Properties of Solid Surfaces

1.1 Introduction

Wear and corrosion of structural materials are ubiquitous reliability and lifetime problems that have existed since the inception of mechanical devices and structures. Additionally, the optical, electrical, and electro-optical properties of solid surfaces were determined by crystalline, compositional, and electrical properties of the bulk solid. Until the advent of surface engineering, these properties belonged to the surface of the bulk materials being used and could be modified to only a limited degree by various metallurgical and plasma surface treatments. Surfaces of bulk materials could be hardened and wear corrosion resistance increased by a number of external treatments, including plasma bombardment, ion implantation, anodization, heat treatment, plasma nitriding, carburizing and boronizing, pack cementation, and ion implantation. They could also be polished or etched to modify optical properties and electrical properties to a limited degree.

Surface Engineering provides additional functionality to solid surfaces, involves structures and compositions not found naturally
in solids, is used to modify the surface properties of solids, and involves application of thin film coatings, surface functionalization and activation, and plasma treatment. It can also be defined as the design and modification of the surface and substrate of an engineering material together as a system, to give cost effective performance of which neither is capable alone.

Surface engineering techniques are being used in the automotive, aerospace, missile, power, electronic, biomedical, textile, petroleum, petrochemical, chemical, steel, power, cement, machine tools, and construction industries. Surface engineering techniques can be used to develop a wide range of functional properties, including physical, chemical, electrical, electronic, magnetic, mechanical, wear-resistant and corrosion-resistant properties at the required substrate surfaces. Almost all types of materials, including metals, ceramics, polymers, and composites can be coated on similar or dissimilar materials. It is also possible to form coatings of newer materials (e.g., met glass. \( \text{b-C}_3\text{N}_4 \)), graded deposits, multi-component deposits, etc.

In 1995, surface engineering was a £10 billion market in the United Kingdom. Coatings, to make surface life resistant to wear and corrosion, was approximately half the market.

In recent years, there has been a paradigm shift in surface engineering from age-old electroplating to processes such as vapor phase deposition, diffusion, thermal spray and welding using advanced heat sources like plasma, laser, ion, electron, microwave, solar beams, pulsed arc, pulsed combustion, spark, friction and induction. Biological materials for self-healing, self-cleaning and artificial photosynthesis are now becoming involved.

It is estimated that loss due to wear and corrosion in the U.S. is approximately $500 billion. In the U.S., there are around 9524 establishments (including automotive, aircraft, power, and construction industries) who depend on engineered surfaces with support from 23,466 industries.

There are around 65 academic institutions world-wide engaged in surface engineering research and education.

Surface engineering can be traced as far back as Thomas Edison in 1900 with the plating of gold films [1]. In 1938, Berghaus was among the first to develop plasma and ion modification of surfaces to improve surface properties and properties of vacuum deposited coatings [2]. The ion plating process, developed in the early 1960’s,
was a significant step forward in plasma-assisted coating deposition [3, 4, 5]. Ion plating was the first true industrial surface engineering process. Because conventional dc-diode sputtering used for ion plating did not provide sufficient levels of ionization to permit deposition of dense ceramic coatings with adequate mechanical properties, post deposition processes such as peening were often required to densify the coating. After the early 1970’s, the history of surface engineering is intimately connected to the development of thin film deposition and plasma processes and closely parallels the history of physical vapor deposition (PVD) coatings and processes (magnetron sputtering, ion assisted deposition), plasma processing, and chemical vapor deposition (CVD) processes in particular.

The majority of surface engineering technology has focused on enhancement of tribological properties (hardness, wear resistance, friction, elastic moduli) and corrosion resistance. The purist might think that surface engineering encompasses only tribological and wear resistant treatments, as initiated by Ron Bunshah as far back as 1961 [6]. Many engineering components need wear or corrosion resistant surfaces as well as tough, impact-resistant substrates. These requirements can be best met by using treatments that alter surface properties without significantly modifying those of the core, or bulk, material. If these principles are applied correctly, surface engineering brings many benefits, including:

- Lower manufacturing costs
- Reduced life cycle costs
- Extended maintenance intervals
- Enhanced recyclability of materials
- Reduced environmental impact

There are, however, many more properties of a solid surface that can be enhanced by application of thin films, plasma treatment, patterning and nanoscale structures. This is reflected in the programs of a number of technical conferences dedicated solely to surface engineering (International Conference on Metallurgical Coatings and Thin Films, for example), starting as early as 1974 [6]. The first conferences focused on modification of the surface of a component to enhance its the overall performance. This area, however, has grown much broader than just this technology, as demonstrated by the symposia presented at the 2010 International Conference
on Metallurgical Coatings and Thin Films (ICMCTF), sponsored by the Advanced Surface Engineering Division of AVS. Conference symposia include:

A. Coatings for Use at High Temperature
B. Hard Coatings and Vapor Deposition Technology
C. Fundamentals and Technology of Multifunctional Thin Films:
D. Carbon and Nitride Materials: Synthesis Structure-Property Relationships: Towards Optoelectronic Device Applications
E. Tribology and Mechanical Behavior of Coatings and Thin Films
F. Characterization: Linking Synthesis Properties and Microstructure
G. Applications, Manufacturing, and Equipment
H. New Horizons in Coatings and Thin Films
TS1. Experimental and Computational Studies of Molecular Materials and Thin Films
TS2. Coatings for Fuel Cells and Batteries
TS3. Bioactive Coatings and Surface Biofunctionalization
TS4. Surface Engineering for Thermal Transport, Storage, and Harvesting

Thus the optical, electrical, magnetic, thermal and even biological properties of a solid surface can also be modified using surface engineering techniques. Glass surfaces can be transformed into highly reflective or high selective reflector, transmitters or emitters. Thin film coatings are applied to glass to reflect heat, transmit heat and create heat. Decorative coatings can change color with viewing angle. The color of a thin film, whether in reflection or transmission, is critical in many applications, including low-e windows, antireflection coatings, hardware, plumbing fixtures, high reflector coatings, jewelry, automotive parts (including paints) and architectural glass. The "color" of a thin film results from its optical properties: transmittance, reflectance and absorption. Thin film coatings can be both wear resistant and colorful, and often must function in extreme environments (such as salt spray, missile domes, underground, windshields). Hard and durable gold, silver, and bronze colored coatings are applied to hardware, jewelry, plumbing fixtures and even auto bodies.
The options for surface engineering are limitless. For example, in addition to wear resistant coatings the following types of coatings are being developed:

