CHAPTER 1: LITERATURE REVIEW

1.1 Hydrogen Embrittlement

Hydrogen Embrittlement (HE) can be defined [20] as the hydrogen caused reduction of the load–bearing and/or the mechanical energy absorption ability of a metallic alloy. Initially found in steels, it is now acknowledged that the mechanical properties of most metals and alloys can deteriorate by hydrogen under certain conditions. [20] For instance hydrogen can be absorbed during metal processing and fabrication, processes as electroplating, solidification, forging and welding. [21] Moreover, hydrogen absorption may occur in specific offshore conditions, such as corrosion and cathodic protection. [22]

Should hydrogen be introduced into the steel structure, it can cause an embrittling effect, as the loss of macroscopic ductility or cracking. [23] Slow strain rate specimens can display a decrease in tensile strength due to HE with a simultaneous reduction in area, but their yield strength is not considerably affected. However, steels with higher yield strength are more susceptible to HE, as even relatively small amounts of hydrogen can deleteriously influence their mechanical properties.

In terms of microstructure, a material is more vulnerable to hydrogen in the following increasing order: (a) lower bainite, (b) quenched and tempered martensite, (c) pearlite or spheroidized structures and (d) martensite. Fine grain and carbide sizes, along with low carbon content, are beneficial to avoid HE. [24]

Slow strain rates and moderately elevated temperatures promote HE suggesting that the phenomenon is under the control of and paced by the lattice diffusion of hydrogen. [21] For low carbon and austenitic steels it has been observed that there is a maximum degree of embrittlement at an intermediate temperature near ambient conditions, from which as the temperature increases or decreases, HE decreases in both cases. (Figure 4) Metallurgical conditions, strain rate and hydrogen content affect the position of this maximum. [25]

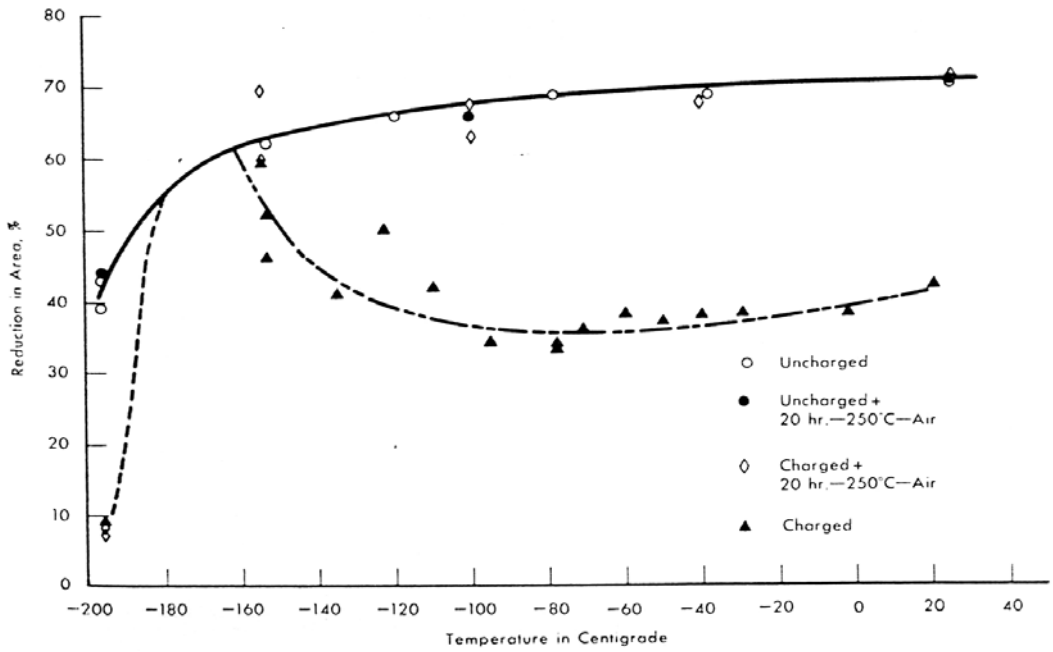

Figure 4. Temperature dependence of HE in steel measured in terms of Reduction in Area (%). [9]

At temperatures in the range of 400–500^{o}C and for extended periods of time sufficient thermal energy is provided for the diffused hydrogen in inclusions and grain boundaries, where hydrogen may reduce carbides and oxide inclusions to form methane gas (decarburisation) or water. Once combined in this way the hydrogen is fixed and cannot be removed. [26] This form of hydrogen damage is known as hydrogen attack and is not a manifestation of HE, as there is a need for thermal energy to induce chemical reactions, which are absent in HE. However, HE and hydrogen attack are not mutually exclusive, for an alloy which may have suffered from hydrogen attack at elevated temperatures may be subject to HE at lower temperatures. [21]

The extent of embrittlement is known so far to be dependent on the strain rate, the temperature and the microstructure. Consequently, the heat treatment used during manufacturing and processing is another parameter affecting the degree of HE, as it substantially affects the final microstructure and its subsequent susceptibility to HE. For example, tempering temperature has to be minimised, for it has effects on tensile strength. (Figure 5) Moreover, post–baking treatments are viewed as necessary to relieve hydrogen–containing steels from hydrogen. Such processes facilitate hydrogen

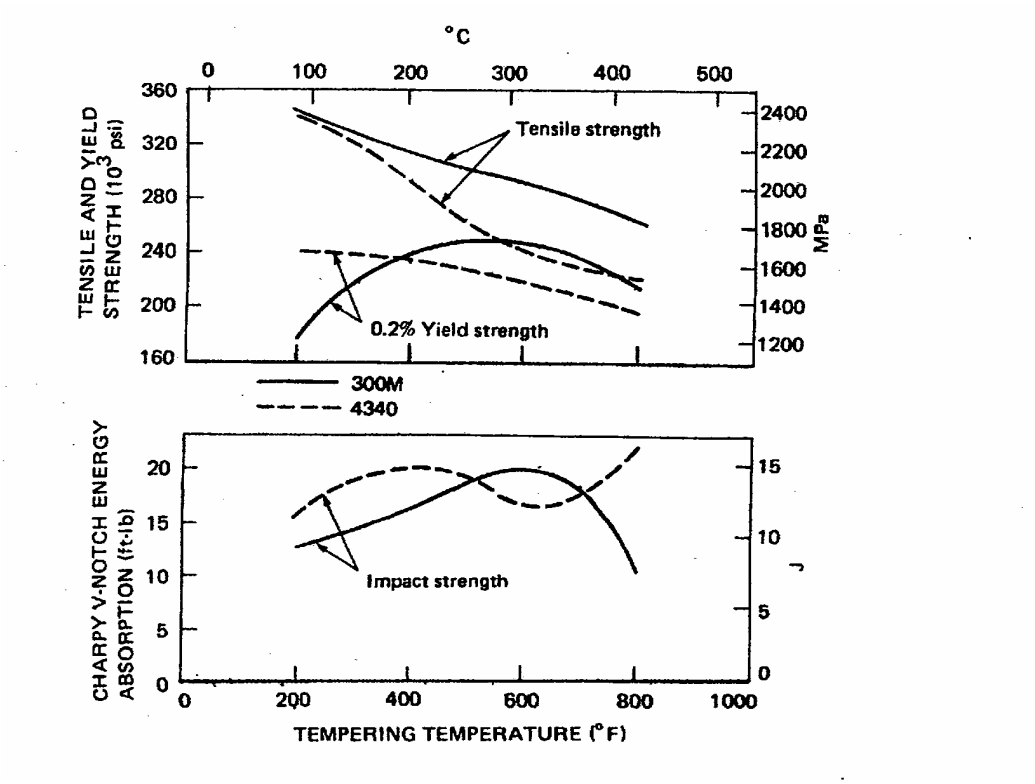

Figure 5. Effects of tempering on tensile properties of 4340 and 300M steels. [9]

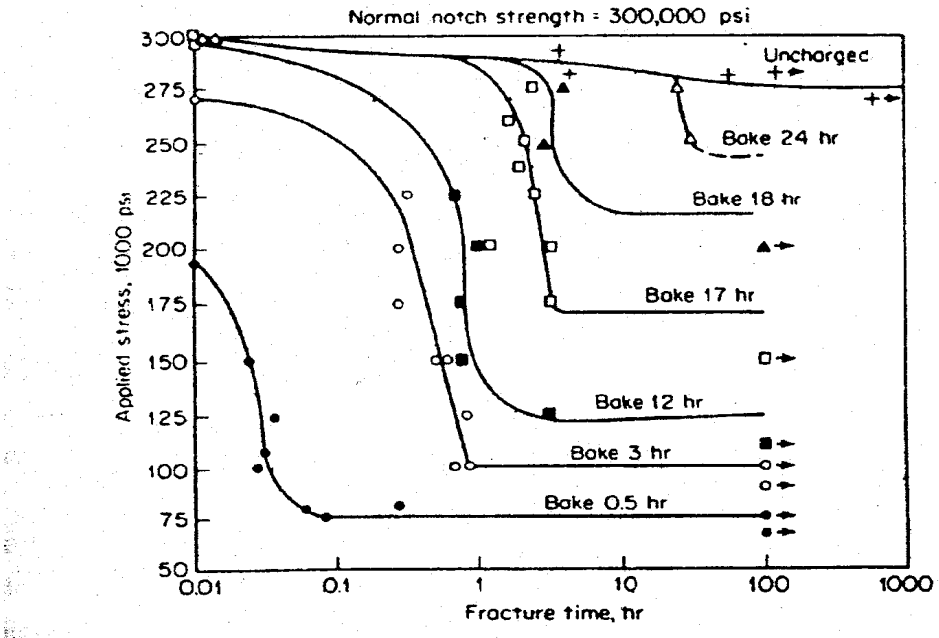

Figure 6. Static fatigue tests on AISI 4340 steel previously charged with hydrogen, after different baking times. [9]

to escape to the atmosphere or diffuse to microstructural traps within the steel, but the temperature and duration of the treatment has to be controlled to avoid loss of strength. Figure 6 shows the effect of baking time on charged AISI 4340 fatigue specimens. Although baking is an effective method for the hydrogen removal, the process is expensive and time consuming.

It is necessary to view the composition, the microstructure and the operating conditions of a steel component to speculate if a failure by HE is possible. This could occur in different ways. There are three main forms of hydrogen damage in metals and alloys. [27] The first form of damage results in internal pores, cracks, blisters or other flaws arising from either the entrapment of hydrogen bubbles during solidification of the melt or the diffusion of hydrogen through the metal lattice to cause the flaws. The trapped gas then creates a pressure, which can cause growth of cracks and voids. The second form of damage is caused by hydrogen forming hydrides at specific lattice positions and, hence, impairing the mechanical properties. The third form of damage comprises the rest of the hydrogen related failures occurring especially in the presence of sustained loads, and it is described by Beachem as Hydrogen Assisted Cracking (HAC). [27, 28]

During the last decades a wide range of materials has been examined in terms of HE. However there hasn't been any consensus on a single theoretical mechanism of HE, as each proposed theory accounts for observed phenomena under specific conditions without being able to explain all cases of HE. This is the result of the complexity of the problem and the multitude of the effects that hydrogen can exert in a defected metallic lattice. [20] Parameters, which must be considered when explaining HE, include hydrogen solubility and diffusivity, hydrogen transport and trapping, interaction of hydrogen with the metal lattice and hydrogen recombination. Despite the abundance of the suggested HE mechanisms, there is a general agreement on a few principles. There might be several mechanisms operating within the same specimen at any time, with the preponderant mechanism not being the same in each case, but varying depending on the different experimental or service conditions. Furthermore, before hydrogen embrittles a metal or alloy, it has to undergo transport to the metal surface, adsorption, absorption and finally transportation into the metal bulk, where it must then contribute to crack nucleation and propagation. The

aforementioned processes are series processes – apart from the at least parallel internal transport processes – that can have an activation barrier and in principle could be rate controlling. Nonetheless, while these processes contribute to the hydrogen degradation, they are not mechanisms on their own but rather part of an overall mechanism.

1.2 Hydrogen Embrittlement Mechanisms

1.2.1 Internal Pressure Theory

The internal pressure theory, suggested by Zapffe and Sims [29], assumes that hydrogen pressure in voids itself is the cause of HE. Embrittlement is attributed to hydrogen concentration in voids, which grow, as atomic hydrogen diffusing through the metal can combine to form molecular hydrogen, and hence create high pressure, stabilise microvoids or microcracks and increase their number. The subsequently caused stress is assumed to be triaxial, which therefore inhibits flow, and only rupture can relieve a superimposed stress.

There were some objections to this mechanism of HE. Hancock and Johnson [30] showed that crack propagation occurs with sub–atmospheric external hydrogen pressure instead of the high internal pressures required for the supersaturated solid solution. In the same context, Hirth and Johnson [24, 31] discount internal pressure theory on the basis of saturation studies showing very low hydrogen levels in a specimen to support this mechanism. Moreover, according to Iino [32], the theory does not account for the embrittlement of high strength steels in low hydrogen gas pressure, or the absence of embrittlement of steel exposed to He dissolving into the steel.

However, the internal pressure theory is still valid in that it is widely acknowledged that this pressure can cause the cracks to grow, with or without additional loads being applied externally. [27] Cases, which are considered to be examples of the internal pressure mechanism, are the blister formation in the absence of external stress and in high fugacity environments, where pressure enhanced void growth can occur. [24] Furthermore, again under very high fugacities, the pressure theory can play an

important role in sour gas pipeline environments that cause hydrogen pressure–induced cracking of comparatively soft steel products.

1.2.2 Hydride Induced Cracking

Westlake [33] proposed a generalised model for HE based on the assumption that the HE of any metal could be accounted for on the basis of localised formation of a phase whose mechanical properties differ from those of the matrix due to hydrogen enrichment. A metal hydride might form ahead of a crack tip, in a region of high triaxial stress. Crack propagation could occur because of the cracking of the brittle hydride phase.

This model had some success with niobium, as it was demonstrated by Gahr et al [34] who showed that hydrides could be stabilised by the hydrostatic stress field at a stressed crack tip due to their increased volume, although they could be unstable without stress. Among other systems, which have been proved to give some evidence of hydride embrittlement, are the Vb group metals, such as Nb, V, Ta, as well as Zr and Ti, along with alloys based on these metals. [35] However, for iron, Baranowski [36] found that no hydride is stable up to the hydrogen pressures of 2 GPa, and consequently, this model has not been considered for iron and steel.

Nonetheless, calculations by Hirth and Carnahan [37] indicate an adequately high level of hydrogen saturation in iron to produce a situation equivalent to hydride formation from a mechanical viewpoint. At lower absolute temperatures the saturation is higher, as it is derived from Boltzmann and Fermi–Dirac equations that show the relationship between the atom fractions of solute in equilibrium, both in defect field and remote from the defect field, and the absolute temperature. The reason for the hydrides not being detected after mechanical failure in iron and steel could lie in the subsequent decomposition of the hydride when the crack begins to open up and its stress field is removed. [35]

1.2.3 Surface Energy (Adsorption Model)

The Griffith cracks have been the bases for many theories and the work resulted in the field of linear elastic mathematical treatments of crack growth. [27] Hydrogen damage theories have been considerably affected by Griffith's suggestions [11], according to which some of the energy of the applied strain is released during crack propagation and is converted to surface energy. Petch and Stables [38] proposed that hydrogen would adsorb onto the crack surface, thus lowering the surface energy and making easier the formation of cracks, as less applied strain energy would be necessary to convert to surface energy. The propagation of the crack is promoted at a speed controlled by the rate of hydrogen supply to the crack tip.

In an atomic scale, hydrogen adsorbs and interacts with strained bonds at the crack tip, causing a reduction in bond strength. [39] (Figure 7) This, as stated by Beachem [27], would require the hydrogen to diffuse through the plastically deforming, as well as entrapping, crack tip to get to the surface. For their model, the fracture path and the microstructure are of little significance, whilst this model is based on energy considerations, neglecting any plasticity effects.

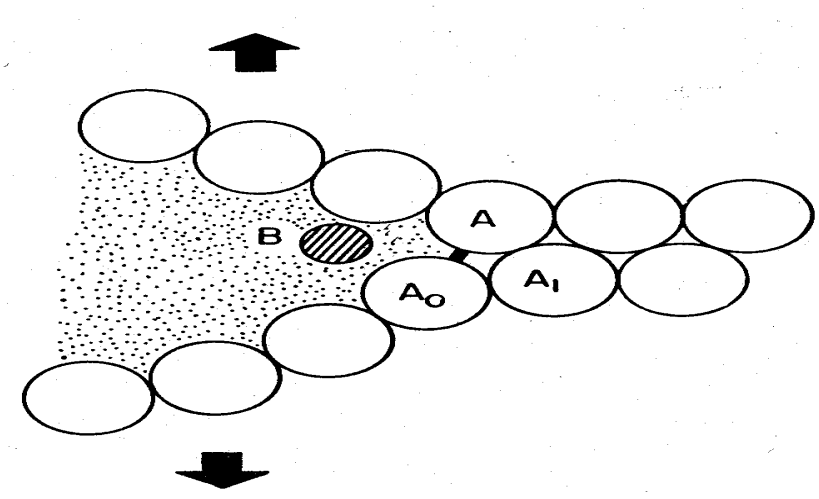

Figure 7. Schematic illustration of the adsorption model. The model requires that a specific ion from the environment B, interacts and reduces the cohesive strength of strained bond A–A_o at the tip of a brittle crack. [39]

Pugh [39], in an attempt to summarise and discuss various failure mechanisms in materials, refers to the surface energy or adsorption model, as well as Uhlig's work, mentioned as stress sorption cracking. Uhlig considered adsorption to occur at dislocations or other mobile imperfections in the vicinity of the tip instead of the strained bonds at the tip of elastic cracks. Moreover, he contemplated that a sharply defined potential exists below which the adsorption of damaging ions, and therefore cracking, does not occur. [40] This was contested by Bockris [41] who supported the view that adsorption begins at the potential of zero charge and increases gradually with shift in potential. He also disputed that adsorption would occur at specific imperfections.

As summarised by Hirth [24], arguments against this theory are that it (a) largely underestimated the work of fracture, (b) cannot explain discontinuous cracking as shown by sonic emissions, (c) cannot account for the reversal of the propensity for delayed failure on stress removal and (d) cannot demonstrate why oxygen fails to promote cracking and inhibits the hydrogen degrading effects.

Concerning the last contention, Hancock and Johnson [30] support that the crack tip surface preferentially adsorbs oxygen rather than hydrogen and water vapour. This is consistent with the affinity of iron for oxygen, which appears as a high absorption heat and as a multilayer coverage of iron by oxygen. It was proposed that this multilayer is mainly an oxide presenting an effective barrier to the access of hydrogen to the crack tip. Finally, another probable omission of this model is any explanatory mechanism as to how the hydrogen facilitates the formation of a new surface, onto which it can afterwards adsorb. [9]

1.2.4 Brittle Crack Tip Theory

Thomson [42] illustrated that in ductile materials the plastic zone could act to shield a brittle crack tip, lying within this plastic zone, from the full force of the external stress field, whilst still permitting brittle propagation. Hydrogen present in steel could adequately reduce the surface energy, so as to deter the blunting of a sharp crack. These ideas were incongruous with the prevalent opinion that fracture in metals is a ductile hole growth phenomenon with crack initiation occurring at the point of

maximum triaxial stress ahead of a well rounded main crack. According to Thomson, the aforementioned mechanism could happen if the crack is unstable against abruptly generated blunting and dislocation formation at the crack tip. Otherwise, this modified brittle fracture is possible in ductile materials and may be accountable for hydrogen embrittlement in high strength steels.

