MSE 104: Microstructure and Properties of Materials

Phase Metallurgy

Dr David Dye
2013–14

Comments and corrections to david.dye@imperial.ac.uk

Lecture notes may be found on Blackboard (http://learn.imperial.ac.uk)
# Contents

1 Introduction .................................................. 1
   1.1 About the Course ........................................... 1
      1.1.1 Syllabus ............................................. 1
      1.1.2 Further Reading ...................................... 2
      1.1.3 Course Support and Assessment ....................... 2
   1.2 Motivation: Structural Aerospace Materials ................. 3
   1.3 Mechanical Properties ..................................... 4
      1.3.1 Elasticity ........................................... 5
      1.3.2 Atomic Basis of Elasticity .......................... 6
      1.3.3 Plasticity .......................................... 7
      1.3.4 Resilience ......................................... 7
   1.4 Phases .................................................... 9
      1.4.1 Observing Microstructures ............................ 11
      1.4.2 Phase Diagrams ...................................... 12

2 Alloys and Phases: Microstructures ............................ 14
   2.1 Single and Multi-Phase Alloys ............................. 14
   2.2 Counting up the atoms: Weight and Atom Fractions ......... 16
   2.3 Formation of a microstructure ............................ 17
   2.4 Solidification ........................................... 18
   2.5 Continuous Solid Solutions; The Lever Rule ............... 18

3 Eutectic Phase Diagrams ....................................... 22

4 Introduction to Thermodynamics ................................ 29
   4.1 Internal Energy $U$ ....................................... 29
   4.2 Enthalpy $H$ ............................................ 30
   4.3 Entropy $S$ ............................................. 31
   4.4 Gibbs Energy $G$ ....................................... 32
   4.5 Relations for equilibrium ................................ 33
   4.6 Helmholtz Energy $F$ ................................... 33
   4.7 Intensive vs. extensive variables ......................... 33
   4.8 Energy of a single phase of constant composition ......... 34
   4.9 Two phases of identical composition ................. ....... 34

5 Thermodynamics of Solutions .................................. 35
   5.1 Enthalpy of mixing $\Delta H_{\text{mix}}$ .................... 35
   5.2 Entropy of Mixing $\Delta S_{\text{mix}}$ ..................... 37
   5.3 Gibbs Energy of Mixing .................................. 38
   5.4 Making Phase Diagrams ................................... 39
Contents

6  Faster Solidification: the Scheil equation 42

7  Diffusion, Nucleation and Growth 45
   7.1  Homogenisation .................................................. 45
   7.2  Processing of Al Alloys ............................................. 46
   7.3  Nucleation and Growth of a New Phase ......................... 46

8  Transformation Kinetics 50
   8.1  Nucleation Rates .................................................. 50
   8.2  Growth of a New Phase ............................................. 51
   8.3  Isothermal Transformation Diagrams ........................... 52

9  Microstructure and Properties of Plain Carbon Steels 54
   9.1  Equilibrium Cooling of Steels .................................... 56
   9.2  Non-Equilibrium Cooling of Steels .............................. 58
   9.3  TTT diagrams ...................................................... 61
Chapter 1

Introduction

1.1 About the Course

The Aim of the course is to examine how microstructures are formed in metals during solidification and heat treatment. To motivate the course, the mechanical properties of structural materials are introduced through the course. A major focus is the thermodynamics of the formation of phases and the construction of phase diagrams. The steels phase diagram is introduced.

1.1.1 Syllabus

MSE104 has five main components;

1. David Dye covers the material on phases and phase equilibria
2. Robin Grimes will introduce crystallography and crystal structures
3. Fraser Wigley will introduce non-metal microstructures
4. John Plummer will introduce microscopic plasticity
5. Mark Wenman will look at the metals life cycle from processing to failure.

The material covered in the module by David Dye is as follows

- Phase Changes, e.g. liquid-solid and latent heat; cooling curves. Phase Diagrams and the lever rule. Interpreting phase diagrams. Continuous solid solutions and eutectics.
- Two phase equilibria; the enthalpy and entropy of mixing; the regular solution model.
• Derivation of phase diagrams using free energy - composition curves.
• Solidification and the Scheil equation. Segregation and homogenisation. Al-Cu.
• Nucleation and growth of a new phase; homogeneous vs. heterogeneous nucleation.
• The Fe-C phase diagram; martensite; transformation kinetics; TTT diagrams

1.1.2 Further Reading

A key skill of a graduate is the ability to extend your knowledge from textbooks, and seeing someone present the material in a different way will help you to resolve difficulties and start to become an independent learner. Therefore reading to support the lectures is very important.


1.1.3 Course Support and Assessment

This section of the course is composed of 9 lectures, 2 tutorials and some of the laboratory sessions. It is examined in the labs (5 marks each), the January exam (10 marks) and the June exam (approx. one quarter of the exam, which is worth 100 marks).
1.2 Motivation: Structural Aerospace Materials

At Sioux City in 1989 we saw how a modern airplane (Lockheed L1011) operated by a well-regarded operator (United) with a widely respected engine manufacturer (GE) could still fail due to a catastrophic fan disc burst. The crack had been present since new and had grown each flight cycle by fatigue crack growth until fast fracture eventually occurred. The Ti-6Al-4V alloy used behaved almost exactly as the textbook data would lead one to expect. The originating defect should not have been present and was due to poor melting and forging practice, which was largely rectified between the time of manufacture (1971) and the incident (1989). The growing crack was missed on multiple inspections in service. The fan disc contained so much energy that on breaking material escaped the engine casing and this severed all three hydraulic control lines. The pilots managed to crash the plane onto a runway using the thrust settings from the two remaining engines, but the airframe broke up on landing leading to the death of just under 40% of the passengers. The were many lessons from this incident for practice, but fundamentally it illustrates how fatigue failures are both common and critical to safety, accounting for around half of all airplane materials failures.

In the Aloha Airlines incident, we saw how an aircraft operated within the manufacturer’s guidelines could nevertheless suffer a catastrophic failure. In this case, the aircraft was subjected to a large number of very short flight cycles and consequently a large number of pressurisation cycles for relatively few airframe hours - the end user was flying a very frequent hopping service between Hawaiian islands. This was also predominantly at a relatively low flight level where there was quite a lot of salt water vapour present. The consequence was that the fuselage failed due to fatigue along rows of rivet holes - stress concentrators. Fortunately, few people died (1), and so this is a good example of the necessity to understand fatigue properly even when simply operating aircraft, in that the normal life and inspection interval rules are made with some flight profile assumptions in mind and if you are using the product outside of those assumptions, then you may need to check that they are correct.

The SR-71 and Concorde, both 1960s airframes, illustrate the use of creep resisting materials in airframes. In this case, supersonic speeds lead to elevated temperatures on the leading surfaces of the wings. Therefore in Concorde an Al alloy first developed for WWII aircraft engines was used, and in the SR71 titanium was used in bulk for the first time. At the time, there was no sizeable western titanium industry and extracting, purifying, casting and forging the metal...
proved so challenging that supplies of this new industrial metal were obtained from the Soviet Union, who had mastered its production, for the first few airframes. Creep is the continuous deformation of a material at constant stress, much like a plastic bag that one uses to carry home groceries (e.g. the handles stretch), and is much more pronounced at high temperatures. Such deformation can occur at quite low stresses, and is the major factor limiting the temperatures that can be used in gas turbine engines and therefore the fuel efficiency that can be achieved. Developments in the alloys used in the hottest part of the jet engine have historically driven improvements in efficiency of air transport, to the extent that most of the improvement of new generations of aircraft is due to the powerplant.

It is worth mentioning that airplanes are designed in the flying configuration, that is, after accounting for the aerodynamic loads. In fact, with the large wingspans present in modern civil aircraft, wing tip deflections on takeoff of $>1$ m are quite common. Therefore design can be modulus limited. Another good example of a stiffness, or modulus, limited structure is a spring like a suspension arm in a car.

In summary, the key mechanical properties associated with airframe and aero-engine materials are their stiffness, strength, creep resistance and their fatigue and fracture resistance.

1.3 Mechanical Properties

In generating a microstructure in a structural material, our aim is to make a material that has good mechanical properties; the job of the mechanical engineer is then to design a safe structure with the material provided by the materials scientist. The most fundamental mechanical properties are measured in a tensile test.

Figure 1.2: Aircraft after landing in the Aloha airlines incident.

Figure 1.3: Both the SR71 and Concorde used fuselage materials designed to resist creep at the wing leading edges.
In a tensile test, we usually measure the load required to generate a given displacement in the sample and grips. If the gripping arrangement in strong and stiff enough, we therefore measure the force-displacement response of the specimen’s parallel gauge length - the region of the sample of lowest area which deforms first. In order to explore this further, we need to first think about the first part of the tensile test - the response when the material behaves like a spring.

1.3.1 Elasticity

At school, you will have met the equation defining the response of an elastic spring:

$$ F = Kx $$

(1.1)

where $F$ is the force applied, $K$ is the spring constant and $x$ is the increase in length, or deflection, of the spring. First lets consider two springs 1 and 2 joined end-to-end, with the same spring constants $K$. In this case, the force on both springs would be the same, $F$, and so therefore would their deflections, so $F = Kx_1 = Kx_2$. The total deflection is equal to $x_1 + x_2$, so the effective spring constant of the two springs joined together is $K/2$. By considering what happens if we join and arbitrary number of springs together, we can therefore find that the spring constant is inversely proportional to the length of the spring. Similarly, if we pulled
on two springs simultaneously, doubling their area, the force would be carried equally split between them, so the spring constant must be proportional to the area of the spring.

In Materials Science and Engineering, we aim to talk about the properties of materials independently of the form of the sample being tested, so that we can extrapolate from testpieces to components. The line of argument above will hopefully persuade you that the right thing to do is therefore to normalise the force by the area of the spring \( A \), giving us a quantity in \( \text{Nm}^{-2} \) or Pascal Pa which we call a stress, and to normalise the deflection by the original length of the spring \( l_0 \) to give a dimensionless (\(-\)) quantity which we call strain. These are defined by

\[
\text{strain } \varepsilon = \frac{x}{l_0} = \frac{l - l_0}{l_0} \tag{1.2}
\]

\[
\text{stress } \sigma = \frac{F}{A} \tag{1.3}
\]

where \( l \) is the current length. To describe the Elastic behaviour of a material we therefore use the following equation, called Hooke’s Law,

\[
\sigma = E\varepsilon \tag{1.4}
\]

where \( E \) is the Young’s modulus (also colloquially called the stiffness) of the material. So therefore in the tensile test we first convert from the load-displacement curve we measured in the test to the test-independent cure of Stress vs. Strain. For the elastic portion of the curve the gradient will be the Young’s Modulus \( E \). This modulus actually varies by a factor of at least 10\(^5\) between materials, Table 1.1. The obvious question this table raises is: what is the source of this variation in the moduli of materials? We will explore this as we go through the course. Note, in particular, that the stiffness of a polymer backbone (C-C bonds) and diamond (also C-C bonds) are apparently different by about a factor of 2 \( \times 10^5 \).

### Table 1.1: Young’s Moduli for a range of engineering materials, in GPa.

<table>
<thead>
<tr>
<th>Material</th>
<th>Modulus E (GPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steel</td>
<td>220</td>
</tr>
<tr>
<td>Cu/Brass</td>
<td>120</td>
</tr>
<tr>
<td>Ti alloy</td>
<td>105</td>
</tr>
<tr>
<td>Al alloy</td>
<td>70</td>
</tr>
<tr>
<td>Mg</td>
<td>42</td>
</tr>
<tr>
<td>Diamond</td>
<td>1200</td>
</tr>
<tr>
<td>Rubber</td>
<td>0.007</td>
</tr>
<tr>
<td>Polythene</td>
<td>1.4</td>
</tr>
<tr>
<td>Plywood</td>
<td>7</td>
</tr>
<tr>
<td>Concrete</td>
<td>17</td>
</tr>
<tr>
<td>Bone</td>
<td>21</td>
</tr>
<tr>
<td>Glass</td>
<td>70</td>
</tr>
</tbody>
</table>

1.3.2 Atomic Basis of Elasticity

In order to consider how modulus arises, we consider the energy released when two atoms bond together. If bonding is favourable, then the energy state will be lowered if we bring two atoms together from infinity. If we push them too close together, the strong nuclear force will dominate, stopping us fusing the atoms, and this would require us to add energy. Therefore we obtain the curve shown in Figure 1.6. The simplest description for this energy variation is the Lennard-Jones potential,

\[
U = 4\epsilon \left[ \left( \frac{\theta}{r} \right)^{12} - \left( \frac{\theta}{r} \right)^6 \right] \tag{1.5}
\]

where \( U \) us the energy, \( r \) the atomic separation, \( 4\epsilon \) characterizes the depth of the potential well (the bonding energy) and \( \theta \) the diameter of the well. If we displace an atom from its natural position at the lowest energy, e.g. the bottom of the well, then the restoring force \( F \) will be given by

\[
F = -\frac{\partial U}{\partial r} \tag{1.6}
\]
since energy = force × distance. This is shown in Figure 1.6 From the spring equation \( F = Kx \), we can therefore deduce that the spring constant \( K \) for each bond will be given by \( \partial F/\partial r \), or the gradient of the force line at the position of the bottom of the well. This is the source of the modulus, so we can say that modulus is a property of the bonds in the material.

Therefore processing of a material to change its microstructure and small alloying additions cannot change the modulus very significantly, because they don’t affect most of the bonds that provide stiffness. In addition, the variation seen in Table 1.1 is due to the bonding of the atoms (although this still doesn’t answer the question about polymers vs. diamond).