- Decorative coatings
- Photocatalytic thin films, such as TiO$_2$, can transform a glass surface into a self cleaning surface
- Piezoelectric thin films can transform a glass or insulating surface into a high frequency transducer
- Polymer/dielectric multilayer films can decrease the water and gas permeation of a plastic surface by six orders of magnitude
- Transparent conductive oxides can make a glass or insulating surface almost as conductive as a metal while still preserving high optical transmission
- Semiconductor thin films can transform a glass, plastic or metal surface into a photovoltaic device
- Oxygen and water permeation barriers for sensitive electronics, plastics, and food packaging
- Organic thin films can transform a glass or plastic surface into a light emitting device

Often an application requires a thin film coating to be multifunctional. Common examples of this are

- Transition metal nitrides (TiN, ZrN, TaN, HfN) applied to hardware, jewelry, and plumbing fixtures for wear resistance and color
- Antireflection, heater coatings
- Wear resistant optical coatings
- Deicing optical filters
- Conductive, wear resistant coatings
- Decorative coatings

This book addresses the fundamentals of modifying and enhancing the tribological, optical, electrical, photo-electric, mechanical, and corrosion resistance of solid surfaces and adding functionality to solids by engineering their surface, structure, and electronic, magnetic, and optical structure. Note that thin film, film, and coating are used interchangeably in this book and all describe structures
with thicknesses ranging from the nanometer (nm) scale to several microns (μm). While thin film applications will be emphasized, the increasing use of low dimensional structures and nanocomposites will integral to this discussion. Adhesion of thin films and elastic properties will not be directly addressed, although they cannot be omitted in some discussions. These topics have been addressed in detail in other books and the reader should refer to these books for in depth discussions [7, 8, 9, 10, 11]. It should be emphasized that, while this book will cover a broad range of surface engineering and engineered materials topics, it is strictly an introduction. Advanced topics and detailed analysis are beyond the scope of this book. Advanced topics, for example, are covered in the Third Edition of the Handbook of Deposition Technologies for Films and Coatings (P M Martin, Ed.) and Handbook of Nanostructured Thin Films and Coatings (Sam Zang, Ed.).

Virtually every advance in surface engineering and engineered materials has resulted from thin film and related technologies. Engineered materials are the future of thin film technology. Properties can now be engineered into thin films that achieve performance not possible a decade ago. Engineered structures such as superlattices, nanolaminates, nanotubes, nanocomposites, smart materials, photonic bandgap materials, metamaterials, molecularly doped polymers, and structured materials all have the capacity to expand and increase the functionality of thin films and coatings used in a variety of applications and provide new applications. New advanced deposition processes and hybrid processes are now being used and developed to deposit advanced thin film materials and structures not possible with conventional techniques a decade ago.

Engineered materials are now being developed for and used in advanced photovoltaic devices, dye sensitized solar cells, quantum cascade lasers, advanced electronics, drug delivery, medical devices, metamaterials, optical photonic bandgap devices, negative refractive index devices, superlenses, artificial magnetism, cloaking devices, thermoelectric power generation, functional biological materials, and much more. Structures and properties not possible in naturally occurring materials are synthesized by a number of lithographic, etching, plasma, and thin film deposition processes.

In order to fully understand how surface engineering technologies improve performance of solid surfaces, it will be useful to review the
properties of solid surfaces that can be enhanced. We will address the following surface properties:

- Wear resistance
- Hardness
- Lubricity
- Corrosion and chemical resistance
- Optical (transmittance, reflectance, emittance)
- Electrical (conductivity)
- Electro-optical (photoconductivity, stimulated emission)
- Photocatalysis
- Surface energy (hygroscopicity and hydrophobicity)
- Temperature stability

1.2 Tribological Properties of Solid Surfaces

Wear and corrosion are ubiquitous problems that affect virtually every type of surface, especially those that come in contact with other surfaces. Tribological coatings are used to mitigate these problems. The word tribology is derived from the Greek word tribos, meaning “rubbing”. Tribology is defined as the science and engineering of interacting surfaces in relative motion, including the study and application of the principles of friction, lubrication, and wear. These materials are concerned with friction, wear, corrosion, hardness, adhesion, lubrication, and related phenomena. They essentially combine the best of hard, wear resistant, and low surface energy coatings. Degradation of a material, tool, or components involves interactions occurring at their surfaces. Loss of material from a surface and subsequent damage may result from tribological interactions of the exposed face of the solid with interfacing materials and environment. Mass loss results from chemical driving forces and wear results from interfacial mechanical forces. Tribological coatings are applied to bearings, engine components, valves, computer disc heads and discs, biological joints, motor bushings, cutting tools, shaving heads, surgical tools, pump components, gears, molds, and the list goes on. Tribological thin film coatings are discussed in detail in chapter 4.
Tribological coatings are applied to surfaces to

- Increase wear resistance and hardness,
- Increase chemical and corrosion resistance
- Improve environmental stability
- Reduce coefficient of friction
- Change color of the surface
- Simulate metals such as gold, brass, copper, bronze surfaces

Three categories of tribological wear behavior can be defined:

- Both friction and wear are low (usually in bearings, gears, cams and slideways)
- Friction is high but wear is low (usually in power transmission devices such as clutches, belt drives, and tires)
- Friction is low and wear of one component is high (machine cutting, drilling, and grinding).

1.2.1 Wear

Before we can address surface engineering techniques used to improve tribological properties, it will be instructive to define and understand each property and associated test procedures in depth. Only then can we understand the significance of improvements in performance.