The brittle crack tip theory is deemed as a thermodynamic theory, where irreversible brittle cracking can locally occur by thermal or athermal activation over the lattice trapping barrier associated with the periodic arrangement at an atomically sharp crack tip. [24]

1.2.5 Lattice Decohesion Theory

The decohesion theory, primarily proposed by Troiano [23], is based on the premise that hydrogen reduces the cohesive strength of iron. This theory is related to the adsorption model which, as previously discussed, lends support to the decrease of surface energy by hydrogen, and hence favours the surface as a potential site of embrittlement, whereas the decohesion model advocates the bulk lattice. [39] Troiano [23] expressed the idea that hydrogen dissolved in steel concentrates in regions of high triaxial stress, forming more concentrated solutions of hydrogen in iron than would occur without stress, and hence weakening the cohesive force between metal atoms and facilitating the nucleation of a microcrack within the plastic enclave. According to this theory, when the local hydrogen concentration reaches a critical level due to stress, the arising brittle crack joins the main front of crack nucleation and, moreover, advances out of the region of hydrogen segregation, where it is blunted by plastic deformation. The whole process is reiterated and the crack propagation is discontinuous and controlled by hydrogen diffusion near the crack tip, as further crack growth must await diffusion of hydrogen to the region of high stress state. (Figure 8)

Troiano [23] speculated how the metal lattice cohesion could be lowered by hydrogen. The electrons of the hydrogen atoms in solution in a transition metal could enter the d bands of the metallic cores, as viewed pictorially in Figure 9. Moreover, the overlapping of the d bands of transition metals, such as iron, cobalt, nickel etc,

causes the repulsive forces between their atoms, which determine their interatomic distances. In this context, an electron concentration increase of these bands may be expected to render an increase in the repulsive forces between the metallic cores, otherwise noticed as a weakening of the lattice cohesion.

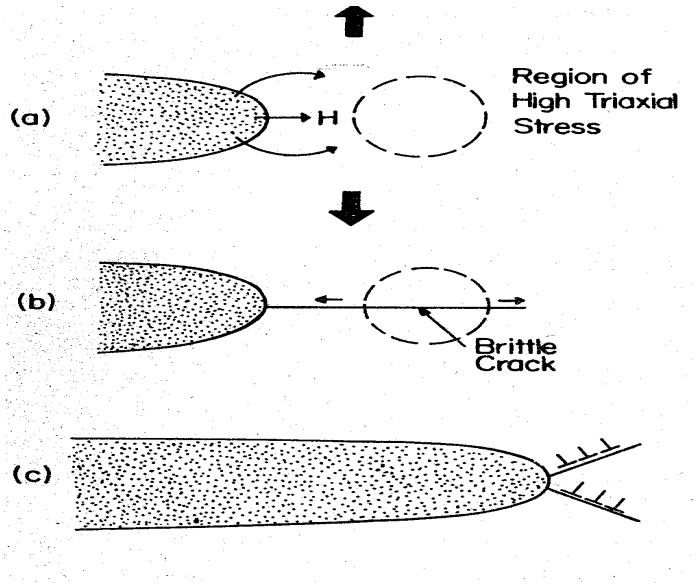

Figure 8. Schematic illustration of the decohesion model for hydrogen embrittlement proposed by Troiano. [39]

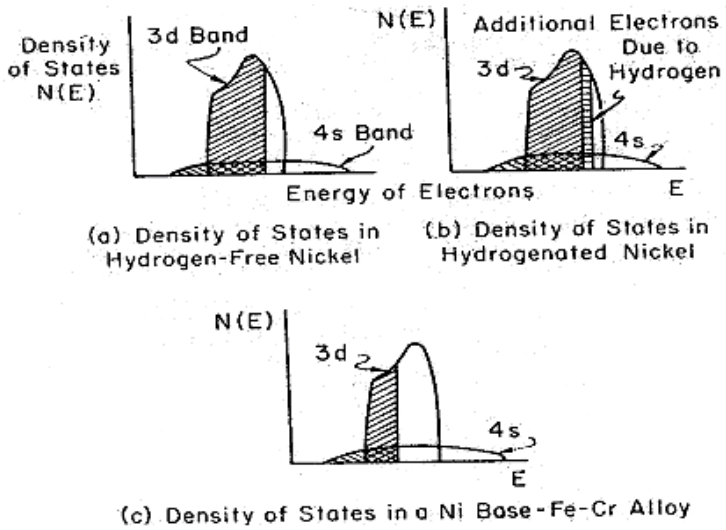

Figure 9. Density of states in hydrogen–free and hydrogenated nickel and nickel base–chromium–iron alloys. [23]

Elaborating Troiano's ideas, Oriani [43, 44] modified the decohesion model in terms of (a) the magnitude of the elastic stresses, necessary to produce adequate hydrogen accumulation, (b) the location of the sites of embrittlement and (c) the mode of crack propagation. With regard to the first point, Troiano [23] supported that the maximum tensile stress ahead of the crack is approximately three times the yield stress. However, Oriani [20, 43] pointed out that a stress of this extent results in a threefold enhancement of hydrogen solubility, $3c_o$, in steel, if c_o denotes the normal stress–free solubility, which in iron at room temperature is roughly 3 atoms of H per 10^8 atoms of iron under 1 atm hydrogen gas pressure. Such a small increase of hydrogen concentration does not account for the observed fracture. Consequently, larger solubilities of dissolved hydrogen must be presumed by a convincing decohesion model. There could be two possible kinds of mutually non–excluding sites in steel, at which sufficiently high stress could be produced, resulting in adequately high and embrittling hydrogen concentrations. The first would be a few angstroms ahead of a sharp crack submitted to opening forces, and the second would be at interfaces, grain boundaries and dislocation cores, where hydrogen is adsorbed, as supported by autoradiography of tritium [45]. These studies demonstrate a heterogeneity of distribution of tritium introduced cathodically in iron, showing effective hydrogen trapping to structural defects.

Concerning the magnitude of the local stresses producing large hydrogen accumulation, Oriani postulated that they could be as large as 10% of the Young modulus E. This is a basic stipulation of Oriani's model, undoubtedly existing at the tips of critical cracks in perfectly brittle solids, but strongly disputed by Pugh [39] for ductile materials, where stresses of E/10 would be expected to be immediately eliminated by plastic deformation.

Another premise of Oriani's theory [43, 44] is that large concentrations of dissolved hydrogen reduce the bonding forces amidst iron atoms so that the maximum cohesive force F_m is lowered. When the local stress equals the maximum cohesive force per unit area at the crack tip, then the crack will propagate in a brittle fashion, i.e. not by plastic tearing, as shown in Figure 10. (Atomic model of crack edge)

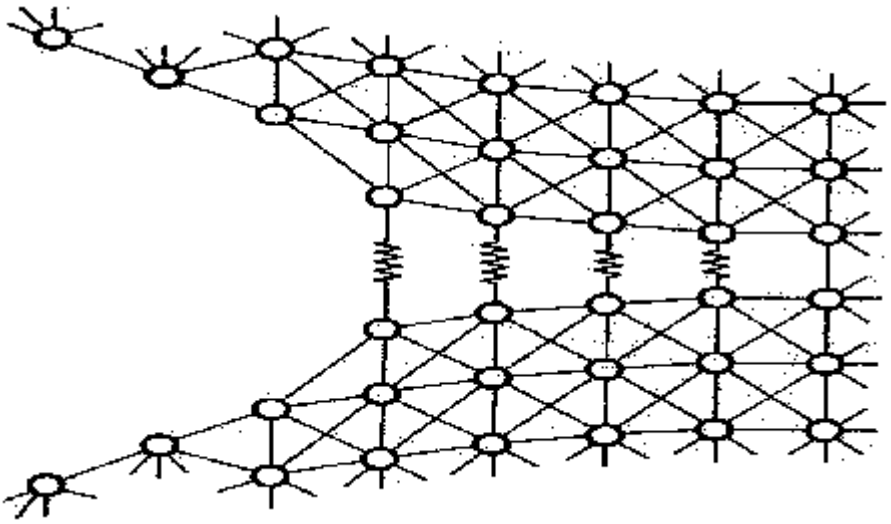

Figure 10. Atomic model of the crack edge. The straight lines represent bonds stretched in the linear stress–strain region. The zig–zag lines represent bonds stretched in the non–linear region. [43]

Figure 11 demonstrates that at any value of separation, ε, between lattice planes, there is a cohesive force F(z) which is different for hydrogen–free straining than for a lattice with hydrogen, and on the same time, the second diagram illustrates the rapid increase of hydrogen concentration with increasing elastic strain, before levelling off towards a maximum. In the first diagram, the area under the curve represents the necessary work for fracture, therefore implying that the presence of hydrogen lowered that work. In other words, hydrogen embrittled the material or reduced its mechanical energy absorption ability, as clearly stated in the definition of HE. [20]

In terms of the mode of crack propagation, Oriani's ideas differentiated from those of Troiano, as the former considered crack propagation as continuous. As in the case of the surface energy theory (adsorption model) he encountered the problem to explain how a brittle crack can propagate at low velocities without becoming blunted by plastic deformation. Moreover, a non–intermittent crack process is not consistent with acoustic emission results. [39] However, Oriani's model differs from the adsorption theory in that the presumed embrittlement occurs a few atomic distances below the surface rather than at the surface itself. Consequently, the latter has the advantage that the hydrogen surface concentration is inherently large, so that any postulation of large elastic stress producing hydrogen accumulation is unnecessary. [39] Finally, as

supported by Oriani [43, 44] the relation between the adsorption theory for HE and the lattice decohesion theory lies in the fact that the hydrogen caused reduction of the bonding forces between the iron atoms implies the reduction of surface energy, and thus the lowering of the thermodynamically established criterion for crack growth. In this context the adsorption theory specifies a necessary condition that hydrogen must satisfy in order to enhance crack propagation, but this is not a mechanism, as the essential phenomenon of the lowering of the cohesive force by hydrogen mechanistically precedes the surface energy reduction.

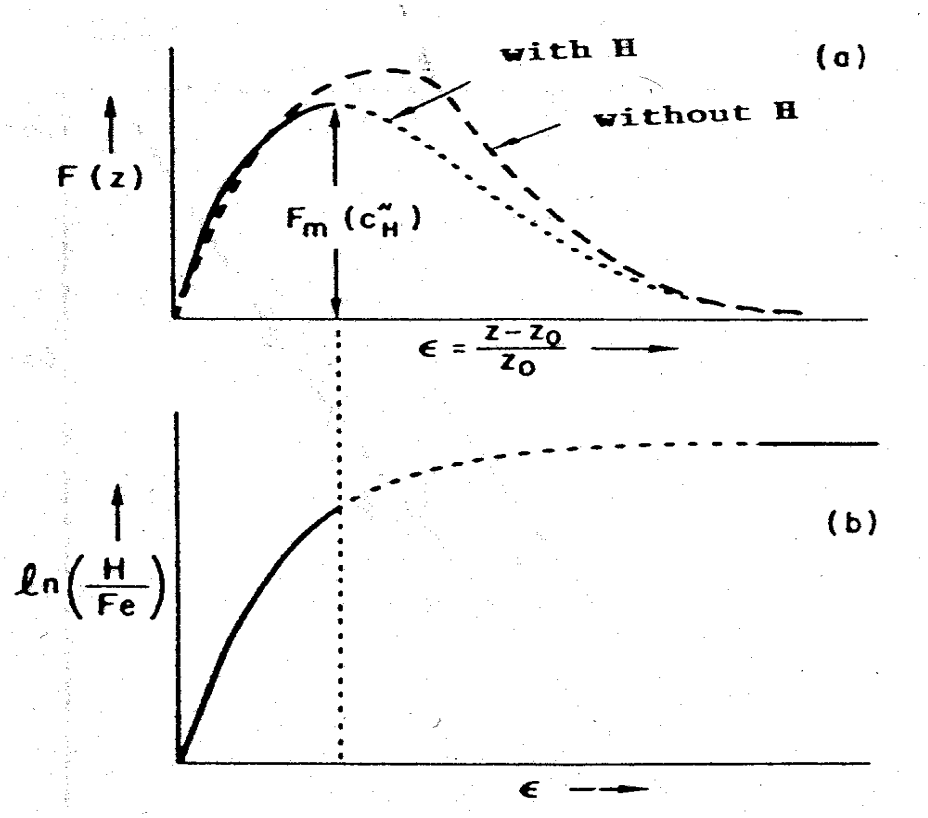

Figure 11. At any value of separation, ε, between lattice planes, there is a cohesive force F(z), shown by schematic diagram (a), which is different for a lattice without hydrogen than for one with hydrogen. Diagram (b) shows schematically the concomitant increase of hydrogen concentration with increasing ε at constant chemical potential of hydrogen. [20]

1.2.6 Localised Slip Model

The first proponent of this mechanism for hydrogen related damage in quenched and tempered steels was Beachem [28]. He presented a model, so as to account for the observations of decreasing microscopic plasticity and changes of fracture modes with decreasing stress intensities at crack tips during hydrogen embrittling phenomena in steels. According to this model, the presence of sufficiently concentrated hydrogen dissolved in the lattice ahead of the crack tip assists whichever deformation processes the microstructure will permit. The microstructure, the crack tip stress intensity and the concentration of hydrogen are the factors, which determine if intergranular, quasicleavage or microvoid coalescence fracture modes will operate.

Based on these proposals, the conspicuous contradiction of a material being embrittled by hydrogen through a mechanism involving plastic deformation disappears, if the plastic deformation is a localised, microscopic effect used to propagate the crack, rather than being an overall macroscopic effect. Moreover, Beachem's model suggests that hydrogen does not inhibit dislocation motion by locking them, but it unlocks and allows them to multiply or move at reduced stresses. Therefore, hydrogen simply permits or coerces the normal fracture processes to become operative at unusually low macroscopic strains. However, it should be noticed that dislocation motion's enhancement is merely a contributing factor to an overall degradation process, and not a mechanism on its own. Finally, Beachem's ideas are not at variance with Troiano's [23] conclusion that hydrogen causes lowering of the true fracture strength of the lattice, if fracture at the crack tips is considered as equal with the microscopic and submicroscopic plastic flow at the crack tips.

The sketches of Figure 12 illustrate the most significant finding of Beachem's research, which is the reduced plastic deformation as a function of decreasing stress intensity and decreasing cracking rate. The stress intensity factor (K) in Figure 12(a) is sufficiently large to cause large volumes of materials to plastically deform at the crack tip provided there is an adequate hydrogen content level. As the arising plastic zone is large enough to include multitudinous inclusions acting as void nuclei, the developing fracture process is microvoid coalescence. As the crack is lengthened, the

K value decreases and the plastic zone is not large enough to incorporate numerous inclusions and thus microvoid coalescence ceases to be sustained.

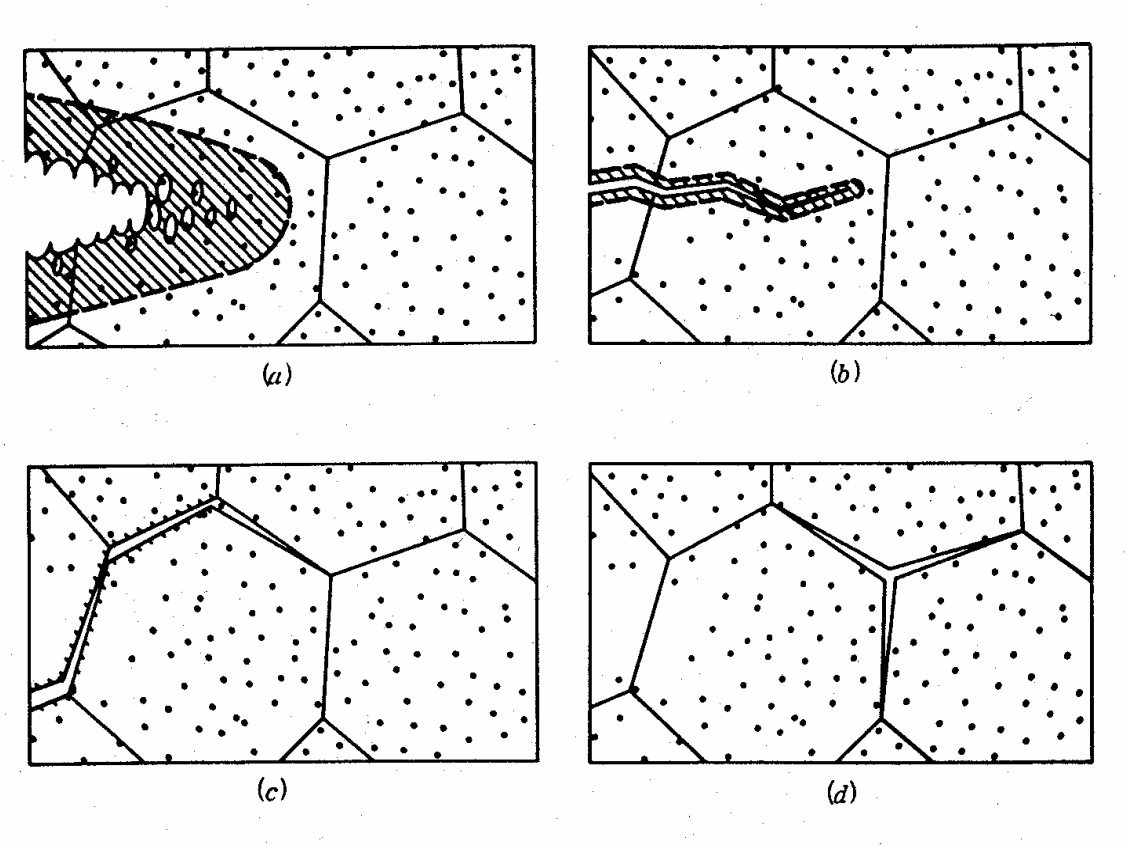

Figure 12. Sketches of microscopic fracture modes observed in experiments as a function of decreasing stress intensity factor and concomitant decreasing cracking rate. (a) high K, microvoid coalescence (b) intermediate K, quasi cleavage (c) low K, intergranular cracking (d) intergranular cracking with an assist from hydrogen pressure [28]

Nevertheless, in Figure 12(b) the K value is big enough to allow forming cracks propagate by plastic deformation required in quasicleavage. In even longer cracks, and consequently, lower K values, neither microvoid coalescence nor quasicleavage can be supported. As a result, the lower energy plastic deformation process of intergranular cracking is viewed to occur, as depicted in Figure 12(c). At even smaller stress intensities and crack lengths, hydrogen diffuses into the metal, and more specifically, at grain boundaries or other weak boundaries with enough pressure to

provide its own driving force to increase the other already existing forces. So, the cracks propagate intergranularly with hydrogen pressure assistance. (Figure 12(d))

Intergranular cracking is considered to be more energetically favourable compared with microvoid coalescence or quasicleavage, which, however, preponderate at higher K values, as they are kinetically faster processes. Slower cracking rates at longer crack lengths and lower K values allow hydrogen to diffuse further ahead of the crack tip, and to concentrate at the grain boundaries, hence making the conditions for grain boundary plasticity more favourable.

Lynch lent support to Beachem's theory, as he observed that identical fractographic results were obtained from Stress Corrosion Cracking (SCC), Hydrogen Assisted Cracking (HAC) and Liquid Metal Embrittlement (LME), as well as, that very small dimples could be found on cleavage surfaces for a range of materials including iron–based alloys. [46–49] The most striking similarity between HAC and LME in low alloy ultra high strength steels, such as D6ac (0.46% C), was that the crack growth in both gaseous hydrogen and liquid mercury produced intercrystalline fractures, whilst fracture in inert environments was transcrystalline with dimpled surfaces. [46, 47]

Lynch believed that LME occurs at a crack tip on the surface of the solid metal. Based on the previously referred fractographic similarities of LME with SCC and HAC, the latter is also postulated to apply to crack tips, in advance of which the formation of microvoids is documented by dimples found on cleavage surfaces. Lynch regarded HE as a surface effect with adsorbed hydrogen rather than dissolved being generally responsible for it, as chemisorption of hydrogen atoms facilitates nucleation of dislocations at crack tips. [46–49] The effect of adsorption on metal surfaces is limited to the first few atomic layers at the surface, inflicting a redistribution of the electron density in these surface layers. Such an effect presumably results in weakening of interatomic bonds at crack tips thereby facilitating either tensile separation of atoms (decohesion) or shear movement of atoms (dislocation nucleation) at crack tips. [48] In this context, it is not coincidental, that Lynch clearly states his agreement with Oriani, that hydrogen decreases the strength of interatomic bonds in the surface or very close to it and agrees with Beachem, that HAC generally occurs by localised plastic flow. [47]

Birnbaum adopted a very similar approach to that one of Lynch, as they both agree that the presence of hydrogen increases the plasticity at the crack tip leading to fracture. However, the proposed by Birnbaum Hydrogen Enhanced Local Plasticity (HELP) model demonstrates the discrepancies between these two investigators, as Lynch considers the phenomenon to be a surface effect, whilst Birnbaum views the hydrogen effect to occur in the volume of the material as well as near the surface.