1.3.3 Plasticity

At some point, the material reaches the elastic limit, termed the yield stress \( \sigma_y \), after which there is a component of permanent plastic deformation. Typically, the gradient of the stress-strain curve is lower after yielding. If at some point during plastic deformation the test is interrupted then the material typically unloads elastically, with the same modulus as before. On reloading, the stress-strain curve will typically continue as before, so the yield stress will have risen. Therefore plastic strain can increase the yield stress of a material, which is called work hardening - if we put work into the material it gets stronger. This is shown in Figure 1.7.

Typical values for the yield stress for aerospace materials are shown in Table 1.2. Also shown is the yield strain \( \sigma_y/E \). We can see that similar alloys can have widely varying strengths - in fact, pure iron with large grains can have strengths as low as 50 MPa, but tool steels can have strengths up to 2000 MPa. So strength is largely a consequence of microstructure - for example, pure metal single crystals are very weak. Data for a mild steel are also shown.

1.3.4 Resilience

For a spring, a key property is how much energy it can store before yielding. The energy stored \( U \), termed the resilience, is the area under the elastic part of the stress-strain curve,

\[
U = \frac{1}{2} \sigma_y \varepsilon_y = \frac{1}{2} \sigma_y^2 / E
\]  

(1.7)
Figure 1.7: Typical Stress-Strain curve for a ductile metal.

Table 1.2: Strengths for typical aerospace materials, and mild steel for comparison.

<table>
<thead>
<tr>
<th>Material</th>
<th>Modulus $E$ (GPa)</th>
<th>Yield Stress $\sigma_y$ (MPa)</th>
<th>Yield Strain $\sigma_y/E$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ti-5553</td>
<td>110</td>
<td>1400</td>
<td>1.3</td>
</tr>
<tr>
<td>Ti-64</td>
<td>105</td>
<td>900</td>
<td>0.86</td>
</tr>
<tr>
<td>Al alloy</td>
<td>70</td>
<td>550</td>
<td>0.78</td>
</tr>
<tr>
<td>CFRP</td>
<td>120</td>
<td>1200</td>
<td>0.75</td>
</tr>
<tr>
<td>High $\sigma_y$ steel</td>
<td>210</td>
<td>960</td>
<td>0.47</td>
</tr>
<tr>
<td>Mild steel</td>
<td>210</td>
<td>250</td>
<td>0.12</td>
</tr>
</tbody>
</table>

We can therefore deduce that this energy varies widely between alloys, which leads to the conclusion that titanium alloys and composites such as CFRP are much more effective spring materials than steels and aluminium alloys, Table 1.4.

Ductility

The tensile test is of course terminated when the material reaches a maximum deformation that it can sustain; the strain at this point is termed the ductility. Ductility can be as high as 30%; generally 5% is considered a minimum and > 10% is desirable in aerospace applications. Ductility is related to the maximum energy adsorbed by the material, i.e. the area under the stress-strain curve. However, the behaviour of the material around a crack tip is quite different and so neither ductility nor the area under the curve are good measures of the crack resistance of a material in fatigue or fracture, which are measured by other properties.
Table 1.4: Specific properties for typical aerospace materials.

<table>
<thead>
<tr>
<th>Material</th>
<th>Young's Modulus $E$ (GPa)</th>
<th>Yield Stress $\sigma_y$ (MPa)</th>
<th>Yield Strain $\sigma_y/E$ (%)</th>
<th>Max Storable Elastic Energy $U$ (MPa)</th>
<th>Density $\rho$ (kg.m$^{-3}$)</th>
<th>$E/\rho$ (MJ.kg$^{-1}$)</th>
<th>$\sigma_y/\rho$ (kJ.kg$^{-1}$)</th>
<th>$U/\rho$ (kJ.kg$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFRP</td>
<td>120</td>
<td>1200</td>
<td>0.75</td>
<td>6.0</td>
<td>1600</td>
<td>75</td>
<td>750</td>
<td>3.8</td>
</tr>
<tr>
<td>Ti-5553</td>
<td>110</td>
<td>1400</td>
<td>1.3</td>
<td>8.9</td>
<td>4500</td>
<td>24</td>
<td>310</td>
<td>2.0</td>
</tr>
<tr>
<td>Ti-64</td>
<td>105</td>
<td>900</td>
<td>0.86</td>
<td>3.9</td>
<td>4500</td>
<td>23</td>
<td>200</td>
<td>0.9</td>
</tr>
<tr>
<td>Al alloy</td>
<td>70</td>
<td>550</td>
<td>0.78</td>
<td>2.1</td>
<td>2700</td>
<td>26</td>
<td>200</td>
<td>0.8</td>
</tr>
<tr>
<td>High $\sigma_y$ steel</td>
<td>210</td>
<td>960</td>
<td>0.47</td>
<td>2.2</td>
<td>7800</td>
<td>27</td>
<td>125</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Specific properties

In aerospace engineering, we care very much about weight because the mass of the aircraft also has to be transported and this has a cost in fuel, so it is instructive to examine the properties per unit mass, termed the specific properties, Table 1.4. Note in particular that the specific modulus is very similar for most metals, but is much better for carbon fibre composites. The resilience of titanium alloys can be quite high, and it is used in automotive suspension springs for this reason, but on a weight-adjusted basis CFRP is superior. In addition, the low yield strength of Al wing spar alloys like 7075 is compensated for by their low density, and the high density of steels is catastrophic for their specific properties in general. CFRP in general appears fantastic on a specific basis, which is why airframe designers favour this material. However, its impact resistance and repairability are a cause for concern that has hindered adoption of these materials. Finally, note that the table below lists aero-grade, best-of-class materials, and that typical bulk structural materials often perform far less well.

These properties of engineering materials depend strongly, therefore, depend strongly on the bonding within the material. However, as we will see, this bonding depends on the phases that form. Moreover, the interfaces between the phases are critical to the yield strength and ductility of the material. This is the critical insight in materials science - that the microstructure of a material is at least as important as its composition (chemistry) in determining its properties. So, a key component is to determine what we mean by phases.

1.4 Phases

You have been dealing with phases all your life; probably the simplest example is for a pure chemical compound; say $\text{H}_2\text{O}$ of water. This has three common phases; vapour or steam, liquid water and solid ice [there are in fact a number of variants of solid ice, but we’ll leave that point for now]. Figure 1.8 shows under what conditions of temperature and pressure each is stable.

When we consider thermodynamics later we will find that a material will always tend to take up its lowest energy state, at least, eventually. In a pure material, this means that under any given external conditions (here, Pressure and Temperature), only one phase is stable. Of course, there are lines at the boundary of stability of each phase where the two coexist; so an ice-water mixture can stably exist at a range of temperatures for pressures above TP in the Figure. At room pressure this is at $0^\circ\text{C}$, of course; similarly water and steam coexist at $100^\circ\text{C}$ at room pressure ($10^5\text{ Pa}$). In this diagram there also exists a single unique point, the Triple Point, where all three phases can coexist.
When we move between the phases, we quite often find there is some latent heat given up; one of the phases seems to have excess energy compared to the other. Why then, doesn’t the transformation happen earlier? This, it turns out it because the lower energy phase is more ordered; lower order states are more energetic. This energy we call *Entropy*.

So what are these phases? In pure materials they seem to be different ways of arranging the atoms or molecules; in ice they take up a regular array, in a liquid they touch but are arranged randomly whilst in a gas they don’t even have to touch at all. When we mix different compounds, we can extend the idea of phases, which is really the subject of this course.

If we examine the microstructure of a two-phase brass, Figure 1.9, we see that the material is composed of two phases. So first we need to define what we mean by a phase.

*A phase is a specific arrangement of atoms in a regular repeating array, with a defined range of atomic compositions on each symmetrically unique site in the array*.

**Figure 1.8:** Pressure-Temperature diagram for pure water. Figure adapted from E. Zimbres [CC-BY-SA-2.5], from Wikimedia Commons.

**Figure 1.9:** Two-phase microstructure of a 60 wt.% Cu - 40 wt.% Zn brass, as used in domestic electricity plugs. The microstructure consists of precipitates of a light phase (CuZn) in a field of a dark phase (Cu). From Doitpoms, University of Cambridge.
... which is quite a mouthful! Therefore there is an element of crystallography in defining phases (Prof Grimes’ lectures). As an example of a crystal structure in 2D, examine Figure 1.10.

If the circles represent atoms (the hard sphere approximation) then Figure 1.10 is the most efficient way to pack in 2D. Here we can identify a repeating unit, called the unit cell, which in this case contains one atom. In 3D, stacking series of these layers turns out to produce the so-called face centred cubic, or fcc, crystal structure, Figure 1.11. Here the unit cell is a cube, with the atoms at the corners and on the face centres. There are eight corners to the cube, and each atom at a corner site is shared with eight other cubes. Thus, there are \(8 \times \frac{1}{8} = 1\) atoms at the corners. There are 6 faces in a cube, and each atom on a face centre is shared with one other cube, so there are \(6 \times \frac{1}{2} = 3\) atoms on the face centres; so there are 4 atoms per unit cell in fcc packing.

At school, you may have met allotropes, which are the different structures taken up by compounds under different conditions of pressure and stress. For example, iron (Fe) is face-centred cubic above 981°C, which is called the \(\gamma\) phase. At lower temperatures it takes on another crystal structure, the body-centred cubic structure (bcc), where the unit cell is a cube with an additional single atom in the centre of the cube. Allotropes are similar in concept to phases, but a phases extend to a range of compositions.

1.4.1 Observing Microstructures

Metals are generally constructed from a large number of crystals (termed grains - like the grains of sand on a beach) composed of one or more phases. These are generally small, on the order of 10’s of \(\mu\)m, but can range from nm to cm in size, and consequently we observe them using microscopes. The simplest of these is the light microscope, which has a resolution limit which is a few times the wavelength of visible light (\(\sim 500\) nm), so around 1 \(\mu\)m. To go further, we can use electrons as the probe, either of the surface of the material in the scanning electron microscope (SEM), or we can make a very thin foil of material, on the order of 100 nm thick, and observe the material using the transmission electron microscope (TEM). Modern TEMs can routinely image the columns of atoms in a material. In both optical and SE microscopes it is common to prepare a polished section of the metal, with a mirror finish, and then to provide
contrast by selectively etching features in the material such as the grain boundaries. This is how the image in Figure 1.9 was obtained, using a light microscope.

If we examine Figure 1.9 afresh, we can see that it is composed of two phases, light and dark. The light phase is fcc $\alpha$-Cu. This phase can contain quite substantial amounts of Zn in its structure, up to around 35%. However, beyond that solubility limit then another phase forms, the $\beta$-CuZn phase.

In titanium, Figure 1.12, the high temperature $\beta$ phase is partially stabilised to room temperature by additions of vanadium (V). The low temperature $\alpha$ phase is present as fine plates, hundreds of $\mu$m long and $\sim 1 \mu$m thick.

1.4.2 Phase Diagrams

In the remainder of the course we will be chiefly concerned with predicting phase diagrams and using them to predict and rationalise microstructures. Phase diagrams show areas of the predominance of phases with varying temperature, composition or pressure. An example of a simple eutectic phase diagram is shown in Figure 6.5.

Here we plot the areas of stability of different phases with composition for alloys composed of two types of atoms, $A$ and $B$, with temperature. By convention, unless the pressure is stated, standard room pressure is assumed. The diagram is interpreted as follows:

- At high temperatures a single phase exists for all compositions, the liquid $L$. This is very typical for molten metals.
- A pure melt of $A$ would then solidify and transform two a solid phase $\alpha$. This phase $\alpha$ has an increasing capacity to take $B$ into its structure as the temperature decreases and its stability relative to the liquid increases - the alloy ‘wants’ to be solid.
- Similarly, a pure melt of $B$ would solidify (at a different temperature to $A$) and form a phase $\beta$.
- Therefore for a dilute alloy of $B$ in $A$ a two-phase region can exist, where solid $\alpha$ co-exists with liquid $L$.
- At some point, it becomes more favourable to form a mixture of $\alpha$ and $\beta$ instead of either $L + \alpha$ or $L + \beta$, and the alloy becomes completely solid (called the eutectic point).

Therefore there are three regions in the diagram where a single phase can exist and three regions where a mixture of two phases exists.

We will spend much of the remainder of the lecture series thinking about how to construct, use and interpret phase diagrams, which will allow us to understand microstructures, the foundation of metallurgical materials science.
Figure 1.13: A simple eutectic phase diagram.
Chapter 2

Alloys and Phases: Microstructures

Previously, we saw how a metal can be composed of grains, which are single crystals, possibly of different phases. Therefore in a single phase material, there are three lengthscales we consider; that of the component, of the microstructure and of the crystal lattice, Figure 2.1. In general, the smaller the grains in a material, the stronger it is, so one of the key questions is how we can make fine-grained materials.

Alloys are mixtures of two or more elements, usually made by mixing the elements together in a liquid. This is because most elements are soluble in each other in the liquid state, allowing a randomly mixed solution to form. However, solid metals often have a restricted capacity to take other elements into solution - they have limited solubility, and so alloys are often composed of two or more different phases. For the rest of this section, we will consider alloys composed of two elements generically termed ‘A’ and ‘B’.

2.1 Single and Multi-Phase Alloys

In the simplest case, if we take an alloy composed mostly of A with limited amounts of B where the two elements are soluble in each other, then we can form a single phase material, called a solid solution, Figure 2.2. This occurs most commonly where the elements have similar atomic sizes and crystal structures. Only a few pairs of atom form a solid solution over the entire range of compositions from 100% A to 100% B but when this occurs each grain or crystal is a random mixture, at the atomic scale, of the two elements. For example, an alloy with 70% A and 30% B would appear the same as an alloy of pure A - the only difference would be that 30% of the sites in the crystal structure would have B atoms rather than A atoms. This is called a substitutional solid solution.