Wear is defined as the erosion of material from a solid surface by the action of another surface. It is related to surface interactions and, more specifically, the removal of material from a surface as a result of mechanical action. It is important to note that mechanical wear requires some form of contact resulting from relative motion, opposed to other processes that can produce similar results. This definition does not include dimensional changes due to mechanical deformation (no removal of material). Impact wear, where there is no sliding motion, cavitation, where the opposing material is a fluid, and corrosion, where the damage is due to chemical rather than mechanical action are also not included [12, 13].

Wear can also be defined as a process in which interaction of the surfaces or bounding faces of a solid with its working environment results in dimensional loss of the solid, with or without loss of material. Working environment here includes loads (such as unidirectional sliding, reciprocating, rolling, and impact loads - pressure),
speed, temperature, type of opposing material (solid, liquid, or gas), and type of contact (single phase or multiphase, in which the phases involved can be liquid plus solid particles plus gas bubbles).

Seven mechanical wear mechanisms are listed in Table 1.1 [14]. These results, however, from only three types of surface to surface interactions: sliding (one surface sliding relative to another over long distances), fretting (one surface oscillates over minute distances relative to the other) and erosion (solid particles impinging on a single surface from an external source). The actual wear mechanism for dry sliding wear depends on a number of variables including, surface finish, surface geometry, orientation, sliding speed, relative hardness (of one surface relative to the other or relative to the abrasive particles between the surfaces), material microstructure, and

<table>
<thead>
<tr>
<th>Classification</th>
<th>Wear Mechanism</th>
<th>Wear Coefficient $K$ (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wear dominated by mechanical behavior of materials</td>
<td>Asperity deformation and removal</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>Wear caused by plowing</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>Delamination wear</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>Adhesive wear</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>Abrasive wear</td>
<td>$10^{-4}$–$10^{-1}$</td>
</tr>
<tr>
<td></td>
<td>Fretting wear</td>
<td>$10^{-6}$–$10^{-4}$</td>
</tr>
<tr>
<td>Wear dominated by chemical behavior of materials</td>
<td>Wear by solid particle impingement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Solution wear</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oxidation wear</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diffusion wear</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wear by melting of surface layer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adhesive wear at high temperatures</td>
<td></td>
</tr>
</tbody>
</table>
more. From these variables, it can be seen that wear rate is not a pure material property and does not always occur uniformly.

The tribological material or coating must address the issues connected with each mechanism. Adhesive friction encompasses cohesive adhesive forces and adhesive wear. Cohesive adhesive forces hold two surfaces together. The atoms and molecules of surfaces in contact actually never touch each other but are separated by atom-atom forces or cluster interactions, i.e., cohesive adhesive forces (thus, adhesion of two contacting surfaces arises from the attractive forces that exist between the surface atoms of the two materials). Adhesive wear occurs when surface features, such as microscopic roughness, are brought into contact under a load. It can be described as plastic deformation of very small fragments within the surface layer when two surfaces slide against each other: asperities (i.e., microscopic high points) found on the interacting surfaces will penetrate the opposing surface and develop a plastic zone around the penetrating asperity. Obviously, adhesive wear increases with increased surface roughness. During initial contact, fragments of one surface are pulled off and adhere to the other, due to the strong adhesive forces between atoms [12]. However, energy absorbed in plastic deformation and movement is the main cause for material transfer and wear. Stress builds up at the microscopic peaks on the surfaces and they deform and tend to weld together. These microscopic welds will eventually shear the crystallites from one surface and transfer them to the other surface, usually from the softer material to the harder material. If further rubbing occurs, particles of the softer material will form and smearing, galling, or seizure of the surfaces will occur. In the case of adhesive wear, the volume of the wear material is proportional to the distance over which the sliding occurs and to the applied load, and is inversely proportional to the hardness or yield stress of the softer material. This can be summarized by the Archard equation [15, 16]:

\[
V = \frac{KFL}{3H},
\]

where \( V \) is the volume of wear material, \( K \) is the dimensionless wear coefficient, \( F \) is the applied force (~pressure), \( L \) is the length of slide, and \( H \) is the yield stress of the soft material. Table 1.1 lists several wear mechanisms and associated coefficient \( K \) [14]. Note
that $K$ can be decreased by many orders of magnitude by lubricating the surfaces or applying a tribological coating.

It is generally not possible to deform a solid material using direct contact without applying a high pressure (or force), and at some time during this interaction, the process must accelerate and decelerate. It is necessary that high pressure be applied to all sides of the deformed material. Also, flowing material will immediately exhibit energy loss and reduced ability to flow if ejected from high pressure into low pressure zones. Thus, once the wear process is initiated, it has the capacity to continue because energy stored when the system is under high pressure will facilitate lower pressure needed to continue sliding.

*Abrasive wear* occurs when a hard rough surface slides across a softer surface [12]. The American Society for Testing and Materials (ASTM) defines it as the loss of material due to hard particles or hard protuberances that are forced against and move along a solid surface [17].

It is also classified according to the type of contact and the contact environment [18]. The type of contact determines the mode of abrasive wear. The two modes are defined as two-body and three-body abrasive wear. Two-body wear occurs when the grits, or hard particles, are rigidly mounted or adhere to a surface, when they remove the material from the surface. Sandpaper is the typical example of two-body wear. Obviously, the smoother the surfaces, the less abrasive wear will occur. In three-body wear, particles are not constrained, and are free to roll and slide down a surface. The contact environment determines whether the wear is classified as open or closed. An open contact environment occurs when the surfaces are sufficiently displaced to be independent of one another.

Three factors that influence abrasive wear and hence the manner of material removal are:

- Plowing
- Cutting
- Fragmentation

Plowing, shown in Figure 1.1, involves the formation of grooves that form when material is displaced to the side, away from the wear particles, and does not involve direct material removal. The displaced material forms ridges adjacent to grooves, which may be removed by subsequent passage of abrasive particles,
i.e., smoothing. Cutting involves separation of material from a surface in the form of primary debris, or microchips, with little or no material displaced to the sides of the grooves, and closely resembles conventional machining. Fragmentation, shown in Figure 1.2, occurs when material is separated from a surface by a cutting process and the indenting abrasive causes localized fracture of the wear material. These cracks then freely propagate locally around the wear groove, resulting in additional material removal by spalling [18].

Figure 1.1 Flowing wear [19, 20].

