THERMODYNAMIC DATA FROM DIFFUSION COUPLES—I

R. R. KOOPER and T. W. EAGAR
Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

(Received 27 June 1990)

Abstract—A non linear least squares based method is presented to estimate activity data from measured interdiffusivity data for isomorphous binary metals. The interdiffusion coefficient is modelled as a product of two functions, one representing the thermodynamic contributions and another the contribution of the phenomenological coefficients. The proposed method factors the interdiffusion coefficient into its two constituent functions. This leads to the existence of non unique solutions and hence yields a set of possible solutions which on an activity-composition diagram exhibit positive and negative deviations from ideality. One additional piece of information, such as the experimentally measured activity at a particular composition or the slope of the Henry’s law line at infinite dilution, is needed to select the correct solution. The method successfully predicted the thermodynamics of the CoNi system from calculations based on the interdiffusion coefficient.

Résumé—On présente une méthode non linéaire basée sur la méthode des moindres carrés pour estimer les données d’activité à partir de données d’interdiffusivité mesurées pour des métaux(binaires) isomorphes. Le coefficient d’interdiffusion est modélisé par le produit de deux fonctions, l’une représentant les contributions thermodynamiques et l’autre la contribution des coefficients phénoménologiques. La méthode proposée factorise le coefficient d’interdiffusion en ses deux fonctions constituant. Ceci conduit à l’existence de solutions non uniques et amène donc une série de solutions possibles pour lesquelles un diagramme composé-activité montre des déviations positives et négatives par rapport au diagramme idéal. Une information supplémentaire, telle que l’activité mesurée expérimentalement pour une composition particulière, ou la pente de la droite de la loi de Henry à dilution infinie, est nécessaire pour choisir la solution correcte. La méthode prévoit avec succès la thermodynamique du système CoNi à partir de calculs basés sur le coefficient d’interdiffusion.


INTRODUCTION

The science of thermodynamics and an understanding of the diffusion process together provide important bases for the current understanding of metallurgy. Hence, there is a large demand for basic data in these fields. Traditionally, metallurgical thermodynamic data have been obtained using calorimetric techniques, wet chemical methods, vapor pressure measurements or electromotive force measurements [1–3] while solid state diffusion data has been generated from diffusion couples [4, 5].

Einstein showed that diffusion was in fact driven by gradients in chemical potential. Thus the Nerst–Einstein relationship which relates mobility and diffusivity, firmly established the link between thermodynamics and diffusion [4]. This was experimentally demonstrated by Darken in his well known uphill diffusion experiment [6]. Other investigators have related the chemical potential to elements of the diffusivity matrix [7–11]. Excellent reviews covering the literature of solid state diffusion and the associated phenomenological theory of irreversible processes are available [11–21].

The main premise of the present investigation was that since diffusion is driven by chemical potential gradients, a diffusion couple must implicitly contain information on chemical potentials (activity-composition data). The main goal of this investigation was to obtain the free energies of mixing (chemical
potentials) from the information available in a diffusion couple. The investigation focuses on isomorphous binary metallic systems and later attempts to extend the treatment to binary systems containing two phases.

The overall thrust of this work is the establishment of new techniques for obtaining thermodynamic data by manipulating diffusion data. The main advantage of such techniques would be the ability to obtain thermodynamic data for metallic systems where it is not feasible to conduct conventional thermodynamic experiments. Examples would include reactive and refractory metal systems. Apart from providing new data, these techniques will also aid in checking the accuracy and consistency of the data already available. Finally, it is hoped that this study will lead to a better appreciation of the link between thermodynamics and solid state diffusion.

Preliminary efforts [22] focused on analyzing the shape of the diffusion profiles in search for activity data. Analysis of the preliminary attempts had shown that it is more fruitful to focus on obtaining chemical potential directly from the variation of the interdiffusion coefficient with composition.

This is the first of two papers covering this investigation. This paper presents the details of a non-linear least squares based algorithm that allows one to compute the activity–composition relationship from a knowledge of the interdiffusion coefficient as a function of composition. The algorithm will be illustrated by using the CoNi system as an example. The second paper applies this algorithm to nine isomorphous binary systems and compares the computed activity data to the activity data available in the literature.

This paper will show that it is possible to decompose or factor the interdiffusion coefficient into a product of two functions, one representing the system thermodynamics and another representing the phenomenological coefficients. Unfortunately, such a decomposition is not unique and hence one ends up with several possible activity curves which are consistent with the given interdiffusivity data. To select the correct activity curve one additional piece of information such as the Henry’s law line at infinite dilution or the measured activity at one composition is needed.

This paper begins by examining various theories in the literature interrelating the chemical interdiffusion coefficient to the system thermodynamics and to the phenomenological (Onsager) coefficients.

LITERATURE REVIEW

Relationship between the diffusivity and the system thermodynamics

The phenomenological theory of irreversible processes expresses the flux of a diffusing species as a linear combination of all driving forces [11, 13–15]. In the case of solid state diffusion, the driving forces are the gradients in chemical potential of the diffusing species. The following equation expresses the phenomenological relationship between fluxes and forces

\[ J_i = \sum_{k=1}^{N} L_{ik} X_k \]

whereas the entropy production rate is given by

\[ \Delta S = \frac{1}{T} \left( \sum_i J_i X_i \right) \]

The earliest atomistic theory relating the interdiffusion coefficient to the system thermodynamics is attributed to Darken [7]. There were two major assumptions in Darken’s analysis. Firstly, Darken assumed that the off-diagonal Onsager coefficients were negligible and secondly he assumed that vacancies were in local equilibrium in the diffusion couple. Darken’s analysis led to the following expression

\[ \bar{D} = (D_A^+ N_A + D_B^+ N_B) \phi(c) \]

where

\[ \phi(c) = \left( 1 + \frac{d \ln \gamma}{d \ln N} \right) \]

In Darken’s theory, the phenomenological coefficients were related to the tracer diffusivities by the following expression

\[ L_1 = c_1 D_A^+ \]

\[ L_2 = c_2 D_B^+ \]

From these expressions it is clear that the phenomenological expressions are strong functions of composition.