Whilst complete solid solubility is rare, most combinations of elements have some solubility in each other - for example A might be able to take only 5% B into its crystal structure with phase α, and conversely B might be able to take only 7% of A into its structure β. This situation is termed partial solubility.
Another possibility is where the A and B atoms have very different sizes - for example iron Fe and carbon C. In a steel, that is, an iron-based alloy, the majority element is iron. The majority element is called the solvent and the minority element the solute. If the solute atoms are very much smaller than the solvent, e.g. C in a solvent of Fe, then it is possible to form a solid solution where the solute atoms fit into some of the gaps in the crystal structure. These gaps are termed the interstitial sites and so this type of solution is called an interstitial solid solution.

In general, we find that solid solutions are very much (3-4×) stronger than pure metals.

In our alloy above with partial solubility, what happens when we make an alloy of A and B that lies outside the solubility range for either the $\alpha$ or $\beta$ phases? If there were no intermediate compounds, an alloy with a B content between 5 and 95% would then be composed of an appropriate mixture of the $\alpha$ and $\beta$ phases, with both phases taking the composition of their solubility limit. So, for example, if we made an alloy with 10% B, then we would form...
a mixture of $\alpha$ phase out of A with 5% B and $\beta$ phase composed of B with 7% A.

Sometimes, intermediate compounds can form in the phase diagram; technologically important examples include Ni$_3$Al, Fe$_3$C, NiTi and Nb$_3$Ti. Ni$_3$Al is used in nickel superalloys, which have the best high temperature mechanical performance of all known alloys and are consequently used in aero-engines and rockets. Fe$_3$C is the basis of plain carbon steels and therefore one of the foundations of modern society. NiTi takes on different crystal structures under varying conditions of temperature and stress; the resulting shape change allows actuators to be made that respond to temperature and stress. Finally, Nb$_3$Ti is superconducting at low temperatures and is used in the majority of superconducting magnets used in the world today, for instance in the hundreds of multi-billion $\ $ particle physics facilities in the world.

Take the example of a phase $\alpha$ with solubility for 7% B and a phase $\delta$ with composition A$_3$B, then an alloy with 15% B would be composed of a mixture of $\alpha$ (A with 7% B) and $\delta$. Conversely, an alloy with only 3% B would be pure $\alpha$.

### 2.2 Counting up the atoms: Weight and Atom Fractions

Alloys are defined according to the fractions of each element they contain. Often, the fractions are defined by weight - the number of grammes of each element they contain, since weight is the quantity that the person making the alloy can most conveniently measure. However, scientifically the atomic bonds are defined by the number of atoms - the number or atom fraction.

For example, consider a molten alloy in a bucket, Figure 2.6. Ignoring the weight of the container, how much will it weigh? Defining each atom type $x$ with atomic mass per mole $m_x$ and atomic fraction $C_x^a$, then $N$ atoms of alloy will weigh

$$\frac{N}{N_A} \sum_x m_x C_x^a$$

(2.1)

where $N_A$ is Avagadro’s number. The weight of only the atomic species $x$ is just $\frac{N}{N_A} m_x C_x^a$. So the weight fraction $C_x^w$ of species $x$ is;

$$C_x^w = \frac{\frac{N}{N_A} m_x C_x^a}{\sum_x \frac{N}{N_A} m_x C_x^a} = \frac{m_x C_x^a}{\sum_x m_x C_x^a}$$

(2.2)

As an example, consider the compound Fe$_3$C. Obviously, the atomic fractions $C_x^a$ must sum to 1, so these are 0.75 for Fe and 0.25 for C. So the weight fraction of carbon in Fe$_3$C is

$$C_C^w = \frac{12.01 \times 0.25}{12.01 \times 0.25 + 55.85 \times 0.75} = 0.0668$$

(2.3)
2. Alloys and Phases: Microstructures

so the atom fraction of carbon in Fe₃C is 6.68%. The inverse operation is to convert from weight fraction to atomic fraction. If we consider 1 kg of alloy in our bucket, that has a weight fraction of atom x of \( C^w_x \), the x atoms will weigh \( C^w_x \) kg. A mole of them would weigh \( m_x \) and therefore there must be \( C^w_x / m_x \) mol of them in our bucket, or \( C^w_x N_A / m_x \) atoms in total. Therefore the total number of atoms in the bucket is just the sum of this quantity over all the elements, and so atomic fraction of \( x \), \( C^a_x \), is

\[
C^a_x = \frac{C^w_x / m_x}{\sum_x C^w_x / m_x} \tag{2.4}
\]

For example, the alloy Ti-10V-2Fe-3Al has weight fractions of 85% Ti, 15% V, 2% Fe and 3% Al. However, the at.%’s are quite different for Fe and Al - in at.% the alloy would be quoted as Ti-9.3V-1.7Fe-5.2Al.

**Exercise:** check for yourself that, if you repeat these two examples working the other way around, you recover the original data.

**Question:** Conventionally, alloy compositions are given in wt.% - is this appropriate or would at.% be better?

### 2.3 Formation of a microstructure

As stated earlier, most engineering alloys are made by mixing together the alloying elements in the liquid state and then solidifying the alloy into a suitable shape, a process known as *casting*. It is possible to cast the alloy directly into a shape at or close to the dimensions of the finished component, but commonly cast materials have large grain sizes, vary in composition within the alloy (segregation) and contain pores (holes) and inclusions (*e.g.* oxide particles). All of these give rise to poor mechanical properties, especially in cyclic loading conditions, called fatigue. Therefore very commonly the cast ingot is first hot worked and heat treated, that is, deformed at high temperatures, to break up the inclusions, close the pores, homogenise the composition and reduce the grain size. These optimization of the material processing sequence is a major topic in materials development, since it concerns the control and optimisation of the microstructure that gives rise to the properties of the material.

![Slab Caster](image)

**Figure 2.7:** Alloys are usually made by mixing metals in the liquid state and then casting them into a solid shape. Often, the material is then deformed at high temperatures (hot working) to make the final microstructure.
In metallurgy we use the physical processes that operate in materials to produce the microstructure we desire, using processing techniques that we can, by and large, control for large quantities (tonnes) of material at a time. This enables us to produce large amounts of material quickly and cheaply with good properties. This is contrast with nanotechnology, where very typically the microstructure is produced step-by-step, layer-by-layer, very laboriously and expensively.

### 2.4 Solidification

Therefore the first process we should aim to understand in the solidification process. In a pure metal, or any pure material, the liquid-solid transition occurs at a single temperature, the melting point, at which latent heat is evolved. This latent heat is the energy difference between the high energy, disordered, liquid state and the lower energy, ordered, solid state - in the liquid the atoms are arranged randomly whereas in the solid they are arranged on a regular crystal lattice. Therefore the solidification reaction is exothermic, i.e. it produces heat.

If heat is continuously extracted from the liquid, therefore, the temperature drops. When the solidification process starts, the latent heat evolved prevents the metal from continuing to cool. The rate of solidification is therefore determined by the rate at which the latent heat can be extracted from the metal. After the solidification process completes, the solid metal then resumes cooling as heat is extracted. If we then wanted to re-melt the material, we would have to supply heat (thermal energy) to overcome the latent heat required for melting.

### 2.5 Continuous Solid Solutions; The Lever Rule

The simplest type of alloy phase diagram is the continuous solid solution, Figure 2.9. Here we have two metals, probably with very similar atomic sizes and crystal structures, that form a single continuous phase in both the liquid and solid states but have different melting points. Therefore there are only two phases - the liquid L and the solid S. A pure solution of either A or B would have a single melting point, $T_m(A)$ or $T_m(B)$. For compositions in between, say at the point identified with composition $C(x)$, on cooling the liquid phase will first cool with a single composition $C(x)$, until it reaches the liquidus line. An alloy of A and B would therefore cooling in the following way: first, it would cool to a temperature below the melting point of pure A, until it reached the liquidus line of the diagram. At this point it would become thermodynamically favourable for the alloy to split into two phases, solid S which was relatively pure in A, and liquid with composition C(x). Initially only very little solid would form, as the overall mixture of the two phases would have to have the same overall composition as the alloy. As the alloy cooled down in the two-phase
field L+S, the alloy would solidify, with both the solid and liquid phases increasing in their overall content of B, but the fractions of phases would change with more solid being formed keeping the overall composition of the system the same. The solid phase would follow the limit of solubility of S with temperature - the *solidus*, while the liquid phase L would follow the *liquidus*.

On the cooling curve the latent heat of solidification would therefore be evolved over the range of temperatures between the solidus and liquidus temperatures for an alloy of that composition, and so the cooling curve would appear as shown in Figure 2.11. In this Figure, we can see that we can identify the liquidus and solidus temperatures for any alloy with composition from pure A to pure B, enabling us to determine the phase diagram in Figure 2.9 experimentally. This diagram shows the phases that will form at equilibrium, that is, at the lowest energy state the system can be in. If the atoms are sufficiently mobile, given enough time equilibrium will always be obtained. Therefore phase diagrams are the most important tool we have for understanding how microstructures form in materials. However, care should be taken when using equilibrium phase diagrams as they tell us only what is favourable to occur; we require an understanding of kinetics (diffusion) to understand what happens and how fast.

In order to calculate the fractions of the phases that form, we need to perform a calculation that embodies our understanding of the physical processes that occur - in this case that the mass of atoms is conserved. At a point in the middle of the two phase field, let us denote the composition of the alloy $C$, of the solid $C_S(T)$ and of the liquid $C_L(T)$. The volume fraction of the phases we will denote $V_S$ and $V_L$. The volume fractions must add to 1, so

$$V_S + V_L = 1$$  \hspace{1cm} (2.5)
The composition of the system will then be given by

\[ C = C_S V_S + C_L V_L \]  \hspace{1cm} (2.6)

because the mass of element B in the alloy is given by the total of the mass in each phase.

Therefore we have two equations and two unknowns \( V \). Since \( V_S = 1 - V_L \) we can substitute for \( V_S \) into the second equation to obtain

\[ C = C_S (1 - V_L) + C_L V_L \]  \hspace{1cm} (2.7)

expanding and collecting terms we obtain

\[ C = C_S + V_L (C_L - C_S) \]  \hspace{1cm} (2.8)

and therefore

\[ V_L = \frac{C - C_S}{C_L - C_S} \]  \hspace{1cm} (2.9)

Therefore the volume fraction of liquid is given by

\[ V_L = \frac{\text{distance between the alloy composition and solidus}}{\text{distance between the solidus and liquidus}} \]  \hspace{1cm} (2.10)

If we had substituted for \( V_L \) instead of \( V_S \) we would instead have obtained

\[ V_S = \frac{C_L - C}{C_L - C_S} \]  \hspace{1cm} (2.11)

The ratio between the fractions is therefore given by

\[ \frac{V_S}{V_L} = \frac{C_L - C}{C - C_S} \]  \hspace{1cm} (2.12)

This is called the \textit{lever rule}, as it is analogous to a lever. Imagine the volume fractions as weights hanging on a lever, where the fulcrum of the lever is placed at the composition of the alloy \( C \) and the lever is the horizontal line joining the compositions. Taking moments we therefore find

\[ V_S (C - C_S) = V_L (C_L - C) \]  \hspace{1cm} (2.13)
which is the same as Equation 2.12. This enables us to determine the phase fractions and compositions that form at any point in a phase diagram, which is amazingly useful - it enables us to use phase diagrams to understand how a microstructure will evolve in a material.

Returning to our alloy cooling curve, we can now depict how the microstructure will evolve during solidification. As the liquid cools, when we first reach the liquidus we can draw a tie line between the liquid and solid and use the lever rule to determine the fraction of solid that forms. Since \( C = C_L \) at this point, \( V_S = 0 \). At a temperature just below this point, \( V_S \) becomes greater than zero and some solid starts to form. As solidification proceeds, \( V_S \) and \( C_S \) increase. Also, \( C_L \) increases, but \( V_L \) decreases. Eventually, as the alloy cools the alloy composition enters the single phase field again where it intercepts the solidus line, and \( V_S \) becomes 1; solidification is complete.

Figure 2.14: Microstructure development during cooling of a solid solution alloy.
Chapter 3

Eutectic Phase Diagrams

In the last lecture, we saw how a metal solidifies for the case of a continuous solid solution. In this lecture, we will examine the situation where the alloying elements do not have complete solubility for each other, but have only a limited capacity to take each other into solution. So again, we return to a material composed of two atoms A and B, but this time the A atoms prefer to form an \( \alpha \) phase and the B atoms a \( \beta \) phase. As is usually the case, complete solubility in the liquid is observed. Therefore we obtain a phase diagram like Figure 3.1. The limit of stability of the liquid phase is the liquidus line, and this decreases in temperature as we move away from the pure metals - the liquid becomes more stable relative to the solid phases. There are two-phase regions where we must form mixtures of both phases, and then there are the solid phase fields themselves. Their limits of stability with respect to the liquid are the solidus lines, and with each other simply the ‘solid solubility limits’.

Geometrically there has to be a point therefore where the liquid transforms to the two solid phases at a single temperature, called the eutectic point. This occurs at the eutectic temperature. We will spend most of this lecture examining how alloys in different regions of the eutectic phase diagram behave.

There are three or four regions we can identify, coloured in Figure 3.2. The blue region at the edges denotes regions that will end up as a single phase at equilibrium at room temperature; the red region denotes those compositions that enter the single phase field at temperature, but exit it again in the solid state. The green region represents the eutectic and near-eutectic compositions. Let us examine each one in turn.

An alloy in the single phase region, Figure 3.3, will solidify just as for the continuous solid solution. Starting from the liquid, when it first reaches the liquidus line on cooling solid crystal of \( \alpha \) start to form. Solidification then proceeds and, for slow cooling under equilibrium condi-
tions, the $\alpha$ phase composition increases in B as the volume fraction increases and temperature decreases. Eventually, when the solidus line is reached, solidification finishes and only solid $\alpha$ phase remains. This phase then has the composition of the alloy. The cooling curve shows no arrest because the latent heat of solidification is evolved over a range of temperatures and so three distinct gradients are observed in the cooling curve - one for the liquid, for the solid and a third during the solidification process.