Birnbaum suggested that internal hydrogen associates around discontinuities in the metal lattice, such as inclusions, point defects and dislocations. Hydrogen transport is facilitated when a dislocation moves, as the hydrogen atmosphere is swept along with it due to hydrogen's high diffusivity. Moreover, hydrogen enhances screw, edge and mixed dislocation mobility and increases their velocities, and it is this enhanced dislocation motion by hydrogen which is the cause of the hydrogen effects on fracture, as investigated in a variety of materials: α–Ti alloys [50, 51], 7075 and 7050 age hardened Al alloys [52], pure Al [53] and Fe [54, 55]. Using Transmission Electron Microscopy (TEM) Birnbaum observed failure in thin, hydrogen free and hydrogen charged specimens. Fracture in hydrogen was similar to that in vacuum, but with the substantial difference that in the presence of hydrogen the plastic deformation and fracture occurred at reduced stresses. Finally, hydrogen was shown to localise the plastic region ahead of the crack tip leading to embrittlement.

The present understanding of hydrogen enhanced dislocation motion is incomplete. Based on a mechanism of "elastic shielding" of dislocations by hydrogen atmospheres, the mobility of hydrogen allows it to diffuse to sites of lowest free energy and to form high concentration atmospheres around dislocations, solutes and precipitates. According to Hirth [24] binding enthalpies for hydrogen are typical in the range between 10 and 50 KJ/mol, and hence, atmospheres are permitted to form over the range of temperatures in which HE is observed. These atmospheres can move with the dislocations, due to high H diffusivity, in response to applied stress over a temperature range. For most studied systems, hydrogen atmospheres are expected to move with the dislocations and continuously adjust their local concentrations to minimise the free energy of the system.

The magnitude of the hydrogen shielding effect increases with an increasing hydrogen concentration in the atmosphere, implying an increasing lattice hydrogen concentration and a decreasing temperature as well. Consequently, "elastic shielding" effects are small at high temperatures, where the hydrogen concentration in the atmospheres is small, and are small at low temperatures, where the atmospheres are not mobile enough to move with the dislocations. At intermediate temperatures and low strain rates, hence low dislocation velocities, the effects are maximised due to atmosphere rearrangements occurring in phase with the dislocation motion. Under these conditions, HE is mostly pronounced. (Figure 4)

Zakroczymski [56] did not find any increase in bulk diffusion of hydrogen during plastic deformation of iron, nickel and stainless steel, as would be expected if dislocation motion was influencing hydrogen transport. On the other hand, generation of dislocations by plastic deformation enhances hydrogen trapping, which is effectively reducing the diffusivity and permeability of hydrogen through the metal, as hydrogen trapping overcomes an increase in hydrogen transport by moving dislocations. Therefore, hydrogen transport in iron is not accelerated by moving dislocations, but, on the other hand, is not totally excluded.

1.3 Hydrogen Evolution and Entry into Metals

The previously described HE mechanisms are unlikely to occur if hydrogen does not enter a metallic material. Therefore, hydrogen has firstly to be transported to the metal surface, then adsorb, absorb and eventually be transported into the metal bulk, where it can contribute to the material's degradation.

When a metal corrodes several reactions can take place apart from the anodic dissolution of the metal. These are the oxygen reduction reaction, which is restricted by the availability of oxygen in the system, and the hydrogen evolution reaction. Both of these reactions can happen in acid, neutral or alkaline conditions. Depending on the pH they can have the following forms; -

Oxygen Reduction Reaction (ORR)

$\frac{1}{2} O_2 + 2H_3O^+ + 2e \rightarrow 3H_2O$ (acid)

$\frac{1}{2} O_2 + H_2O + 2e \rightarrow 2OH^-$ (neutral, alkaline)

Hydrogen Evolution Reaction (HER)

$H_3O^+ + e \rightarrow \frac{1}{2} H_2 + H_2O$ (acid)

$H_2O + e \rightarrow \frac{1}{2} H_2 + OH^-$ (neutral, alkaline)

The understanding of the hydrogen entry problem depends on an understanding of the characteristics of the HER and the hydrogen absorption into the metal. Even a simple reaction as the HER occurs in a number of steps, which are depicted according to McCright [57] in Figure 13. Firstly a hydrated proton H_3O^+ is transported to the double layer of the metal surface. During the subsequent desolvation step the water of the hydration shield is lost in the surrounding area of the double layer, followed by adsorption of the proton to the electrode surface. Afterwards, an electronation (discharge) of the proton occurs forming an adsorbed hydrogen atom. The following step is that of combination, which can occur in two ways, chemically or electrochemically. In the former case, two adjacent adsorbed hydrogen atoms can combine to form a hydrogen molecule with a possibility of surface migration between the discharge and recombination site. In the latter case, an adsorbed hydrogen atom

can combine with a proton – which is reduced by an electron, hence forming a hydrogen atom – and consequently form a hydrogen molecule. Eventually, either desorption of hydrogen molecules, evolving as bubbles of hydrogen gas (HER), occurs, or absorption of hydrogen atoms and diffusion into the metal substrate takes place.

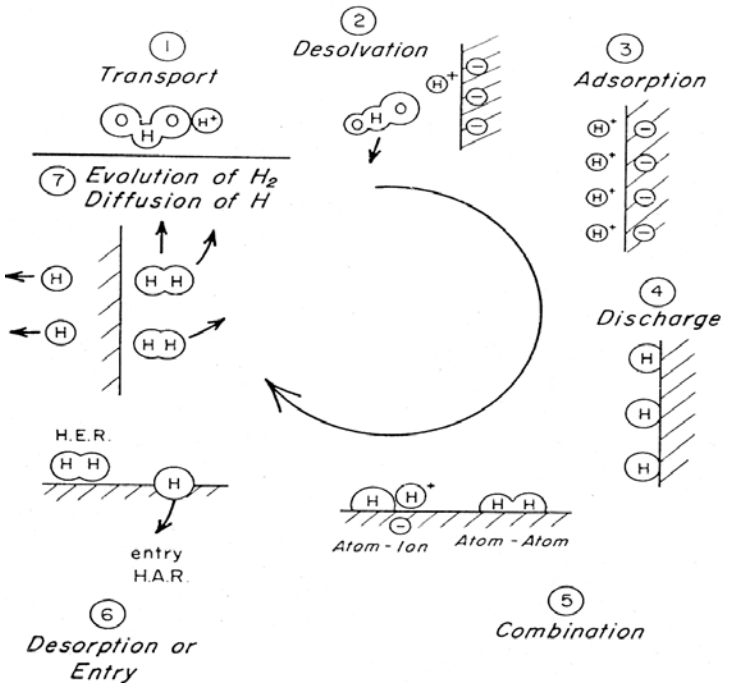

Figure 13. Processes of hydrogen evolution and absorption [57]

The steps (1) and (7) in Figure 13 are seldom restrictive, whilst steps (2) and (3) are thought not to affect the reaction rate on the whole. On the other hand the combination and desorption steps are usually combined because of the hydrogen's weak affinity to chemisorb. Hence, the two important steps, in which the HER occurs, are the discharge step and one of the chemical or electrochemical combination steps. Either one of these steps may be the rate–determining step or they may proceed in a coupled discharge–recombination reaction. Moreover, according to some investigators [58], the chemical and the electrochemical recombination reactions on iron occur simultaneously, but often the preponderance of one of these recombination reactions strongly depends on the composition of charging solution and the cathodic current density or overpotential. In acidic environments chemical recombination predominates at high current densities, while in neutral or alkaline solutions the

electrochemical recombination reaction prevails at comparatively low current densities. The main reason for this trend lies on the different exchange current densities of the electrochemical recombination reaction with hydrogen ions or with water molecules.

Concerning hydrogen entry into the metal, two models have been proposed. The first one considers hydrogen entry in the atomic state, in which hydrogen exists on the surface. In this case the absorption and the chemical recombination processes are competing. According to the second model, hydrogen entry occurs from a directly discharged proton without passing through the intermediate absorbed phase and thus, the HAR is competing with the HER for the same discharged protons. [57] (Figure 14)

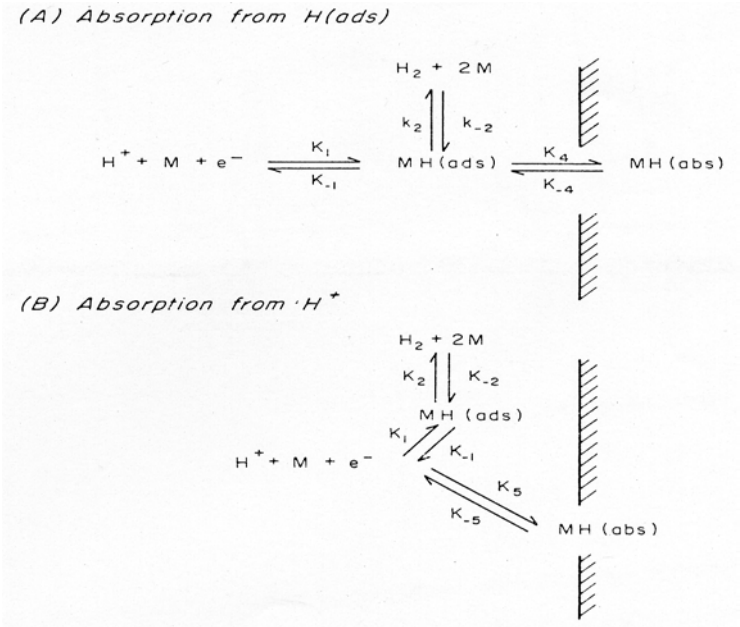

Figure 14. Models for hydrogen entry into metals: (A) Absorption from atomic hydrogen and (B) absorption from protons [57]

The previously referred models of hydrogen entry are correlated with two models of hydrogen chemisorption. (Figure 15) The firstly described model of hydrogen entry in the atomic state is related to the r–type adsorption model, according to which hydrogen adsorbed atoms are the slightly negative members of the dipoles they form with metal atoms in a largely covalent bonding. Each hydrogen atom is located over

and outside a corresponding metal atom's electronic cloud. From measurements of electrical resistivity on films of nickel and platinum, r–type adsorption is concluded to dominate at low hydrogen coverage, while the s–type adsorption model preponderates at higher hydrogen coverage and it is related with the secondly described model of hydrogen entry from discharged protons. According to the s–type adsorption, the hydrogen behaves more like a proton dissolved in the surface layer, whilst being the positive member of a more ionic bonding.

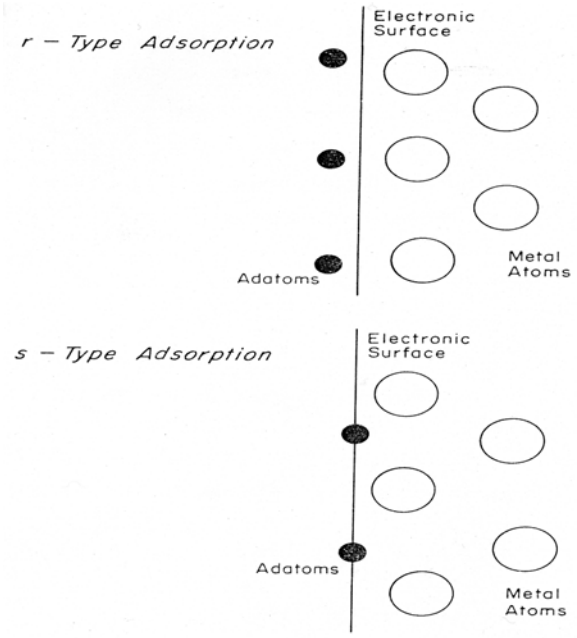

Figure 15. Models for chemisorbed hydrogen. [57]

The previous models for hydrogen evolution and entry into metals contributed in a more qualitative manner to the understanding of these issues. However, a mechanistic model, proposed by some investigators [59–61] with the name IPZ model, attempted for the first time to quantitatively correlate hydrogen surface coverage and concentration, as well as hydrogen discharge, adsorption, absorption and recombination reaction rate constants. This was an advance compared to some previous work [58, 62, 63], where some, but not all of these kinetic parameters, were addressed. Some aspects of the IPZ model will be presented in the following sections, combined with the study of hydrogen permeation, which is going to be viewed after a short discussion about some basic concepts of diffusion.

1.4 Hydrogen Diffusion

Diffusion is the process by which matter is transported from one part of a system to another as a result of the random and chaotic motion of particles. Consequently, it is impossible to predict a velocity for each particle individually. However, in the case of solutions, where there is a net excess of solute particles at one extremity of the solvent, there will be a clear diffusion from the region of higher concentration towards the region of lower concentration. As large numbers of particles are involved in such situations, the prediction of the transport of one substance through another becomes possible at a specific temperature.

Concerning the mathematical interpretation of the diffusion problem, there are many analogies between the heat conduction through a solid and the mass transfer, which were first recognised by Fick in 1855 who applied the heat conduction equations derived by Fourier in 1822. Hence, there is a range of heat conduction equations, the solutions of which are directly applicable to diffusion. [64]

Fick's first law of diffusion considers the situation in an isotropic medium, where the flux j (moles cm^{-2} s^{-1}) is directly proportional to the concentration gradient $\partial C/\partial x$ (moles cm^{-3} cm^{-1}), and the diffusion coefficient D (cm^2 s^{-1}) is assumed to be constant at a given temperature.

Equation 1:

$$j = -D\frac{\partial C}{\partial x}$$

Under non–steady state, the Fick's second law is applied and Equation 2 is its fundamental differential equation, the exact solution of which varies for different problems depending on the boundary conditions imposed by the physical situation.

Equation 2:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2}$$

1.5 Hydrogen Permeation

Devanathan and Stachurski [65] established in 1962 a new sensitive electrochemical technique that allows the recording of the instantaneous permeation rate of electrolytic hydrogen through metallic membranes by using a simple double–cell arrangement, the depiction of which, as well as its electrical circuit, are presented in Figure 16.

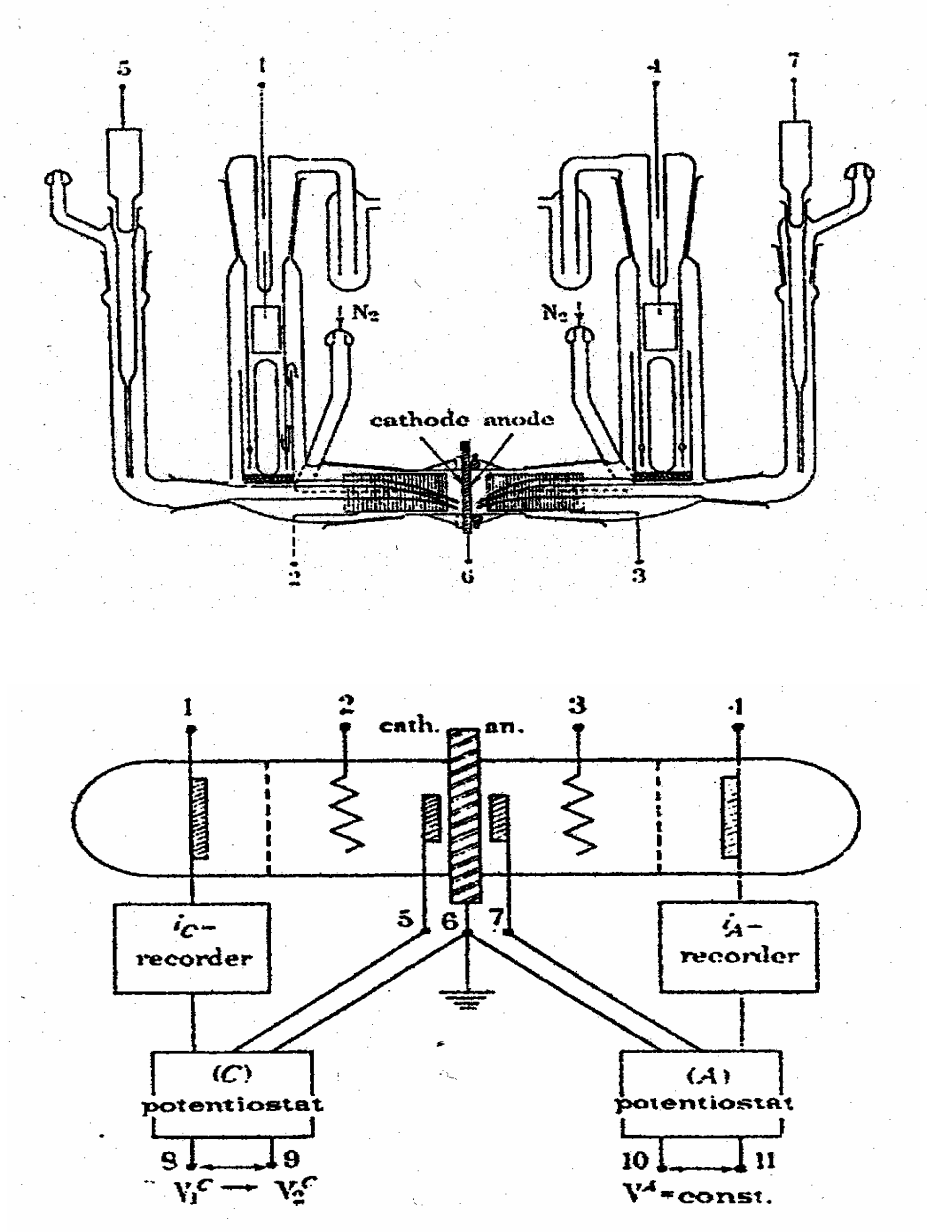

Figure 16. Double permeation cell arrangement and its electrical circuit diagram. [65]

The two cells are separated by a thin membrane electrode, made of the test material, the one side of which acts as the cathode in one cell (cathodic), whilst the opposite side acts as the anode for the other cell (anodic). As viewed from the hydrogen evolution and entry reactions, hydrated protons or water molecules are reduced from aqueous solution at the cathodic side of the cell. Most hydrogen reduced in this way combines to form hydrogen gas bubbling away from the electrode surface, while a small fraction of hydrogen that is absorbed at the membrane surface diffuses through the steel to the far side of the membrane, where it comes under control of the anodic cell. At the anodic surface of the membrane the electrochemical conditions are established to oxidise the emerging hydrogen atoms and remove them from the membrane, thus producing a current that is recorded as a function of time. This measured current is by Faraday's laws a direct estimation of the instantaneous rate of permeation of hydrogen, a continuous recording of which is obtainable with all the sensitivity associated with current measurement. A 0.2 N NaOH solution is used in the anodic cell, in order to keep steel in the passive condition and prevent an iron oxidation current from increasing the measured current. The oxidation of hydrogen at the anodic cell imposes a zero hydrogen concentration at the anodic surface, and with both surfaces at fixed concentrations the hydrogen concentration profile through the membrane approaches gradually a steady state

The double-cell permeation technique uses an intentionally produced concentration gradient and Fick's first law to equate each electron of the oxidation current to a hydrogen atom permeating the membrane. Consequently, if F represents the Faraday constant (96500 coulombs moles^{-1}), then the permeation current J (A/cm^2) is given by the Equation 3.

Equation 3:

$$J = -DF\frac{\partial C}{\partial x}$$

During permeation experiments the first hydrogen atoms diffuse through the membrane to the exit surface after the elapsing of a specific period of time, termed the breakthrough time t_b. Afterwards, the permeation current steadily increases until the establishment of steady state conditions, where an equilibrium flux or plateau current

value is reached (J_∞) which permits the surface hydrogen concentration to be evaluated from the combination of Fick's first law and Faraday's laws. Should the hydrogen diffusion coefficient D ($cm^2\ s^{-1}$) be known and from the thickness of the membrane L (cm), as well as the steady state flux J_∞ ($\mu A\ cm^{-2}$), the surface hydrogen concentration C_o (mol cm^{-3}) can be calculated as follows; -

Equation 4:

$$C_o = \frac{J_\infty L}{FD}$$

The diffusion coefficient D can be determined from either a measurement of the breakthrough time t_b (s) or the time lag t_L (s) from the permeation current transients, given the thickness L of the membrane.