Subsequently, Howard and Lidiard [9], Manning [10], and Kirkaldy [11] modified Darken’s treatment to include the effect of non-zero off-diagonal Onsager coefficients. By considering a quaternary \( A^{-}A^{-}B^{-}B^{-} \) system and assuming that the vacancies were in local equilibrium, they arrived at the following expression

\[ \bar{D} = \left[ (D_A^+ c_2 + D_B^+ c_1) + \frac{kT}{N} \left( \frac{X_B}{X_A^*} L_{AA^*} + \frac{X_A}{X_B^*} L_{BB^*} \right) \right] \phi(c) \]

where

\[ \phi(c) = \left( 1 + \frac{d \ln \gamma}{d \ln c} \right) \]

In equations (3a) and (4), the thermodynamic factor \( \phi(c) \) provides a complete description and embodies the thermodynamics of the system. Knowledge of the function \( \phi(c) \) completely defines the solution behavior of the alloy system.

It has been estimated [11] that the additional term in equation (4) leads to a difference of around 5% in the computed interdiffusion coefficient when compared with equation (3) for isomorphous binary alloys and an error of up to 28% in the measurements.
of the marker velocity. Since traditional measurements of interdiffusion coefficients often contain as much as 25% error, it is not possible to detect the contribution of the additional term in equation (4) by measurement of the diffusion coefficient alone. More recent studies using STEM have greatly improved the spatial resolution and the precision of diffusion profile data, thereby reducing the error associated with inter-diffusion measurements.

Experimentally, the interdiffusion coefficient is obtained as a function of composition by subjecting the measured diffusion profile data to the Boltzmann–Matano treatment [11, 12, 23]. This treatment is embodied in the following expression

\[ \dot{D}(c) = \left( \frac{-1}{2t} \right) \left( \frac{1}{dc} \right) \right] \int_0^c [x(c) - x_m(c)] dc \tag{5} \]

Crank [5] and Balluffi [24] have modified the Boltzmann–Matano treatment to account for the volume change of mixing. The correction is small for isomorphous binary alloys, which exhibit small changes in molar volume. Hence this effect will be neglected in the present investigation.

It should be noted that Darken’s theory [equation (3a)] and the theory proposed by Howard and Lidiard [9], Manning [10] and Kirkaldy [11] [equation (4)] differ only in the number and functional form in which the phenomenological coefficients contribute to the interdiffusion coefficient. Darken’s theory calls for two independent phenomenological coefficients whereas the theory proposed by Howard and Lidiard, Manning and Kirkaldy calls for five independent phenomenological coefficients to completely describe binary diffusion under the assumption that the vacancies are in local equilibrium. Despite these differences, it is important to note that both theories predict an identical dependence of the interdiffusion coefficient on the system thermodynamics [equations (3a) and (4)]. Both treatments express the diffusivity in the following general form

\[ \dot{D} = f(c) \left( 1 + \frac{\delta \ln y}{\delta \ln c} \right) = f(c)\phi(c) \tag{6} \]

where \( f(c) \) is a function of the phenomenological coefficients and the alloy composition. Hence, \( f(c) \) depends on the diffusion coefficient being considered [equations (3a) and (4)]. It is clear from equation 6 that the interdiffusion coefficient is a very strong function of the thermodynamics of the system. Borovskii [12] has provided several examples to illustrate this point effectively.

As far as the present investigation is concerned, the main point to be noted is that the interdiffusion coefficient is related to the system thermodynamics through the function \( \phi(c) \) in equation (6) and that the diffusion coefficient is a product of two terms.

\^ Later in the paper a discussion is presented to help deduce such functional forms.

PROPOSED ALGORITHM

The proposed scheme relies on the following assumptions: (i) the molar volume is invariant with composition, (ii) the system is isothermal and (iii) vacancies are in local equilibrium everywhere in the diffusion couple.

Algorithm overview

The basic idea underlying the proposed algorithm is to factor or decompose (de-convolute) the diffusivity data into its constituent functions, \( f(c) \) and \( \phi(c) \). This is accomplished by a non-linear least-squares fit to the isothermal interdiffusion data. One begins by obtaining a discrete representation (several sets of data-points \( \{c_k, D_k\} \)) of the interdiffusion coefficient as a function of composition. Next one selects two mathematical functions, one to represent the variation of \( f(c) \) with composition and another to represent the variation of \( \phi(c) \) with composition. The product of these two functions yields a new mathematical function that represents the variation of the interdiffusion coefficient with composition. This new mathematical function is used to conduct the least squares fit to the diffusivity data.

The mathematical models for \( f(c) \) and \( \phi(c) \) are defined in terms of unknown coefficients. The least squares fitting process provides numerical values of the unknown coefficients that constitute the model functions chosen for \( f(c) \) and \( \phi(c) \) for a certain alloy. Knowledge of the numerical values of the variables that constitute the mathematical function representing \( \phi(c) \) allows one to compute \( \phi(c) \). Knowledge of \( \phi(c) \) is equivalent to knowing the thermodynamics of the system since one can integrate \( \phi(c) \) to obtain the \( y-c \) relationship for the alloy and from it the activity–composition data. With this general overview in mind, one may next examine the details of the proposed computational scheme.

Algorithm details

It is clear from equation (6) that to obtain the chemical potential [or equivalently, the function \( \phi(c) \)] from the interdiffusion coefficient, demands knowledge of the function \( f(c) \), which is related to the Onsager phenomenological coefficients. Since \( f(c) \) is not known a priori, an analytical solution of the \( \phi(c) \) function cannot be obtained. Hence additional assumptions must be made regarding the nature of the functions \( f(c) \) and \( \phi(c) \). Clearly an important first step is the selection of appropriate functional forms for \( f(c) \) and \( \phi(c) \).