For an alloy in the two-phase region, Figure 3.4, the same solidification process is observed, but in the solid state the alloy reaches the solubility limit of the $\alpha$ phase. At this temperature, $\beta$ phase crystals start to form, a process termed precipitation. This precipitation will often occur at the high-energy features in the structure where there are nucleation sites for the new phase, such as at grain boundaries.

At a temperature in the two-phase field $T_5$, we can then use the lever rule to determine the fraction of each phase;

$$V_\beta = \frac{C - C_\alpha}{C_\beta - C_\alpha}$$

(3.1)

As usual, this is the opposite ‘small distance’ divided by the ‘big distance’ in the phase diagram. An alloy on the $\beta$ side of the diagram would behave similarly.

An alloy with the eutectic composition $C_E$, Figure 3.5, behaves in another manner entirely. On cooling, the liquid phase is stable to the lowest temperature possible, the eutectic temperature,
Figure 3.4: Solidification and solid-state precipitation in a two-phase alloy.

Figure 3.5: Solidification of the eutectic alloy.
at which point it transforms into a mixture of $\alpha$ and $\beta$. The cooling curve therefore has a eutectic arrest at $T_E$ where the latent heat of solidification is evolved, plus regions of cooling of the liquid and solid phases. Often, this solidification results in a very fine-scale structure composed of lamellae (plate-like structures) of $\alpha$ and $\beta$ phase. The fraction of each phase in the eutectic mixture we can again calculate using the lever rule:

$$V_\beta = \frac{C_E - C_\alpha}{C_\beta - C_\alpha}$$  \(3.2\)

We are now prepared to look at the most complicated case, that of a near-eutectic alloy, Figure 3.6. In this figure initially as the alloy cools from the liquid crystals of $\alpha$ form. As the temperature decreases in the $L + \alpha$ region, the solute content in the $\alpha$ increases and the volume fraction of $\alpha$ increases, whilst the liquid enriches in $B$. Eventually, the liquid becomes so enriched in $B$ that the eutectic point is reached at a composition $C_E$. Over this first part of solidification, the latent heat is evolved over a range of temperatures. The volume fraction of the eutectic liquid is then

$$V_L = V_E = \frac{C - C_\alpha}{C_E - C_\alpha}$$  \(3.3\)

The liquid then transforms to solid at the eutectic point, forming the eutectic lamellae where the liquid was in between the grains of $\alpha$ that had already formed. The composition of both the $\alpha$ lamellae and primary grains is then $C_\alpha$ and the $\beta$ lamellae in the eutectic is then $C_\beta$. The volume fraction of $\beta$ in the eutectic is then given by the lever rule:

$$V'_\beta = \frac{C_E - C_\alpha}{C_\beta - C_\alpha}$$  \(3.4\)

The overall volume fraction of beta is still given by

$$V_\beta = \frac{C - C_\alpha}{C_\beta - C_\alpha}$$  \(3.5\)
Figure 3.7: (a) A Hypoeutectic alloy (40wt.% Sn), showing dendrites of primary Pb in a Pb-Sn eutectic, (b) a eutectic Pb-Sn alloy, (c) a hypereutectic alloy (80wt.% Sn) with dendrites of primary Sn in a Pb-Sn eutectic, (d) the Pb-Sn binary eutectic phase diagram [micrographs from Richard Dashwood; diagram from ASM handbook].

which is equal to the volume fraction of the eutectic liquid \( V_E \times V_\beta \). During this final part of solidification the eutectic arrest is again observed in the cooling curve. Note that all eutectic microstructures will be the same, since they all form from liquid of the same composition.

The lead-tin Pb-Sn phase diagram is shown in Figure 3.7. Primary lead (dark phase) compositions are termed hypoeutectic (hypo=less than), whilst tin-rich compositions forming primary tin (light phase) are termed hypereutectic. In your labs, you also examine the Al-Si eutectic phase diagram, given in Figure 3.8. Notice that the primary phases that form during initial solidification usually have a dendritic structure, whilst the eutectic lamellae are very much finer in scale.

Often, intermetallic phases \( A_xB \) form in the phase diagram, and very often these intermetallic compounds have no solubility range, Figure 3.9. This phase diagram can be divided in two,
Figure 3.8: The Al-rich end of the Al-Si binary eutectic phase diagram [redrawn from the ASM handbook].

Figure 3.9: Schematic eutectic phase diagram with intermediate line compound $A_xB$; it can be divided into two separate, non-interacting diagrams.
one eutectic diagram between $\alpha$ and $A_xB$ and another eutectic diagram between $A_xB$ and $\beta$. These separated phase diagrams can then be treated in the same way as the normal eutectic diagrams we have already seen.

The situation is slightly more complicated in the case where the intermetallic has some solubility range, such as the phase $\gamma$ shown in Figure 3.10. However, despite the fact that there is no neat division, we can still use this diagram using the simple approach of stepping through the solidification sequence that we have already used.

An example of this type of more complicated phase diagram is the Al-Cu system, which is widely used as the basis for aerospace Al alloys, e.g. in wing skins. Here we see a pure Al phase and a $\theta$ phase that form a simple eutectic. At richer copper compositions the situation is more complicated because of the peritectic transformations that occur, which you will meet later in the degree programme.

Figure 3.10: Intermetallic compounds can themselves have some solubility range, such as the phase $\gamma$ depicted here.

Figure 3.11: The Al-Cu binary eutectic phase diagram [from the ASM handbook].
Chapter 4

Introduction to Thermodynamics

In the last lecture, we examined how to interpret and use eutectic phase diagrams to predict and interpret microstructures. Over the next few lectures, we will examine how these phase diagrams arise from thermodynamic considerations.

The Laws of Thermodynamics make some axiomatic statements that we shall use. Our contention will be that systems tend towards equilibrium given the choice, based upon a blend of arguments from the laws, i.e. they will attempt to minimise their energy.

First, let us consider a ball in a valley, Figure 4.1. It is stable at point E at the bottom of the valley, i.e. for small perturbations a force will act that restores it to its original position. Anywhere else on the hill it is in a state of unstable equilibrium where it is unstable to perturbations. It is also possible for the ball to be in a metastable equilibrium, at point M. Here the ball is stable to small perturbations, but for larger perturbations a lower equilibrium state can be found. Given that, above absolute zero, there will always be thermal vibrations and excitations, a material will tend to be in a state of stable equilibrium. It will be useful from herein to define some thermodynamic quantities.

4.1 Internal Energy $U$

The first law says that the change in the internal energy is equal to the heat and work that it does or are done by it;

$$\partial U = \partial q - \partial w$$

(4.1)

where we define the internal energy $U$ as the sum of the kinetic, potential and electrical energy of the atoms in the material, $q$ is the heat energy and $w$ is the work done. Generally we consider thermodynamics as a system of atoms in a container under a regime of pressure and temperature.

If the system expands the container, it performs work on the surroundings, equal to a force
times and distance or, equivalently, a pressure times the volume change of the system. Therefore

\[ \partial w = P \partial V \]  \hspace{1cm} (4.2)

Similarly, the change in temperature of the system is given by

\[ \partial q = C \partial T \]  \hspace{1cm} (4.3)

where \( C \) is the heat capacity per mole of atoms and \( T \) the temperature. So we obtain

\[ \partial U = C \partial T - P \partial V \]  \hspace{1cm} (4.4)

At constant volume therefore

\[ \partial U = C_v \partial T \]  \hspace{1cm} (4.5)

where the subscript \( v \) reminds us that the heat capacity \( C \) is the heat capacity at constant volume, also called the volumetric heat capacity. We can therefore find the internal energy as the integral of the heat capacity, together with the latent heats \( L \) due to the \( i \) phase changes and the energy at \( T = 0 \text{ K}, U_0; \)

\[ U(T) = U_0 + \int_0^T C_v \partial T + \sum_i L_i \]  \hspace{1cm} (4.6)

Of course, this implies that measuring \( U \) is impossible because we can never measure \( U_0 \). In practice what we do is we measure \( U_0 \) relative to a standard reference material at a standard reference state, and then measure changes from that reference.

### 4.2 Enthalpy \( H \)

Enthalpy is defined as

\[ H = U + PV \]  \hspace{1cm} (4.7)

This is the constant pressure version of the internal energy, which more closely represents the conditions under which experiments are performed. Therefore Enthalpy tends to be preferred by experimentalists, whilst theorists find constant volume conditions easier to work with and prefer internal energy. However, for condensed phases (liquids and solids), the differences will be small. Taking the total differential of \( H \) we obtain

\[ dH = dU + (PdV + VdP) \]  \hspace{1cm} (4.8)

But, since \( dU = dq - dw = dq - PdV \) we find that

\[ dH = dq + VdP \]  \hspace{1cm} (4.9)

and the change in heat \( dq = CdT \). At constant pressure therefore we obtain

\[ dH = C_p dT \]  \hspace{1cm} (4.10)
where $C_p$ defines the heat capacity at constant pressure. We can then measure the enthalpy is a similar manner to before;

$$H(T) = H_0 + \int_0^T C_p dT + \sum_i L_i$$  \hspace{1cm} (4.11)

Where $H_0$ is the potential energy at 0 K. Enthalpy changes correspond to the heat required when heating a material at constant pressure. Alternatively, we can measure the enthalpy relative to a standard by measuring the heat evolved when dissolving the sample in a suitable high-temperature liquid flux.

![Figure 4.3: Obtaining the enthalpy content by integrating $C_p$ from absolute zero.](image)

The enthalpy change is therefore the amount of heat liberated or consumed when a reaction takes place; if heat is required the reaction is termed *endothermic* and if heat is released it is termed *exothermic*.

### 4.3 Entropy $S$

The direction of natural changes involves entropy, the degree of disorder in the system. The third law of thermodynamics states that

*As a system approaches absolute zero, all processes cease and the entropy of the system approaches a minimum value.*

In an ideal thermodynamic system at equilibrium the entropy would be zero at absolute zero. The second law of thermodynamics also states that

*The entropy of an isolated system which is not in equilibrium will tend to increase over time, approaching a maximum value at equilibrium.*

So natural changes in the system tend to increase $S$, $dS > 0$. This does not preclude one part of the system becoming more ordered whilst another becomes less ordered. In a reversible process, the entropy change in the universe is zero, whereas an irreversible process is associated with an increase in entropy; the third case, processes where the entropy of the universe decreases, is impossible. The concept of entropy naturally gives rise to the idea of *the arrow of time*.

Entropy turns out to be a very important concept and has been worked on by very many well-known scientists in history, including Boltzmann, Maxwell, Fermi, Clausius, Kelvin and
Planck, von Neumann, Kelvin and Carnot. Often, it is felt to be dismal in its outlook on life; but science isn’t an outlook, just observation about the natural world.

The second law can also be restated to define the entropy, $S$, as

$$dS \geq \frac{dq}{T}$$

where $dq$ is the heat absorbed and $T$ is the temperature. For a reversible process, which we will most frequently consider, $dq = TdS$ whereas for irreversible processes, $dq < TdS$. Since we can measure $dq$ from the heat capacity, then we can measure $S$:

$$S = S_0 + \int_0^T \frac{C_p}{T}dT + \sum_i \frac{L_i}{T}$$

\[\text{(4.13)}\]

\[\text{Figure 4.4: Obtaining the entropy content } S \text{ by integrating } \frac{C_p}{T} \text{ from absolute zero.}\]

### 4.4 Gibbs Energy $G$

The Gibbs Energy (pre-1988: ‘Gibbs Free Energy’) is a quantity defined to be useful for defining equilibria; the Gibbs energy is minimised when a system reaches equilibrium at a given pressure and temperature. It is useful in that it allows us to do this by considering only the properties of the system, not its surroundings. $G$ is defined as

$$G = H - TS$$

\[\text{(4.14)}\]

The total differential gives

$$dG = dH - SdT - TdS$$

\[\text{(4.15)}\]

substituting for $dH$ from Equation 4.9 gives

$$dG = VdP - SdT + dq - TdS$$

\[\text{(4.16)}\]

For changes under isothermal, isobaric (pressure) conditions, $dT = dP = 0$, then

$$dG = dq - TdS$$

\[\text{(4.17)}\]

Since $TdT \geq dq$ we immediately find that $dG \leq 0$. For reversible changes $dG = 0$. Therefore we see that $G$ never increases, so that

$G$ tends to a minimum at equilibrium.
We also have the result that for two phases at equilibrium (at a particular T and P on a phase boundary), the free energy of both phases is identical (e.g. L → S in a pure metal). We can obtain \( G \) from measurements of \( H \) and \( S \).

In traditional use, the moniker 'free' was attached the Gibbs Energy, denoting the energy available to do work; the Gibbs energy is the maximum amount of non-expansion work that can be extracted from a closed system (which can only be obtained by a reversible process). However, the concept is often felt to be misleading and therefore the term has been dropped by IUPAC.