Equation 5:

$$t_b = \frac{L^2}{15.3\,D}$$

Equation 6:

$$T_{lag} = \frac{L^2}{6\,D}$$

The time lag is the time for the permeation rate to reach 0.63 times the steady state value J_∞. [65] Moreover, the hydrogen diffusion coefficient D is dependent on the microstructure and trapping capability of a material, as well as the experimental temperature. Reduction of the experimental temperature reduces the hydrogen flux and calculated hydrogen concentration values will reflect the lower diffusion coefficients. On the other hand, D is independent of applied stress, but hydrogen permeability increases with higher applied stress. [66]

Figure 17 depicts a typical transient showing the flux of hydrogen permeating a metallic membrane with a constant hydrogen concentration on the charging side. A steel membrane held at a constant potential is expected to behave in this manner.

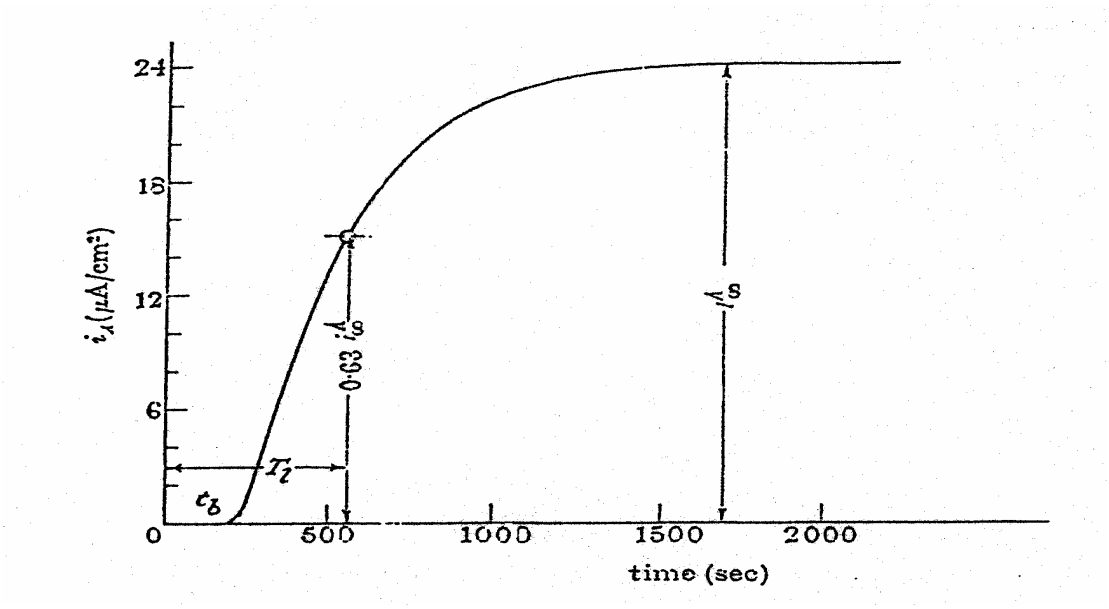

Figure 17. A typical permeation transient

Before the establishment of steady state conditions in the permeation process the phenomenon is under non–steady state. The equation that describes the above case is derived from the second Fick's law.

In the case of flow through a membrane [64], if one face (x = 0) of the membrane is kept at a constant concentration C_o, and the other x = L at C_L, and the membrane is initially at a constant concentration C_i, there is a finite interval of time during which the steady state condition is established. During this time the concentration changes according to the Equation 7.

Equation 7:

$$C = C_o + (C_L - C_o)\frac{x}{L} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{C_L\cos n\pi - C_o}{n}\sin\frac{n\pi x}{L}\exp(-\frac{Dn^2\pi^2 t}{L^2})$$
$$+\frac{4C_i}{\pi}\sum_{m=0}^{\infty}\frac{1}{2m+1}\sin\frac{(2m+1)\pi x}{L}\exp\{\frac{-D(2m+1)^2\pi^2 t}{L^2}\}$$

In the commonest experimental arrangement both C_i and C_L are zero, i.e. the membrane is initially at zero concentration, and the concentration at the face through which diffusing substance emerges is maintained effectively at zero level. As a result, the previous relationship is written now as follows; -

Equation 8:

$$C = C_o - C_o \frac{x}{L} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{-C_o}{n} \sin \frac{n\pi x}{L} \exp(-\frac{Dn^2 \pi^2 t}{L^2})$$

As t approaches infinity the terms involving the exponentials vanish and we have simply a linear concentration distribution. This can be shown in the following graph (Figure 18).

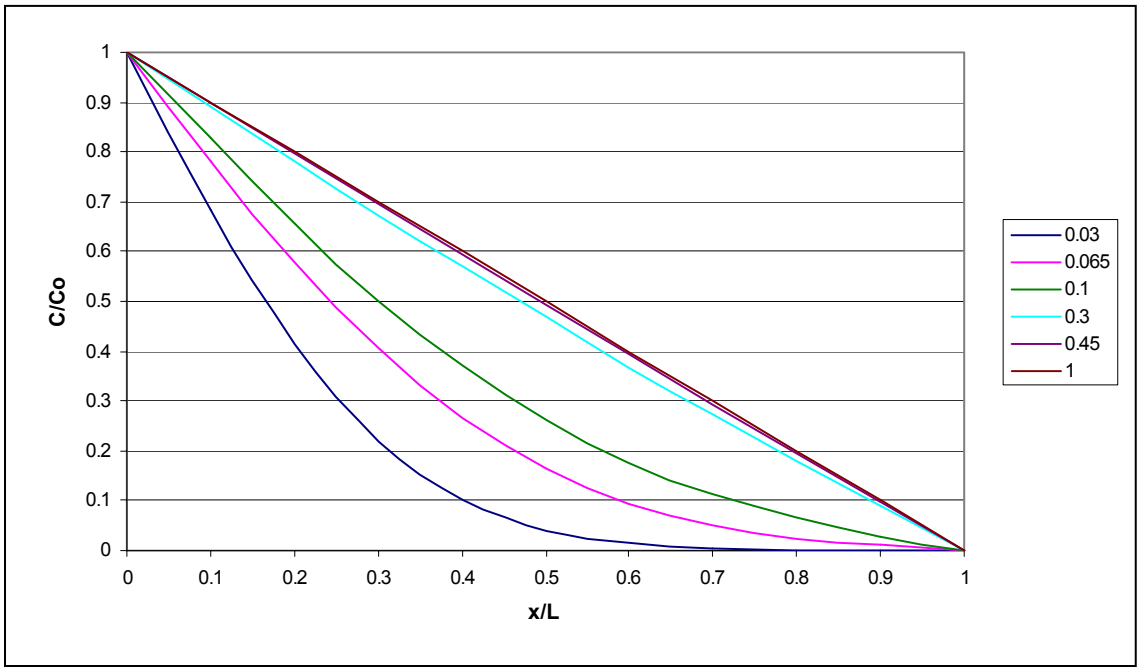

Figure 18. Concentration distributions at various times (values for Dt/L^2).

The rate at which the hydrogen emerges from unit area of the face x = L of the membrane is given by $-D(\partial C/\partial x)_{x = L}$. [64] By integrating with respect to t, we obtain the total amount of diffusing substance Q_t (hydrogen) which has passed through the membrane in time t.

Equation 9:

$$Q_t = D(C_o - C_L)\frac{t}{L} + \frac{2L}{\pi^2}\sum_{n=1}^{\infty}\frac{C_o\cos n\pi - C_L}{n^2}\{1-\exp(-\frac{Dn^2\pi^2 t}{L^2})\}$$
$$+\frac{4C_i L}{\pi^2}\sum_{m=0}^{\infty}\frac{1}{(2m+1)^2}\{1-\exp(-\frac{D(2m+1)^2\pi^2 t}{L^2})\}$$

Under the conditions of our case (C_i=0 and C_L=0) we find

Equation 10:

$$\frac{Q_t}{LC_o} = \frac{Dt}{L^2} - \frac{1}{6} - \frac{2}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\exp(-\frac{Dn^2\pi^2 t}{L^2})$$

This, as t tends to ∞, approaches the line

Equation 11:

$$Q_t = \frac{DC_o}{L}(t - \frac{L^2}{6D})$$

This has an intercept on the t–axis, which is known as the time lag, and given by $T_{lag} = L^2/6D$. This time lag has been used as the basis of a method to obtain the diffusion constant for the steel membrane. [64]

A graph of Q_t/LC_o as a function of Dt/L^2 is shown for the case $C_i = C_L = 0$ in Figure 19. It is observed that the steady state is achieved when $Dt/L^2 = 0.45$ approximately. Moreover, the first hydrogen permeates the membrane when $Dt/L^2 = 0.0654$. As a result, the breakthrough time is $t_b = L^2/15.3D$, which is another option for the establishment of a value for D.

The flux of hydrogen permeating the membrane can be recorded throughout the experiments and used to calculate the total amount of hydrogen absorbed by the steel substrate.

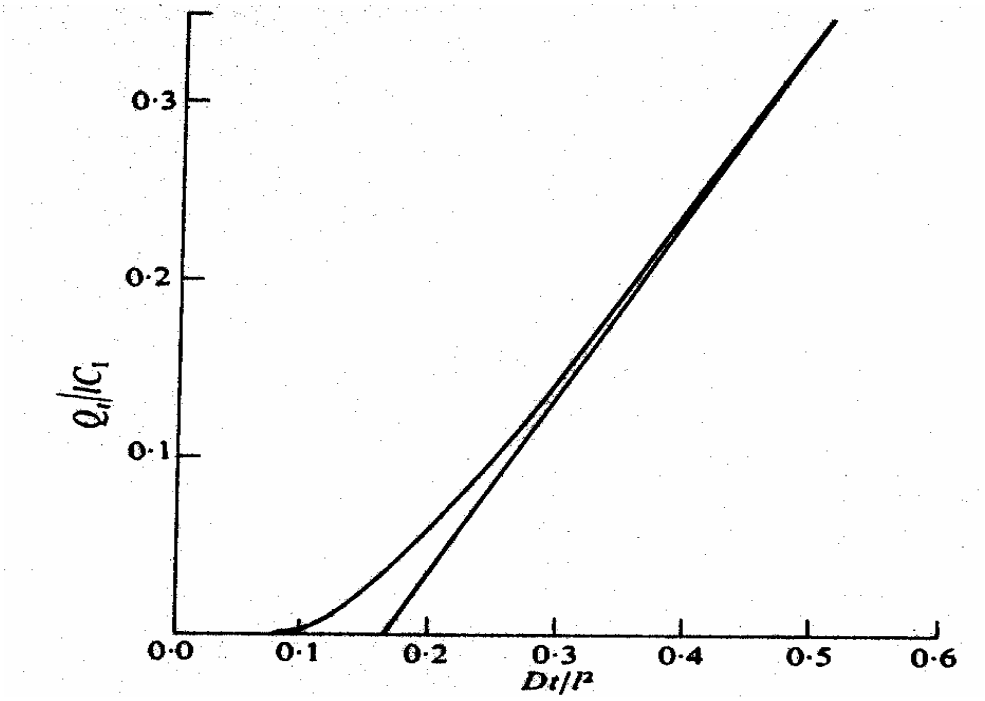

Figure 19. Approach to steady–state flow through a plane sheet. [64]

Various electrochemical methods for studying diffusion, permeation and solubility of hydrogen into metals were considered by Boes and Züchner [67]. The flexibility of these techniques is based on the copious possibilities of experimentally varying the boundary conditions at both the entrance and the detection sides of the sample. Some of the different possible combinations of boundary conditions led to diverse methods, such as the non–steady–state galvanostatic time–lag method, the step method, the non–steady–state potentiostatic time–lag method, the pulse method, the enforced oscillation method and the self–excited oscillation method. For each of these methods equations were derived to predict the theoretical permeation transients at the exit side of a permeation membrane, as well as the hydrogen concentration profiles throughout the membrane.

For hydrogen permeation carried out under potentiostatic and galvanostatic conditions the hydrogen concentration at the exit surface of the membrane is in both conditions maintained at zero level, while at the entry surface simple static boundary conditions are produced. Potentiostatic (P) conditions were thought to maintain a constant hydrogen surface concentration with a variable flux at the entrance surface (C case),

whilst galvanostatic (G) charging was assumed to maintain a constant flux at the charging (entry) surface with a variable potential (F case). [68]

The non–steady state potentiostatic conditions were just examined in the double permeation cell by Devanathan and Stachurski. Hence, the hydrogen concentration distribution is described by Equation 8 and illustrated in Figure 18. Moreover, the equation which predicts the rising permeation transient, depicted in Figure 17, is derived from the Equations 3, 4 and 8 as follows; -

Equation 12:

$$J(t) = -DF(\frac{\partial C}{\partial x})_{x=L} = J_{\infty}\{1 + 2\sum_{n=1}^{\infty}(-1)^{n}\exp(-\frac{Dn^{2}\pi^{2}t}{L^{2}})\}$$

For the non–steady state galvanostatic conditions Boes and Züchner [67] give Equation 13 for the rising permeation transient, while the eventually observed steady state permeation flux J_{∞} at the exit side of the membrane is the same as the constant flux of hydrogen J_0 entering the membrane at all times. Moreover, the hydrogen concentration distribution through the membrane thickness is given by Equation 14.

Equation 13:

$$J(t) = J_{0}\{1 - \frac{4}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^{n}}{2n+1}\exp(-\frac{(2n+1)^{2}\pi^{2}Dt}{4L^{2}})\}$$

where:

J_0 Entrance current density

J(t) Efflux or permeation current density

n Integer; 0,1,2,3…

Equation 14:

$$C = \frac{J_{0}(L-x)}{FD} - \frac{8J_{0}L}{FD\pi^{2}}\sum_{n=0}^{\infty}\frac{(-1)^{n}}{(2n+1)^{2}}\sin\{\frac{2n+1}{2L}\pi(L-x)\}\exp(-\frac{(2n+1)^{2}\pi^{2}Dt}{4L^{2}})$$

After the establishment of steady state conditions the hydrogen concentration at the cathodic surface of the membrane, C_o, is given by Equation15.

Equation 15:

$$C_0 = \frac{J_0 L}{FD}$$

Bearing in mind the dimensionless variables for time, distance, concentration and current density, provided by Equations 16, 17, 18 and 19, respectively, Figure 20 shows the normalised concentration profiles and permeation fluxes for both the non-steady steady state potentiostatic and galvanostatic conditions. [68]

Equation 16:

$$\tau = \frac{Dt}{L^2}$$

Equation 17:

$$\chi = \frac{x}{L}$$

Equation 18:

$$\sigma(\chi, \tau) = \frac{C(x,t)}{C_0}$$

Equation 19:

$$\iota(\tau) = \frac{J(t)}{J_\infty}$$

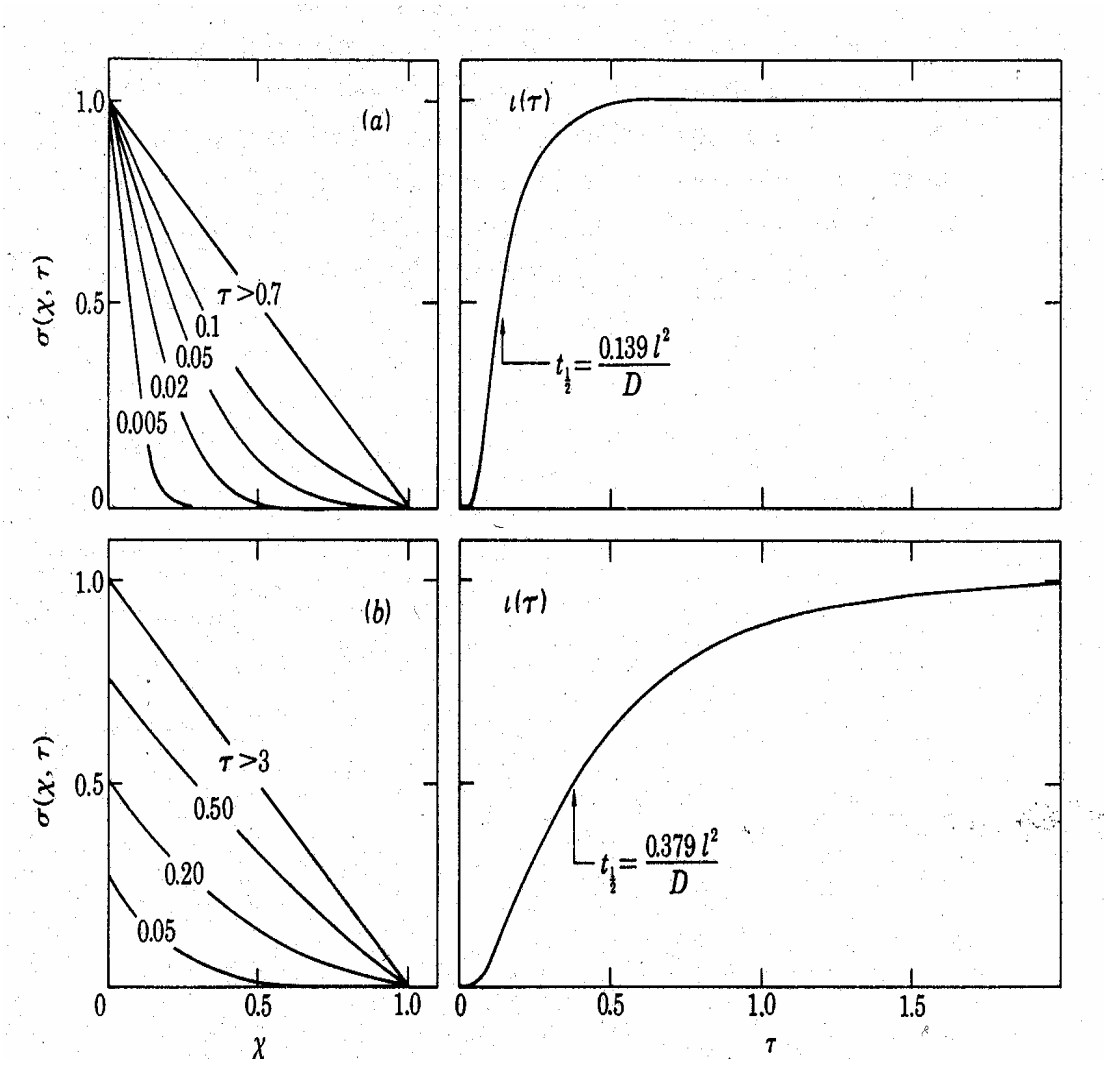

Figure 20. Normalised concentration profiles σ(χ,τ) and permeation currents ι(τ) for (a) the C boundary condition and (b) the F boundary condition. Half rise times $t_{½}$ are indicated as well. [68]

Archer and Grant [68], based on experimental work with palladium and nickel foils, examined the widely accepted assumption that potentiostatic (P) conditions lead to a constant hydrogen surface concentration at the entrance surface (C case), while galvanostatic (G) conditions guarantee a constant flux at the charging side of the membrane. (F case) They disputed that and lent support to the view that a produced transient is possible to be in–between the two forms of boundary conditions, as the overall hydrogen entry depends not only on the electrochemical method by which hydrogen is introduced to the membrane, but also on the relative rates of the various steps involved in hydrogen evolution at the entrance face and its diffusion through the

membrane. Moreover, since the C boundary condition necessitates a small fraction of electrolytically generated hydrogen to enter the membrane, whilst the F boundary condition requires that all of it enters the membrane, then the two boundary conditions are unlikely to be both applicable to the same membrane. Undeniably, if generated hydrogen divides rather uniformly between metal and solution, neither boundary condition will apply strictly.

Allcock [69] compared experimentally determined permeation transients, obtained under galvanostatic conditions, with theoretically generated transients. (Figure 21) He found that the experimental transient shape was in very close agreement with the theoretical transients for constant surface coverage instead of the anticipated constant flux model.