For the present, let us assume that the functional forms of the two functions \( f(c) \) and \( \phi(c) \) are known a priori even though the numerical values of \( f(c) \) and \( \phi(c) \) are not known. Further assume that \( f(c) \) is being modelled by a function of \( "n" \) variables, \( \tau_1, \tau_2, \ldots, \tau_n \) whereas \( \phi(c) \) is being modelled as a function of \( "m" \) variables \( \tau_{m+1}, \ldots, \tau_{m+n} \). This
implies that the interdiffusion coefficient is being modelled as a function of \( m + n \) variables
\[
\overline{D} = f(t_1, t_2, \ldots, t_n) \phi(t_{n+1}, \ldots, t_{n+k}). \tag{7}
\]

The mathematical functions chosen to represent \( f(c) \) and \( \phi(c) \) are presumed to possess sufficient degrees of freedom (free variables) to adequately model the quantities they represent. In other words, given a real, metallic, isomorphous, binary alloy system there exists at least one set of real values for the variables \( t_1, t_2, \ldots, t_{n+k} \) such that the resulting function \( \phi(c) \) adequately represents the solution behavior of the alloy system.

Further, assume that the interdiffusion coefficient has been measured at several discrete compositions over the entire range from 0 to 100% solute. This may be regarded as the input data, comprising several sets of points \( (c_i, \overline{D}_i, i = 1, k) \) say \( k \) sets where \( k \) is presumed to be a large number of the order of 100. The diffusivity data are presumed to have been generated during a prior diffusion experiment.

To conduct a least squares fit to the diffusivity data one needs to define a least squares error function (sometimes referred to as the chi squared function). In the present case this may be expressed as
\[
E^2 = \sum \left( w_i \left[ D_i - \left[ f(t_1, \ldots, t_n) \times \phi(t_{n+1}, \ldots, t_{n+k}) \right] \right]^2 \right). \tag{8}
\]

The non linear least squares estimation process amounts to searching an \( n + m + 1 \) dimensional space for a minimum. By systematically varying the parameters \( t \), one attempts to search the \( E^2 - t \) space for a minimum. At the minimum, the numerical value of the least squares error function, \( E^2 \) is “small” but not necessarily zero. The numerical value of \( E^2 \) at the minimum is termed the residual and is a measure of the goodness of fit.

The proposed least squares analysis will yield values of the parameters \( t_1, t_2, \ldots, t_{n+k} \) which minimize the least square error between the measured data points and the product of the proposed model functions for \( f(c) \) and \( \phi(c) \). The parameters \( t_1, t_2, \ldots, t_{n+k} \), define the functions \( \phi(c) \) and \( f(c) \), and thereby provide an estimate of the thermodynamics of the system.

There are two major problems when least squares analysis is applied to a function \( \overline{D} \) that is a product of two other functions \( [f(c) \text{ and } \phi(c)] \). Firstly, there is no guarantee that the space spanned by \( E^2 \) and the \( m + n \) unknown variables \( t_i \) contains only a unique minimum, rather there is a possibility of multiple minima. Secondly, there is a possibility that the two functions \( \phi(c) \text{ and } f(c) \) are correlated.† These problems are related and undermine the deconvolution efforts. Therefore they must be carefully addressed.

The term correlated refers to the possibility that the two functions may be mathematically similar and hence indistinguishable. As an extreme example of two completely correlated functions consider two polynomials. If a product of two polynomials was taken, then it would be impossible to deduce the coefficients of the original polynomials given only the functional form of their product. In the present case correlation effects arise due to the inherent non uniqueness of the problem wherein one is attempting to compute two functions knowing only their product.

The problem of multiple minima in the \( E^2 - t \) space stems from the non-uniqueness of the problem. Since only the product of the two functions (i.e. the interdiffusion coefficient) is known, one can make compensating errors in \( f(c) \) and \( \phi(c) \) while still obtaining the correct product of the two. Thus it may be seen that the decomposition can lead to several sets of function \( [f(c), \phi] \). The product of each set will provide a good fit to the diffusivity data but only one set of functions correctly represents the alloy system.

Thus after computing the several sets of possible or feasible solutions to the problem one still needs to select the correct solution from the set of possible solutions. It will be shown that this demands one additional piece of information such as the slope of the Henry’s law line at infinite dilution or the measured activity at one composition.

From the preceding arguments it is clear that the choice of functions to model the term \( f(c) \) and \( \phi(c) \) in equation (7) should consider the possibility of correlation and of multiple solutions.

\[\text{CONSTRAINTS ON THE THERMODYNAMIC FUNCTION } \phi(c)\]

It is appropriate to examine some properties of the function \( \phi(c) \). Firstly, it should be noted that for equilibrium conditions, \( \phi(c) \) is always positive,‡ otherwise \( \overline{D}(c) \) would be negative. Alternately one may regard the requirement for \( \phi(c) \) to be positive as arising from the intrinsic thermodynamic stability criterion \( [3, 26] \delta\mu/\delta c > 0 \), which must hold true for all isomorphous binary systems at equilibrium. For ideal alloys the function \( \phi(c) \) is identically unity at all compositions; for alloys exhibiting positive deviation \( \phi(c) < 1.0 \), and for alloys exhibiting negative
deviations from ideality \( \phi(c) > 1.0 \). Hence one may summarize

\[ 0.0 < \phi(c) < 1.0 \quad \text{positive deviation} \]
\[ 1.0 = \phi(c) \quad \text{ideal alloy} \]
\[ 1.0 < \phi(c) \quad \text{negative deviation.} \quad (9) \]

In addition, \( \phi(c) \) tends to unity as the composition tends to zero or to unity. In other words

\[ \phi(c) \rightarrow 1 \quad \text{as} \quad c \rightarrow 0 \quad \text{or as} \quad c \rightarrow 1. \]

This is more transparent when \( \phi(c) \) is expressed as follows

\[ \phi(c) = \left(1 + \frac{c \, \partial \gamma}{\gamma \, \partial c}\right) \]

As \( c \rightarrow 0 \quad \frac{c \, \partial \gamma}{\gamma \, \partial c} \rightarrow 0 \) hence \( \phi(c) \rightarrow 1 \)

whereas

\[ \text{as} \quad c \rightarrow 1 \quad \frac{c \, \partial \gamma}{\gamma \, \partial c} \rightarrow 0 \quad \text{hence} \quad \phi(c) \rightarrow 1. \quad (10) \]

The restrictions embodied in equations (9) and (10) will be referred to later in the paper. It should be noted that these constraints are based upon thermodynamic knowledge and are true for all binary alloy systems. Hence the proposed algorithm must not violate these constraints.