Figure 1.2 Fragmentation wear.
Wind and water erosion are familiar examples of erosive wear. Sand and bead blasting are common industrial processes that involve this type of wear. Filtration and sealing are the two main methods used to minimize this type of wear. Erosive wear is shown in Figure 1.3. Also, as demonstrated by wind and sand erosion, gases can be a very powerful method of mass removal [21]. The rate of erosion is very dependent on incidence angle. Particles impinging on the surface of a brittle material remove material faster at higher angles of incidence, as shown in Figure 1.4. The inverse occurs for ductile materials. Components that are strongly affected by erosion are mining equipment, gas turbines, and electrical contacts in motors and generators.

Note that a Taber Abrasion Test is commonly used to assess abrasive wear. Fretting is the cyclical rubbing of two surfaces and fretting wear, shown in Figure 1.5, is the removal of material from one or both surfaces as a result of fretting. This type of wear can occur during the contact of round, cylindrical, or spherical surfaces, such as roller bearings. Fretting fatigue results from fretting wear. Cracks in one or both of the surfaces in contact form as a result of fretting fatigue. This process is accelerated if particulates are ejected by the contact. These particulates can oxidize and further abrade the surfaces. Fretting corrosion will occur if there are vibrations in the system. Pits or grooves and oxide debris can are typical of this type of damage, which is typically found in machinery, bolted assemblies, and ball or roller bearings. Contact surfaces

![Figure 1.3 Erosive wear.](image)
Figure 1.4 Relative erosion rates in brittle and ductile materials as a function of particle impact angle [22].

Figure 1.5 Fretting wear.

Exposed to vibration during transportation are exposed to the risk of fretting corrosion.

1.2.2 Coefficient of Friction: Lubricity

The coefficient of friction (COF) between two surfaces is a major factor that affects wear rate. We are interested here in kinetic friction (as opposed to static friction), which generates a force directed
opposite to the velocity \( (v) \) of one surface over the other. This force is defined as

\[
F_k = -\mu_k N v
\]

where \( \mu_k \) is the coefficient of kinetic friction. Note that \( \mu_k \) is independent of \( v \) (for small \( v \)) and that \( F_k \) is independent of contact area. COF is characteristic of the physical and chemical states of the two surfaces and is also affected by temperature, humidity, and environment. Kinetic friction involves dissipation of mechanical energy into heat and eventually wearing away of the softer surface. Frictional wear involves breaking of interfacial bonds on one part of the surface and creation of new bonds on another part of the surface. Obviously, if a bond cannot be broken, friction will be less. The work needed for bond destruction is supplied by the relative motion of the two surfaces. Creation of new bonds releases energy in the form of solid excitations (phonons = heat). The energy created into heat can be expressed as

\[
E_h = \int \mu_k F_k \, dx
\]

1.2.3 Hardness

Hardness is vaguely defined as "Resistance of metal to plastic deformation, usually by indentation" [23], that is, as the measure of how resistant a solid surface is to various kinds of permanent shape change when a force is applied. Hardness is related to the minimum stress needed to produce irreversible plastic deformation to the surface of a solid [24]. Hardness and wear resistance are intimately connected, although there is some evidence that elastic modulus also plays a role. Hardness, however, is not a basic property of a material, but rather a composite with contributions from the yield strength, work hardening, true tensile strength, elastic modulus, and others factors. In many cases, hardness is defined by the type of test used to measure it.

The greater the hardness of the surface, the greater resistance it has to deformation. We are interested primarily in micro and nano-hardness of surfaces, which is generally characterized by strong intermolecular bonds. Hardness is characterized in various forms, including scratch hardness, indentation hardness, and rebound
hardness. The values of these types of hardness are reflected in specific types of tests. Hardness is dependent on ductility, elasticity, plasticity, strain, strength, toughness, viscoelasticity, and viscosity.

A number of tests have been developed to quantify hardness of a solid and compare the hardness of various solids:

- Mohs
- Brinell
- Rockwell hardness
- Rockwell superficial hardness
- Vickers
- Knoop
- Scleroscope and rebound hardness
- Durometer
- Barcol
- Microindenter
- Nanoindenter

Note that in most cases there is no one-to-one correspondence between the above tests, and hardness measurements must be compared using the same test methodology. Also note that virtually every hardness measurement made on thin films uses a micro or nanoindenter technique.

Scratch hardness defines the resistance of a surface to fracture or plastic (permanent) deformation due to friction from a sharp object. The scratching material is obviously harder than the surface under test. The most common scratch hardness test is Mohs scale.

Indentation hardness measures the resistance of a surface to permanent plastic deformation due to a constant compression load from a sharp object, usually a stylus or indenter tip. Hardness is deduced from critical dimensions of an indentation left by a specifically dimensioned and loaded indenter. Common indentation hardness scales are Rockwell, Vickers, and Brinell. Most modern techniques give hardness in gigapascals (GPa).

Rebound hardness, or dynamic hardness, is related to elasticity and measures the height of the "bounce" of a diamond-tipped hammer dropped from a fixed height onto a material. Two scales that measure rebound hardness are the Leeb rebound hardness test and Bennett hardness scale.

Hardness and tensile strength ($\sigma_{TS}$) are proportional to each other, but ultimately hardness is related to bonding and microstructure of the solid [25].
The Mohs hardness test is typically used only to identify a mineral and is arguably one of the oldest tests [26]. This test involves observing whether a material's surface is scratched by a substance of known or defined hardness. The Mohs scale involves 10 minerals, shown in Figure 1.6, and is used to provide a numerical "hardness" value. The hardness of each of the 10 minerals is given an arbitrary value. Minerals are ranked along the Mohs scale, which is composed of 10 minerals that have been given arbitrary hardness values. This test, therefore, is not suitable for accurately gauging the hardness of industrial materials such as steel or ceramics.

