HYPERBOLIC TANGENTS and alloys of iron

by Harry Bhadeshia*

Mathematical modelling can be an effective aid to the design and optimisation of alloys, here illustrated with iron alloys.

In terms of bonding, an atom of iron is essentially an argon core, with an additional eight electrons which delocalise to form bonds. These bonding electrons can be distributed in a variety of ways to yield one, two and three bonding electrons per atom respectively, corresponding to the body centered cubic (*bcc*), hexagonal close packed (*hcp*) and cubic close packed (*ccp*) structures in iron. The crystal structures of many metallic elements can be rationalised in this way – an electron:atom ratio less than 1.5 gives *bcc*, 1.7-2.1 *hcp* and larger ratios *ccp*.

Satisfying patterns like these stimulate thought, but to predict the stable crystal form requires a model based on quantum mechanics. An example is illustrated for iron at 0 K in figure 1, which shows the cohesive energy as a function of the density and crystal structure.

Of all the test structures, *hcp* iron is found to show the highest cohesion and therefore should represent the most stable form. The result contradicts experience, because *bcc* iron (ferrite) occurs naturally at low temperatures. This discrepancy arises because the model ignores magnetic effects – it is ferromagnetism which stabilises *bcc* iron over *hcp* iron.

We therefore have ferromagnetism to thank for civilisation as it exists, for what would life be without *bcc* iron? Ruthenium, which is an iron analogue, does not exhibit ferromagnetism and has the expected *hcp* crystal structure at low temperatures.

This example illustrates that all models begin with intentional or unintentional simplifying assumptions and yet can be extremely useful. The electronic theory not only highlights the importance of magnetic terms, but in addition makes it possible to study crystal structures of iron which have yet to be achieved in practice.

The diamond form of iron would have a density of only 5gcm^{-3}.

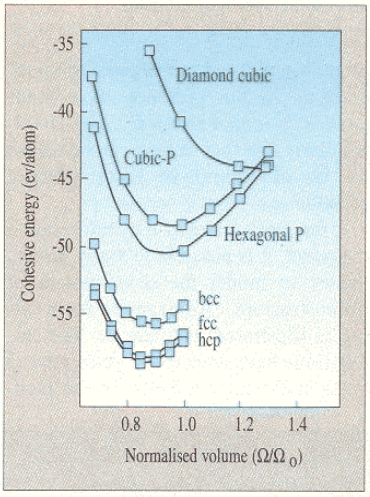

Figure 1 Plot of cohesive energy vs the normalised volume per atom for a variety of crystal structures of iron – hexagonal-P and cubic-P are primitive structures (after Paxton, Methfessel and Polatoglou)

Unfortunately, the calculations show that the difference in energy between the diamond and bcc forms is so large that it is improbable that the *bcc* to diamond transformation can be induced by alloying.

Although this particular example is rather esoteric, the methodology is routinely being used to see whether certain intermetallic compounds can be forced to occur into more desirable crystal structures.

Perhaps the best way to radically reduce the density of iron is to use a method invented in the Ukraine, involving a high pressure eutectic reaction in which liquid metal containing dissolved hydrogen decomposes into a eutectic mixture of solid metal and gaseous hydrogen. The bubbles can occupy up to 75% of the final volume! A model for morphological stability is being used there to produce many varieties of bubble dispersions and bamboo structures. It is easy to imagine applications in vibration damping materials, catalysis etc.

The purpose of this article is to illustrate the methodology of modelling as an aid to the design and optimisation of alloys, using experience on iron an steel.

Thermodynamic equilibrium

In fact, the role of ferromagnetism in stabilising *bcc* iron has been known for some fifty years, long before electron theory became prominent in metallurgy. Zener factorised the free energy of iron into the magnetic and non-magnetic components in order to explain the peculiar stability of *bcc* iron at high and low temperatures, interrupted by the occurrence of austenite.

This eventually explained why austenite has a larger thermal expansion coefficient than ferrite, even though the two phases have the same Debye temperature and electronic specific heat coefficients.

Thermodynamic analysis of this kind has undoubtedly been the most successful of all modelling techniques – multicomponent multiphase calculations are now routine precursors to alloy design programmes in both universities and industry. The thermodynamic method is useful particularly because it 'links together many variables so that they can be seen to be a consequence of a few'. Elements lose their identity – they simply contribute to an overall free energy.