**SELECTION OF MODEL FUNCTIONS**

This section explains how appropriate functions to model \( f(c) \) and \( \phi(c) \) were selected.

*Models for thermodynamic term*

Five isomorphous binary alloy systems were selected; one of these exhibited ideal solution behavior (Co-Ni), two systems exhibited positive deviations from ideality (Au-Ni and Cu-Ni) and two systems exhibited negative deviations from ideality (Ag-Au and Cu-Au). Isothermal, activity–composition data and activity coefficient–composition \( (\gamma-c) \) data were obtained for each of these alloys from the handbook by Hultgren et al. [25].

It should be noted that the function \( \phi(c) \) is usually not tabulated in most handbooks, since it is not a popular means of describing the system thermodynamics. Hence to evaluate the function \( \phi(c) \) it was necessary to obtain not only the activity coefficient as a function of composition but also its slope with respect to composition. Since the slope \( (\partial \gamma / \partial c) \) is not measured, it was decided to fit the available activity coefficient–composition \( (\gamma-c) \) data with simple functions such as polynomials or exponentials and then use the fitted functions (and their derivatives) to evaluate the function \( \phi(c) \).

Curve fitting routines from the RS/1 package [27] were utilized in fitting functions to the raw \( \gamma-c \) data. Polynomials of increasing order (up to fourth order) and exponential functions were successively tried in order to obtain a reasonable fit to the data. The data from Hultgren [25] et al. was available at 10 compositions, evenly spaced on the composition interval \([0, 1]\). To ensure that the coefficient \( \gamma \), tended to unity as the composition tended to unity, it was necessary to add certain fictitious points (in excess of \( c = 1 \)). These points were weighted preferentially during the fitting process. Figure 1 depicts the \( \gamma-c \) data from Hultgren [25] along-with the fictitious points and the final fitted functions. The CoNi system is not depicted in Fig. 1 since it is known to be ideal [25].

Using the functional approximations to the \( \gamma-c \) relationship for all 5 alloys, \( \phi(c) \) data were generated at 100 evenly spaced points on the interval \([0, 1]\). The generated \( \phi(c) \) data for all five alloy systems is depicted in Fig. 2. It is clear from Fig. 2 that the constraints embodied in equation (10) are not strictly
obeyed at $c = 1$ despite the additional fictitious points that had been added to the $\gamma$--$c$ data and the preferential weighing of these points. Nonetheless, these data were taken to represent the true $\phi(c)$ data for the purpose of selecting a generic functional form for $\phi(c)$.

Efforts were next directed towards locating a generic function that would adequately model the $\phi(c)$ data. To accomplish this, several trial functions (polynomials and $\exp(P_5)$) were fitted to the $\phi(c)$ data.† After several attempts it was concluded that the group of functions, $f(c) = \exp(P_5)$, (where $P_5$ refers to polynomials of order 5), could provide a reasonable fit to the $\phi(c)$ data. Hence for all future work the function $\exp(P_5)$ was chosen to model the $\phi(c)$ function. It should be noted that the functional form chosen for $\phi(c)$ ensures that $\phi(c)$ is always non-negative, which is consistent with the thermodynamic requirement that $\phi(c) > 0$ for equilibrium.

Thus the generic function for $\phi(c)$ was selected based on the thermodynamic data for five isomorphous binary systems. The activity composition relationships for these systems are somewhat simple in that the general behavior of the system is unchanged with composition. In other words, these systems either show positive deviation over the entire composition range or show negative deviation over the entire composition range. In none of these cases do the activity-composition curves cross over from positive to negative deviation (or from negative to positive deviation). This may lead to the erroneous belief that the selected functional form for $\phi(c)$ cannot model systems for which the activity curves do cross over. However, it will be shown in the following paper that the selected functional form for $\phi(c)$ can adequately model the thermodynamics of systems for which the activity curves cross over.

Similarly, it should be noted the selection of the functional form for $\phi(c)$ is based on examination of isomorphous binary systems. Thus such functional forms may not be valid for other systems that possess intermetallic compounds or contain miscibility gaps. For example, it is known that the activity coefficient approaches singularity near a critical point. For such a case the modelling function for $\phi(c)$ may need to be modified. Such a modification can be easily accomplished and one may proceed with the new function in a manner similar to that described here for isomorphous binary systems. The interested reader is referred to the second paper in this series for a full discussion on the extension of the present treatment to non-isomorphous binaries.

Models for the function $f(c)$

Since Darken's theory and the theory proposed by Howard and Lidiard result in less than a 5% difference in the interdiffusion coefficient, it was decided to use Darken's equations to estimate the function $f(c)$ for the five isomorphous binary systems chosen earlier. It should be noted that Darken's theory does not underlie the current algorithm but that it was used only to obtain an initial estimate of the function $f(c)$. This initial estimate of $f(c)$ was helpful in choosing an appropriate functional form for $f(c)$.

Consideration of equation (3), indicates that an estimation of the function $f(c) = (D^f + D^gN_1 + D^gN_2)$, requires knowledge of the variation of the tracer diffusivities as a function of the composition. Hence, the literature was scanned to locate tracer diffusivities for each of the five isomorphous binary systems selected earlier. However, only the tracer diffusivity data for two systems, AuNi [28, 29] and AgAu [30, 31] could be obtained. The tracer diffusivity data for these systems was available at a temperature different from the temperature at which Hultgren et al. [25] report the activity data. Nonetheless it was felt that the general functional behavior of the $D^i(c)$ data could be studied by using the limited data available.

Since exponential functions had been chosen to represent the function $\phi(c)$, functions other than exponentials had to be chosen to represent the tracer diffusivities [i.e. the functions $f(c)$] in order to avoid the possibility of correlation.‡ Based on several attempts at curve fitting the tracer diffusivity data, it was concluded that second order polynomials could adequately model the variation of tracer diffusivities as a function of composition. From equation (3) this implied that the function $f(c)$ could be adequately modelled by a cubic polynomial.