The Brinell hardness test, developed in 1900, uses a desktop machine to apply a specified load to a hardened sphere of a specified diameter. The Brinell hardness number (Brinell), is obtained by calculating the ratio of the load used (in kg) and measured surface area of the surface indentation (mm²), in square millimeters. Figure 1.7 shows the test geometry. This test is often used to determine the hardness metal forgings and castings that have a large grain structures, and provides a measurement over a fairly large area that is less affected by the coarse grain structure of these materials than micro and nanoindentation tests. A typical test uses a 10 mm diameter steel ball as an indenter with a 3,000 kgf (29 kN) force. For softer materials, a smaller force is used; for harder materials, a TiC ball is substituted for the steel ball. The indentation is measured and hardness calculated as:

$$BHN = \frac{2P}{2\pi \left[ D - (D^2 - d^2)^{1/2} \right]}$$ \hspace{1cm} (1.2)

Where \(P\) = applied load (kg), \(D\) = diameter of indentor, \(d\) – diameter of indentation. A number of standards have been developed for this test [27, 28].

All subsequent tests described here are variations of the indentation test. The Rockwell Hardness Test employs an instrument to apply a specific load and then measure the depth of the resulting indentation. The indenter is either a steel ball of a specified diameter or a spherical diamond-tipped cone of 120° angle and 0.2 mm tip radius, called a braille. A minor load of 10 kg is first applied, which causes a small initial penetration to seat the indenter and remove the effects of any surface irregularities. Next, the load measurement is zeroed and a major load is applied. Upon removal
of the major load, the depth reading is taken while the minor load is still on. The hardness number may then be read directly from the scale. The indenter and the test load used determine the hardness scale that is used (A, B, C, etc). In order to get a reliable
reading the thickness of the test-piece should be at least 10 times the depth of the indentation and standards [29, 30, 31].

This test has the following scales:

A - Cemented carbides, thin steel and shallow case hardened steel
B - Copper alloys, soft steels, aluminum alloys, malleable iron, etc.
C - Steel, hard cast irons, pearlitic malleable iron, titanium, deep case hardened steel and other materials harder than B 100
D - Thin steel and medium case hardened steel and pearlitic malleable iron
E - Cast iron, aluminum and magnesium alloys, bearing metals
F - Annealed copper alloys, thin soft sheet metals
G - Phosphor bronze, beryllium copper, malleable irons
H - Aluminum, zinc, lead
K, L, M, P, R, S, V - Bearing metals and other very soft or thin materials, including plastics.

For example, 50 HRB indicates that the material has a hardness reading of 50 on the B scale.

The Rockwell Superficial Hardness Tester is used to test thin materials, lightly carburized steel surfaces, or parts that might bend or crush under the conditions of the regular test. All indenters and test geometry are the same except that loads are reduced. A lighter minor 3 kg load is used and the major load is either 15 or 45 kg, depending on the indenter used. Additionally, a 0.0625" diameter steel ball indenter is used. "T" is added (= thin sheet testing) to identify the superficial hardness designation. For example, a superficial Rockwell hardness of 23 HR15T, indicates the superficial hardness as 23, with a load of 15 kg using a steel ball.

Vickers and Knoop hardness are macrohardness and microhardness tests developed specifically to measure the hardness of thin films and surfaces [32, 33]. The indenter used in these tests is a small diamond pyramid, shown in Figure 1.8, which is pressed into the sample under loads significantly less than those used in the Brinell test. A typical indentation is also shown in the figure. Figures 1.9 and 1.10 show the geometry of these two tests. Basically, the only
Figure 1.8 Indenter used in Vickers hardness tests and typical indentation.

Vickers hardness test

Load F (1 - 120 kgf)
Pyramid-shaped diamond of 136° angle
D (diagonal of indentation)

Micro-Vickers
Load F (1 - 1000 gf)
h > 1.5D
D (diagonal of indentation)

www.substech.com

Figure 1.9 Geometry of Vickers hardness test.

difference between these tests is the shape of the diamond pyramid indenter. The Vickers test uses a square pyramidal indenter which is prone to crack brittle materials while the Knoop test uses a rhombic-based (diagonal ratio 7.114:1) pyramidal indenter. For
equal loads, Knoop indentations are about 2.8 times longer than Vickers indentations. The yield strength of a material can be approximated by

$$\sigma_y \sim \frac{H_v}{C}$$  \hspace{1cm} (1.3)

c is constant ranging between 2 and 4.

Vickers hardness can be calculated using [34]

$$HV = \frac{1.8554F}{d^2} \text{(kg/mm}^2)$$  \hspace{1cm} (1.4)

Here F is the indenter load and d is the width of the pyramidal indentation.
Knoop hardness can be found using the following [35]:

\[ HK = \frac{F}{A} = \frac{F}{C_p L^2} \]  

(1.5)

With \( A \) = area of indentation, \( L \) = length of indentation and \( C_p \) = correction factor for shape of indenter (~0.070729).

The Scleroscope test involves dropping a diamond tipped hammer, which free falls inside a glass tube from a fixed height onto the test surface [36]. The height of the rebound travel of the hammer is measured on an arbitrarily chosen 100-unit graduated scale, called Shore units. One hundred Shore units represents the average rebound from pure hardened high-carbon steel. Because there are harder materials than the calibration material, the scale is continued higher than 100. The Shore Scleroscope measures hardness in terms of the elasticity of the material; the hardness number is related to the hammer’s rebound height. The harder the material, the higher the rebound.

The Rebound Hardness Test Method employs the Shore Scleroscope for testing relatively large samples [37]. Typically, a spring is used to accelerate a spherical, tungsten carbide tipped mass towards the surface of the test surface, producing an indentation on the surface which takes some of this energy from the impact body. The harder the surface, the smaller the indentation: the indenter will lose more energy and its rebound velocity will be less when testing a softer material. The velocities of the impact body before and after impact are measured and the loss of velocity is related to Brinell, Rockwell, or other common hardness value.

