon. *Ultramicroscopy* 110(3): 190-194.

Winter, H., Oster, P. (1990) Influence of lubrication on pitting and micropitting resistance of gears. *Gear Technology* 7(2): 16--23.

Winter, H., Weiss, T. (1981) Factors influencing the pitting and slow speed wear of surface hardened gears. *Journal of Mechanical Design* 103: 499-505.

Winter, H., Knauer, G. (1989) White etching areas on case hardened gears. *Gear Technology*: 19-45.

Winter, H., Weiss, T. (1981) Some factors influencing the pitting, micropitting(frosted areas) and slow speed wear of surface hardened gears. *Journal of Mechanical Design* 103: 499-505.

Yoshida, B., Taki, T. (2008) Micropitting generation mechanism for gears. *International Journal of Automation Technology* 2(5): pp 341-342.

Zener, C., Wert, C., (1950). Interference of growing spherical precipitate particles. *Journal of Applied Physics 21*(1): 5-8.

Zhou, R. S., Cheng, H. S., Ura, T. (1989) Micropitting in rolling and sliding contact under mixed lubrication. *Journal of Tribology* 111(4): 605-613.

Appendix 1

```
&CONTROL
calculation = 'scf' ,
restart_mode = 'from_scratch' ,
outdir = '../tmp-martensite' ,
pseudo_dir = '../../pseudo' ,
prefix = 'MarFe27C-40' ,
tstress = .true. ,
tprnfor = .true. ,
/
&SYSTEM
ibrav = 7,
celldm(1) = 16.19343994,
celldm(3) = 1.034285714,
nat = 28,
ntyp = 2,
nspin = 2,
starting_magnetization(1) = 0.7,
ecutwfc = 40 ,
ecutrho = 320 ,
occupations = 'smearing' ,
degauss = 0.005D0 ,
smearing = 'cold' ,
/
&ELECTRONS
conv_thr = 1.D-6 ,
mixing_beta = 0.7D0 ,
mixing_ndim = 8,
diagonalization = 'david' ,
/
ATOMIC_SPECIES
Fe   55.84500  Fe.pbe-sp-van.UPF
C    12.01070  C.pbe-van_bm.UPF
ATOMIC_POSITIONS crystal
Fe       0.000000    0.000000    0.000000
Fe       0.000000    0.000000    0.333333
Fe       0.000000    0.000000   -0.333333
Fe       0.000000    0.333333    0.000000
Fe       0.000000   -0.333333    0.000000
Fe       0.333333    0.000000    0.000000
Fe      -0.333333    0.000000    0.000000
Fe       0.000000    0.333333    0.333333
Fe       0.000000   -0.333333    0.333333
Fe      -0.333333    0.000000    0.333333
Fe       0.333333    0.000000    0.333333
Fe       0.000000    0.333333   -0.333333
Fe       0.000000   -0.333333   -0.333333
Fe       0.333333    0.000000   -0.333333
Fe      -0.333333    0.000000   -0.333333
Fe       0.333333    0.333333    0.000000
Fe      -0.333333   -0.333333    0.000000
Fe      -0.333333    0.333333    0.000000
Fe       0.333333   -0.333333    0.000000
Fe       0.333333    0.333333    0.333333
Fe      -0.333333   -0.333333    0.333333
Fe      -0.333333    0.333333    0.333333
Fe       0.333333   -0.333333    0.333333
Fe      -0.333333    0.333333   -0.333333
```

Fe 0.333333 -0.333333 -0.333333
Fe 0.333333 0.333333 -0.333333
Fe -0.333333 -0.333333 -0.333333
C 0.000000 0.500000 0.500000
K_POINTS automatic
8 8 8 0 0 0