It may be argued that the choice of a functional form for $f(c)$ was based on data from two alloy systems only and hence may not be representative of tracer diffusivities in general. This is certainly true, and modifications to the functional form may be called for at a later stage. However, these are relatively straightforward to incorporate into the algorithm. Hence for illustrative purposes one may assume that $f(c)$ is well represented by a cubic polynomial.

Functional model for the interdiffusion coefficient

Since $f(c)$ was modelled by a cubic polynomial $(P_5)$ and $\phi(c)$ was modelled by the function $\exp(P_5)$, it

†Initially, it was believed that functions symmetric about $c = 0.5$ would adequately model $\phi(c)$. A symmetric $\phi(c)$ function would greatly simplify the mathematics involved in addition to limiting the possibility of correlation. However these functions were found to be too restrictive since they impose additional constraints on the solution behavior. Even simple models of solution behavior such as the quasi-regular solution model [3] do not lead to symmetric $\phi$--$c$ functions. This is easily proven by using the defining equations for a quasi-regular solution and imposing the condition $\phi(c) = \phi(c + 0.5)$ for $0 < c < 0.5$. It will be found that such a condition leads to additional restrictions on the parameters that constitute the quasi-regular solution model.

‡The term “correlated” in the present context is very distinct from the correlation effects that are discussed under the atomistic theory of diffusion. It is important that this distinction be clearly understood.
follows from equation [3] that the interdiffusion coefficient may be represented by the function \([P_1, \exp(P_2)]\). This function possesses 10 degrees of freedom which implies searching a chi squared space that is eleven dimensional. Clearly the possibility of multiple minima is very strong.

**ACQUISITION OF INTERDIFFUSION DATA**

Data describing the variation of the interdiffusion coefficient with composition for the five isomorphous binary systems (AuNi, CuAu, CoNi, CuNi and AgAu) were obtained from the text by Borovskii [12]. The curves drawn by the authors were magnified and digitized at approximately 100 points. The data were then fitted with a fourth order spline (cubic spline) using a linear least squares approximation approach. Least squares spline fitting splits the interval of interest (in this case \([0, 1]\)) into several subintervals and fits a piecewise polynomial over each subinterval ensuring that several derivatives of the interpolant stay continuous at the boundaries between two intervals [32]. Details on spline fitting are available in a separate publication [33].

After fitting the raw digitized data with a spline, the diffusivity was obtained at 100 evenly spaced points on the interval \([0, 1]\), by evaluating the spline at each of the 100 points. These interpolated data were used as input for further calculations.

**PRELIMINARY RESULTS**

The interpolated diffusivity data were fit with the function \([P_1, \exp(P_2)]\) using the unconstrained non linear least squares (NLLS) routines from the RS/1 package [27]. An unconstrained method was used initially to determine whether there was any need to expend the additional effort in incorporating the constraint functions \([\phi(c = 0) = 1; \phi(c = 1) = 1]\). A good fit was obtained to the \(\bar{D} - c\) data. However, examination of the \(\phi(c)\) function revealed that it was continuously increasing with concentration over the entire interval \([0, 1]\), and that it violated the constraint that \(\phi(c = 1) = 1\). This is a clear example of the non-uniqueness problem alluded to earlier. In this case, a set of functions \(f_i(c)\) and \(\phi_i(c)\) were computed such that their product represents the \(\bar{D} - c\) data, but the function \(\phi_i(c)\) violates the requirements for the \(\phi(c)\) function. The problem occurred because the fitting routines did not incorporate the constraints embodied in equations [9] and [10].

Thus it is clear that one must choose an algorithm that permits the specification of constraints in the non-linear least squares (NLLS) fitting process. The theoretical framework for accomplishing such a task is presented next.

**THEORETICAL BASIS FOR CONSTRAINED NON LINEAR LEAST SQUARES**

Since solution of the least squares problem (linear and non linear) demands the minimization of the least squares error function [equation (8)], it is appropriate to examine the mathematical basis underlying the general problem of minimizing a given function. There are several excellent treatises available on this topic [34–37], the present discussion is based on the book by Gill et al. [34].

The non linear constrained problem NCP is defined as follows [34]

\[
\begin{align*}
\text{minimize } & F(x) \\
& x \in \mathbb{R}^n
\end{align*}
\]

subject to

\[
\begin{align*}
c_i(x) &= 0 \quad i = 1, 2, \ldots, m' \\
c_i(x) &\geq 0 \quad i = m' + 1, \ldots, m
\end{align*}
\]

where \(F(x)\) is the function being minimized and \(c_i\) are the various constraints (equality and inequality constraints). For sake of brevity only the conditions that must be satisfied at the minimum of the objective function are presented. Details of the algorithms for locating a minimum are found in the standard reference [34].

The simplest example of a minimization problem is the unconstrained minimization of a univariate function. From basic calculus it is known that if the function possesses a minimum then at the minimum \((x^*)\) the following conditions hold

\[
\begin{align*}
f'(x^*) &= 0 \\
f''(x^*) &\geq 0.
\end{align*}
\]

In the multivariate case these conditions are replaced by their multivariate equivalents

\[
\|g(x^*)\| = 0
\]

\(G(x^*)\) is positive definite.

There exist several popular algorithms to solve multivariate minimization problems [34]. In the non linear case these are necessarily iterative in nature and may be classified according to the amount of information used to locate the minimum in the \(\chi^2\) space. Some methods use only repeated function evaluations to arrive at the minimum while first order methods use the gradient, and second order methods use the Hessian in addition to both the gradient and the function. In general the higher the order of the method, the better its convergence rate.

The multivariate case is more complex than the univariate case. Firstly, in the multivariate case, the algorithm must choose an appropriate search direction from the current point and secondly it must choose a step length along the chosen direction. In contrast, in the univariate case there is no ambiguity in regard to the search direction.

The situation is further complicated by the addition of constraints. Constrained minimization problems are usually classified based on the type of constraints imposed. Thus one distinguishes between-
linear equality constraints, linear inequality constraints and non linear constraints. In each case the conditions for a minimum to exist differ as do the solution schemes.