Hardness testing of plastics, polymers, elastomers, and resins can be tricky. The Durometer and Barcol tests are used specifically for these elastic materials. A durometer uses a calibrated spring to apply a specific pressure to an indenter, which can be either cone or sphere shaped. Depth of indentation is measured. A durometer employs several scales, depending on the type of material:

- ASTM D2240 type A and type D scales (A - softer plastics, D - harder plastics)
- ASTM D2240-00 consists of 12 scales, depending on the intended use; types A, B, C, D, DO, E, M, O, OO, OOO, OOO-S, and R. Higher values correspond to a harder material.
The Barcol hardness test obtains a hardness value by measuring the penetration of a sharp steel point under a spring load. The sample is placed under the indenter of the Barcol hardness tester and a uniform pressure is applied until the dial indication reaches a maximum. The governing standard for the Barcol hardness test is ASTM D 2583.

Nanoindentation testing is used to measure hardness and elastic constants of a small area and volumes and thin films [24, 38]. This measurement generally employs an atomic force microscope (AFM) or a scanning electron microscope (SEM) to image the indentation. This test improves on macro and micro indentation tests discussed above by indenting on the nanoscale with a very precise tip shape, high spatial resolutions to place the indents, and by providing real-time load-displacement (into the surface) data while the indentation is in progress.

Small loads and tip sizes are used in this test, and as a result, the indentation area may only be a few µm² or nm². The problem here is that the indentation is not easy to find. A typical indentation is shown in Figure 1.11. Instead of trying to image the indentation, an indenter with a geometry known to high precision (usually

![Figure 1.11 Typical nanoindentation.](image-url)
a Berkovich tip, which has a three-sided pyramid geometry) is employed. Instead, the depth of penetration is recorded, and then the area of the indent is determined using the known geometry of the indentation tip. The response of the film to indentation, shown in Figure 1.12, is obtained for loading and unloading of the tip; load and depth of penetration are measured. Figure 1.12 shows a load-displacement curve, which is used to extract mechanical properties of the material [38].

This test can determine two types of hardness:

- Single hardness value
- Hardness as a function of depth in the material

Again, the above relation for hardness \( H = \frac{P_{\text{max}}}{A_f} \) is also valid for micro and nanoindenters. As mentioned above, some nanoindenters use an area function based on the geometry of the tip, compensating for elastic load during the test. Use of this area function provides a method of gaining real-time nanohardness values from a load-displacement graph. Figure 1.13 demonstrates this type of measurement, plotting hardness of a nanolaminate against load [39]. Software has been specifically designed to incorporate all test parameters and analyze the displacement vs. load curve [40, 41]. Nanoindenters are often incorporated into AFMs.

![Load displacement curve for nanoindentor](image)

**Figure 1.12** Load displacement curve for nanoindentor [38].
A number of tests have been developed to measure hardness of solid surfaces and thin films. Indentation tests work best for thin films. Most modern methods use nanoindentation coupled with some form of computer and AFM analysis. Because of the types of indenter used, it is difficult to cross-compare hardness values obtained by the various indentation techniques (Vickers, Knoop, nanoindentation). Most modern tests give hardness values in GPa, while a few decades ago, this unit was not used. It is, therefore, judicious to compare hardness values obtained by the same test method.

1.3 Optical Properties of Solid Surfaces

Unlike many other properties, modification of the optical properties of a surface can totally change the optical properties of the substrate. However, in addition to modifying the optical properties, in many cases they serve two or more surface engineering
functions, such as improving abrasion and wear resistance, electromagnetic shielding, and as gas, chemical and water permeation barriers. Surface engineering of optical properties is presented in more detail in Chapter 5. Optical properties of a surface can be modified using:

- Polishing
- Thin films
- Patterned structures
- Etching
- Low dimensional structures
- Photonic crystals
- GLAD coatings

Electromagnetic radiation (which includes light) when incident on a solid surface can be transmitted, reflected, absorbed, or scattered. Light can also be diffracted and refracted. The film/substrate interface is important for many optical coatings since it contributes to the overall optical properties of the entire optical system and will be addressed in Chapter 5. The optical properties of a surface depend on a number of factors:

- Surface quality (roughness, pits, digs, scratches)
- Optical constants (refractive index and extinction coefficient)
- Absorptance and emittance

Transmittance \( T \), reflectance \( R \), and absorptance \( A \) are directly related to the optical constants of the material's real and imaginary parts of the refractive index \( n \) & \( \kappa \). Transmittance, reflectance, and absorptance (see Figure 1.14) at an air interface \( (n_{air} = 1) \) are related to optical constants by

\[
n = n + i\kappa \text{ (complex refractive index)} \quad (1.6)
\]

\[
T = 1 - R - A \quad (1.7)
\]

\[
R = \frac{[(n - 1)^2 + \kappa^2]}{[(n + 1)^2 + \kappa^2]} \quad \text{(normal incidence)} \quad (1.8)
\]
\[ A = -\log_{10} T \]  
\[ \alpha = \frac{4\pi k}{\lambda} \text{ (absorption coefficient)} \]

Figure 1.14 Relationship between $T$, $R$, and $A$.

Here $\lambda$ is the wavelength of light.

All surfaces scatter incident electromagnetic radiation at some level. Scattering causes attenuation of a light beam similar to absorption. The intensity of light due to scattering cross section $\sigma_s$ is

\[ I(z) = I_0 \exp(-\sigma_s N z) \]

where $N$ is the number of scattering centers and $z$ is the propagation direction through the solid. Note that scattering only gets worse at shorter wavelengths

\[ \sigma_s \sim \frac{1}{\lambda^4} \]

Except for reducing scattering (by polishing), there is very little that can be done to modify the optical properties of a bulk surface without some external process (such as application of thin films, smoothing layers, etching, and patterning). Because the optical constants are ultimately properties of the lattice and bonding in materials, stress and strain can also affect their behavior.
Methods to modify and enhance the optical properties of solid surfaces will be discussed in more detail in Chapter 5. Treatments listed above can

- Change a simple transparent ceramic window into a reflector, EMI shield, heat mirror, cold mirror, emissive display, laser shield, or switchable optical device
- Increase transmission of ceramic, semiconductor and semi-opaque substrates
- Modify the optical band structure of a surface
- Increase or decrease optical scattering of a surface
- Enable read-write capability on a substrate

Color is another important optical property of the solid surface and can be modified over wide ranges by surface engineering techniques. Color is critical in decorative coating applications. One of the primary applications of thin films is to change the color or reduce the color of a surface, window, optic, or object. Color is an intrinsic property related to transmission, reflectance, and emissivity of a surface. Color is defined as the visual perceptual property corresponding in humans to the categories called red, green, blue, and others. We all “see” the color of an object differently. Color results from the visible light spectrum interacting in the eye with the spectral sensitivities of the light receptors located in the retina. Color categories and physical specifications of color associated with objects, materials, light sources, etc., are based on optical properties such as visible light absorption, transmission, reflection, or emission spectra. Several attempts have been made to quantify color by identifying numerically their coordinates in a color space model [40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50].