Appendix 2

```
&CONTROL
calculation = 'relax' ,
restart_mode = 'from_scratch' ,
outdir = '../tmp-martensite' ,
pseudo_dir = '../../pseudo' ,
prefix = 'Fe27C-g01' ,
tstress = .true. ,
tprnfor = .true. ,
etot_conv_thr = 1.0D-4,
forc_conv_thr = 1.0D-3,
verbosity = high,
/
&SYSTEM
ibrav = 7,
celldm(1) = 16.134398241,
celldm(3) = 1.040516200,
nat = 28,
ntyp = 2,
nspin = 2,
starting_magnetization(1) = 0.7,
ecutwfc = 65 ,
ecutrho = 520 ,
occupations = 'smearing' ,
degauss = 0.005D0 ,
smearing = 'cold' ,
/
&ELECTRONS
conv_thr = 1.D-5 ,
mixing_beta = 0.7D0 ,
mixing_ndim = 8,
diagonalization = 'david' ,
electron_maxstep = 1000 ,
/
&IONS
ion_dynamics = 'bfgs',
upscale = 1000.D0,
bfgs_ndim = 2,
trust_radius_max = 0.8D0,
trust_radius_min = 1.D-3,
trust_radius_ini = 0.5D0,
/
ATOMIC_SPECIES
Fe   55.84500  Fe.pbe-sp-van.UPF
C    12.01070  C.pbe-van_bm.UPF
ATOMIC_POSITIONS crystal
Fe      0.000000000   0.000000000   0.000000000   0   0   0
Fe      0.000000000  -0.009950756   0.331895907
Fe      0.000000000   0.009950756  -0.331895907
Fe      0.000000000   0.331895907  -0.009950756
Fe      0.000000000  -0.331895907   0.009950756
Fe      0.341846663  -0.009950756  -0.009950756
Fe     -0.341846663   0.009950756   0.009950756
Fe      0.000000000   0.296780704   0.296780704
Fe      0.000000000  -0.337865169   0.337865169
Fe     -0.336259053   0.000000000   0.336259053
Fe      0.333031218  -0.001450428   0.331580791
Fe      0.000000000   0.337865169  -0.337865169
```

Fe	0.000000000	-0.296780704	-0.296780704
Fe	0.336259053	0.000000000	-0.336259053
Fe	-0.333031218	0.001450428	-0.331580791
Fe	0.333031218	0.331580791	-0.001450428
Fe	-0.333031218	-0.331580791	0.001450428
Fe	-0.336259053	0.336259053	0.000000000
Fe	0.336259053	-0.336259053	0.000000000
Fe	0.324268662	0.337865169	0.337865169
Fe	-0.333031218	-0.335386991	0.331580791
Fe	-0.341846663	0.331895907	0.331895907
Fe	0.333031218	-0.331580791	0.335386991
Fe	-0.333031218	0.331580791	-0.335386991
Fe	0.341846663	-0.331895907	-0.331895907
Fe	0.333031218	0.335386991	-0.331580791
Fe	-0.324268662	-0.337865169	-0.337865169
C	0.000000000	0.500000000	0.500000000

K_POINTS automatic

8 8 8 0 0 0

Appendix 3

```
                    Dst01_01
                    &CONTROL
       calculation      = 'relax'
     restart_mode     = 'from_scratch'
   outdir           = '../tmp-martensite'
     pseudo_dir       = '../../pseudo'
     prefix           = 'MarFe27C-cij'
         tstress          = .true.
         tprnfor          = .true.
         etot_conv_thr    = 1.0D-3
         forc_conv_thr    = 1.0D-2
                     /
                  &SYSTEM
        ibrav            = 0
    celldm(1)        = 16.130246601
         nat              = 28
         ntyp             = 2
         nspin            = 2
    starting_magnetization(1) = 0.7
         ecutwfc          = 65
        ecutrho          = 520
     occupations      = 'smearing'
       degauss          = 0.005D0
       smearing         = 'cold'
                  /
               &ELECTRONS
        conv_thr         = 1.D-4
        mixing_beta      = 0.7D0
          mixing_ndim      = 8
       diagonalization  = 'david'
                  /
                &IONS
        ion_dynamics     = 'bfgs'
       upscale          = 1000.D0
          bfgs_ndim        = 2
        trust_radius_max = 0.8D0
        trust_radius_min = 1.D-3
        trust_radius_ini = 0.5D0
                  /
             ATOMIC_SPECIES
      Fe   55.84500  Fe.pbe-sp-van.UPF
       C    12.01070  C.pbe-van_bm.UPF
          ATOMIC_POSITIONS (crystal)
Fe     0.0000000000    0.0000000000    0.0000000000
Fe     0.0000000000   -0.0097889370    0.3319486500
Fe     0.0000000000    0.0097889370   -0.3319486500
Fe     0.0000000000    0.3319486500   -0.0097889370
Fe     0.0000000000   -0.3319486500    0.0097889370
Fe     0.3417375870   -0.0097889370   -0.0097889370
Fe    -0.3417375870    0.0097889370    0.0097889370
Fe     0.0000000000    0.2970778650    0.2970778650
Fe     0.0000000000   -0.3378040870    0.3378040870
Fe    -0.3361467420    0.0000000000    0.3361467420
Fe     0.3328650410   -0.0014201780    0.3314448650
Fe     0.0000000000    0.3378040870   -0.3378040870
Fe     0.0000000000   -0.2970778650   -0.2970778650
Fe     0.3361467420    0.0000000000   -0.3361467420
```

```
Fe    -0.3328650410    0.0014201780   -0.3314448650
Fe     0.3328650410    0.3314448650   -0.0014201780
Fe    -0.3328650410   -0.3314448650    0.0014201780
Fe    -0.3361467420    0.3361467420   -0.0000000000
Fe     0.3361467420   -0.3361467420    0.0000000000
Fe     0.3243908260    0.3378040870    0.3378040870
Fe    -0.3328650410   -0.3356890940    0.3314448650
Fe    -0.3417375870    0.3319486500    0.3319486500
Fe     0.3328650410   -0.3314448650    0.3356890940
Fe    -0.3328650410    0.3314448650   -0.3356890940
Fe     0.3417375870   -0.3319486500   -0.3319486500
Fe     0.3328650410    0.3356890940   -0.3314448650
Fe    -0.3243908260   -0.3378040870   -0.3378040870
C      0.0000000000    0.5000000000    0.5000000000
                K_POINTS automatic
                   6 6 6  0 0 0
              CELL_PARAMETERS (cubic)
   0.4996965908   -0.5015427570    0.5176521555
   0.4966895360    0.4964486870    0.5156433442
  -0.4992952346   -0.4985357022    0.5201602143
```