### 4.5 Relations for equilibrium

For equilibrium situations where \( dq = TdS \) we find from equation 4.16

\[
dG = VdP - SdT
\]

(4.18)

Therefore we obtain the following relations

\[
S = - \left( \frac{\partial G}{\partial T} \right)_P
\]

(4.19)

\[
V = \left( \frac{\partial G}{\partial P} \right)_T
\]

(4.20)

\[
C_P = \left( \frac{\partial H}{\partial T} \right)_P = T \left( \frac{\partial S}{\partial T} \right)_P = -T \left( \frac{\partial^2 G}{\partial T^2} \right)_P
\]

(4.21)

### 4.6 Helmholtz Energy \( F \)

The Helmholtz energy \( F \) is the equivalent of \( G \) for changes at constant volume, defined as

\[
F = U - TS
\]

(4.22)

### 4.7 Intensive vs. extensive variables

We can distinguish two groups among the thermodynamics variables we have been defining here. \( T, P \) and density are intensive variables that don’t change when two identical systems are combined. \( V, U, H, S, G \) and \( F \) are all extensive variables that are (i) proportional to the mass of the system, (ii) additive when two identical systems are combined and (iii) are functions of state - i.e. can be used to define the state of the system. Typically, we quote these extensive variables per mole in the system, i.e. for an energy, in kJmol\(^{-1}\).
4.8 Energy of a single phase of constant composition

We take as our starting points Equations 4.16 and 4.17, that

\[ dG = VdP - SdT + dq - TdS \]  

(4.23)

and the condition of reversible equilibrium, such that \( dq = TdS \). At constant pressure \( dP = 0 \) and so

\[ \left. \frac{dG}{dT} \right|_P = -S \]  

(4.24)

Therefore, since \( S \) increases with temperature, \( G \) must continuously decrease with temperature, always following the general form below. The intercept at \( 0 \text{K} \) is \( H_0 \), which is equal to \( G(0 \text{K}) \). The gradient is \( -S \), and becomes increasingly negative as \( T \) increases.

4.9 Two phases of identical composition

Each phase has its own \( G \) vs. \( T \) curve, as shown below. Where they intersect the two phases may coexist at equilibrium, at \( T_E \). Below \( T_E \) only one phase will exist in stable equilibrium, but the other phase could exist in a metastable state; above \( T_E \) the roles are reversed. When one phase transforms into another, even though \( \Delta G = 0 \) at \( T_E \), there are discontinuous changes in entropy and enthalpy associated with the release of latent heat of transformation. Therefore the transformation is termed first order.

If we define \( \Delta \) to be the change in an energy between two different states, such as ice \( \leftrightarrow \) water, we define;

\[ \Delta G = G_L - G_S = 0 \]  

(4.25)

\[ \Delta S = S_L - S_S \geq 0 \]  

(4.26)

\[ \Delta G = \Delta H - T\Delta S = 0 \]  

(4.27)

the last obtained if we impose the condition that \( \partial T = 0 \), i.e. isothermal conditions. Therefore

\[ \Delta H = T\Delta S \]  

(4.28)

Since \( \Delta S \) must be positive, we must supply latent heat to make the transformation ‘go’.

Figure 4.5: Form of variation of \( G \) with \( T \).

Figure 4.6: Form of variation of \( G \) with \( T \) for the phase change from solid ice to liquid for water.
Chapter 5

Thermodynamics of Solutions

Having examined how pure materials transform in terms of their G-T curves, we are now in a position to consider how dilute solutions behave.

If we had a complete theory of alloys it would be possible to deduce, e.g., $U_0$ from first principles quantum mechanical calculations. However, such calculations would need to be rather more precise than is commonly attainable since, because of the nature of the metallic or free electron bond, the free energies of formation tend to be rather similar. This makes it difficult to predict reliably which alloy structure would be stable. Even the best theoreticians can often only predict trends in metals. However, by making a small number of very simple assumptions we can reproduce the form of the Gibbs energy curves and therefore the form of the phase diagrams we observe.

Let us consider when we make of system that is a substitutional solid solution of $N$ atoms, with fraction $x_A$ of A atoms and fraction $x_B$ of B atoms ($x_A + x_B = 1$), by mixing together the pure metals. We will define the $G$ of the alloy as the sum of the $G$'s of the components plus the energy of mixing, $\Delta G_{\text{mix}}$, such that,

$$G = G_A x_A + G_B x_B + \Delta G_{\text{mix}}$$

(5.1)

This energy of mixing will be composed of (i) the enthalpy of mixing associated with the A-B bonds, $\Delta H_{\text{mix}}$ and (ii) an entropy of mixing $\Delta S_{\text{mix}}$ associated with the random mixing of A and B atoms. Therefore we will write, following Equation 42,

$$\Delta G_{\text{mix}} = \Delta H_{\text{mix}} - T \Delta S_{\text{mix}}$$

(5.2)

5.1 Enthalpy of mixing $\Delta H_{\text{mix}}$

Let us assume that only the potential energy part of $H$ changes significantly on mixing, so that $\Delta H_{\text{mix}} = \Delta U_{\text{mix}}$, i.e. there is no volume change on bringing together the system of A and B atoms. Then, if we define $w_{AA}$ as the interaction energy of an A-A bond, $w_{BB}$ as the energy of a B-B bond and $w_{AB}$ as the energy of an A-B bond, we will be able to find the total energy by adding up the number of bonds. All of these energies $w$ will be negative, as the zero in potential energy if conventionally for infinite separation and bringing atoms together generally is an exothermic process. We assume that each atom is co-ordinated with $z$ neighbouring atoms, so the total number of nearest neighbour pairs, or bonds, is $\frac{1}{2} N z$.  

35
If the alloy is randomly mixed the probability of finding an A atom \((x_A)\) with another A atom (also \(x_A\)) will be

\[
\text{Prob. of A–A neighbours} = x_A^2
\]  
(5.3)

Similarly for B atoms,

\[
\text{Prob. of B–B neighbours} = x_B^2
\]  
(5.4)

For A–B and B–A atoms, similarly

\[
\text{Prob. of A–B neighbours} = x_A x_B
\]  
(5.5)

\[
\text{Prob. of B–A neighbours} = x_B x_A
\]  
(5.6)

so

\[
\text{Prob. of unlike neighbours} = 2 x_A x_B
\]  
(5.7)

The total probability must be 1, so we can check this;

\[
x_A^2 + 2x_A x_B + x_B^2 = x_A^2 + 2x_A(1-x_A) + (1-x_A)^2 = x_A^2 + 2x_A - 2x_A^2 + 1 - 2x_A + x_A^2 = 1
\]

So, we can find the total interaction energy of the solution by adding up the energies and probabilities of all of the bonds;

\[
H_s \approx U = \frac{1}{2} N z (x_A^2 w_{AA} + x_B^2 w_{BB} + 2x_A x_B w_{AB})
\]  
(5.8)

The pure end member phases A and B will have energies of

\[
H_A = \frac{1}{2} N z w_{AA}
\]  
(5.9)

\[
H_B = \frac{1}{2} N z w_{BB}
\]  
(5.10)

Hence the enthalpy of mixing \(\Delta H_{\text{mix}}\) on forming the solution from a mechanical mixture of the pure phases will be

\[
H_s - x_A H_A - x_B H_B = \Delta H_{\text{mix}}
\]

\[
\frac{1}{2} N z (x_A^2 w_{AA} + x_B^2 w_{BB} + 2x_A x_B w_{AB} - x_A H_A - x_B H_B) =
\]

\[
\frac{1}{2} N z \{x_A(x_A - 1) w_{AA} + x_B(x_B - 1) w_{BB} + 2x_a x_B w_{AB} \} =
\]

\[
\frac{1}{2} N z \{x_A(-x_B) w_{AA} + x_B(-x_A) w_{BB} + 2x_A x_B w_{AB} \} =
\]

\[
\frac{1}{2} N z x_A x_B (2w_{AB} - w_{AA} - w_{BB}) =
\]

Defining the work of mixing per mole to be \(W_H = \frac{1}{2} N_A z (2w_{AB} - w_{AA} - w_{BB})\), we find the molar energy of mixing to be

\[
\Delta H_{\text{mix}} = x_A x_B W_H
\]  
(5.12)

where \(N_A\) is Avagadro’s number.
This is a very useful result; it implies that the energy of mixing will be parabolic (an $x^2$ function). Furthermore, we can distinguish three cases;

1. $W_H = 0$. Here A–B interactions are equivalent to A–A and B–B interactions, so $\Delta H_{\text{mix}} = 0$, which is termed an *ideal solution*.

2. $W_H > 0$. Here, there is an energy penalty for forming A–B bonds; A–A and B–B bonds are more favourable and so the solution will have a tendency to cluster into A- and B-rich regions.

3. $W_H < 0$. Here, A–B bonds are more favourable than A–A or B–B, so solutions are favourable and there will be a tendency to form intermetallic compounds.

### 5.2 Entropy of Mixing $\Delta S_{\text{mix}}$

In order to complete our calculation of the Gibbs energy of mixing, we need to be able to calculate the second component, which is the Entropy. In statistical mechanics the configurational entropy is given by Planck’s postulate

$$S = k \ln \Omega$$

where $k$ is Boltzmann’s constant and $\Omega$ is the number of possible configurations of the system. There are other contributions to $S$, most notably that due to thermal vibrations. However, the thermal vibrations should be fairly independent of the atoms present on the lattice and hence we ignore this term when calculating the change in $S$, $\Delta S$ that occurs on mixing our elements together.

Prior to mixing ($p$), A atoms are on all sites in A and B atoms are on all sites in B. So there is only only possible configuration and we can say that

$$\Omega_p = 1$$

For a solution $s$ of $N$ atoms containing $n$ A atoms with the remainder B we then need to calculate the number of possible configurations. The first A atom can be placed on any of the $N$ sites, and the second on any of the $N - 1$ remaining sites. Hence there are $N(N - 1)$ ways of arranging the first two atoms. However, the A atoms are indistinguishable, so there are $N(N - 1)/2$ configurations. Placing the third atom gives $N(N - 1)(N - 2)/3!$, and so on until we obtain

$$\Omega_s = \frac{N(N - 1)...(N - n + 1)}{n!} = \frac{N!}{n!(N - n)!}$$

The remaining $N - n$ B atoms then fill the remaining empty sites. Using $n = x_A N$, $N - n = N(1 - x_A) = x_B N$ and Stirling’s approximation that $\ln s! = s \ln s - s$ for $s > 10$, we can then
obtain
\[ \ln \Omega_s = N \ln N - n \ln n - (N - n) \ln (N - n) + (N - n - N + n) \]
\[ = N \ln N - x_A N \ln x_A N - x_B N \ln x_B N \]
\[ = N \ln N - x_A N (\ln x_A + \ln N) - x_B N (\ln x_B + \ln N) \]
\[ = \ln N (N - x_A N - x_B N) - x_A N \ln x_A - x_B N \ln x_B \]
\[ = -x_A N \ln x_A - x_B N \ln x_B \]

So we can find \( \Delta S_{\text{mix}} \):
\[
\Delta S_{\text{mix}} = -kN(x_A \ln x_A + x_B \ln x_B) - k \ln 1
\]
\[ = -R(x_A \ln x_A + x_B \ln x_B) \quad (5.17) \]

Since for 1 mole of atoms \( k N_A = R \), the Gas Constant, and \( \ln 1 = 0 \), where \( S \) in is units of \( \text{J mol}^{-1} \text{K}^{-1} \). The Enthalpy of mixing is a positive number since the \( \ln \) of a number less than 1 is negative. The curve is infinitely steep at \( x_A = 0 \) and 1, which implies that making pure substances should be very hard, no matter how strong the Enthalpy driving force for separation, which is in fact found to be the case. However, it should be noted that modern society is very dependent on our ability to make very pure steels, nickel and titanium alloys, sputtering targets and Si in microelectronics.

\[ \text{Figure 5.2: Dependence of } \Delta S_{\text{mix}} \text{ on composition.} \]

### 5.3 Gibbs Energy of Mixing

We can combine our results to obtain the Gibbs energy of mixing:
\[
\Delta G_{\text{mix}} = \Delta H_{\text{mix}} - T \Delta S_{\text{mix}}
\]
\[ = x_A x_B W_H + RT(x_A \ln x_A + x_B \ln x_B) \quad (5.18) \]

So we can consider two cases;

1. \( W_H \leq 0 \): \( \Delta G_{\text{mix}} \) is negative for all \( T \), 5.3

2. \( W_H > 0 \): \( \Delta H_{\text{mix}} \) is positive but at high \( T \), \( \Delta S_{\text{mix}} \) will dominate. Hence at low temperatures there will be a maxima at \( x = 0.5 \) and there will be two intermediate minima. So between the two minima, we can find the lowest overall energy by forming a mixture of the two phases. In contrast, at high temperatures, a complete solid solution will be favoured.

The immediate consequences of this curve are that both pure metals and 50/50 solutions are difficult to achieve, so very many materials will prefer to form partial solid solutions at low temperatures. This is a consequence on the one hand of Entropy making pure metals difficult to produce but of most atoms preferring to bond to themselves than to unlike atoms.
5. Thermodynamics of Solutions

Figure 5.3: Dependence of $\Delta G_{\text{mix}}$ on composition where (left) $W_H \leq 0$ (solutions favourable) and (right) $W_H > 0$ (unmixing favoured).

5.4 Making Phase Diagrams

We then add in the contribution to $G$ from the original pure materials to obtain

$$G = x_A G_A + x_B G_B + \Delta G_{\text{mix}}$$

We can then plot the curve for $G$ at different temperatures, draw tie lines where required to obtain the overall energy of the minimum energy mixture, and plot out the phase diagram.

At temperatures where there are two minima, the mixture will separate into two phases. We obtain their compositions by drawing the lowest energy line possible across the curve, which is along the tangent of identical gradient - the common tangent construction. For our regular solution we obtain the following:

In Figure 5.5, in the single phase region the lowest $G$ is obtained at the system composition. At lower temperatures, a two-phase region exists where we can lower the overall energy by forming a mixture of the two phases, using the lever rule. Here, changing temperature changes both the proportions and compositions of the phases.

The model we have developed in the equations above is symmetric in $x_A$ about $x = 0.5$ - in general this is an oversimplification. This is because we have ignored interactions beyond the nearest neighbour i.e. second-neighbour effects. Here we observe that at high temperatures there is only a single minima but that as the temperature is lowered two minima and a central maxima begin to develop. At these temperatures the material will separate into two compositions. At temperatures above the solvus, there is a minima at $x = 0.5$, below the solvus a this maxima. So the solvus is reach when $d^2G/dT^2 = 0$ at $x = 0.5$.
which is therefore a criteria for this phase separation. The model also predicts that there is no solid solubility at $T = 0\text{K}$, because there is no entropy to favour solutions.