A number of models have been presented to quantify color perception, which will be discussed in Chapter 5. These models relate color coordinates to the actual color experience by the human eye. Figure 1.15 summarizes color vision characteristics and associated models [46, 50]. The figure diagrams color perception factors and how they are mapped into a chromaticity model, discussed in Chapter 5. The entire picture ends up in the bottom chromaticity diagram and includes:

- Vision characteristics
- Color perception
Figure 1.15 Mapping of human color perception [46].

- Color mixing
- Color measurement
- Spectral power distribution
- Chromaticity model

Chapter 5 summarizes several chromaticity models and relates these models to colors and change in color of reflective and transmissive surfaces.

1.4 Electric and Opto-electronic Properties of Solid Surfaces

The electrical and opto-electronic properties of a solid are determined by the transport and excitation of electrons and holes in the solid and can be attributed to

- Energy band structure
- Lattice structure:
  - Short range order or lack of short range order
  - Crystalline phase composition
- Wide variations in density
- Grain boundary effects
- Surface and interface effects
- Quantum effects

Many of these factors are interdependent, e.g., energy band structure is derived from composition, lattice structure, and defects. In this introduction, we will address classical and quantum mechanical models for electrical conduction in solids and relate these to electrical conduction in thin films and low dimensional structures (superlattices, nanowires, quantum dots, nanotubes). We briefly review the classical theory of electrical conduction. The basic relationship between current density \( J(r) \) and electric field \( E(r) \)is given by [25]

\[
J(r) = \sigma E(r)
\]  

(1.13)

Here \( J, E \) and \( r \) are vectors and \( \sigma \) is the electrical conductivity tensor. If the medium is isotropic, \( \sigma \) is a scalar quantity and \( J \) and \( E \) are parallel. Drude theory defines conductivity in terms of an electron's charge \((e)\), free electron density \((n)\), mass \((m)\) and relaxation time between collisions \((\tau)\):

\[
\sigma = ne^2 \tau \quad \text{or in terms of the mobility (}\mu)\quad \sigma = ne\mu
\]

(1.14)  

(1.15)

Table 1.2 displays conductivity for some common metals. It will be useful to refer back to this table when addressing conductivity of thin films. We will also see that all the quantities shown above become dependent on energy band structure in the quantum world and are functions of the shrinking dimensions of thin films and low dimensional structures and defect density; i.e., properties of smaller structures are more severely affected by factors that would not significantly affect bulk materials.

Without electron (or other charge carrier) collisions \( \sigma \) and \( \tau \) would be infinite and the material would be a perfect conductor. We will see in Chapter 6 that collisions also have dramatic effects on the conductivity of thin films and low dimensional structures.
Table 1.2 Conductivity parameters for some common metals [25].

<table>
<thead>
<tr>
<th>Metal</th>
<th>Conductivity (10^5 S/cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag</td>
<td>6.28</td>
</tr>
<tr>
<td>Al</td>
<td>3.65</td>
</tr>
<tr>
<td>Au</td>
<td>4.92</td>
</tr>
<tr>
<td>Cu</td>
<td>5.96</td>
</tr>
<tr>
<td>Ni</td>
<td>1.43</td>
</tr>
<tr>
<td>Fe</td>
<td>1.0</td>
</tr>
<tr>
<td>Ti</td>
<td>2.38</td>
</tr>
<tr>
<td>C</td>
<td>1.43</td>
</tr>
<tr>
<td>In</td>
<td>1.14</td>
</tr>
<tr>
<td>K</td>
<td>1.39</td>
</tr>
<tr>
<td>Li</td>
<td>1.07</td>
</tr>
<tr>
<td>Mg</td>
<td>2.23</td>
</tr>
</tbody>
</table>

Optoelectronic properties result from the interaction of photons with electrons in solid surfaces, as is the basis for photovoltaics, lasers, photoconductivity, photon detectors, and photodiodes. Solids can absorb and emit photons, both of which involve movement of an electron between atomic or molecular energy levels. An electron is excited into a higher energy level when absorbed by atom, and relaxes to a lower energy level as a result of photon emission. The band structure of solids will be discussed in more detail in Chapter 7. Figure 1.16 shows energy transitions of an electron when a photon is either absorbed or emitted [51]. When a photon is absorbed, the electron is excited to a higher energy (E_i → E_f), essentially absorbing the photon's energy (= Planck’s constant X frequency = hν). A photon is emitted when the electron decays from E_f → E_i. As shown in Figure 1.17, assuming the energy of the photon is greater than the band gap (E_g) of a semiconductor, electrons can also be excited from the valence band (E_v) into the conduction band (E_c) to create an electric current, i.e., photoconductivity.
Figure 1.16 Electronic transitions due to the absorption or emission of a photon [51].

Figure 1.17 Electron excitation from the valence band into the conduction band of a semiconductor [51].

Light emitting diodes (LED) and organic light emitting diodes (OLED) function on the emission of a photon when electron and electron-hole pairs are excited in semiconductors and conductive polymers. In this case, charge carriers are excited to higher energy levels by application of a voltage (energy) and decay by giving off photons [51, 52]. Electrons can be excited to higher energy levels thermally. Thermal energy can excite electrons into higher energy levels and photons emitted when they decay to lower energy levels.