Linear equality constraints are usually expressed as a set of equations

\[
\text{minimize } F(x) \text{ subject to } Ax = b. \tag{14}
\]

The matrix \( A \) contains the coefficients corresponding to each constraint. One next defines a matrix \( Z \), the columns of which form a basis for all vectors \( p \) satisfying the condition that \( AP = 0 \). This condition defines all feasible search directions \( p \) from a given feasible point. At the minimum, the following conditions hold

(i) \( Ax^* = b \)

(ii) \( Z^Tg(x^*) = 0 \) or equivalently \( g(x^*) = \hat{A}^Tz^* \)

(iii) \( Z^TG(x^*)Z \) is positive definite. \tag{15}

The quantities, \( \lambda \), are termed the Lagrange multipliers. Solution of a linear equality constraint problem (LEP) with \( "t" \) constraints in \( "n" \) unknowns is equivalent to solving an unconstrained problem in \( n-t \) dimensions. In essence the equality constraints reduce the dimensionality of the problem. All solution algorithms first locate an initial feasible point (i.e. one that satisfies the constraints) and then search for the minimum by moving away from the current feasible point along feasible directions only (i.e. along the vectors \( p \) defined earlier).

In the case of linear inequality constraints \( (Ax \geq b) \) there exists one restriction in addition to the conditions expressed in equation \( [15] \). The additional restriction demands that all the Lagrange multipliers by strictly positive. The popular solution scheme utilized is termed the active set method wherein a subset of the inequality constraints is determined. All the constraints in this subset are binding constraints, i.e. the condition \( Ax = b \) is exactly satisfied for these constraints. Such a set of constraints is termed as the active set or working set. Once an active set has been determined the solution scheme is very similar to that for the non linear equality problem.

A recurring theme in algorithms for minimization is the modelling of the objective function \( F(x) \) by a quadratic function

\[
F(x) \approx g^T(x^*) + \frac{1}{2} x^*^T G x^*. \tag{16}
\]

Equation (16) can be derived by expanding \( F(x) \) as a Taylor series about \( x = x^* \) and neglecting the terms higher in order than the second order term. Clearly such an approximation is local, i.e. it holds only close to the minimum. In practice it is found that iterative schemes usually do converge to the minimum even when the initial point is far from the minimum.

CONstrained FITTING OF DIFFUSIVITY DATA

Earlier in this paper it was decided that the function \( \exp (P_3) \) would be used to model the quantity \( \phi (c) \). In itself, the chosen function \( \exp (P_3) \) only ensures that \( \phi (c) \) is positive but does not satisfy the constraints of equation (10). In order to satisfy the constraint, \( \phi (c = 0) = 1 \), the constant term in the polynomial \( P_3 \) must be set to zero. Hence the function used to represent \( \phi (c) \) must be modified to read as follows

\[
\phi (c) = \exp (\tau_1 c^3 + \tau_6 c^4 + \tau_7 c^3 + \tau_4 c^2 + \tau_9 c) \tag{17}
\]

where \( \tau_i \) are the unknown parameters that constitute the mode and \( c \) is the composition, i.e. \( c \) ranges over the interval \([0, 1]\).

Thus the imposition of the constraint \( \phi (c = 0) = 1 \) reduces the dimensionality of the problem to 9 variables. Similarly the constraint \( \phi (c = 1) = 1 \) leads to the following constraint on the parameters

\[
\sum_{k=3}^{k=9} \tau_k = 1. \tag{18}
\]

There are no constraints on \( f(c) \).

Thus the diffusivity was modelled by the following function

\[
h(x) \approx \left( \sum_{j=1}^{j=9} (\tau_j c^{\ell-1}) \right) \left[ \exp \left( \sum_{k=5}^{k=9} (\tau_k c^{p-k}) \right) \right] \tag{19}
\]

and was subject to the constraint embodied in equation (18).

During the fitting process the data were weighed to reflect the fact that the absolute uncertainty in each data point was identical. Hence the weights were selected as the reciprocal of the absolute value of the ordinate of the data. Based on all of the above, the least squares error function may be defined as follows

\[
E^2 = \sum_{i=1}^{i=100} (\{\bar{D}_i - h_i\} w_i)^2 \tag{20}
\]

where, \( h_i \) is given by equation (19) and \( w_i = 1.0/\bar{D}_i \).

The NAG [37] fortran library subroutines E04UAF and E04VCF were used to locate the minimum. E04UAF uses a sequential augmented Lagrangian method [34] coupled with the Quasi Newton method [34]. Since the method does not utilize the gradient or the Hessian of the objective function, its convergence rate is very slow.\(^\dagger\) In contrast, the subroutine E04VCF uses a sequential quadratic programming algorithm wherein the search direction is the solution of a quadratic programming problem [34]. Since this is a comprehensive routine that utilizes information about the gradient of the objective function the convergence rate is very rapid.\(^\ddagger\) In addition E04VCF makes it easier to monitor the progress of the minimization process.

Since the objective functions \( E^2 \) is non linear in the unknown parameters \( (\tau_i) \) the process of minimization is iterative and commences at an initial guess

\(^\dagger\) A typical run took 2-3 hours on a Micro Vax II.

\(^\ddagger\) A typical run required 7-10 min on a Micro Vax II.
which must be provided. For all the computations conducted in this investigation the initial guess was taken to be the unit vector, i.e. $\tau_i = 1$ for all $i$. This initial guess is clearly not consistent with the constraints imposed. However, the subroutine E04VCF, is powerful enough to locate an initial feasible point† within the first iteration. Subsequent iterations always move from one feasible point to another. With each iteration the subroutine attempts to minimize the value of the least squares error function ($E^2$). When no improvement is possible and the gradient of the objective function is "small", the minimum is said to have been attained and the program terminates. At the minimum, the least squares error, $E^2$, is usually non-zero and is termed the residual. The numerical value of the residual is a measure of the goodness of fit. This is particularly useful in cases where two computed solutions are close to each other. The solution with the smaller residual provides a better fit to the diffusivity data and is therefore selected as the correct solution.