Appendix 4

```
!*********************************************************************
*****
      PROGRAM dhkl
!*********************************************************************
*****
! dhkl calculates interplanar distances in crystals
! Author: Adrian Oila, Newcastle University
! Copyright (C) Adrian Oila
! gfortran dhkl.f -o dhkl
!*********************************************************************
*****
      CHARACTER(20)::inputfile
      INTEGER::h,k,l,bravais,n,m,i,j,m1,m2,m3,t
      DOUBLE PRECISION::a,b,c,alpha,beta,gama,d_hkl,invsq
      DOUBLE PRECISION::e1,e2,e3,k1,k2,k3,k4,k5
      DOUBLE PRECISION:: PI=3.14159265358979323846264338327950D0
!*********************************************************************
*****

      radian=PI/180.0D0


      CALL GET_COMMAND_ARGUMENT(1,inputfile)

      IF (LEN_TRIM(inputfile) == 0) THEN

      WRITE(*,*)'No input file'

      ELSE

      OPEN(UNIT=9,FILE=TRIM(inputfile),ACTION="READ")

      OPEN(10,FILE='dhkl.dat')
      OPEN(11,FILE='hkl.dat')

        READ(9,*)a,b,c,alpha,beta,gama
        READ(9,*)m
        WRITE(10,'(6F12.6)')a,b,c,alpha,beta,gama

        DO i=0,m
           m1=i
           DO j=0,m
              m2=j
              DO k=0,m
                 m3=k
        WRITE(11,"(3I4.1)")m1,m2,m3
        END DO
        END DO
        END DO
        CLOSE(11)

        IF ((alpha.EQ.90).AND.(beta.EQ.90).AND.(gama.EQ.90)) THEN
         IF (a.EQ.c) bravais=1
         IF (a.NE.c) bravais=2
         IF (a.NE.b) bravais=3
        ELSE
```

```
         IF ((alpha.EQ.90).AND.(beta.EQ.90).AND.(gama.EQ.120))
bravais=4
         IF ((beta.NE.90).AND.(a.NE.b).AND.(a.NE.c)) bravais=5
         IF ((alpha.NE.90).AND.(beta.NE.90).AND.(gama.NE.90))
bravais=6

        END IF

        IF (bravais.EQ.1) WRITE(10,'(1I4.1,A8)')bravais,'cubic'
        IF (bravais.EQ.2)
WRITE(10,'(1I4.1,A13)')bravais,'tetragonal'
        IF (bravais.EQ.3)
WRITE(10,'(1I4.1,A15)')bravais,'orthorhombic'
        IF (bravais.EQ.4)
WRITE(10,'(1I4.1,A16)')bravais,'hex./trigonal'
        IF (bravais.EQ.5)
WRITE(10,'(1I4.1,A13)')bravais,'monoclinic'
        IF (bravais.EQ.6) WRITE(10,'(1I4.1,A12)')bravais,'triclinic'
        IF ((bravais.NE.1).AND.(bravais.NE.2).AND.(bravais.NE.3)
     &      .AND.(bravais.NE.4).AND.(bravais.NE.5)
     &      .AND.(bravais.NE.6)) WRITE(*,*)'Unknown symmetry!'

        WRITE(10,'(1I4.1)')m

        e1=radian*alpha
        e2=radian*beta
        e3=radian*gama

        t=(m+1)*(m+1)*(m+1)
      OPEN(11,FILE='hkl.dat')
      DO i=1,t
      READ(11,*)h,k,l

        IF (bravais.EQ.1) THEN
        invsq=(h**2 + k**2 + l**2)/(a**2)
        d_hkl=sqrt(1/invsq)
        END IF

        IF (bravais.EQ.2) THEN
        invsq=((h**2 + k**2)/a**2) + (l**2/c**2)
        d_hkl=sqrt(1/invsq)
        END IF

        IF (bravais.EQ.3) THEN
        invsq=(4/3)*((h**2 + h*k + k**2)/a**2)+(l**2/c**2)
        d_hkl=sqrt(1/invsq)
        END IF

        IF (bravais.EQ.4) THEN
        invsq=((h*2/a**2) + (k**2/b**2) + (l**2/c**2))
        d_hkl=sqrt(1/invsq)
        END IF

        IF (bravais.EQ.5) THEN
        k1=(h**2)/(a**2*sin(e2)*sin(e2))
        k2=k**2/b**2
        k3=(l**2)/(c**2*sin(e2)*sin(e2))
        k4=(2*h*l*cos(e2))/(a*c*sin(e2)*sin(e2))
        invsq=k1+k2+k3+k4
        d_hkl=sqrt(1/invsq)
```

```
   END IF

   IF (bravais.EQ.6) THEN
   k1=(h**2)/(a**2*sin(e1)*sin(e1))
   k2=(2*k*l/b*c)*(cos(e2)*cos(e3)-cos(e1))
   k3=(k**2)/(b**2*sin(e2)*sin(e2))
   k4=(2*h*l/a*c)*(cos(e1)*cos(e3)-cos(e2))
   k5=l**2/(c**2*sin(e3)*sin(e3))
   k6=(2*h*k/a*b)*(cos(e1)*cos(e2)-cos(e3))
   k7=1-cos(e1)*cos(e1)-cos(e2)*cos(e2)-cos(e3)*cos(e3)
&     +2*cos(e1)*cos(e2)*cos(e3)
   invsq=(k1+k2+k3+k4+k5+k6)/k7
   d_hkl=sqrt(1/invsq)
   END IF

   IF (d_hkl<100D0) THEN
 WRITE(10,"(3I4.1,1F11.6)")h,k,l,d_hkl
   END IF
 END DO
 END IF

 CLOSE(9)
 CLOSE(10)
 CLOSE(11)

 END PROGRAM dhkl
```