The friendly output of the such calculations consists of phase fractions and compositions as a function of the overall alloy composition and variables such as temperature and pressure. Scientists have become so used to appreciating the quantitative data for alloys with large numbers of compo-

*Dr H K D H Bhadeshia is in the Department of Materials Science and Metallurgy at the University of Cambridge

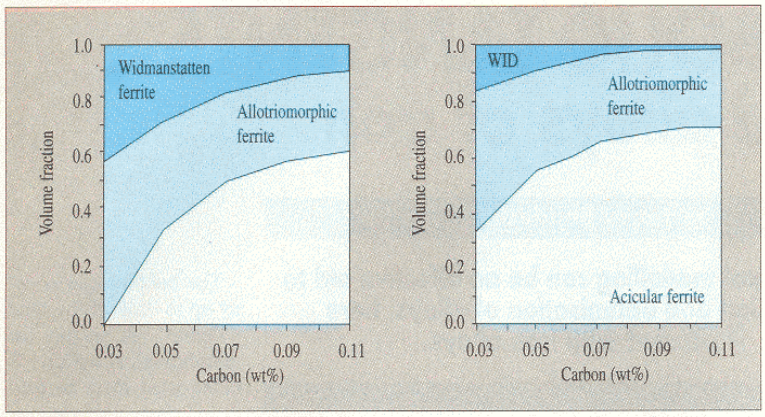

Figure 2 Calculated microstructures of manual metal arc weld deposits as a function of chemical composition – the microstructure of the boron-free alloy (left) is very different to the microstructure of the alloy with a boron concentration of 20ppmw (right)

nents and phases that it is no longer necessary to teach the tortuous geometries of ternary and quarternary phase diagrams. In this country, the National Physical Laboratory's *MTDATA* system is commercially available for realistic alloys of much higher order.

A system is in thermodynamic equilibrium when there is no perceptible change no matter how long one waits. Calculated phase diagrams rely on thermodynamics alone – they can deal with equilibria, metastable equilibria (which are defined by local minima) or equilibria which are constrained in some way (for example when some of the atoms are mobile and others are not). Most useful microstructures cannot be estimated from phase diagrams because they are not at equilibrium, but the thermodynamic parameters are essential inputs to kinetic theory which describes the approach to equilibrium.

Rate

The simplest assumption in kinetic theory is to take a 'flux' to be proportional to a 'force'. The flux could represent an interface velocity, an electrical current or a heat flux, with the driving force, the electromotive force and temperature gradient representing the corresponding forces. The assumption of proportionality is probably reasonable for small departures for equilibrium. The proportionality can be generalised, so that a given flux can be expressed as a function of a combination of forces.

In a ternary Fe–Mn–C alloy, the diffusion flux of carbon depends not only on the gradient of carbon, but also on that of manganese. Consequently, a uniform distribution of carbon will tend to become inhomogeneous in the presence of a manganese concentration gradient. Similarly, the flux of heat can be driven by an electromotive force (Peltier effect).

If the mechanism of a transformation is known, it is possible to apply kinetic theory to model the development of microstructure. An example for welding is illustrated in figure 2. The calculations have over many years proved to be sufficiently reliable for the models to be used routinely in industry, both for alloy design and to address customer queries with confidence. The experimental determination of the data illustrated in figure 2 would cost in excess of £100,000.

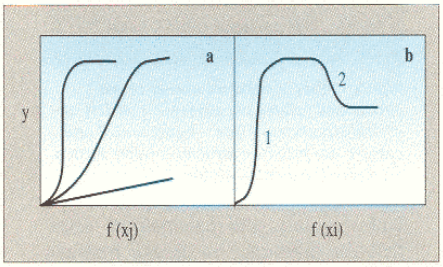

Figure 3 **a** *Three different hyperbolic tangent functions – the 'strength' of each depends on the weights* **b** *A combination of two hyperbolic tangents to produce a more complex model*