**RESULTS OF CONSTRAINED FITTING**

The salient features and results from the attempts at constrained fitting will be highlighted by discussing the Co–Ni system in detail. Results for the other systems will be presented in the following paper. Throughout this discussion the term **original diffusivity data** will refer to the data obtained by interpolating the spline approximation to the raw digitized diffusivity data.

Diffusivity data for the CoNi system were obtained at 1356°C. Figure 3 compares the fitted and original diffusivity data. The subroutine E04UAF was used for this fit and the only constraint imposed was that embodied in equation (18). From Fig. 3 it is clear that a good fit was obtained to the original diffusivity data. Figure 4 depicts the $\phi(c)$ function corresponding to the fit. Figure 4 suggests that the CoNi system exhibits a positive deviation from ideality. However, it is known, from Hultgren [25], that the CoNi system is ideal, so $\phi = 1$. This implies that the chemical potential results in Fig. 4 are spurious, despite the fact that the computation yielded an excellent fit to the $D$–$c$ data. This is clearly a manifestation of the non-uniqueness problem which had been alluded to earlier. In the present case the product of the computed functions $f(c)$ and $\phi(c)$ provides a good fit in the diffusivity data but the computed function $\phi(c)$ does not represent the system thermodynamics.

The non-uniqueness problem

The problem of non-unique solutions undermines the attempts to compute chemical potentials by leading to multiple solutions. Non unique solutions arise since the problem at hand is intrinsically non unique.

†A feasible point is one where all the constraints are satisfied.
a strategy is presented to pick the correct minimum from the set of computed minima.

Obtaining a set of feasible minima

No general mathematical technique exists for simultaneously locating all the minima of the non linear least squares problem. Therefore, it was decided to search subsets of the $E^2$-c space for the minima. However, since the main goal of this investigation was to obtain thermodynamic data it was decided to focus on subsets of the $E^2$-c space instead of the $E^2$-c space. Thus, the following iterative strategy was adopted. The algorithms were modified to incorporate additional constraints, the general form of which was as follows

$$l_i \leq \phi(c_i) \leq u_i \quad i = 1, \ldots, 10.$$  \hspace{1cm} (21)

These constraints were imposed at ten compositional points evenly spaced on the interval [0, 1]. The lower bounds at each compositional point were identical, i.e. $l(c = 0.5) = c = 0.7) = l(c = 0.9)$ as were the upper bounds. By varying the lower and upper bounds, $l_i$ and $u_i$ respectively, it was possible to constrain the $\phi(c)$ function to different regions of the two dimensional $\phi(c)$-c space. For each set of constraints a complete non linear least squares fit (in $E^2$-c space) was carried out. After each such run the goodness of the fit to the original diffusivity data was examined. If a satisfactory fit was obtained then the corresponding $\phi(c)$-c function was added to the set of possible minima. In this manner by varying the constraints, $l_i$ and $u_i$, a set of possible minima was obtained. In essence, the additional constraints in equation (21) restrict the search to a subset of the $\phi(c)$-c space. By searching several subsets of the space one ends up with the set of possible minima. It should be noted that the constraints, $l_i$ and $u_i$, must be consistent with the “natural” constraints on $\phi(c)$, which are embodied in equation (10). In other words the choice of $l_i$ and $u_i$ must be such that the function $\phi(c)$ tends to unity at either end of the interval [0, 1]. Hence a choice of $l_i = 2.0$ and $u_i = 4.0$ would be incorrect.

Figure 5 depicts several such minima for the CoNi system. The figure plots $\phi(c)$ as a function of c for different sets of upper and lower bounds. The legend in the figure corresponds to the upper and lower bounds for each computation. In each case an excellent fit to the original diffusivity data were obtained; similar to the fit depicted in Fig. 3. The goodness of fit can be judged from the table of residuals (Table 1).

The $\phi(c)$-c data in Fig. 5 was integrated to yield $\gamma$-c data. Numerical integration was carried out using a fourth order Runge–Kutta–Nyström technique [38], and the boundary condition that $\gamma = 1$ at $c = 1$. The step size used was 0.0001 and the data were plotted every 2%. Using this integration scheme the activity composition curves depicted in Fig. 6 were generated. In this manner a set of the minima were obtained.

---

**Table 1. Results of least squares fitting for the CoNi system using the modelling function $P(c \cdot \exp(P(c)))$**

<table>
<thead>
<tr>
<th>Bounds</th>
<th>Final activity curve</th>
<th>Residual at minima</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0.00, 1.00]</td>
<td>Positive deviation</td>
<td>$0.5961 \times 10^{-2}$</td>
</tr>
<tr>
<td>[0.20, 1.00]</td>
<td>Positive deviation</td>
<td>$0.7657 \times 10^{-3}$</td>
</tr>
<tr>
<td>[0.40, 1.00]</td>
<td>Positive deviation</td>
<td>$0.6332 \times 10^{-3}$</td>
</tr>
<tr>
<td>[0.60, 1.00]</td>
<td>Positive deviation</td>
<td>$0.6966 \times 10^{-3}$</td>
</tr>
<tr>
<td>[0.80, 1.00]</td>
<td>Positive deviation</td>
<td>Non convergence</td>
</tr>
<tr>
<td>[0.90, 1.00]</td>
<td>Positive deviation</td>
<td>$0.7753 \times 10^{-2}$</td>
</tr>
<tr>
<td>[0.99, 1.10]</td>
<td>Nearly ideal</td>
<td>$0.7804 \times 10^{-2}$</td>
</tr>
<tr>
<td>[1.00, 1.125]</td>
<td>Nearly ideal</td>
<td>$0.7844 \times 10^{-3}$</td>
</tr>
<tr>
<td>[1.00, 4.00]</td>
<td>Negative deviation</td>
<td>$0.6192 \times 10^{-2}$</td>
</tr>
<tr>
<td>[1.00, 6.00]</td>
<td>Negative deviation</td>
<td>$0.6192 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

---

Fig. 5. Complete set of feasible solutions for the CoNi binary at 1356°C on a $\phi(c)$-c diagram. Each solution corresponds to a minimum in the $\phi(c)$-c space. Legend in the figure refers to the upper and lower bounds imposed on $\phi(c)$ to compute each minimum. The solutions for the bounds [1, 4] and [1, 6] are coincident. The solution computed on imposing the bounds [0.99, 1.10] was taken to represent the thermodynamics of the CoNi system.