The interaction of photons with electrons is also responsible for optical properties such as the plasma frequency, surface plasmons, evanescent waves and, of course, photosynthesis. The plasma frequency is particularly important in metals and transparent conductive oxides in that it determines the frequency dependent electrical conductivity, and at the wavelength of light that high reflectivity initiates. The reflectivity of metals is totally dependent on the plasma frequency [53].
The frequency dependent conductivity can be expressed as [54]

\[ \sigma(\omega) = \left( ne^2 \frac{\tau}{m} \right) \left( \frac{1 + i\omega\tau}{(1 + \omega^2\tau^2)} \right) \]  (1.16)

where \( \omega \) is the frequency of the electromagnetic wave (light in this case), \( \tau \) is the relaxation time between electron collisions, and \( m \) is the mass of the electron. Thus, as \( \omega \) decreases, \( \text{Re}[\sigma(\omega)] \) increases and we get essentially the dc conductivity \( ne^2\tau/m \). If the frequency dependent permittivity is

\[ \varepsilon(\omega) = \varepsilon_0 + \frac{i\sigma(\omega)}{\sigma(\omega)} \]  (1.17)

The frequency dependent dielectric function is given by

\[ \varepsilon(\omega) = \frac{1 - \omega_p^2}{\omega} \left( \frac{1}{\omega + \frac{i}{\tau}} \right) \]  with the plasma frequency \( \omega_p^2 = \frac{ne^2}{m\varepsilon_0} \).  (1.18)

![Figure 1.18](image.png) Dependence of the electrical conductivity of a metal with frequency, showing the plasma frequency.
The plasma frequency is very significant in that it is the natural frequency of the electron gas in the conduction band. The plasma frequency in metals generally occurs at ultraviolet and short visible wavelengths, and separates the frequency domains that are either absorbing \( (\omega \sim \omega_p) \), reflecting \( (\omega < \omega_p) \), or transparent \( (\omega \gg \omega_p) \). Optical reflectivity of metals is thus derived from this relation. Figure 1.18 shows the behavior of the electrical conductivity with frequency. The imaginary part peaks at the plasma frequency.

We can thus relate electrical conductivity and other physical properties to energy transitions of electrons in a solid due to a number of energy sources (light, heat, electrical). This is complicated in thin films by the fact that defects, lack of long range order, and compositional variations can affect and degrade their band structure [54]. These factors can totally smear out band structure in thin films, not to mention size effects. Quantum theory also predicts magnetic properties, magnetoresistance, thermal conductivity, thermoelectric effects, and optical phenomena.

Thin films are also applied to non-piezoelectric surfaces to provide a number of piezoelectric and ferroelectric properties. Materials such as AlN, ZnO, Si\(_3\)N\(_4\), BaTiO\(_3\), PbTiO\(_3\), LiNbO\(_3\) and KNb\(_3\)O\(_5\) possess excellent piezoelectric performance, if the microstructure is correct [55]. Such materials can be used for high frequency transducers, sensors, actuators and motors.

### 1.5 Corrosion of Solid Surfaces

In addition to wear, corrosion causes significant damage and economic losses and reliability problems. Corrosion is the degradation and removal of an engineered material's surface into its constituent atoms due to chemical reactions (oxidation, sulfidation, chlorides/salts) with its environment [10]. It is essentially the electrochemical oxidation of metals in reaction with an oxidant such as oxygen or sulfur. The severity of corrosion is determined by chemical kinetics, which can have a strong dependence on temperature. Formation of Fe\(_2\)O\(_3\), or rust, on an iron surface is due to oxidation of the iron atoms in solid solution and is a common example of electrochemical corrosion. Corrosion can also occur in other materials than metals, such as ceramics and polymers. Corrosion can also be self-limiting when a protective scale forms on the surface of the metal.
Corrosion can be classified as dry, wet, or stress enhanced corrosion [10]. Dry corrosion is a chemical process that involves interaction of a gaseous environment with a solid surface, for example, oxidation in air. Gaseous sulfides and halides can also be strongly corrosive to metal surfaces. Wet corrosion involves exposing the surface to liquids, which may or may not function as electrolytes (wet galvanic or nongalvanic). Ionic charge transport in an electrochemical redox process occurs in galvanic corrosion, which involves an aqueous medium with dissolved salts. If a deposit does not form, corrosion can proceed until the entire object or component is consumed. In wet nongalvanic corrosion, the surface is dissolved by the aqueous medium in a non-redox reaction. Again, it is possible for the entire solid structure to be dissolved.

Stress enhanced corrosion is characterized by corrosion concentrated locally to form a pit or crack, or it can extend across a wide area more or less uniformly corroding the surface. Because corrosion is a diffusion controlled process, it occurs on exposed surfaces. Passivation and chromate conversion can increase the corrosion resistance of a surface. Applied stress or fatigue can enhance localized corrosion, even in otherwise benign conditions. Stress-corrosion cracking occurs when a surface is simultaneously exposed to a corrosive environment and applied stress. Intergranular regions and intragranular regions are extremely susceptible to this type of corrosion. Corrosion fatigue occurs when loads and stresses are variable.

A number of surface treatments are possible to mitigate the above types of corrosion, including:

- Corrosion resistant thin film coatings
- Reactive coatings
- Galvanization
- Anodizing
- Cathodic protection

Corrosion resistant thin films have the following general characteristics:

- Inert to the environment
- Wide range of thicknesses
- High density
- Continuous over the entire area (prevent permeation)
• Low permeability
• Very low porosity
• Low stress

This completes our brief overview of many of the properties that a solid surface can display. All these properties can be significantly modified by application of thin films, low dimensional structures, and other surface engineering techniques.

Each chapter of this book is dedicated to a specific area of surface engineering or materials engineering. Thin film deposition processes are described in chapter 2. Structure and microstructure of the thin film depend almost entirely on the deposition process, and are addressed in chapter 3. Chapters 4–6 address tribological, optical, electric, and opto-electronic aspects of surface engineering. Films and coatings are often multifunctional, and to this end, chapter 7 is an introduction to how materials and physical properties are engineered onto surfaces and reflect the most up-to-date aspects of surface engineering science. Hybrid surface engineering applications are described in chapter 8. We look ahead to the future in chapter 9 with emphasis on functional biological materials.

References

27. International (ISO) and European (CEN) Standard
35. ASTM D1474.
50. See HyperPhysics: Color Vision.
52. See Hyperphysics.com.