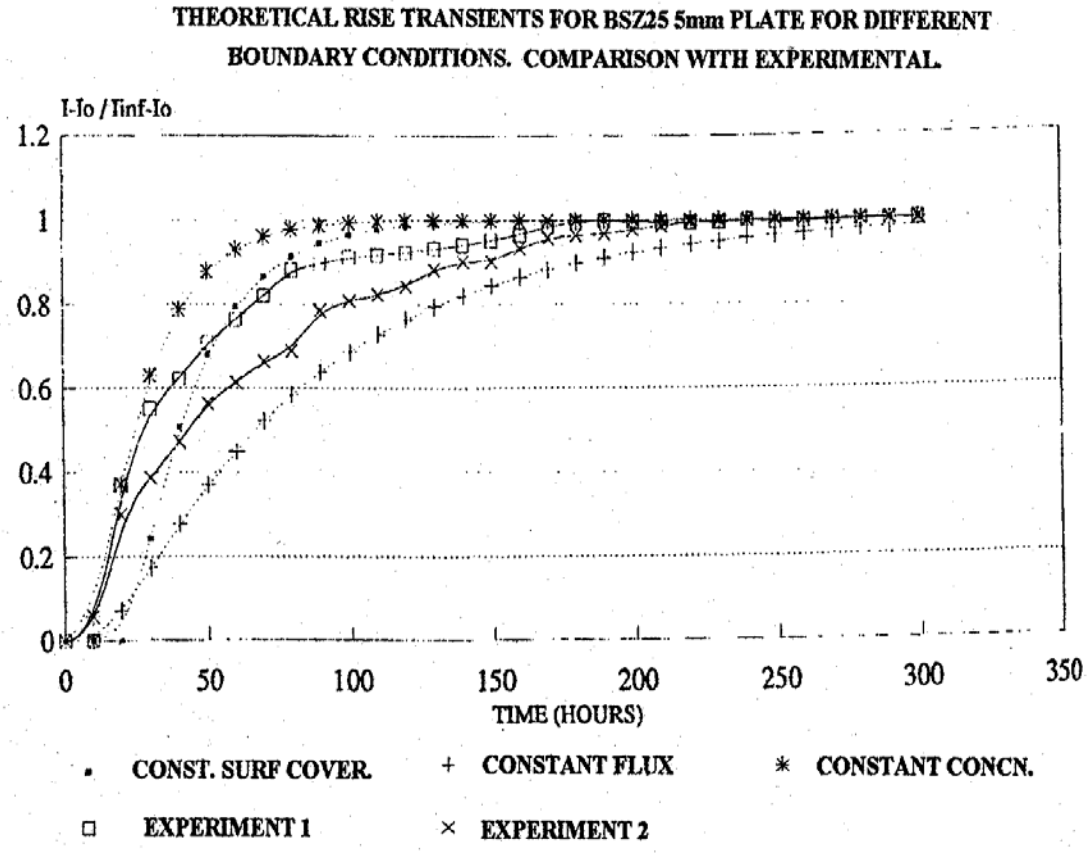

Figure 21. Comparison of theoretical and experimental transients [69]

A constant surface coverage is predicted by the IPZ model [59–61], which takes into account the hydrogen permeation into the metal on the hydrogen evolution reaction, and where surface coverage is considered, along with rate constants for discharge, adsorption, absorption and recombination. These constants can be experimentally estimated from the steady state hydrogen permeation current, obtained at various cathodic charging currents, using a series of interconnected equations developed by Iyer, Pickering and Zamanzadeh.

Some of the conditions of this model are: 1) Hydrogen discharge involves only a single electron transfer reaction and hydrogen evolution occurs by the chemical recombination reaction. It is also assumed that $\eta >> RT/F$ and hence, the recombination step is not rate determining so that the backward reaction of hydrogen oxidation can be neglected. 2) The intermediate adsorption–absorption reaction is in local equilibrium. 3) $i_c = i_\infty + i_r$ where i_c is the charging current density, i_r the steady state desorption flux or hydrogen recombination current density and i_∞ the steady state hydrogen permeation flux. From the above, only steady state conditions are considered, as by the end of the transient stage of hydrogen permeation all the traps in the metal will be saturated. The assumption of steady state allows calculations of the hydrogen recombination current density from the measured i_c and i_∞ values in hydrogen permeation experiments.

When a graph of i_∞ against $\sqrt{i_r}$ is plotted to find a linear relationship, then hydrogen partition determination is possible. Otherwise, an absence of linearity would be an indication that steady state was not achieved in the experiments or that the backward reaction of the discharge step is significant or that the discharge–recombination mechanism is not operating. An alternative way to check the previously mentioned assumptions is the linearity between the variable $f_{i\eta}$ against i_∞ where

Equation 20:

$$f_{i\eta} = i_c e^{a\alpha\eta}$$

η hydrogen overpotential

a $= F/RT$

α hydrogen evolution reaction transfer coefficient

The hydrogen adsorption and absorption rate constants describe the surface and subsurface kinetics for the metal electrode. Based on permeation data for different membrane thicknesses a quantitative description of the surface hydrogen coverage can be provided using Equation 21,

Equation 21:

$$\frac{1}{k''} = \frac{k_{ads}}{k_{abs}} + \frac{D_1}{k_{abs}} \frac{1}{L}$$

where k'' is the equilibrium absorption–adsorption constant.

Should the plotted graph 1/k'' against 1/L be linear, then the rates of absorption and adsorption can be determined.

At this point it should be noticed that all the discussion about hydrogen permeation so far is valid for constant thickness, composition and diffusion coefficient. However, this is not the case during electroplating, as fresh deposit is added to the thickness of the membrane, hence complicating the mathematics of diffusion. However, if hydrogen is considered to be supplied to the steel substrate by the electrodeposit, a constant hydrogen surface concentration can be approximated for short periods of time, allowing estimation of the hydrogen concentration at the electrodeposit-steel surface. The previously stated remarks are cited for completion concerning the hydrogen entry and permeation problems, but they are not further developed, as no electroplating procedure was involved in the experimental part of this research work.

1.6 Trapping of Hydrogen in Steels

1.6.1 Introduction

All real solids are not perfect crystals, for they contain defects, which can play a crucial role in the uptake and transport of hydrogen in metallic materials and their subsequent embrittlement. These microstructural features usually attract hydrogen and are characterised as traps. [69, 70] Traps are termed by Bernstein and Pressouyre [71] as localised regions, where a hydrogen atom spends more time than in a normal interstitial lattice site. Typical traps for hydrogen include vacancies, voids, crack tips, grain boundaries, carbide interfaces, inclusions and dislocations. Solute atoms can also act as traps with those to the left of iron in the periodic table forming attractive traps, whereas atoms to the right of iron causing repulsion.

Hydrogen traps are sites in the material's bulk, which are identified by the deepening of the associated potential well. The consequence of trapping is a decrease in the rate of transport of hydrogen through the metal because there is a finite probability for the hydrogen atoms to jump into trap sites and because the residence time in the trapping site is longer than in a normal lattice diffusion site. [72] Moreover, hydrogen trapping affects some fundamental parameters for the understanding of hydrogen transport in steel, such as the hydrogen distribution in the microstructure, the hydrogen diffusion coefficient, the quantity of hydrogen desorbed from the steel at room temperature and the hydrogen embrittlement susceptibility of a material. [69]

Two clearly different kinds of traps are present to a varying extent in materials; these are reversible and irreversible traps. Reversible traps are those, at which hydrogen atoms have a limited residence time at a specific temperature, associated with low interaction energy. On the contrary, at irreversible traps hydrogen has, for the same temperature, a negligible probability of being released from the trapping site. [11, 69]

Figure 22 [11, 69, 73] shows pictorially the previously distinguished types of traps in terms of their energy levels. Hydrogen atoms reside at the energy levels of either interstitial sites, A, or trapping sites, B. The evolution rate of hydrogen from a trapping site is a function of the trap activation energy Et, whilst En and Es are the

energy levels for absorption from the surface to the bulk and from the bulk to a trapping site, respectively. The potential well of a trapping site is represented by the energy Eb, the magnitude of which determines the nature of a trapping site. If Et >> Eb, then the trap is characterised as irreversible.

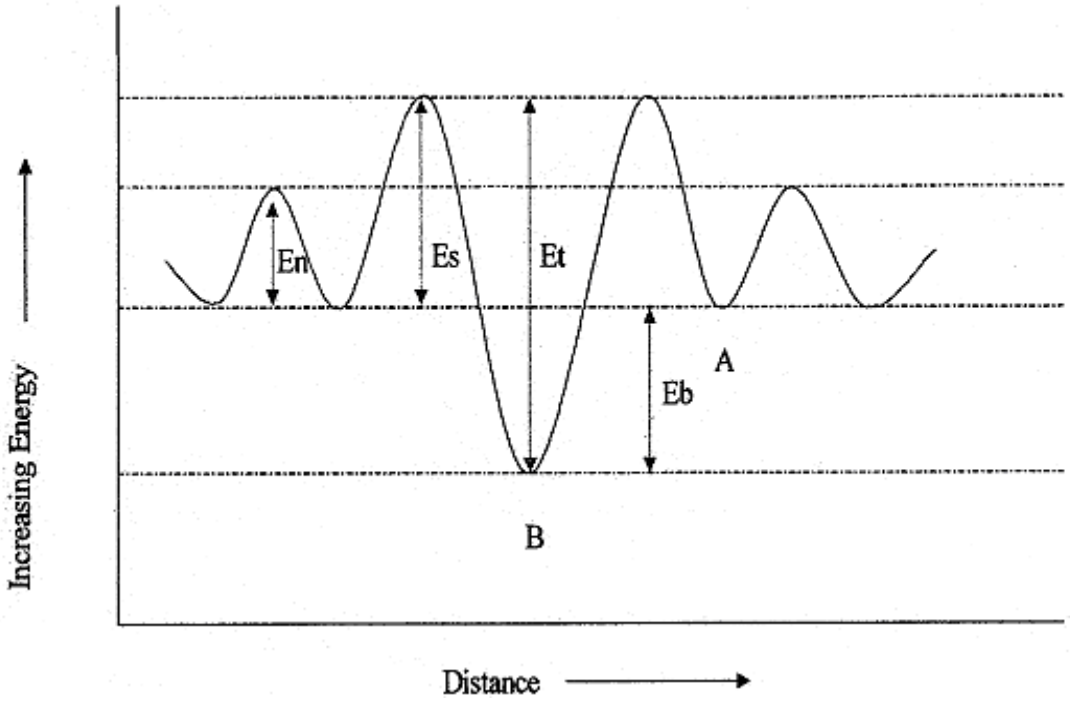

Figure 22. Trapping site model showing hydrogen energy levels in the trap vicinity [11, 69, 73]

Measurements of hydrogen solubility at high pressures and discrepancies between diffusion theoretical and experimental studies at low temperatures led to a more detailed examination of the effect of trapping on the hydrogen diffusion coefficient. Several investigators [74–76] examined trapping from a theoretical point of view and considered the influence of different types of traps in different lattice sites on the hydrogen diffusion coefficient in steel.

McNabb and Foster [74] recognised that the diffusion coefficient D is not only a function of temperature, but also depends on some other neglected variables, in some way connected with the work hardening experienced by the specimen. Therefore, they doubted about the validity of Fick's laws and described solutions of the diffusion

equation for simple geometrical shapes taking into account the effect of trapping. The value for the apparent diffusion constant for thick specimens is given in Equation 22,

Equation 22:

$$D_{app} = \frac{D_{latt}}{1 + \frac{Nk}{p}}$$

where D_{latt} is the trap free lattice diffusion coefficient, N the number of trap sites per unit volume of material, p the release rate parameter which is dependent on the temperature and nature of the trap, but independent of the local concentration of trapped and diffusing hydrogen, and k is the trapping parameter.

Oriani [75] has introduced a limiting solution of McNabb and Foster's model for the case of rapid local equilibrium between trapped and mobile populations of hydrogen atoms (or trapping and normal diffusion sites). He concluded that the interfaces, as traps, are more important for hydrogen trapping in non-cold-worked steels, although dislocations provide trapping sites. Finally, Oriani's expression for the apparent hydrogen diffusion coefficient is given in the following equation,

Equation 23:

$$D_{app} = D_{latt} \frac{C_L}{C_L + C_T(1 - \theta_T)}$$

where C_L is the lattice concentration of hydrogen, C_T is the hydrogen concentration in trapping sites and θ_T is the fraction of occupied trapping sites.

Koiwa [76] contested Oriani's assumption for local equilibrium between the mobile and trapped populations of hydrogen, which was based on a constancy of the activation energy for diffusion right up to the trapping site. His contention was that, since diffusion controls the rate of approach to equilibrium, then the apparent diffusion coefficient cannot be derived from a hypothesised equilibrium condition. Therefore, Koiwa considered an expression for the apparent diffusion coefficient based on a model, which takes into account a change in the activation energy for

diffusion near traps, ΔE. If $\Delta E = 0$, then his expression coincides with Oriani's. Finally, Koiwa examined differences in trapping and escape parameters for tetrahedral or octahedral site occupancy in a body centred cubic (bcc) lattice.

In the previous models there was no referred distinction between reversible and irreversible traps. The first who analysed the combined effects of reversible and irreversible trapping were Leblond and Dubois [77], although the obtained solutions were only for low fractional occupancy of reversible trap sites. They included in their model the effect of stresses and temperature, as well as the possibility of non–uniform solubilities.

Turnbull et al. [72] proposed an extension of the previous model by incorporating the combined effects of reversible and irreversible trapping, for varying degrees of occupancy, on hydrogen transport in metals. Their theoretical analysis was accompanied by repetitive electrochemical permeation measurements to quantify the diffusion and trapping parameters associated with a 13% Cr martensitic stainless steel used for oil production tubing in low H_2S containing environments. One of the recommendations of their analysis is to avoid the use of an effective diffusion coefficient as a quantitative parameter for the characterisation of hydrogen transport in systems, where irreversible trapping preponderates. Citation of the development of their mathematical diffusion model, as well as the involved trapping parameters, is beyond the scope of this work.

1.6.2 Trapping Sites in Materials Permeable to Hydrogen

Accumulation in Voids

Luppo and Ovejero–Garcia [78] studied the influence of the microstructure on the trapping and diffusion of hydrogen in a low carbon steel. In that attempt, they deliberately produced quenched vacancies, where they observed hydrogen accumulation. Using permeation method (electrochemical and gas charging), mechanical tests and hydrogen microprint technique (or silver decoration technique), they demonstrated that trapping in microvoids retarded hydrogen diffusion. They concluded that at room temperature, the predominant trapping process is low energy

trapping at vacancies, dislocations and perhaps lath interfaces, as in the fresh martensite, which displays the lowest hydrogen diffusivity implying a trapping process.

Von Ellerbrock, Vibrans and Stüwe [79] assumed that cavities or voids containing precipitated hydrogen in annealed or deformed iron or steel could be the cause for the substantial deviation of hydrogen diffusivity and solubility at lower temperatures compared to values extrapolated from high temperature behaviour, as reversible trapping in microvoids can occur. Moreover, they used published information of outgassing experiments performed at room temperature from several investigators, to fit experimental data that they obtained by calculating the release rate of hydrogen.

Kumnick and Johnson [80] performed electrochemical permeation measurements to study the effect of deformation and annealing on hydrogen transport through Armco iron. From hydrogen permeation transients, they concluded that deformation–induced microvoids could not be removed by annealing, explaining the thickness–dependent hydrogen diffusivities observed with annealed iron membranes that were previously mechanically deformed. On the other hand, iron membranes, which were chemically reduced in thickness prior to annealing, exhibited higher diffusivities compared to the previously referred procedure.

Accumulation in High Strain Fields

Numerous theories have lent support to the view that hydrogen preferentially segregates into regions of high triaxial stress and hence assists in crack growth. Cottrell and Bilby [81], in an attempt to account for a proposed theory of yielding and strain ageing of iron, they considered how interstitial atoms (carbon in their case) could interact with strain fields around dislocations and concluded that atmospheres of interstitials could strongly interact with edge dislocations, as they relieve hydrostatic stresses by entering the expanded region below the dislocation centre.

Cochard, Schoek and Wiedersich [82] calculated the interaction energies of impurity atoms (carbon interstitials) with the strain fields about both edge and screw dislocations in bcc metals. It was not exactly known how the iron lattice is distorted

round a carbon atom, and hence, for simplification reasons, it was assumed that a unit cell containing a carbon atom is deformed as in martensite.

Grain Boundary Traps

It has long been acknowledged that grain boundaries are sinks for impurities and that hydrogen can be absorbed on the interfaces between precipitates and the metal lattice. Podgurski and Oriani [83] studied the nitrogenation of Fe–Al alloys and showed that hydrogen is reversibly absorbed to the surface of aluminium nitride particles at room temperatures, but only at sites not pre-empted by preadsorbed nitrogen. At temperatures higher than 500^{o}C AlN–ferrite interface reversibly absorbs nitrogen that can block sites to later hydrogen absorption. Other examples of hydrogen absorption at interfaces between iron and metallic carbides or sulphide inclusions in steels are incorporated in the Bernstein and Thompson's review [84] on the effect of metallurgical variables on environmental fracture of steels.

Dislocation Trapping

Kumnick and Johnson [80] found also other types of traps, apart from the earlier stated microvoids, for the deformed Armco iron, which could anneal in the recovery range of temperature between 100 and 400^{o}C. These traps were explained on the basis of the cellular dislocation structure caused by deformation.

Donovan [85] investigated the accelerated evolution of hydrogen from various metals, including Armco iron and nickel, during their plastic deformation. He used tritium, a radioactive tracer, in his experiments, so that low level releases can be determined. The release rate was found to be strain–dependent with a rapid increase when plastic deformation begins. The enhanced release of hydrogen during plastic deformation was attributed to hydrogen segregation to and movement with, dislocations going out of the specimen.

Generally, the most affected microstructural defect by cold working is the dislocation density, which increases as a rapid multiplication of dislocations occurs with increasing plastic strain. In this context, it is not coincidental that the main trapping

site of hydrogen for cold worked AISI 4340 steel was found to be dislocations. Similarly, Huang and Shaw [86] found that cold work decreased hydrogen diffusivity and increased hydrogen absorption for hot–rolled type 1020 steel, as cold work increased dislocation density and hydrogen trapping.

Finally, Iino [32] refers to the interaction of dissolved hydrogen with defects, such as point, line and surface defects, and he also mentions the role of dislocations in hydrogen trapping. Dislocations are not only trapping sites, but there are also moving dislocations carrying hydrogen atmospheres that are likely to meet the previously cited defects and exchange hydrogen according to the relative strength of hydrogen–defects interactions, the binding energy of which must be precisely known.

Impurity Trapping

Iino [87] described precipitates or second phases in steel as trap sites that possibly control hydrogen induced embrittlement, and he attempted to describe the relationship between hydrogen and substitutional foreign atoms in terms of binding energies. This was feasible, as the thermodynamics and kinetics of hydrogen bound to impurities are easier to specify thanks to the simpler geometry of impurity trapping compared to other traps, such as voids, dislocations, boundaries etc., where the geometry cannot be specified completely.

1.6.3 Trap Ranking

From the aforementioned types of traps, dislocation and grain boundaries traps can be characterised as reversible traps with similar hydrogen binding energies, while the other stronger traps can be considered as irreversible. [72] Reversible traps are more significant for hydrogen embrittlement susceptibility, as hydrogen can diffuse from these sites to defect sites within the material, while irreversibly trapped hydrogen cannot. The connection between hydrogen trapping and embrittlement is developed in the following paragraph.

1.6.4 Effect of Trapping on Embrittlement

Generally, trap type and its efficiency to contain hydrogen are considered to be mainly a microstructural factor, which is related to the observed embrittlement of a metallic system. [11] Luppo and Ovejero–Garcia concluded that there is a direct relationship between HE susceptibility and the quantity of desorbed hydrogen, suggesting that diffusible hydrogen causes embrittlement but strongly trapped does not. [78] In other words reversibly trapped hydrogen is more damaging than that which is held irreversibly. Moreover, when hydrogen accumulates to reversible traps, it is able to migrate to a neighbouring incipient crack initiation site in sufficient quantities for embrittlement to occur. The above are illustrated in Figure 23, which compares the amounts of hydrogen released during outgassing from precharged low carbon steel. [78] The greatest quantity of released hydrogen was observed for as quenched martensite and increasing tempering temperature decreased the released hydrogen, displaying the same trend as for embrittlement susceptibility.