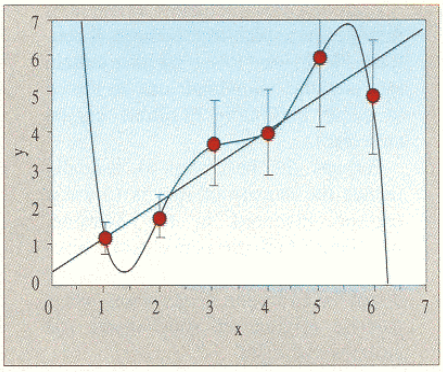

Figure 4 A complicated model may fit the data, but in this case, a linear relationship is all that is justified by the noise in the data

Empirical models

The physical models described above are capable of predicting entirely new phenomena. But there are difficult problems where the general concepts might be understood but which are not as yet amenable to mathematical treatment. We are at the same time told that good engineering has the responsibility to reach objectives in a cost and time effective way.

Any model which treats a small part of the required technology is therefore unlikely to be treated with respect. Empiricism can in these circumstances be extremely useful in filling in any gaps whilst striving for longer term solutions.

Most people are familiar with regression analysis, where data are best-fitted to a specified relationship which is usually linear. The result is an equation in which each of the inputs x_j is multiplied by a weight w_j. The sum of all such products and a constant θ then gives an estimate of the output $y = \Sigma_j w_j x_j + \theta$. It is well understood that there are dangers in using such relationships beyond the range of fitted data.

A more general method of regression is neural network analysis. As before, the input data x_j are multiplied by weights, but the sum of all these products forms the argument of a hyperbolic tangent. The output y is therefore a non-linear function of x_j, the function usually chosen being the hyperbolic tangent because of its flexibility.

The exact shape of the hyperbolic tangent can be varied by altering the weights, fig 3a. Further degrees of non-linearity can be introduced by combining several of these hyperbolic tangents, fig 3b, so that the neural network method is able to capture almost arbitrarily non-linear relationships. It is well known that the effect of chromium on the microstructure of steels is quite different at large concentrations than in dilute alloys. Ordinary regression analysis cannot cope with such changes in the form of relationships.

A potential difficulty with the use of powerful regression methods is the possibility of overfitting data, fig

4. For example, it is possible to produce a neural network model for a completely random set of data. To avoid this difficulty, the experimental data can be divided into two sets, a training dataset and a test dataset. The model is produced using only the training data. The test data are then used to check that the model behaves itself when presented with previously unseen data.

Neural network models in many ways mimic human experience and are capable of learning or being trained to recognize the correct science rather than nonsensical trends. Unlike human experience, these models can be transferred readily between generations and steadily developed to make design tools of lasting value. These models also impose a discipline on the digital storage of valuable experimental data, which may otherwise be lost with the passage of time.

An example of the use of a neural network model in predicting the strength of steel weld deposits, as a function of a very large number of variables, is illustrated in figure 5. The inputs included the detailed chemical composition (almost twenty elements), heat treatment and welding parameters. Models like these have been trained for many complex phenomena for which physical models do not exist – the fatigue strength of nickel-base superalloys, the formation of austenite during the heating of steels, the toughness of arc welds etc. They are particularly powerful in serving as a bridge between microstructure and mechanical properties.

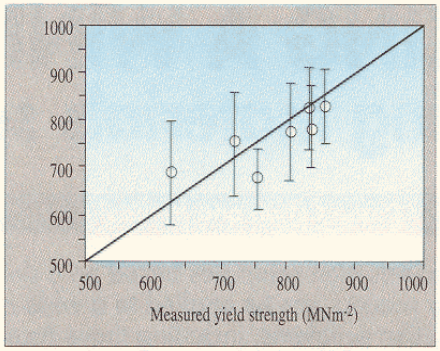

Figure 5 The neural network-predicted vs measured yield strength for a variety of welds previously unseen by the model

Progress

'Modelling' has now become a very prominent subject in materials science. There are many UK university departments which specialise in the subject and there are numerous examples of similar work in industry. It is a very attractive area for collaboration between universities and industry because it combines the excitement of science with the focus on complex technological issues.

The subject requires detailed documentation and The Institute of Materials has already published a number of books in an expanding series on 'Materials Modelling'. There are moves afoot at the National Physical Laboratory to establish materials software and data banks and to define software documentation standards.

The growth of the subject will only be sustained in the long term if it is seen to produce results which are useful to the entire community, not just one partner in the industry/university alliance. ◆