We can now draw out the free energy curves schematically that generate the other types of phase diagram we have seen.

In Figure 5.6 we see that at high temperatures the liquid $L$ is the lowest energy phase at all compositions, and vice versa at low temperatures. At some temperature in between, the lines cross over. They first do so at either pure A or pure B, and then as they cross over a two phase region is formed.

In the eutectic phase diagram there are Gibbs energy curves for each of the three phases, Liquid, solid $\alpha$ and solid $\beta$. The $\alpha$ phase has very high energies for compositions far from pure A, meaning it only forming relatively close to pure A. The $\beta$ phase is the B-atom analogue of $\alpha$. The liquid, having complete solubility for both elements, has a relatively flat variation of $G$ with composition. The two phase regions are just as we have already seen; the only difference is that at some point logically the three-phase common tangent must be possible.

**Figure 5.5:** Plotting out the phase diagram for the phase separation predicted at low temperatures by the regular solution model derived above.
Figure 5.6: Plotting out the phase diagram and Gibbs energy curves for complete liquid and solid solubility.

Figure 5.7: Gibbs energies and the eutectic phase diagram.
Chapter 6

Faster Solidification: the Scheil equation

So far, we have considered solidification to proceed in equilibrium conditions only, where slow cooling conditions apply and the solid and liquid compositions are those given by the tie line in the liquid + solid two-phase region. However, diffusion in solid phases is generally much slower than diffusion in liquids. Therefore another approach to solidification is to consider what happens when diffusion is limited in the solid but infinitely fast in the liquid, which corresponds to the industrial situation where the cooling rate can be quite rapid. Expressing the diffusion coefficient as $D$, this is equivalent to considering $D_s = 0$ and $D_L = \infty$. Then, the first solid to form would enrich the liquid in solute by removing relatively pure material from the system. The liquid would then move to a more enriched composition, and so the next solid to form would form from a relatively enriched liquid. The situation will be like that depicted in Figures 6.1–6.2.

We can equate the areas in Figure 6.1, which are amounts of solute, considering the change in liquid composition that occurs due to the solidification of some fraction of liquid; 

$$(C_L - C_S)dV_S = V_LdC_L \quad (6.1)$$

From the phase diagram, we can define the partition coefficient $k$ to be

$$k = \frac{C_S}{C_L} \quad (6.2)$$

Figure 6.1: Incremental change in liquid composition with change in fraction solid.
so \( k < 1 \). In addition, mass must be conserved so

\[
V_S + V_L = 1 \quad (6.3)
\]

Therefore we can obtain

\[
C_L(1 - k)dV_S = (1 - V_S)dC_L \quad (6.4)
\]

which is a first order differential equation. It will be subject to the boundary condition

\[
C_L = C \quad \text{at} \quad V_S = 0 \quad (6.5)
\]

\( i.e. \) that the liquid composition is the alloy composition at the beginning of solidification. We can rearrange the differential equation to obtain

\[
\frac{dC_L}{C_L} = (1 - k) \frac{dV_S}{1 - V_S} \quad (6.6)
\]

integrating;

\[
\int_C^{C_L} \frac{dC_L}{C_L} = \int_0^{V_S} (1 - k) \frac{dV_S}{1 - V_S} \quad (6.7)
\]

\[
[\ln C_L]_C^{C_L} = (k - 1) [\ln (1 - V_S)]_0^{V_S} \quad (6.8)
\]

so

\[
\ln \frac{C_L}{C} = (k - 1) [\ln (1 - V_S) - \ln(1)] \quad (6.9)
\]

so

\[
\ln \frac{C_L}{C} = (k - 1) \ln(1 - V_S) \quad (6.10)
\]

and finally

\[
C_L = C(1 - V_S)^{k-1} \quad (6.11)
\]

or, for the composition of the solid

\[
C_S = kC(1 - V_S)^{k-1} \quad (6.12)
\]

This is called the Scheil Equation. It describes, for a continuous solid solution, the evolution of the solid and liquid compositions for the case where diffusion in the solid is zero and liquid mixing is fast. This is the situation that produces the most extreme compositional variation in the final solid, compared to the slow cooling or equilibrium assumption that there is no variation in the solid.

We can use the Scheil equation in several ways; to describe the maximum volume fraction of eutectic that might form during solidification or to describe the evolution of composition of the solid during casting. Note in addition that, compared to equilibrium cooling, the freezing range may be greater for Scheil solidification.

**Figure 6.2**: Evolution of solidification where there is no diffusion in the solid - (a) at the beginning of solidification, (b) mid-way and (c) at the end.
This variation in the solid composition is called *segregation*. In Al-Cu alloys, often this is visible in an etched micrograph as the phenomenon of *coring*, where the solidified grains or dendrites appeared to be cored.

The Scheil equation describes the situation where there is no eutectic reaction in the phase diagram. Where a eutectic exists, when $C_L$ reaches the eutectic point (as it always will since $C_L \to \infty$ as $V_S \to 1$), then the eutectic reaction will occur, with the two solid phases being precipitated out of solution as the eutectic mixture. The graph will therefore look more like the one below, where once the eutectic is reached the two phases of solid are at each of their solubility limits.

The fact that we don’t see eutectics in all solidified alloys in eutectic systems therefore implies that some diffusion indeed occurs in the solid. Therefore the Scheil Equation describes an extreme - an upper bound - to the amount of segregation that occurs, whereas equilibrium slow cooling describes a lower bound (no segregation).

**Figure 6.3:** Scheil’s equation for the evolution of fraction solid and liquid during rapid solidification.

**Figure 6.5:** Eutectic observed in an industrial Al-4wt.%Cu casting. Under equilibrium cooling conditions, no eutectic should be observed, which indicates that diffusion in the solid was restricted in this case.

**Figure 6.4:** (a) Coring in a cast AA7020 aluminium alloy (courtesy D Garcia), and (b) close-up view of coring.
Chapter 7

Diffusion, Nucleation and Growth

7.1 Homogenisation

The consequence of this alloy segregation is that the composition varies across the grains of the material, which often take the shape of the dendrites (tree-like grains) shown in Figure 6.5. At the end of the solidification sequence thermal expansion will mean that the first solid to solidify will have shrunk, which in the absence of a good feed of liquid metal to the terminal eutectic will mean that shrinkage pores may appear in the casting.

These two issues - the variation in composition and the appearance of casting pores - mean that as-cast materials usually have poor mechanical properties. The pores act as small cracks in the material, initiating fatigue cracks that grow to early failure. Therefore heat treatment and further mechanical processing is often applied to even up to composition variations (homogenisation) and close the pores.

Homogenisation is typically performed at temperatures all the way up to the melting point, for times varying from a few hours to a whole day. It proceeds by the diffusion of atoms in the material. Atoms move down the gradient in internal energy, which is typically down the gradient in composition. We relate the flux $J_B$ of B atoms - the number passing per unit area per unit time - to the composition gradient $\partial C_B / \partial x$ through Fick’s First Law

$$J = -D_B \frac{\partial C_B}{\partial x}$$  \hspace{1cm} (7.1)

It is also interesting to consider the diffusion of an individual atom jumping from point to point in the lattice, a process known as a random walk. It can be shown that after $n$ jumps an atom will, on average, have moved a net distance of $\sqrt{\alpha n}$, where $\alpha$ is the jump distance. In diffusion problems, the characteristic distance is $D$, so we find that the distance $r$ a diffusing species moves is typically

$$r \propto \sqrt{Dt}$$  \hspace{1cm} (7.2)

By considering the rate at which atoms enter and leave a control volume it is fairly easy to derive from Fick’s First Law the following equation

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2}$$  \hspace{1cm} (7.3)

which is Fick’s Second Law. It says that the rate of change of the composition is proportional to the curvature of the composition curve.
These themes will be picked up and developed more fully in the rest of the course.

### 7.2 Processing of Al Alloys

Aluminium is an extremely abundant, light, and corrosion resistant material. Unfortunately in its pure form it is extremely soft and therefore of no engineering use. In general most high strength aluminium alloys are strengthened by alloying aluminium with other elements. One particularly popular alloying addition is Cu. When Cu is added to aluminium (normally no more than 4wt.%) then strength can be increased by up to \(10\times\). The process by which this occurs is known as precipitation hardening.

In general high strength aluminium alloys are processed in the following way. The aluminium is melted and any alloying elements added. The solidification of the melt is normally conducted by semi-continuous DC (direct chill) casting. The molten metal is cast onto a hydraulic ram which is slowly lowered though a water cooled mould. This process produces ingots that are several meters long and can weigh up to 25,000 kg. This process cools the metal relatively rapidly, and results in a 100% equiaxed dendritic microstructure. This is aided with the addition of a grain refiner to the melt such as TiB\(_2\) to provide heterogeneous nucleation sites for the solid aluminium crystals. However the rapid cooling also gives rise to microsegregation, which means that after solidification the ingot has to be given a homogenisation treatment. After the microstructure has been homogenised the ingot will be worked into its near final shape by processes such as hot rolling, extrusion or forging. This will normally take place at around 300–500°C. After this processing a very soft microstructure is obtained, that contains \(\alpha\) grains with coarse second phase particles at the grain boundary. Prior to being used the alloy will be heat treated (solution treated and aged) to achieve full strength. This occurs by precipitation of fine second phase particles within the \(\alpha\) grains.

![Figure 7.1: The casting of Al ingots for rolling. In a semi-continuous process, the Al ingot is solidified in a water-cooled mould.](image-url)

### 7.3 Nucleation and Growth of a New Phase

Previously, we have considered the thermodynamics of phase transformations without any consideration of the effects that the interfaces between phases have on the situation. Naturally, one can imagine that there will be an energy associated with a surface or interface - for example,
the dangling bonds at the surface between a metal and vacuum, or more generally at any surface or interface. Where a transformation is favourable (reduces the overall energy), then some of the energy given up on transformation is available to be expended in creating the free surface.

The first step will therefore be to consider what energy is available to produce interfaces - the driving force.

**Driving Force for Transformation**

For a first-order phase transformation between two phases of the same composition then we can write

\[ \Delta G = \Delta H - T \Delta S \]  

(7.4)

When considering the driving force, we take \( \Delta G \) to be positive for a natural change. Then, at the equilibrium temperature \( T_e \),

\[ \Delta G = 0 \]  

(7.5)

and so

\[ \Delta H = T_e \Delta S \]  

(7.6)

We can therefore substitute for \( \Delta H \) into the first equation to obtain

\[ \Delta G(T) = T_e \Delta S - T \Delta S \]

\[ = \Delta S(T_e - T) \]  

(7.7)

Note that we have assumed here that \( \Delta H \) and \( \Delta S \) are independent of \( T \), which will be true when both phases have the same \( C_P \), which is a good approximation for condensed matter, as long as \( T \) is close to \( T_e \).

We refer to \((T_e - T)\) as the supercooling \( \Delta T \). The implication of this calculation is that the driving force for transformation - the excess energy that is available to create new surfaces, for example - is proportional to \( \Delta T \).

**Nucleation of a New Phase**

We can now consider what happens when we first start forming a new phase during a first order phase transformation between two phases of identical composition. The simplest case is where the new phase grows from an old phase with no heterogeneities. Such flaws in the crystal structure - such as precipitates, voids, vacancies, grain boundaries or dislocations - are features with raised internal energy that can acts as nucleation sites. Here we ignore these - the case of homogeneous nucleation.

First we consider the driving force for transformation, the \( \Delta G \) from Equation 7.7, which we express per unit volume as \( \Delta G_v \) (in kJm\(^{-3}\) - we do this by multiplying \( \Delta G \) in kJmol\(^{-1}\) by the molar mass \( m_r \) in kgmol\(^{-1}\) and the density \( \rho \) in kgm\(^{-3}\) so \( \Delta G_v = \Delta G m_r \rho \)).

Defined as a the Gibb’s energy of the old phase minus that of the new, \( \Delta G_v \) is a positive quantity for a natural transformation.

We take the new interface created between the two phases to have an energy of \( \sigma \) per unit area, which is again a positive quantity. The interface contributes an excess energy to the system and
is the energy associated with the defect in packing that must occur between the perfect crystal lattices of the two phases. We assume that the interfacial energy $\sigma$ is essentially independent of $T$.

Now consider the case of a spherical nucleus of radius $r$. We assume that the formation of this nucleus of new phase does not induce any stresses in the surrounding material, which is the situation that occurs, for example, for the nucleation of solid in a liquid.

The work $W$ of forming the nucleus is then the balance of the energy gained from the transformation (a volume term) and the energy required to form the surface.

$$W = -\frac{4}{3}\pi r^3 \Delta G_v + 4\pi r^2 \sigma$$  \hspace{1cm} (7.8)

The curve for $W$ has a maxima, $W^*$ at a critical radius $r^*$. Particles smaller than $r^*$ in the figure can therefore reduce the energy of the system by shrinking, and particles larger than $r^*$ by growing. Hence

for $r < r^*$, a nucleus will shrink

for $r > r^*$, a nucleus will grow

We can evaluate $r^*$ by finding the maxima in the curve;

$$\frac{dW}{dr} = -4\pi r^2 \Delta G_v + 8\pi r \sigma$$  \hspace{1cm} (7.9)

setting $\frac{dW}{dr} = 0$ at $r = R^*$ we obtain

$$0 = -4\pi r^2 \Delta G_v + 8\pi r \sigma$$

$$0 = -\Delta G_v r^* + 2 \sigma$$  \hspace{1cm} (7.10)

so

$$r^* = \frac{2\sigma}{\Delta G_v}$$  \hspace{1cm} (7.11)

We can evaluate $W^*$ by substituting for $r^*$ in Equation 7.8 to obtain

$$W^* = -\frac{4}{3}\pi \left(\frac{2\sigma}{\Delta G_v}\right)^3 \Delta G_v + 4\pi \left(\frac{2\sigma}{\Delta G_v}\right)^2 \sigma$$

$$= \frac{\sigma^3}{\Delta G_v^2} \pi \left(16 - \frac{32}{3}\right)$$

$$= \frac{16\pi \sigma^3}{3 \Delta G_v^2}$$  \hspace{1cm} (7.12)

From Equation 7.7 we know that as $T \rightarrow T_e$, $\Delta G_v \rightarrow 0$, so $r^* \rightarrow \infty$ and $W^* \rightarrow \infty$, so nucleation becomes impossible.