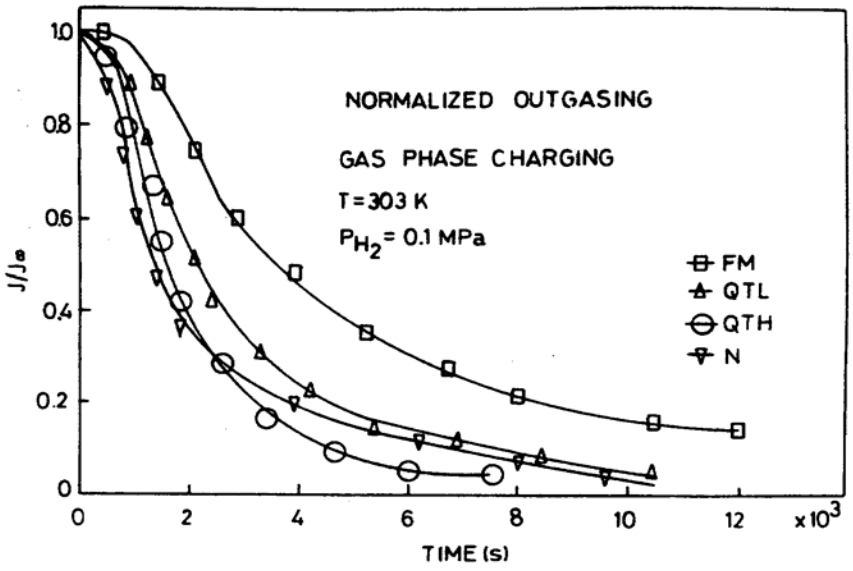

Figure 23. Comparison of the amounts of hydrogen released from different microstructures in 0.12% C steel following gas phase charging. (N) normalised, (FM) fresh martensite, (QTL) quenched and tempered at low temperature, (QTH) quenched and tempered at high temperature [78]

Pressouyre and Bernstein [88], based on work with a Fe–Ti–C system, proposed a model which assumes that hydrogen induced cracks would initiate on a given defect once enough hydrogen has been trapped on this location. In other words, embrittlement will occur when the quantity of trapped hydrogen at a given microstructural defect exceeds a critical amount for the material. This critical concentration probably corresponds to a sufficient hydrogen pressure and/or adequate induced or applied stresses to overcome the cohesive strength. In addition, they classified traps as good or bad, according to their ability to increase the time to reach a critical concentration or to scatter the hydrogen so that a critical concentration is not reached.

Tiwari et al. [89] agreed with the critical concentration concept, but in a study of internal HE for mild and maraging steels, they expressed the view that some critical stress concentration besides a critical hydrogen concentration must be achieved at some trapping sites in order to nucleate a void for the ductile failure. In their study, such stress concentration was achieved primarily at the particle–matrix interface by dislocation pile up.

As earlier stated, Pressouyre and Bernstein's model was examined for the Fe–Ti–C system. Ti acts as a reversible trap when it is the sole solute atom in the iron lattice, while combined with carbon to form TiC it behaves as an irreversible trap and is considered to be more beneficial to the control of embrittlement. 1.5% Ti additions to iron were shown to cause trapping that both diminished the rate of intergranular cracking and lowered the hydrogen permeation by an order of magnitude. Strong interactions between hydrogen and Ti were responsible for the reduction of the available amount of hydrogen for crack nucleation and propagation, as they acted as innocuous traps.

One of the most substantial conclusions drawn for the influence of trapping on the embrittlement phenomenon is that not all traps are potential sites for crack initiation. The desired properties for traps, so as to be beneficial in the control of embrittlement for a material, are the following [70]:

(a) A high critical amount of hydrogen to instigate crack initiation.
(b) They should be adequately plentiful so that they occupy the hydrogen that would otherwise go to more dangerous traps.
(c) They should be irreversible so that they do not release their hydrogen and accelerate initiation.
(d) They should rather be homogeneously distributed in the matrix.

The previously stated parameters must be considered when manipulating the microstructure to create a fine dispersion of trapping sites within a metal, so as to improve the hydrogen embrittlement performance of a particular steel. Applying de–embrittlement baking treatment to hydrogenated specimens, in order to help the steel regain its non–hydrogenated properties is a basic idea to remove as much potentially mobile hydrogen from steel as economically viable in a set time. [11] However, the restoring of non–hydrogenated properties is not guaranteed, as there are cases, where the retained hydrogen at high concentrations in the austenite phase does not totally diffuse from that phase during the de–embrittlement (or outgassing) treatment, because of the low diffusivity and high solubility of hydrogen in γ phase. Consequently, the austenite acts as a hydrogen reservoir with hydrogen slowly diffusing from the austenite to the ferrite when the material is under stress, causing a catastrophic failure. [90]

Robinson and Sharp [91] showed in another study that baking dramatically increased the embrittlement susceptibility of a 0.8% C steel wire. This unexpected finding was accounted for on the premise that temperature affects reversible and irreversible hydrogen trapping. Baking is believed, firstly, to provide sufficient thermal energy for the innocuous traps to become reversible and, secondly, to release hydrogen from relatively low energy innocuous traps and to allow it to concentrate in traps with higher interaction energies. These sites act as significant stress concentrators, upon reloading, leading to more rapid failure.

1.7 Corrosion

1.7.1 Introduction

Corrosion is the spontaneous reaction between a metal and its environment, where that environment may be a solid, liquid or gas. [92, 93] The corrosion product is usually more stable than the metal that is approaching thermodynamic equilibrium. The study of corrosion is clearly divided into aqueous corrosion and oxidation. [92] In aqueous environments corrosion proceeds by an electrochemical mechanism involving the anodic reaction of the metal oxidation (Equation 24) and the simultaneous cathodic reaction of the water reduction (Equation 25) that causes hydrogen evolution. It is this source of hydrogen that is responsible for the re–embrittlement of steel substrates as a consequence of the corrosion of protective metallic coatings, as viewed in the introduction. (Figure 3) In the case of aerated conditions another cathodic reaction takes place, referred to as the oxygen reduction. (Equation 25)

Equation 24:

$Fe \rightarrow Fe^{2+} + 2e$

Equation 25:

$H_20 + e \rightarrow H + OH^-$

$O_2 + 2H_20 + 4e \rightarrow 4OH^-$ (aerated conditions)

Corrosion, as any chemical or electrochemical reaction, is a phenomenon influenced by thermodynamic, kinetic and physical factors. From a thermodynamic point of view, the tendency for a metal to corrode is expressed by the electromotive force (emf) of a corrosion cell. The emf is related to the thermodynamic quantity of the Gibbs free energy through Equation 26, where ΔG is the Gibbs free energy change, n the number of electrons participating in the reaction, F the Faraday constant, and E the electrode potential. [94]

Equation 26:

$\Delta G = - n F E$

If ΔG is positive, it can then be stated with certainty that corrosion will not occur under the described conditions. However, a negative ΔG will demonstrate the tendency of corrosion to occur without indicating its rate.

In Nernst equation (Equation 27) the emf of a cell is defined in terms of the activities of reactants and products, as well as the temperature dependence of the system. [94]

Equation 27:

$$E = E^{o} - \frac{RT}{nF} \ln\left[\frac{a}{b}\right]$$

where:

E electrode potential

E^{o} standard electrode potential (when all the activities of reactants and products are equal to unity)

R gas constant

T temperature (in K)

n number of electrons participating in the electrochemical reaction

F Faraday constant

a chemical activity of products

b chemical activity of reactants

Based on thermodynamic data, Pourbaix developed potential–pH diagrams that correlate the electrochemical and corrosion behaviour of a metal in water, and demonstrate the reaction state of a metal at a particular combination of pH and potential. Pourbaix diagrams do not give any corrosion rates, and assume a single pH value throughout an electrochemical system, which is not true. They are considered as a useful research tool, but they cannot on their own solve the problem of corrosion prediction. For a more reliable corrosion prediction the galvanic series (Figure 2) are considered as a more useful corrosion guide. In Figure 24 is given the Pourbaix diagram for the iron–water system at 25^{o}C and considering as solid substances only Fe, Fe_3O_4 and Fe_2O_3. [95]

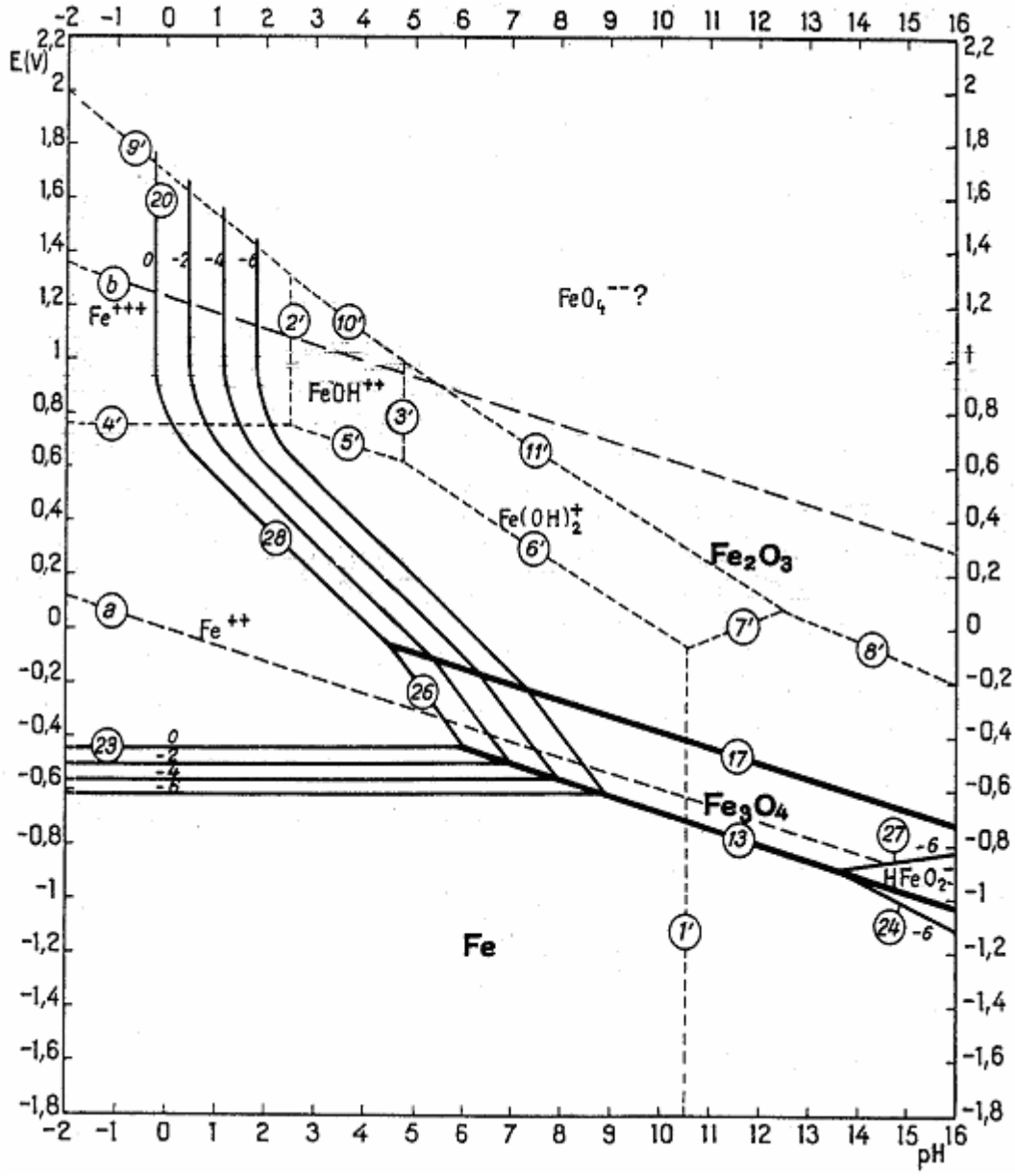

Figure 24. Potential–pH equilibrium diagram for the iron–water system at 25°C considering as solid substances only Fe, Fe_3O_4 and Fe_2O_3. [95]

1.7.2 Galvanic Corrosion

A form of aqueous corrosion that is a concern in this work is galvanic corrosion, which occurs when metals with different electrochemical potentials are in direct electrical contact in a conducting electrolyte. [96] The one being lower in the galvanic series (the more active) will act as the anode, whilst the other (the more noble) will act as the cathode. (Figure 2) In this useful guide, the galvanic series [8], metals and

alloys are arranged in a list according to their relative open–circuit potentials in a particular environment. Moreover, the several states in which a substance can appear are taken into account. As a result, comparison of metals in a specific environment can direct to the correct material selection for a certain task.

However, the use of galvanic series provides only qualitative information about the direction of galvanic corrosion in a specific environment, and it should not be used for quantitative predictions of galvanic corrosion rates. The reason is that there are more factors that affect galvanic corrosion apart from the corrosion potentials. For example, anode to cathode area ratio, electrolyte conductivity, distance between coupled metals, shielding of metal surfaces by marine growth and sediments, localised electrolyte concentration changes in shielded areas and polarisation characteristics of the metals involved have their effect on the eventually observed galvanic corrosion rate. [8]

Galvanic corrosion will accelerate the corrosion rate of the more active (more electronegative) metal with a simultaneous deceleration, but not elimination, of the corrosion rate of the more noble (more electropositive) metal. [8, 97] Figure 25 depicts the corrosion behaviour of galvanically coupled metals A and B in the case of charge transfer control, i.e. when there is no control of the electrochemical phenomena by any diffusion mechanism. As seen, when two different corroding metals are coupled electrically in the same electrolyte, both metals are polarised so that each corrodes at a new rate. When the metals A and B are uncoupled, they corrode at the corrosion rates $i_{corr,A}$ and $i_{corr,B}$, respectively. When equal areas of metals A and B are coupled, the resultant mixed potential of the system E_{AB} is at the intersection, where the total oxidation rate equals the total reduction rate. The rate of oxidation of the individual coupled metals is such that metal A corrodes at a reduced rate i'_A, and metal B corrodes at an increased rate i'_B.

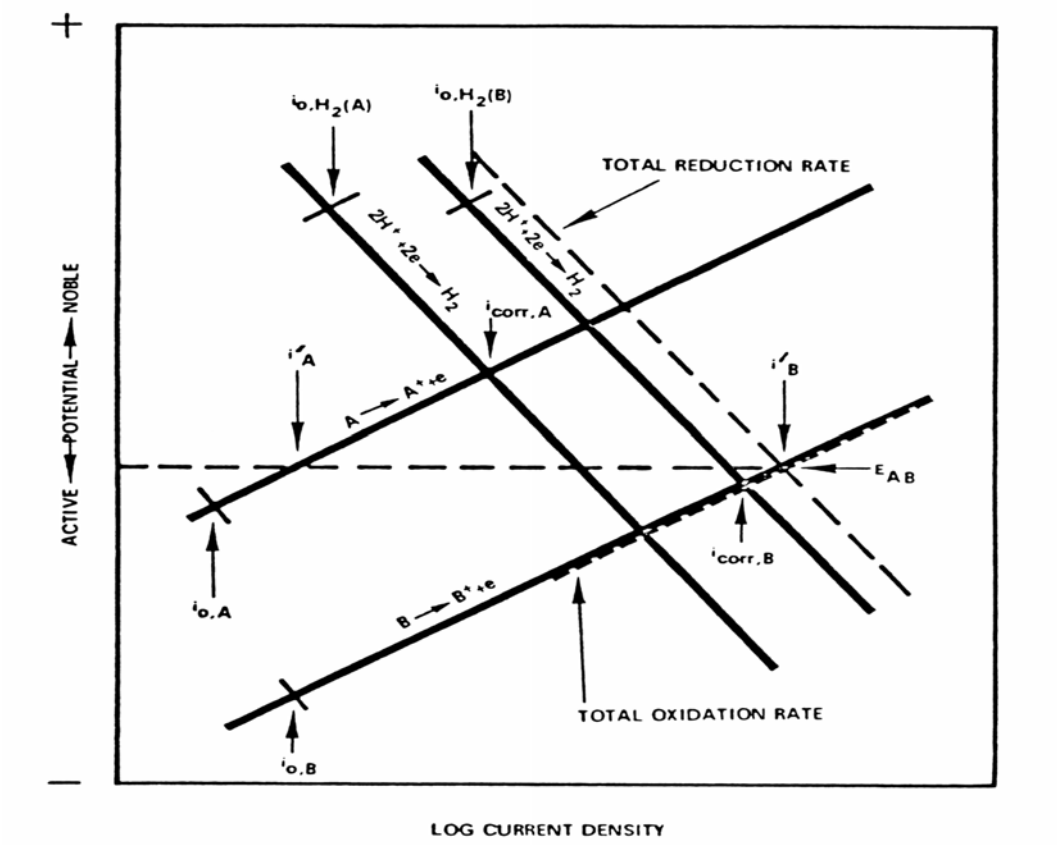

Figure 25. Galvanic coupling of metals A and B in the case of charge transfer control. Depicted corrosion rates before ($i_{corr,A}$ and $i_{corr,B}$) and after coupling. (i'_A and i'_B) [8]

Figure 26 [9] displays the effects of the polarisation of the metals of a couple on galvanic corrosion rate. The coupled potential and the galvanic corrosion rate will depend on the type of control, which could be mixed, anodic or cathodic. If there is mixed control, then the two metals are equally polarised, and the coupled potential lies in the middle of the open–circuit potentials of the two metals. If there is cathodic control, polarisation occurs mostly at the cathode, and as a result the coupled potential lies closer to the potential of the more active metal (anode). Finally, if there is anodic control, polarisation occurs mostly at the anode, and thus, the coupled potential lies closer to the potential of the more noble metal (cathode). [94]

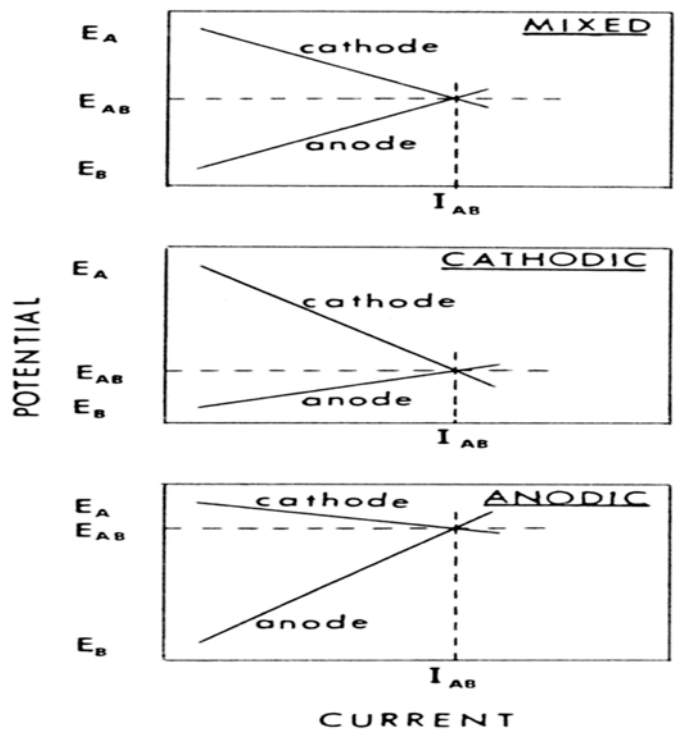

Figure 26. Different types of polarisation characteristics of the metals involved in a galvanic couple [9]

1.7.3 Corrosion Control

Cathodic Protection

Although the corrosion process is in most cases inevitable and its prevention difficult, its control is feasible. [93] The basic principle in the protection of a metallic structure from corrosion is to find a way to stop its oxidation, i.e. its loss of electrons. Bearing in mind Le Chatelier's principle in Equations 24 and 25, if electrons are supplied to the metal, the electron consumption of the cathodic reaction will accelerate while the electron release of the anodic reaction will slow down. Consequently, the hydrogen evolution rate will increase, while the iron dissolution will slow down and the electrode potential will fall. This is the principle of cathodic protection which is a means of electrochemical control upon corrosion.

Cathodic protection can be implemented by two methods, through power impressed current or sacrificial anodes. In the first case the protected structure is made

electrically negative so that it acts as a cathode, whilst a second electrode is made electrically positive and completes the circuit as an auxiliary anode. The structure to be protected and the auxiliary anode should be both in electronic and electrolytic contact for the cathodic protection to be attained.