Appendix 5

```
!*********************************************************************
*****
      PROGRAM dhklcomp
!*********************************************************************
*****
! dhklcomp compares theoretical and experimental d_hkl
! Author: Adrian Oila, Newcastle University
! Copyright (C) Adrian Oila
! gfortran dhklcomp.f -o dhklcomp
!*********************************************************************
*****
!     Define Variables
!*********************************************************************
*****
      INTEGER,PARAMETER::Lc=1000
      CHARACTER(20)::inputfile
      CHARACTER(10)::symmetry
      INTEGER::bravais,n,m,i,j,t,measure
      DOUBLE PRECISION::a,b,c,alpha,beta,gama
!     DOUBLE PRECISION::e1,e2,e3,k1,k2,k3,k4,k5
!*********************************************************************
*****
!     Derived data types: elements
!*********************************************************************
*****
      TYPE interplane
        INTEGER::h !Miller
        INTEGER::k !Miller
        INTEGER::l !Miller
        DOUBLE PRECISION::Dth !dhkl theoretic
        DOUBLE PRECISION::Dex !dhkl experimental
      END TYPE interplane

      TYPE(interplane)::D(Lc)
!*********************************************************************
*

      CALL GET_COMMAND_ARGUMENT(1,inputfile)

      IF (LEN_TRIM(inputfile) == 0) THEN

      WRITE(*,*)'No input file'

      ELSE

      OPEN(UNIT=9,FILE=TRIM(inputfile),ACTION="READ")

        WRITE(*,*)'Number of measurements='
        READ(*,*)measure

      DO i=1,measure
      READ(9,*)D(i)%Dex
        END DO


        OPEN(10,FILE='dhkl.dat',ACTION="READ")
      OPEN(11,FILE='dhklcomp.dat')
```

```
         READ(10,*)a,b,c,alpha,beta,gama
         READ(10,*)bravais,symmetry
         READ(10,*)m
!         WRITE(11,'(6F10.6)')a,b,c,alpha,beta,gama
         WRITE(11,'(3A3,2A12,1A7)')'h','k','l','d_hkl(th)',
     &                              'd_hkl(exp)','delta'

         t=(m+1)*(m+1)*(m+1)-1
       DO i=1,t
       READ(10,*)D(i)%h,D(i)%k,D(i)%l,D(i)%Dth

         DO j=1,measure
              delta=ABS(D(i)%Dth-D(j)%Dex)
              IF (delta<0.3D0) THEN
              WRITE(11,'(3I4.1,3F10.6)')D(i)%h,D(i)%k,D(i)%l,D(i)%Dth,
     &                                  D(j)%Dex,delta
              END IF
           END DO
         END DO

         END IF
       CLOSE(9)
       CLOSE(10)
       CLOSE(11)

       END PROGRAM dhklcomp
```