The maxima in the curve for $W$ is termed the barrier to nucleation. In effect, a particle of size $r^*$ must spontaneously arise by atomic vibrations and diffusion in order for there to exist
a driving force for particle growth. If \( r^* \) is very large, then this will not occur and therefore quite significant undercooling must occur for precipitation of the new phase to take place.

An example of this is purified and filtered water, where significant undercooling must be imposed before homogeneous nucleation and hence freezing takes place.

**Effect of Strain Energy**

![Figure 7.3: Increase in \( W^* \) and \( r^* \) when strain energy is considered.](image)

A volume change is likely to accompany the nucleation of one solid in another, which we can consider to be a strain. Of course, there will be some elastic accommodation between the elastic matrix and elastic precipitate phases, but the total strain energy will scale with the volume of the nucleus and the difference in the lattice parameters in the unstrained state. This acts to adsorb some of the Gibbs Energy driving force, so the work of nucleation is modified;

\[
W = -\frac{4}{3}\pi r^3(\Delta G_v - CK\gamma^2) + 4\pi r^2\sigma \tag{7.13}
\]

where \( C \) is a constant, \( K \) is the bulk modulus and \( \gamma \) is the strain or fractional difference in molar volume of the two phases. As shown in the Figure below, the effect is to increase \( r^* \) and \( W^* \), possibly to the point where nucleation becomes impossible.

**Heterogeneous Nucleation**

Nucleation may be made easier by the presence of particles in the original phase. This will be discussed more in the Spring term, but here it will suffice to realise that the amount of new surface that must be created can be reduced by precipitating the new phase on a favourable surface. This reduces the size of the energy barrier \( W^* \) and therefore reduces \( r^* \), making nucleation easier and reducing the amount of undercooling that must take place for precipitation of the new phase to occur.

We will consider the rate of nucleation in the next lecture.
Chapter 8

Transformation Kinetics

Our consideration of thermodynamics is now essentially complete and it remains only to briefly introduce some ideas in kinetics.

It is often said that thermodynamics describes what ‘wants’ to happen, at infinite time, \( i.e. \) it is the study of driving forces, in which case kinetics is the study of what actually happens in a world with restricted atomic mobility. We have already seen how a nucleation barrier can lead to the (temporary) appearance of metastable phases that do not appear in the equilibrium phase diagram, a situation which is not particularly unusual.

The first situation to consider is the rate at which nucleation can occur.

8.1 Nucleation Rates

Previously, we determined that there would be an energy barrier of size \( W^* \) to the nucleation of a precipitate particle. We consider the occurrence of critical nuclei to be given by an Arrhenius relationship;

\[
\text{population of critical nuclei } \propto \exp \left( -\frac{W^*}{kT} \right) \tag{8.1}
\]

This equation can be motivated from the kinetic theory of gases of Maxwell and Boltzmann, from which it turns out that the fraction of molecules with kinetic energy greater than \( E \) is proportional to \( e^{-\frac{E}{kT}} \).

We consider the rate of addition of atoms to a critical cluster \( (r = r^*) \) to make it post-critical, \( i.e. \) \( dG/dr < 0 \) such that it grows spontaneously, to be a diffusional process with activation energy \( Q \);

\[
\text{rate of addition to a critical nucleus } \propto \exp \left( -\frac{Q}{kT} \right) \tag{8.2}
\]

The nucleation frequency \( I \) (number per m\(^3\) per s - m\(^{-3}\)s\(^{-1}\)) is then given by multiplying these two together;

\[
I = C_n \exp \left( -\frac{(W^* + Q)}{kT} \right) \tag{8.3}
\]
Figure 8.1: Variation of nucleation rate (Equation 8.3) with T, for $T_e = 980^\circ$C.

Where $C_n$ is a constant. However, $W^*$ is itself temperature dependent, as we established before;

$$W^* = \frac{16\pi \sigma^3}{3 \Delta G_v^2} = \frac{16\pi \sigma^3}{3 \Delta S^2 (T_e - T)^2} \quad (8.4)$$

since $\Delta G_v = \Delta S(T_e - T)$. This means that the nucleation rate will approach zero as $T \to T_e$, because then $W^* \to \infty$ and so $I \to 0$. Similarly, as $T$ decreases, then the diffusion element $e^{-Q/kT}$ decreases, so $I$ decreases. Therefore we have a situation where the nucleation rate has the following form with temperature;

The point is that (i) some undercooling is required for nucleation to occur, (ii) the nucleation rate increases with undercooling, but also (iii) that eventually the decrease in diffusivity takes over and the rate drops.

### 8.2 Growth of a New Phase

Once a particle has nucleated, the rate of growth is obtained by considering the rates of atomic jumps at the interface between the new and old phases. When there is a driving force for transformation, the energy barrier is asymmetric, taking the form shown in Figure 8.2.

To jump into the favourable new phase, an atom merely has to overcome the activation energy for diffusion, $Q$. However, to jump out of the new phase it must also overcome the driving force for transformation, which is $\Delta G_v$ per unit volume or, if $v$ is the volume of an atom, $v\Delta G_v$ per atom. Therefore the jump rates are

$$\text{Forward jump rate, old} \to \text{new} \propto \exp \left( \frac{-Q}{kT} \right)$$
$$\text{Backward jump rate, new} \to \text{old} \propto \exp \left( \frac{-(Q + v\Delta G_v)}{kT} \right) \quad (8.5)$$
The rate of movement of the interface, which is the growth rate \( \frac{dr}{dt} \), must be proportional to the difference between these jump rates, and is therefore of the form:

\[
\frac{dr}{dt} = C_g \left( e^{-\frac{Q}{kT}} - e^{-\frac{(Q + v\Delta G_v)}{kT}} \right)
\]

(8.6)

where \( r \) is the particle size, \( t \) the time and \( C_g \) the growth rate constant. Again, because \( G_v \) has a temperature dependence then the growth rate can be estimated using \( \Delta G_v = \Delta S(T_e - T) \). This produces the following functional form:

Therefore growth can occur if pre-existing nuclei are present, but close to equilibrium (\( T_e \)) the growth rate will still be small because the rate at which the new phase dissolves into the old phase will be similar to the rate at which the old dissolves into the new; at equilibrium they will be equal.

8.3 Isothermal Transformation Diagrams

A first-order phase transformation involves nucleation and growth; therefore multiplying the terms for the rates of nucleation and growth together will give an estimate of the overall transformation rate.

Experimentally, we measure this by instantaneously quenching a sample to a hold temperature \( T \) below the equilibrium temperature, holding it at that isotherm for a time \( t \) and then quenching to room temperature for inspection. The volume fraction of each phase can then be measured, and the time required for a volume fraction \( V_i \) of phase \( i \) to form estimated. An example is sketched in Figure 8.5.

These diagrams are termed Time-Temperature-Transformation (TTT) diagrams. Approximately, nucleation limits the transformation as high temperatures and growth is restricted by low diffusivity at low temperatures, although as we have seen both the nucleation and growth rates possess maxima due their dependence on the driving force \( \Delta G_v \) and thermal activation \( e^{-Q/kT} \).

Figure 8.3: Variation of growth rate (Equation 8.6) with \( T \), for \( T_e = 980^\circ C \).

Figure 8.4: Variation of overall transformation rate (nucleation \( \times \) growth) with \( T \), for \( T_e = 980^\circ C \).
Therefore, by rapid quenching the transformation may be avoided. Although these diagrams are only strictly valid for isothermal conditions, we can estimate the cooling rate required to avoid the transformation by avoiding the nose in the transformation C-curve. If the C-curve is for a low transformation fraction (say 1%), then the critical cooling rate $\dot{T}$ will be

$$\dot{T} = \frac{dT}{dt} = -\frac{\Delta T_{nose}}{t_{nose}}$$  \hspace{1cm} (8.7)

Figure 8.5: Schematic TTT diagram for a typical phase transformation.

Figure 8.6: Avoiding formation of the new phase during cooling.
Chapter 9

Microstructure and Properties of Plain Carbon Steels

This Chapter covers the second half of lecture 8 and the whole of lecture 9. In years where the course is over ten sessions, it also forms the first half of the 10th session, followed by a review.

Iron-based alloys are a foundation-stone of our civilization. Almost everything we do is dependent on cheap, light weight, strong steels in a huge variety of forms, from forgings to sheet to extrusions. The iron-carbon phase diagram is incredibly rich, with a number of phases and this allows us to attain great variability in the microstructures we form in steels, which allows us to attain strengths anywhere from 200 MPa to 2 GPa. Most often, other alloying elements are added to further alter the situation, but most of the metallurgy can be understood using just the Fe-C binary diagram.

At this stage in the course, we have examined most of the concepts required to understand steels, so this is the point at which we put everything together and use our new understanding.

On solidification at 1534°C pure iron initial forms as the bcc phase δ. Pure iron undergoes two polymorphic changes, first at ≈ 1400°C to the fcc austenite γ phase, and then at 910°C to the bcc ferrite α phase, which is the stable form at room temperature. The carbon atom is smaller than the iron atom and forms interstitial solid solutions with all three forms of iron. The maximum solubility of C in δ, γ, and α is 0.1, 2.0 and 0.03 wt.%, respectively.

The Fe-C diagram shown is not a complete diagram, as it is only constructed for C contents of less than 6.67wt.% (25 at.%), the composition of Fe₃C, an extremely hard and brittle intermetallic phase. Alloys with C contents greater than 6.67 are of little commercial interest, and therefore this region of the phase diagram is rarely presented. Alloys with a carbon content less than 2.06 (the maximum solid solubility of C in austenite), are referred to as steels (and it composition in this range we will discuss in this course), whilst those greater than 2.06wt.% C are referred to as cast irons. The majority of commercial steels have a carbon content of less than 1wt.%, with the greatest production being in the range 0.2 to 0.3 wt.%C.

On cooling from the liquid, low C content Fe alloys undergo a peritectic transformation, and enter the single phase austenite field, whilst steels with a higher C content solidify directly to austenite. As it can be assumed that on cooling sufficient time is available to permit diffusion, and bring about a homogeneous solid solution, the influence of the peritectic reaction on subsequent lower temperature transformations can be ignored. Therefore by the time a steel of
any composition has reached 1200°C we can consider it as being a homogeneous solid solution of C in γ-Fe (austenite).

The most important transformations that occur in steel are those that occur as we cool austenite to room temperature. Looking at the phase diagram we can see that on cooling, austenite will undergo a transformation that appears to be very similar to a eutectic transformation. In fact the only difference between this reaction and a eutectic is that austenite is a solid not a liquid. Such a solid state eutectic transformation is called a eutectoid transformation, and is completely analogous to its liquid state counterpart. It is this eutectoid reaction which we will now study in more detail. The term ‘Hypo-eutectoid’ is used to describe steel with a C content less than 0.8wt.%; the eutectoid composition, whilst steels with a C content greater than 0.8wt.% are referred to as ‘Hyper-eutectoid’.

Figure 9.1: The iron-carbon phase diagram.
9.1 Equilibrium Cooling of Steels

Eutectoid Composition

![Diagram showing equilibrium cooling of steels]

Figure 9.2: Formation of pearlite at the eutectoid temperature.

When a steel of composition of 0.8%C (eutectoid composition) is cooled to 723°C austenite will decompose to form ferrite and cementite, which form as alternate plates of each phase, to yield a structure called pearlite. The ferrite fraction in pearlite can be calculated using the lever rule

\[
\text{fraction of ferrite} = \frac{6.7 - 0.8}{6.7 - 0.02} = 88\%
\]  

(9.1)

As the density of ferrite and cementite are similar the plates (lamellae) of ferrite and cementite have a ratio of widths approximately 8:1 (which gives a ferrite fraction of 88%). The decomposition of austenite to form pearlite occurs by nucleation and growth, with nucleation occurring heterogeneously at grain boundaries. The formation of pearlite is depicted in Figure 9.3.

![Diagram showing co-operative growth of pearlite]

Figure 9.3: Co-operative growth of pearlite from a γ grain boundary.

Nucleation occurs by the formation of a lamella of cementite that forms at an austenite grain boundary and grows into an austenite grain. As this plate grows, it removes carbon from the adjacent regions of austenite. As a result the concentration of carbon in the austenite in contact with the cementite falls, until it reaches a value where ferrite is nucleated. As ferrite can only contain very low levels of carbon, the growth of ferrite results in a build up of carbon adjacent to the ferrite. This build-up continues until a new layer of cementite nucleates, the process then repeats itself, and the alternating lamellae of a pearlite colony are formed. Growth of pearlite colonies occurs not only by the nucleation of additional lamella, but also by the advance of ends of the lamellae. Nucleation of a pearlite

![Microstructure of a eutectoid steel - 100% pearlite]

Figure 9.4: Microstructure of a eutectoid steel - 100% pearlite.
colony can also occur at the interface of an existing colony and the parent austenite. Pearlite colonies grow unimpeded until they impinge on adjacent colonies.

**Hypo-eutectoid Steels (< 0.8 wt.% C)**

![Figure 9.5](image_url)

**Figure 9.5:** Development of the microstructure of a hypo-eutectoid steel during equilibrium cooling.

Let us now look at the transformations that will occur when we slowly cool a hypo-eutectoid steel (for example with a composition 0.3 wt. % C) from the austenite phase field to room temperature (point a to point d in the Figure). Initially we have grains of pure austenite, and all the carbon is in solid solution. On cooling no transformation will take place until the temperature reaches point b. At this point the transformation begins, with ferrite being nucleated at the austenite grain boundary. As the solubility of C in ferrite is very low this transformation is associated with the rejection of C into the remaining austenite. As the temperature is further reduced the amount of ferrite present increases and the concentration of C in the austenite increases.