In the second case two dissimilar metals are electrically connected in an electrolyte solution, and a current flows from the more electronegative metal, which becomes the anode, to the metal with the more electropositive potential, which becomes the cathode. The current flowing between the two metals accelerates the dissolution of the anode, which is therefore sacrificed. Hence, in order to apply cathodic protection through sacrificial anodes, the anode must possess a potential that is more electronegative than that of the protected structure. When connected, the structure is polarised cathodically and the sacrificial anode is polarised anodically by both reaching the same potential E_{SC} providing the resistance of the electrolyte R is sufficiently low. The ensuing galvanic current is the short–circuit current I_{SC}. (Figure 27) [98]

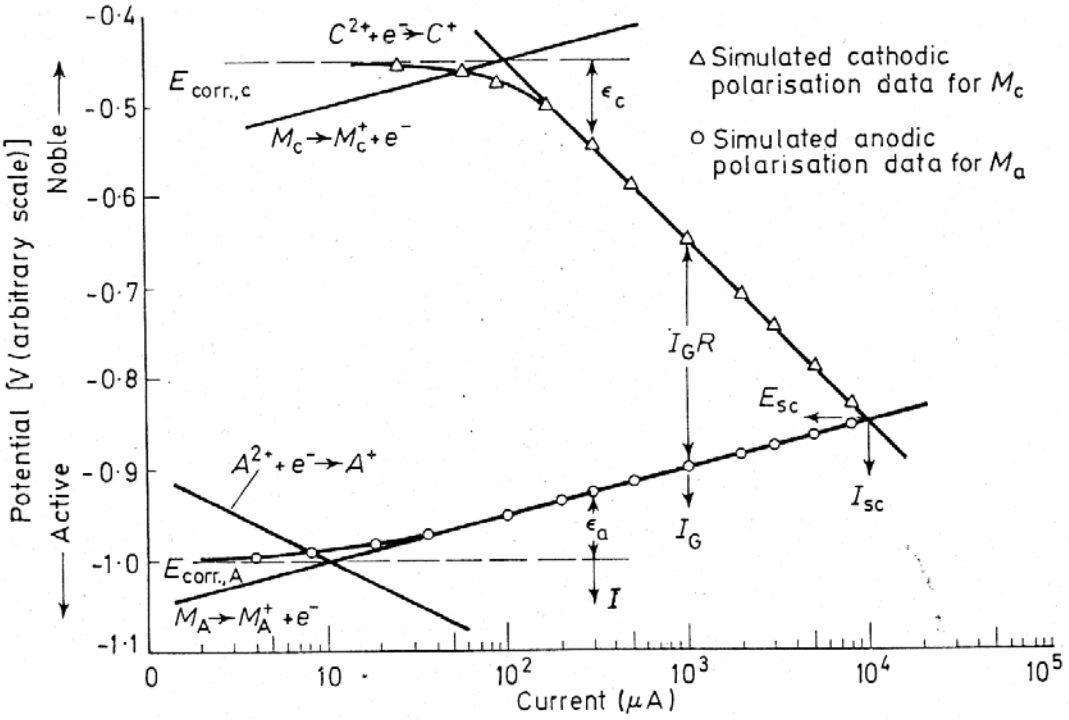

Figure 27. Polarisation in a galvanic couple showing the behaviour of the sacrificial anode in cathodic protection. [98]

Cathodic protection is applicable either by each of the previously described methods separately or by combining both in certain situations. On the other hand, it should be emphasised that cathodic protection has to be applied with care, in order to balance its protection effectiveness against the deleterious effects of hydrogen embrittlement as a result of hydrogen generation caused by overprotection in very negative potentials. Finally, another means of electrochemical corrosion control is anodic protection which is cited merely for completion, although it is irrelevant to the objective of this study. Anodic protection uses a low level impressed current to maintain the passive condition of a metal. Careful planning and design of the system it will be applied is of paramount importance, as it is often more complicated to implement and monitor. [11]

Coatings

Other systems of corrosion protection include metal alloying e.g. stainless steel alloyed with chromium, covering with an impervious layer of enamel, ceramic, rubber or paint, and coating or cladding with a material that will preferentially corrode e.g. zinc on iron as shown on the galvanic series. [11] In the light of the current work, non–corrosion resisting steels are protected by coatings that will preferentially corrode if the substrate to be protected, i.e. the steel, is exposed to a corrosive environment. Thus, it is not coincidental that cadmium or zinc coatings are selected for the cathodic protection of steel, as they are more electronegative in the galvanic series and will act as sacrificial anodes.

Coatings can be applied to a substrate material by a wide range of methods. The application of coatings follows a pretreatment on the substrate, and this procedure could be mechanical, chemical or electrochemical, in order to remove existing natural oxide films, corrosion products, greases or oils. Methods of coating application include among other electroplating, hot dipping, diffusion, metal spraying and physical vapour deposition. [98]

When investigating suitable systems for an effective corrosion protection on a material, two coating properties should be considered, and these are the barrier and sacrificial properties. The barrier properties of a coating underline its ability to

successfully isolate the substrate from the aggressive environment, and it is this particular coating property that will initially offer corrosion protection on a steel substrate. For a coating which forms a perfect barrier layer, i.e. is completely pore–free and level in nature, the time for red rust formation on the steel substrate will be directly proportional to the self–corrosion rate of the coating, which can be estimated by performing corrosion rate measurements on detached coatings. [10]

If for any reason the coating is penetrated, then its sacrificial properties will determine the effectiveness of corrosion protection on the steel substrate. In case that a more noble coating is applied to the steel substrate, then small areas of the more active steel substrate will be exposed to the corrosive environment, and corrosion will be rapid as a result of high galvanic current densities produced on small anodic areas. In contrast, if gaps occur in a more active coating, then the substrate can remain cathodically protected, as the more active coating will preferentially corrode. [92] Both previously referred cases are shown pictorially in Figure 28. Finally, the sacrificial corrosion properties can be investigated by measuring the galvanic current flow between coating and substrate material. [10]

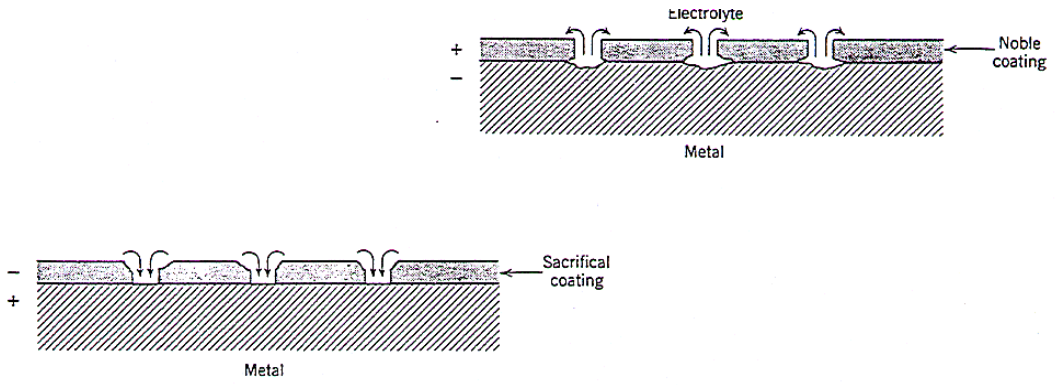

Figure 28. Effectiveness of corrosion protection on a metal exposed to an electrolyte because of coating defects. Active coating offers sacrificial corrosion protection to the metal substrate, while noble coating does not prevent its corrosion. [94]

1.8 Protection Requirements for Aerospace Applications

1.8.1 Corrosion of Aircraft

Corrosion of aircraft structure usually occurs when protective coatings are damaged or succumb to the action of a hostile environment. According to the map of global corrosivity, depicted in Figure 29, locations near the sea are particularly severe environments. Moreover, changes in atmospheric temperature and pressure have a significant effect on corrosivity, as aircraft may experience various climates between their initial take off point and their final destination. For example, an aircraft could take off from an area of high ground temperature and humidity, climb through industrial pollutants, experience temperatures below –50°C and land at a hot humid airport. This can result in ideal conditions for the formation of condensation, as well as the accumulation of corrosive salts on skins, crevices and exposed areas of the aircraft. [99, 100]

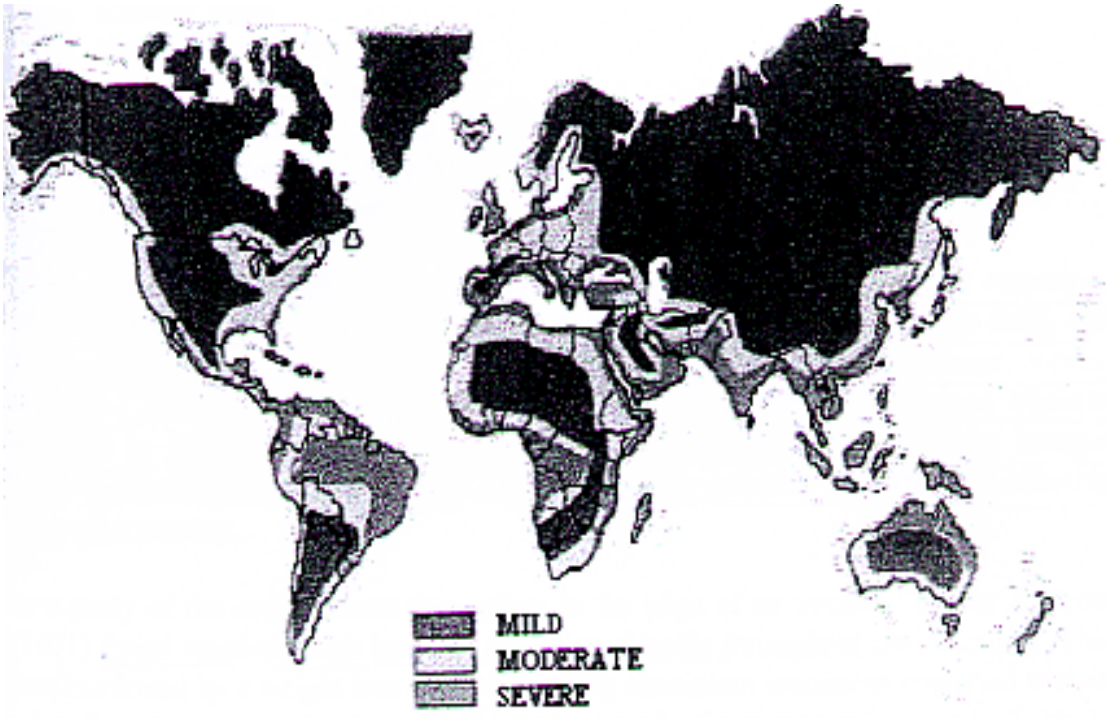

Figure 29. Global corrosivity map showing severity of marine exposure. [99]

Conditions and contaminates that are known to cause corrosion in a base airport and operation environment could incorporate climate and humidity, location relative to salt water, airborne salts and industrial impurities, condensation from passengers and livestock, runway salts and contaminants, as well as sunlight and ozone affecting rubber and plastic materials. [99, 100]

The present study is more focused on aircraft components and fasteners manufactured from non–corrosion resisting high strength steels. These components, as part of the entire aircraft structure, are inevitably experiencing the previously described changes in climatic conditions and environments, and for this reason, particular corrosion protection strategy for these components is required.

1.8.2 Corrosion Protection Requirements

Electroplated cadmium has been the 'traditional' material for the sacrificial corrosion protection of steel in aerospace applications, as rigorous testing is required before an alternative could be accepted. On the other hand, it is more active than steel, but not too electronegative, thus leading to smaller galvanic interaction with steel than zinc with steel. Moreover, cadmium, being more noble than zinc, is less likely to cause re–embrittlement of the steel substrate because of its corrosion. Cadmium is also undoubtedly preferred instead of nickel as a potential sacrificial coating, for nickel is less active than steel leaving steel to preferentially corrode. (Figures 1 and 2) [1, 8]

Moreover, cadmium is also galvanically compatible with aluminium alloys used in aircraft construction, as their potential is similar to that of cadmium and hence, there is little driving force to promote galvanic corrosion between them. Furthermore, other advantages of cadmium, as outlined in the introduction, include the fact that cadmium does not produce voluminous corrosion products, is self–lubricating and has a low coefficient of friction making it ideal for plating threaded components like fasteners. [1, 2] Any embrittling effects of its electroplating process are countered by a subsequent de–embrittlement procedure, during which cadmium plated high strength steel components are baked at a temperature between 190 and 230^{o}C for 24 hours. [2–4]

Smith [2] describes in detail the standard protection scheme for non–corrosion resisting steels on aircraft, according to which the steel substrate is cadmium plated to a thickness up to 15 microns and then chromate passivated to promote paint adhesion. The application of an epoxy primer and a polyurethane or acrylic top–coat completes the protective scheme. (Figure 30)

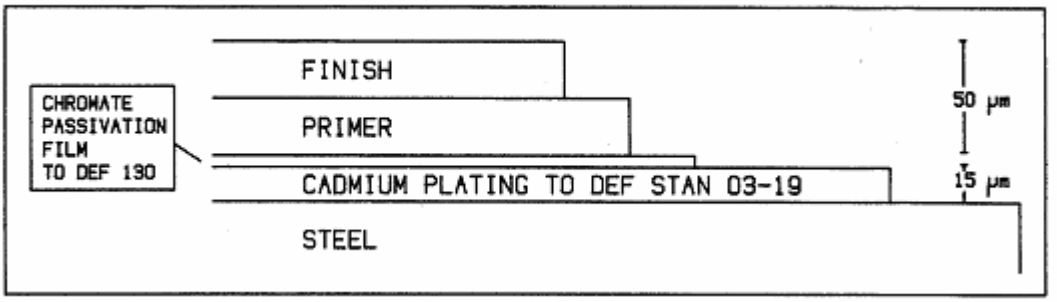

Figure 30. Standard protective treatment used on steel components. [2]

However, cadmium and its corrosion products are toxic and can pose a serious risk to human health. In addition, the effluent from the cadmium electroplating process contains cadmium cyanide in a concentrated NaOH solution, and only low cadmium levels are allowed to be disposed of in the environment. Consequently, there are restrictions on the use of cadmium worldwide, including European legislation, as well as limitations in countries like USA and Japan. [1]

Any replacement candidate coating for the protection of steel aircraft fasteners and undercarriage components must fulfil certain criteria. Firstly, it must be sacrificial to steel, but not too electronegative, as this would lead to a high corrosion rate and a strong galvanic interaction with steel. Secondly, galvanic compatibility with high strength aluminium aircraft alloy is necessary. Finally, it should not pose the risk of hydrogen embrittlement either as a result of its application process or its post–application corrosion (hydrogen re–embrittlement). The re–embrittlement issue should be particularly dealt with, as it is a phenomenon occurring in service, where intervention to protect components in operation from failure is not an easy task.

A wide range of materials has been proposed so far as possible alternatives to cadmium for the needs of the aerospace industry. These candidate coatings include zinc, zinc alloys, aluminium and aluminium alloys. Emphasis was placed in recent studies at Cranfield on zinc alloyed with more noble elements, in order to achieve a slightly more electronegative potential than that of steel, so as to avoid rapid galvanic corrosion of zinc. Iron group elements, like nickel and cobalt were chosen. However, Zn–Ni alloys are believed to react with aircraft fuel, stored in the wings, thus being considered as inappropriate for use on fasteners. On the other hand, Zn–Co electroplating impairs the mechanical properties of high strength steels and following de–embrittlement baking treatments for either 24 or 48 hours at 200^{o}C prove to be inadequate for the full recovery of the high strength steel tensile properties. [11]

1.8.3 Aluminium–Based Coatings

Aluminium–based coatings are another category of materials considered as cadmium substitutes. Manipulating their composition could lead to slightly more active potentials than that of steel. (Figure 2) On the other hand, they are more likely to be galvanically compatible with aircraft aluminium alloys, reducing the risk of galvanic interaction with steel and aluminium alloy components. However, as they are more active than cadmium (Figure 2), they pose a higher risk of re–embrittling the steel substrate when they corrode in service. (Figure 3) For this reason, as earlier stated in the general cadmium replacement requirements, hydrogen re–embrittlement should be carefully examined along with other concerns, such as corrosion behaviour of the coatings and the risk of any hydrogen embrittlement as a result of their application process.

Various methods are employed for the application of aluminium–based coatings. These include electrodeposition from non–aqueous electrolytes as in the case of the Alcotec Galvano–Aluminium [14, 15], paint spraying as in the case of aluminium pigmented metallic–ceramic coatings, e.g. SermeTel CR962 and SermeTel CR984–LT [5], ion vapour deposition (IVD) [16], and unbalanced magnetron sputtering applied in the case of aluminium–magnesium coatings [17]. Other referred application techniques comprise hot dipping, electrophoretic deposition, chemical deposition, cladding or mechanical bonding and casting. [98]

Information about some of the previously cited aluminium–based coatings can be found in the international literature. [5, 15–17] However, most sources refer to issues of galvanic compatibility with steel and aluminium alloys, corrosion performance and hydrogen embrittlement, without raising the problem of hydrogen re–embrittlement for the substrate to be protected.

For example, producers of aluminium pigmented metallic–ceramic coatings describe SermeTel CR984–LT as a suitable substitute of cadmium, for it possesses excellent corrosion resistance according to neutral salt spray testing. Furthermore, it is considered as non–embrittling for the substrate as a result of its paint spraying, and it is compatible with paint systems. [5] However, no information is provided on the detrimental effects of the corrosion of SermeTel CR984–LT on the mechanical properties of steels.

Similarly, Alcotec Galvano–Aluminium is presented by its producers as displaying excellent corrosion performance in neutral salt fog tests, and as not causing any embrittlement to the steel substrates during its electrodeposition process. [14, 15] Again, vital information on the re–embrittling effect of an active coating, such as aluminium, is absent.

In addition, aluminium–magnesium coatings are considered to exhibit the optimum of their corrosion performance at 20% (wt) magnesium content, and that adding more magnesium shifts the Al–Mg alloy potential to more active values. [17] This would pose a higher risk for the re–embrittlement of any steel substrates as a result of the corrosion of the aluminium–based coatings.

Re–embrittlement issues were more emphasised on the work of Hinton et al. [16] concerning IVD aluminium coatings for the corrosion protection of steel. According to this investigation, IVD Al coatings protect steel in aggressive chloride environments, with the degree of corrosion protection depending on the porosity of these coatings. Their corrosion performance can be improved by glass beading and chromate conversion, and a low porosity IVD aluminium coating appears to offer the same amount of protection when performing low HE cadmium plating. However, these investigators underlined the fact that during coating breakdown sufficient

hydrogen can be admitted to the steel and cause embrittlement. This hydrogen can be made available through the growth of an active pit, as well as when the steel is galvanically protected by the aluminium.

From the previously described cases of aluminium–based coatings it has been made clear that the examination of SermeTel CR984–LT and Alcotec Galvano–Aluminium as possible replacements for cadmium needs not only the investigation of their corrosion behaviour, but also the checking of any embrittlement risks to the steel from both their application (HE) and their post–application corrosion (hydrogen re–embrittlement). Based on these concepts, the aims of this research were set in the Introduction with the implemented methods described in the Experimental part of this work.

1.9 Mechanical Testing

There are various mechanical testing methods for the assessment of HE or hydrogen re–embrittlement of metals in various environments. These methods contribute to the observation of the change in mechanical properties of specimens charged with hydrogen. Such properties comprise ductility, elongation, tensile strength, reduction in area or fracture toughness.

Turnbull [101] describes several methods of measuring environmentally assisted cracking, under which he classifies HE. Moreover, differences between these methods and the interrelation between their results are referred to, as well as their advantages and disadvantages. Mechanical testing may incorporate both smooth and pre–cracked specimens depending on the objective of the test and the nature of the application.

The testing of pre–cracked specimens recognises the difficulty of ensuring a complete absence of crack–like defects introduced during either manufacture or subsequent service. In addition, the existence of such defects can cause a susceptibility to stress corrosion cracking which in some materials, like titanium, may not be apparent from tests under constant load on smooth specimens. [101] Furthermore, it has to be underlined that precracked specimens were introduced as a method of controlling crack growth, limiting experimental scatter and separating between the initiation and propagation phases of cracking [11, 102]. Specimen geometry can be used to obtain critical combinations of crack length and external applied stress using a single threshold stress intensity factor K_{th}, below which crack initiation and propagation cannot occur. K_{th} values depend on many variables, one of which is the diffusible lattice hydrogen content that was found by Lucas and Robinson [103] to be inversely related to the threshold stress intensity for hydrogen–assisted cracking of carbon–manganese steel [104]. The testing with pre–cracked specimens will not be developed further, as it is outside the scope of this research, where smooth specimens were used.

Two of the most accepted methods for measuring the extent of hydrogen embrittlement on smooth specimens are the constant load test and the slow strain rate test (SSRT). These techniques allow time for the mobile hydrogen to diffuse within a specimen and hence, the embrittling effect of hydrogen can be detected.

1.9.1 Constant Load Test

The characteristic feature of this test is that the load is maintained constant during the test. The simplest way to provide a constant load consists of a dead weight suspended on one end of a uniaxial specimen, restricting in that case the applicability of this method to wire specimens only. However, lever systems are more suitable for large diameter specimens. [101]

Constant load specimens are usually in the form of a cylindrical bar, wire or C–ring supporting a load equivalent to a fixed percentage (between 50% and 90%) of the breaking load of the specimen. A range of specimen shapes and sizes can be used together with wire, plate, rod, tubing or welded parts. [9, 11]

The time to total failure is used to estimate cracking susceptibility and is commonly evaluated over a range of critical applied stresses to assess the threshold stress below which total failure is not occurring. It is usual to terminate a constant load test after an arbitrary chosen time if failure has not occurred, as such likelihood exists. In that case, the number of cracks per unit length of material can be used for a comparative estimation of susceptibility with respect to the initiation of stress corrosion cracks. Crack velocity can be approximately estimated by the measurement of the length of the largest crack on the fracture surface divided by the time of testing. The assumption in this approach of crack velocity estimation is that crack initiation occurred at the start of the test and that the longest crack also initiated at that time which is not necessarily true, and could be a misleading hypothesis. [101]

1.9.2 Slow Strain Rate Test

Introduction

The slow strain rate method of testing has been used extensively over the past 30 years to examine the effect of stress corrosion cracking and hydrogen embrittlement of high strength steels. Essentially, the method consists of subjecting a tensile specimen to an increasing strain at a rate that is slow enough to allow hydrogen to diffuse to flaws within the steel microstructure, where one or more of several possible

embrittlement mechanisms can occur. Ideal strain rates for observing the effects of hydrogen embrittlement are in the range of 10^{-6} to 10^{-5} s^{-1}. [101]

Slow Strain Rate Testing (SSRT) is a quick, comparative technique based on time to failure after elongation that is the most commonly quoted measure of embrittlement in these tests, as HE is a diffusion controlled process and, moreover, time to failure is directly proportional to ductility through the strain rate. Moreover, if failure occurs before the ultimate tensile strength (UTS) the choice of reduction in area (RA) may not be the most apposite criterion, as the change in RA prior to UTS is usually small. Furthermore, tests may be performed not only in tension, as in our case, but also in bending on a variety of specimen types, such as plain, notched or even pre–cracked specimens. [101]

Advantages of the SSRT method

One of the most significant advantages of the SSRT compared to the constant load test is that it always causes rupture of the test specimen, as the specimen is continuously strained in tension until fracture. [105, 106] As a result, there is no need to specify an arbitrary test time as in the constant load technique and, in addition, the SSRT technique is by far less expensive compared to fracture mechanics crack propagation tests, as the geometry of the specimens and the test procedure are simpler. [106] Moreover, as the plotting of a stress–strain diagram is more feasible, information about the tensile strength of the specimen and the shape of this diagram can be useful when comparing embrittled with non–embrittled specimens.

The SSRT method has been shown to be highly effective in assessing Stress Corrosion Cracking (SCC), hydrogen damage and related phenomena. According to Kim and Wilde [105], the major advantage of the dynamic slow strain test procedure, as the SSRT is termed by them, is that it is much more aggressive in producing SCC than conventional constant–strain or constant–load tests. Hence, the testing time is reduced considerably and there is no need to follow other methods, as in other types of mechanical testing, in order to minimise or eliminate the incubation stage that often occupies 90% of the test duration. For instance, the use of pre–cracked specimens encourages acidification of the environment due to crevice corrosion. On the other

hand, dynamic straining plays an important role during crack propagation by maintaining localised corrosion at an active crack tip, while the sides of the crack are rendered inactive, as they are covered by films. This active crack condition depends on both electrochemical conditions at the crack, as well as the rate at which fresh metal is exposed to the corrosive environment by dynamic straining. In the light of the previous observations, the SSRT method is ideally suitable for rapid identification of environment / metal combinations producing SCC and the method can be used for quantitative ranking of SCC properties for steels of similar microstructure and strength levels.

Parkins [107], who was the first to promote the concept of using dynamic strain in a SCC test, incorporates in his work about strain rate testing some of the previously mentioned benefits of using SSRT in SCC. He supports the view that SSRT is proven to be eventually not only a rapid sorting test, but also a method with more realistic significance to service failures than thought in the past. SSRT frequently facilitates cracking in cases where, at constant load, cracking is not observed, displays poor reproducibility or lasts an excessively long time.

Amidst other investigators mentioning the usefulness of SSRT, McIntyre et al. [108] support the view that the SSRT method overcomes many of the disadvantages of more traditional test methods, like the NACE tensile test, when evaluating materials in high–pressure hydrogen sulphide environments. SSRTs allow a relatively quick ranking of materials in realistically simulated sour service conditions. Additionally, the method is more flexible than other ones because corrosion–resistant alloys can be evaluated along with low alloy steels.

Pollock [18, 109] attempted to quantitatively evaluate the degree of HE produced in low alloy high strength steels by plating and baking procedures, and he also investigated the effects of paint strippers in causing re–embrittlement. He used the SSRT method on notched tensile specimens and characterised the technique as a viable standard method for the assessment of the degree of HE in plated and baked high strength steel, as it is a rapid way of quantifying the relative severity of test procedures defined in various standard specifications, as well as, because its severity can be controlled by altering the strain rate.

Limitations of the SSRT method

Apart from the advantages of the SSRT technique, which overcome drawbacks of other methods, there are some limitations associated with its implementation, essential to be considered before attempting to draw any conclusions based on SSRT results. Various investigators [105–108, 110, 111] have cited different concerns arisen from the application of the SSRT method.

For a specific material / environment combination the two most important parameters controlling the SSRTs are the strain rate and the potential. [106] With regard to the rate of strain, there are concerns that crack propagation during testing can lead to partial relaxation of the specimen and, more considerably, specimen necking can cause a local increase of the strain rate in the necking region by as much as an order of magnitude. [107] Hence, the heavy plastic deformation in the late stages of a SSRT can incur pronounced ductility loss. [108] Such an alteration can cause the strain rate to move out of the critical range for the detection of HE in test tensile specimens. This point may be particularly substantial when the strain rate is altering from test to test.

From the above it is apparent that strain rate is such an important experimental parameter in the SSRT, that the critical strain rate for cracking should be identified for each investigated material / environment system either from the literature or from preliminary SSRTs. [106] Kim and Wilde presented two different cases for the selection of an appropriate strain rate range, so as to detect SCC and Hydrogen–Induced Cracking. [105]

For the detection of SCC, when the test is performed at an extremely high strain rate, the specimens fail by tensile rupture, as there is not sufficient time for SCC to develop. At high strain rates, even if SCC has started, the cracks become blunted due to dynamic straining before any significant amount of localised corrosion can occur at the crack tip, and thus no SCC can be observed. On the other hand, when a very low strain rate is applied, SCC susceptibility is suppressed because the corrosion reaction is hindered by the film formation, and the active condition at the crack tip cannot be preserved if the film rupture rate is below a critical value. Therefore, the SSRT does not generate SCC at either extremely high or low strain rates, but will produce SCC in

a critical range of strain rates characteristic for each material / environment system. For instance, for an ASTM A517 Grade F steel tested at various strain rates in air contaminated liquid ammonia, the most severe SCC was produced at a strain rate of approximately 10^{-6} s^{-1}. [105]

A different dependency of cracking susceptibility on strain rate was observed for Hydrogen–Induced Cracking. The reason is that Hydrogen–Induced Cracking does not require any critical value for film rupture rate, as in the SCC, but it is caused as a result of the occluded hydrogen in the steel structure. During the corrosion reaction for hydrogen cracking, the cathodic reaction generates hydrogen on the steel surface, some of which is adsorbed by steel up to the saturation limit. By decreasing the strain rate the testing period is increased and the specimen becomes more susceptible to hydrogen cracking. Based on this varying response to strain rate, the SSRT can be applied to determine whether the failure mechanism in a given material / environment system is SCC or Hydrogen–Induced Cracking. [105] (Figure 31)

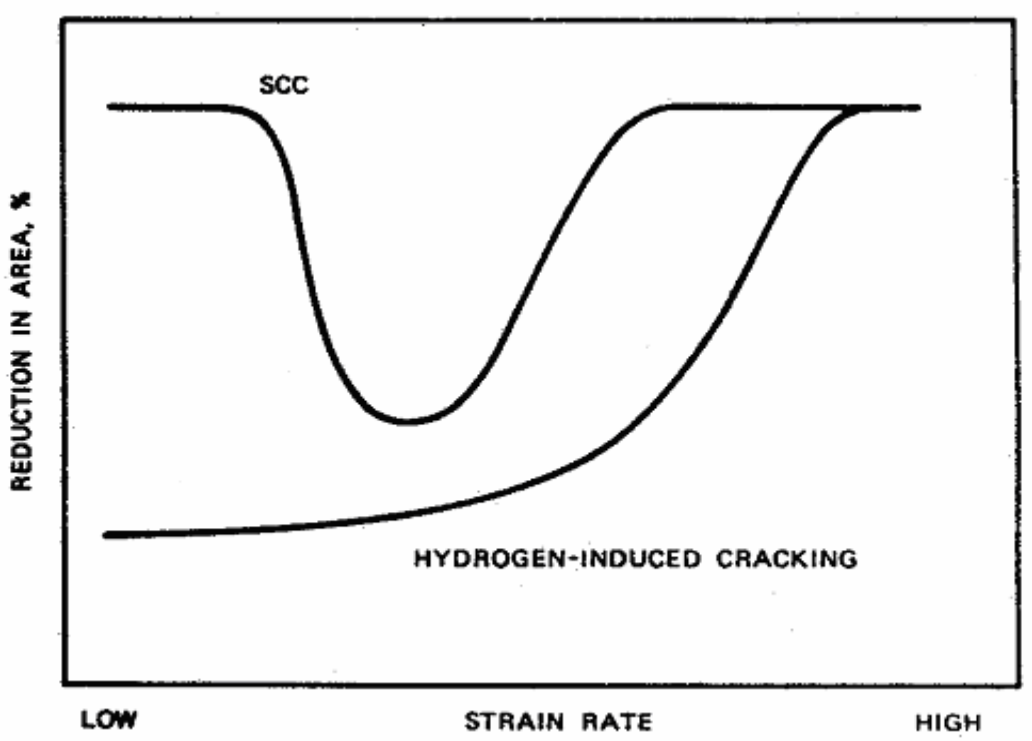

Figure 31. Schematic presentation of the effect of strain rate on SCC and Hydrogen–Induced Cracking. [105]

Other investigators corroborated some of the previous conclusions by providing data for 304 stainless steel, the HE susceptibility of which increases with the reduction in strain rate [110]. For low alloy steel 90 MnV 8 in spheroidised condition, the elongation to fracture does not vary with strain rate for tests performed in air, whereas HE increases with decreasing strain rate when such specimens are tested and simultaneously charged with hydrogen at –200 A/m^2 in a 0.1 N NaOH solution. [111] (Figure 32)

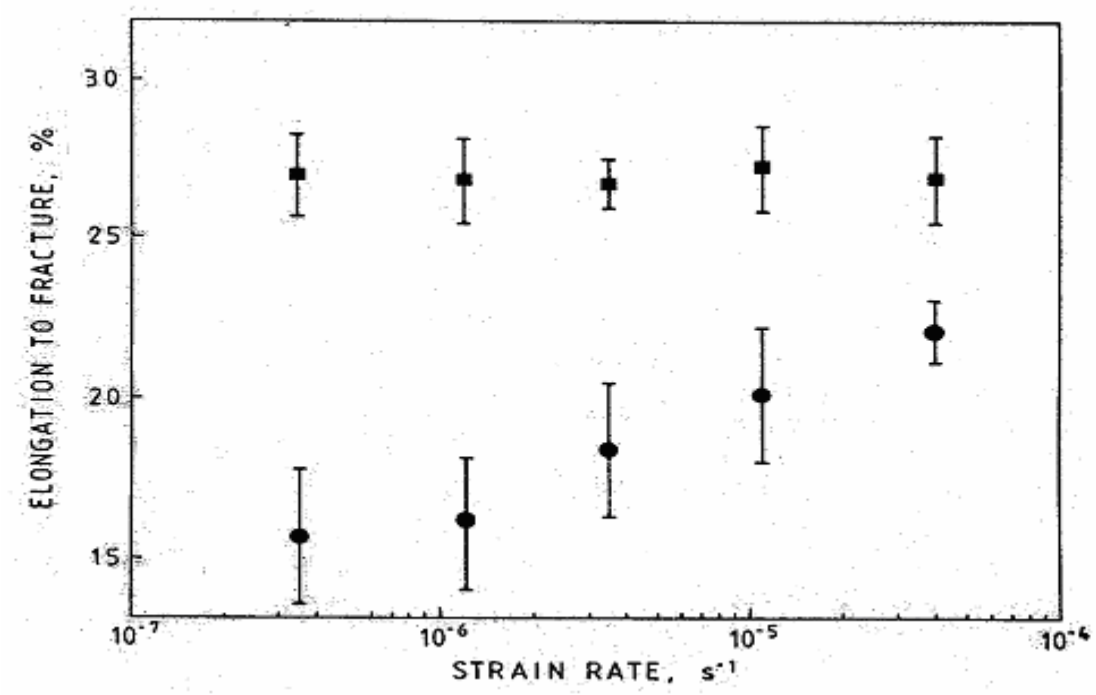

Figure 32. Elongation to fracture as a function of strain rate for low alloy steel 90 MnV in spheroidised condition. (■) Tests in air. (●) Tests with a simultaneous charging at a current density of –200 A/m^2 in a 0.1 N NaOH solution. [111]

The other most important parameter controlling the SSRTs in a given combination of material and environment is the potential that should be controlled or monitored in all SSRTs. If the critical potential range for the detection of embrittlement is not known, a series of SSRTs at different potentials should be performed covering the range of expected potentials in service. The previously cited remarks of Beavers and Koch [106] were based on an investigation upon SSRT experience in various combinations of materials and environments, including carbon steels. In the light of the dependency of the SSRTs on strain rate and potential, it is not coincidental that most reported

cases of anomalous behaviour when implementing the SSRT can be ascribed to inadequate control or measurement of one or both of these parameters. Therefore, comparison of different materials or environments at a single strain rate or at a single potential could produce erroneous results due to the specific dependency of the critical strain rate or potential for cracking.

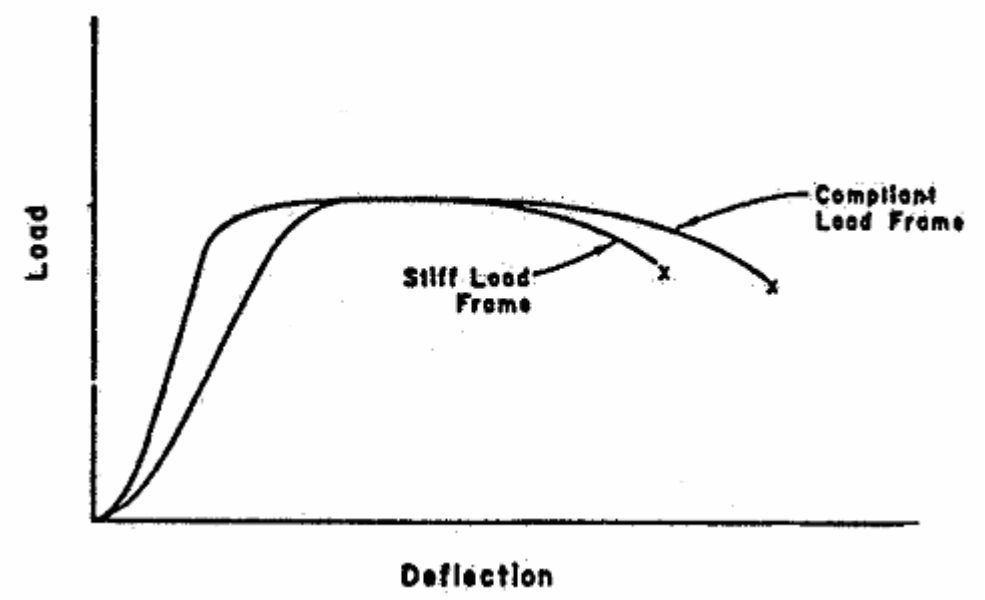

Figure 33. Evident difference in time to failure with the same plastic elongation, as a result of differences in load frame stiffness. [108]

A different issue implicated with the use of SSRT could be the stiffness of the frame of the testing machine, which can affect the time to failure measurements. [108] When a specimen is loaded in a SSRT, a small component of the measured elongation stems from the elastic deformation of the load frame. Consequently, two load frames with different compliances are likely to give slightly different total elongation on specimens with the same actual elongation, as depicted in Figure 33. However, once plastic deformation commences, all the elongation in the load train can be safely assumed to occur in the specimen's gauge length. Therefore, the load frame stiffness effect can be eliminated by reporting only plastic elongation rather than total elongation. However, this approach should not be followed with materials of widely different strengths and microstructure, where a quantitative ranking of their SCC properties based on the use of the SSRT method is difficult. [105] Moreover, even if

the same material is concerned, it could behave in different ways under various conditions depending on how embrittling these conditions are. Thus, there could be specimens failing before the yield point or before UTS, and under these circumstances no comparable plastic elongation can be reported. Nevertheless, for a pure ranking purpose, all these previously stated thoughts and concerns about the effect of the load frame compliance on the credibility of the SSRT results should not be a problem, provided the parameters of the testing machine remain unaltered.

1.9.3 Statistical Analysis Techniques

Time to failure measurements from the SSRTs are renowned for scatter, and consequently the quantification of the hydrogen effects on the steel strength is difficult to evaluate. The scatter can derive from changes in factors such as charging conditions, defect size or location in a specimen, and surface imperfection. Many statistical evaluation techniques have been implemented for the analysis of experimental data, including times to failure. [9, 11]

The problem in general lies under the case of identical components subjected to identical environmental conditions failing at different and unpredictable times. Some of the continuous probability distributions playing a role in this sort of problem are the gamma and exponential distributions. Another statistical approach, extensively used to deal with problems related to random phenomena, is the Weibull distribution [112], introduced by the Swedish physicist Waloddi Weibull in 1951. The major usefulness of this distribution is that it affords an excellent approximation to the probability law of many random variables. One important area of application has been as a model for time to failure in electrical and mechanical components and systems, and for this reason it was implemented in the mechanical testing part of this study. [113, 114]

Weibull analysis [112] was initially developed to deal with the scatter of times to failure for the brittle failure of ceramics and glass. Fracture of these materials depends upon the size and distribution of imperfections throughout the material. Yokobori (1965) adapted the technique to a variety of phenomena including the brittle fracture of steel. [115] He treated the scatter in failure times as a stochastic process, defined as

a probability system generated by time, where the state of the system at time t is dependent upon chance. Strecker, Ryder and Davies (1969) used the previously referred approach to an investigation for hydrogen–induced delayed failure in unnotched specimens of 0.9% C steel strip. [116] They showed that Yokobori's analysis is valid for precharged high carbon unnotched wires tested to failure at a constant applied load. Although the use of unnotched specimens is prone to results with scatter, the absence of a notch allows failure to be strongly affected by structure, and not by geometric effects, thus permitting the influence of the former variable to be examined.

Moreover, Robinson and Sharp (1985) slightly modified Yokobori's analysis for specimens being charged and simultaneously loaded at constant applied loads. [91] In each case, they found that a minimum incubation time t_i exists, corresponding to the time (in the most susceptible specimen tested) for hydrogen to diffuse to a depth where a critical crack length is generated. This approach was successfully implemented to the study of the embrittlement of nickel electroplated or cathodically charged and, simultaneously subjected to a constant load, steel specimens. [117, 118] More recently, Tsai and Shih (1996) applied the Weibull failure distribution on the time to failure assessment of high strength low alloy steel plates tested at constant loads in hydrogen sulphide containing environments. [119] These investigators attempted to adapt their Weibull plots to one of the initially proposed forms of the Weibull model.

More discussion about the Weibull continuous probability distribution and its different modifications by various investigators in the last 50 years is available in the Appendix, where there is also information about the Student's t–test, a statistical technique that compares two sets of results and gives an indication if these sets belong to the same population. Pollock [19] used this statistical technique to compare sets of SSRT results on notched specimens, which were cadmium plated, plated and baked, or exposed to paint stripper.