When the temperature reaches point c, the concentration of C in austenite reaches 0.8% (the eutectoid composition). The proportion of ferrite and austenite at this point can be calculated using the lever rule, *i.e.*

\[
\text{fraction of ferrite} = \frac{0.8 - 0.3}{0.8 - 0.02} = 64\%
\] (9.2)

On cooling the alloy below the eutectoid temperature (to point d) all the remaining austenite is transformed, via the eutectoid reaction, to produce pearlite. The fraction of ferrite in the pearlite is the same as before, as the composition of the austenite forming the pearlite is always the same. Continued cooling to room temperature (point e) causes no visible change in the microstructure. Since the solubility of carbon in ferrite decreases with T a change should occur, but it is extremely small.
Hyper-eutectoid steels (> 0.8 wt.% C)

Let us now consider the cooling of a hyper-eutectoid steel (for example with a composition of composition 1.2 wt.% C) from the austenite phase field (point f). Transformation of the austenite begins when the temperature falls to point g. At this point cementite starts to form. As cementite is rich in C the concentration of C in the austenite decreases as the temperature falls, and more cementite is formed. By the time the temperature reaches point h (just above the eutectoid temperature), the concentration of carbon in the austenite falls to 0.8% (the eutectoid composition). At this point the fraction of cementite is

\[
\text{fraction of } \text{Fe}_3\text{C} = \frac{1.2 - 0.8}{6.7 - 0.8} = 6.7\% \quad (9.3)
\]

Below the eutectoid temperature all the remaining austenite transforms to pearlite. One of the distinguishing features of hyper-eutectoid steels is the high percentage of pearlite present, in this case \(\sim 93.3\%\).

### 9.2 Non-Equilibrium Cooling of Steels

In the previous section we have looked at the microstructures of steels that are obtained when they are cooled under equilibrium conditions from the austenite phase field to room temperature. To achieve equilibrium conditions in reality entails placing the steel into a furnace, heating the steel into the austenite phase field and then turning the furnace off and letting the steel cool down in the furnace. This process takes many hours and is known as a ‘full process anneal’. This is the softest state that a plain carbon steel will exist in. The strength of the steel can be improved by a process called normalising. In this case the steel is again heated into the austenite phase field, but in this case once the temperature of the steel has equilibrated the steel is removed from the furnace and the allowed to cool down to the room temperature in still air. The microstructures of normalised steels are similar to those shown above, however the scale of the microstructure is refined (the ferrite grain size is reduced and the interlamellae spacing in the pearlite is reduced), leading to a higher strength steels.
The two previous cooling paths are relatively slow and allow plenty of time for the eutectoid reaction to take place. That is to say that as the temperature falls below the eutectoid temperature there is sufficient time for the carbon atoms to diffuse out of the austenite to form Fe₃C, and then the austenite to transforms to ferrite. However if the steel is cooled very quickly from the austenite phase field (i.e. by water quenching) then a new structure known as martensite can be formed. Martensite is simply a supersaturated solid solution (i.e. it contains more C than it should have) of carbon in iron. The transformation of austenite to martensite is a diffusionless transformation, and can be thought of as the fcc Fe structure trying to transform to bcc, but due to the presence of the C atoms (which cannot diffuse away) this cannot occur, and a body centred tetragonal (bct) structure is formed instead (see figure below). It is referred to as a displacive transformation, as distinct from the reconstructive transformation that occurs when diffusion results in complete reconstruction of the crystal lattice. The bct structure is highly stressed due to the entrapment of the C atoms and as a result martensite is often hard and brittle.

As martensite is extremely brittle it is rarely used in engineering applications, except where the steel has been alloyed to produce a softer form of martensite. However, in plain carbon steels by controlling the cooling of a steel component it is possible to create a surface layer of martensite with a softer core of ferrite/pearlite. This yields a tough component with a heard, wear-resistant surface. Where the transformation to martensite is useful is in a process known as the quenching and tempering of steels. In order to increase their strength, steels can be quenched, from the austenite phase field, to room temperature, in order to form martensite. However steels in the martensitic condition are too hard and brittle for normal service. However the properties are improved by re-heating (tempering) the steel to a temperature below the eutectoid temperature. During tempering carbon is removed from solution by the precipitation of iron carbides, and the matrix reverts to the bcc structure.
Martensite contains a lot more carbon in solid solution than iron can normally accommodate. As a result the driving force for the removal of carbon from solid solution and the formation of Fe\textsubscript{3}C is extremely high. The only factor that stops this happening at room temperature is the fact that diffusion of carbon cannot take place. As a result martensite is in a metastable state (i.e. it is thermodynamically unstable, but kinetically stable). If the structure is heated to a temperature where carbon can now diffuse the structure will try and revert to the equilibrium condition of ferrite and Fe\textsubscript{3}C. As the driving force for this transformation is very high this will result in the formation of a large number of small crystals of Fe\textsubscript{3}C (known as precipitates). Once the carbon has been removed from solid solution the remaining martensite transforms to ferrite. Whilst the tempering process results in a reduction in the strength of the steel compared to the martensitic condition, the ductility and toughness are much improved.

If we cool a steel too quickly to form pearlite but too slowly to form martensite we can get yet another microstructure known as bainite. Bainite is a very refined microstructure containing ferrite and Fe\textsubscript{3}C (as is the case with pearlite). The transformation starts with the nucleation of ferrite plates at the austenite grain boundaries. The growth of these plates is similar to the formation of martensite, i.e. it is displacive. However in this case as the temperature is high enough for C to diffuse out of the lattice, the fcc structure can fully transform to bcc. As the ferrite plates form the remaining austenite becomes richer in C. Eventually Fe\textsubscript{3}C precipitates in the remaining austenite layers in between the ferrite platelets. Bainite has a good combination of strength and toughness, because it is very fine in scale but does not contain brittle martensite.

Figure 9.11: Formation of a fine microstructure by quenching to martensite and the tempering, and the effect on properties.

Figure 9.12: Microstructure of bainite compared to pearlite - the bainite microstructure is both strong and tough, and is very much finer-scale than that of pearlite.
9.3 TTT diagrams

When steels are rapidly cooled from the austenite phase field a number of different microstructures are obtained depending on the rate at which the material is cooled.

The transformation of austenite to pearlite occurs by nucleation and growth (i.e. before pearlite can grow it has to be nucleated). That is to say that the overall rate of transformation of austenite to pearlite is dependent on the combined rate of both nucleation and growth. If either the nucleation rate or growth rate is slow then the overall transformation rate will be slow. The rate of nucleation depends on the driving force for the formation of pearlite. Above the eutectoid temperature austenite is stable and pearlite will not form and hence there is no driving force for pearlite to form. As we lower the temperature below the eutectoid temperature austenite becomes unstable and wants to transform to pearlite. As we keep on lowering the temperature the austenite becomes more and more unstable and hence the driving force for pearlite formation becomes greater and greater. What this means is that as we lower the temperature below the eutectoid temperature that rate of nucleation of pearlite will increase, and the driving force for its formation increases. At temperatures close to the eutectoid temperature the time required to effect complete transformation of the austenite will be large due to the fact that the nucleation rate will be low. Decreasing the temperature will have the effect of increasing the nucleation rate, and initially the overall transformation rate will increase. Further decrease in the temperature leads to a decrease in diffusion rate (which controls the growth rate), hence the growth rate will slow down leading to a decrease in the overall transformation rate. This is summarised in the TTT diagram below which has two curves, representing the time required for the onset and completion of the transformation.

As there is no long range movement of atoms involved in the transformation from austenite to martensite this transformation is referred to as a diffusionless transformation. Therefore the transformation from austenite to martensite must be independent of time and is dependent only on temperature; it is purely thermodynamic in nature. The temperature at which the transformation occurs depends on the carbon content, but is generally near to room temperature. Therefore in order to form martensite we must prevent the eutectoid reaction from occurring whilst cooling austenite to room temperature.

Effect of C on Martensite Formation

During the transformation of austenite to pearlite, decreasing the transformation temperature has the effect of decreasing the inter-lamellae spacing in the pearlite. Therefore at high temperatures coarse pearlite is formed, whilst at lower temperatures fine pearlite is formed. If the transformation occurs at a temperature below \( \sim 500^\circ\text{C} \) the product of the austenite transformation is bainite. If, on cooling, austenite remains present at even lower temperatures it will
transform to martensite. The extent of the transformation is independent of time, but depends on the temperature to which it is cooled. There is a martensite transformation start temperature ($M_s$) and decreasing the temperature below this temperature will result in an increasing amount of martensite being formed.

The martensite transformation start and end temperatures are strongly influenced by the carbon content and decrease with increasing carbon content. For plain carbon steels containing more than 0.7 wt.% C the martensite finish temperature $M_f$ is below room temperature. This means that if we quench a eutectoid or hyper-eutectoid steel to room temperature not all the austenite will be transformed to martensite and some austenite will be retained. The retention of austenite affects the hardness of quenched steels. The hardness of quenched steels containing up to 0.7 wt.% C the increasing proportion of retained austenite, which is relatively soft, balances the increasing hardness of martensite, and as a result the hardness of the steel remains essentially constant.

**Example: Using a TTT diagram for a steel**

Let us look now at the TTT diagram for a commercial eutectoid steel, Figure 9.15, and consider some arbitrary time-temperature paths along which we will cool austenitic samples to room temperature.

- **Path 1:** The specimen is rapidly cooled to 160°C and left there for 20 minutes.
  
The rate of cooling is too rapid for pearlite to form at the higher temperatures, therefore the steel remains in the austenite phase until the $M_s$ temperature is passed, where martensite begins to form athermally. Since 160°C is the temperature at which half the austenite transforms to martensite, the direct quench converts 50% of the austenite to martensite. Holding at 160°C forms only a very small quantity of additional martensite because in simple carbon steels isothermal transformation to martensite occurs only to a very limited extent.

- **Path 2:** The specimen is rapidly cooled to 250°C, held at this temperature for 100 s, then quenched to room temperature.
  
  Again the cooling is too rapid to form pearlite, and holding at 250°C for 100 s is not a sufficiently long time to form Bainite, such that on quenching to room temperature a martensitic structure is developed.

- **Path 3:** The specimen is rapidly cooled to 350°C, held at this temperature for 200 sec, then quenched to room temperature.
  
  The isothermal hold at 350°C produces a structure composed of half bainite and half austenite. Quenching this structure to room temperature results in the transformation of
the remaining austenite to martensite. Therefore the final structure is half bainite, half martensite.

- Path 4: The specimen is held at 600°C for $10^4$ s or 3 hours, then quenched to room temperature.

At 600°C the transformation of austenite to pearlite is completed within $\sim 8$ s. The pearlite microstructure is reasonably stable, and therefore continued heating at this temperature results in no change in the pearlite structure. As the sample now contains no austenite, rapid cooling to room temperature does not yield any martensite. Therefore the final structure is fine pearlite.

**Quenching and Tempering of Steels**

In order to increase their strength, steels can be quenched from the austenite phase field to room temperature in order to form martensite. However steels in the martensitic condition are too hard and brittle for normal service. The properties are improved by re-heating (tempering) the steel to a temperature below the eutectoid temperature. During tempering carbon is removed from solution by the precipitation of iron carbides, and the matrix reverts to the bcc structure. At higher tempering temperatures some spheroidisation and coarsening of the carbides occurs. Whilst the tempering process results in a reduction in the strength of the steel (as compared to the martensitic condition) the ductility and toughness are much improved.
This hardening process is dependent on producing a martensitic structure on quenching the material from the austenite phase field. In the general TTT curve for a eutectoid steel, cooling curve (a) represents a high enough cooling rate for this to occur. However cooling curve (b) on the same diagram passes through the nose of the transformation curve, therefore some austenite will be transformed, and only a partial martensitic structure will result. With large components, this can be a problem, with the surface of the component being completely transformed to martensite, whilst the centre (which cools at a slower rate), will only be partially transformed. The ability of steel to form martensite on quenching is termed the hardenability. In order to increase the hardenability, the rates of proeutectoid, pearlitic and bainitic reactions must be displaced to longer times, and this can be achieved by the addition of alloying elements. With the exception of Co, all the common alloying elements in steel decrease the reaction rates, and increase the hardenability. Alloying additions enable us to obtain steels in the martensitic condition more easily, but they also affect the tempering characteristics.

Firstly the precipitation and coarsening of the iron carbides may be altered by the presence of an alloying addition. For example Si slows down the precipitation and coarsening of the cementite, and this can reduce the fall off in hardness during tempering. More important is that some alloying elements, called carbide formers (Cr, Mo, Ti, V, & W), form alloy carbides such as $V_4C_3$ and $Mo_2C$ which are more stable than the iron carbides. Due to the slow rate of diffusion of these substitutional alloying additions, the alloy carbides do not form until tempering temperatures in the range $450 - 600\, ^\circ C$ are reached. The precipitation of the alloy carbides results in an increase in hardness, termed secondary hardening, at high tempering temperatures. Secondary hardening enables us to produce steels with a high hardness and strength, but with a better ductility that that exhibited by a plain carbon steel tempered to the same hardness by lower temperature tempering. Also, because of the slow diffusion of the carbide forming elements, their carbides do not readily coarsen, consequently alloy steels so hardened may be used for high temperature applications.

Figure 9.16: TTT diagram for a commercial steel with 0.76% Mn and 0.79%C, subjected to two cooling curves (a) and (b).

Figure 9.17: Effect of tempering temperature (amount of tempering) on the properties of a quenched-and-tempered 0.55wt.%C steel.