Fig. 6. Complete set of feasible solutions for the CoNi binary at 1356°C on an activity composition diagram. This data were obtained by integrating the data in Fig. 5. The solutions for the bounds [1, 4] and [1, 6] are coincident. The solution computed on imposing the bounds [0.99, 1.10] was taken to represent the thermodynamics of the CoNi system.
It should be noted that this scheme does not determine all the minima in \( E^2-c \) space. As mentioned earlier, there is no guarantee that all of the minima have been computed. However, the scheme does locate several minima in the \( \phi(c)-c \) space, which is the space of interest. The possibility does exist that there are additional minima in the \( \phi(c)-c \) space that have not been located. For some of these minima the activity-composition curves will lie very close (within the error associated with activity measurements) to the set of curves determined by a previous computation and hence one will not commit significant errors. On the other hand there are likely to be solutions that are far away from the actual activity curve. There is little point in expending the effort to compute these solutions.

**Selection of the appropriate minimum**

From Fig. 6 it is clear that if the true activity were known at any one composition then one could easily select the correct activity composition curve from the set of possible minima. This implies that a single experimental point in conjunction with the \( D-c \) curves and the proposed computational scheme can yield the entire \( c-c \) curve for the particular alloy. The set of minima may be examined prior to the experiment to choose an appropriate composition range within which an experiment could be performed. For example, for the CoNi system [Fig. 6], one might choose to conduct an experiment in the composition range \([0.35, 0.55]\) since the various activity-composition curves exhibit the maximum separation in this range of compositions. One might also accomplish this if the slope of the activity composition curve was known at the dilute end. It will be recalled [3, 26] that the slope of the activity composition curve corresponds to the Henry law line for the solution.

In the present investigation, it was decided to use the Henry's law line in picking the correct solution. The slope of Henry's law can be easily estimated from data on heats of solution. Heats of solution for liquid binary metal systems have been theoretically estimated by Miedema et al. [40].

Miedema's model is atomistic and arrives at the heat of formation by modelling the contributions of the change and discontinuities in electron charge density (that occurs on alloying) to the heat of formation. Based on this model, Miedema et al. [40] have estimated the heats of formation at infinite dilution for several liquid binary metal systems. A good match was obtained between their predictions and the measurements reported in the literature.

Although Miedema's model works well for liquids it is less applicable to solids since it does not account for the strain energy [41]. For solids, Russell [41] has proposed a simple scheme that permits one to estimate the enthalpy of formation in the solid state knowing the latent heat of fusion of the solute, its melting temperature and the binary phase dia-

gram. Russell's [41] scheme leads to the following expression

\[
\Omega^b = \Omega^b + L^B \left( 1 - \frac{T}{T^B} \right) + RT \ln \left( \frac{y}{c} \right)
\]

where \( \Omega^b \) represents the partial molar enthalpy of formation in the solid state, \( Q^b \) represents the partial molar enthalpy in the liquid state.

It was decided to utilize Miedema's theoretical predictions in conjunction with Russell's technique to estimate the slope of the Henry's law line in order to select the correct minimum from the set of possible minima (Fig. 6). It was hoped that at the very least a reasonable estimate of the Henry's law line could be obtained. It was further assumed that the entropy of solution at infinite dilution was ideal. The activity coefficient was computed using the following equation

\[
\Omega^b = -RT \ln y^b.
\]

By using this approach, the Henry's law constant was estimated for the CoNi system.

For the CoNi system Miedema's data predicts that the heat of formation at infinite dilution is \(-1\) kJ/mol for Ni in liquid Co. Since Miedema et al. did not report the temperature to which their predictions correspond, it was assumed that in each case the \( \Omega^b \) data corresponded to the melting point of the solvent. Based on this premise and using equations (22) and (23), the Henry's constant for Ni in Co was computed to be 0.95 at 1356°C. This line has been plotted in Fig. 6 for comparison. Referring next to Fig. 6, it is seen that the solution closest to the estimated Henry's law line is the one which was computed using the bound: \(0.99 < \phi(c) < 1.10\). Hence the other solutions (sets of coefficients) may be discarded as spurious minima. From Fig. 6 it is clear that the chosen set of coefficients predict that the CoNi system is nearly ideal which is in excellent agreement with the experimental data reported by Hultgren et al. [25]. This is clearer from Fig. 7 which plots \( \ln(y) \) as a function of composition for the selected solution.

**DISCUSSION**

Thus by considering the CoNi system as an example, the proposed scheme for calculation of activities from interdiffusivities has been illustrated. The only other piece of information used was the theoretical predictions of the heats of formation. One may also have used the experimentally measured activity at one particular composition. This paper has presented the background, theoretical basis and the proposed computational scheme for estimating thermodynamic activities from interdiffusion coefficients. The next paper discusses the results of similar computations conducted on ten additional systems. For these 10 systems the computed activities will be compared with the experimental data presented by Hultgren [25]. These ten systems will help
Acknowledgements—The authors are grateful to Professor J. F. Elliott, Professor K. C. Russell and Professor N. L. Trefethen for helpful discussions and constructive criticism. The authors also wish to thank Professor J. S. Kirkaldy, McMaster University, and Dr. J. Cahn, NIST for reviewing early drafts of this paper. Financial support was provided by an NSF creativity extension, under grant number DMR-8502411.

CONCLUSIONS

A non linear least squares based strategy was successfully implemented to estimate activity composition data from interdiffusion data. The proposed scheme requires the $\tilde{D} - c$ data and one additional piece of information, such as the experimentally determined activity at a particular composition or the slope of the Henry's law line at infinite dilution, to estimate the entire activity-composition curve. Due to its inherent non-uniqueness, the scheme leads to multiple solutions. The correct solution can be selected from the possible set of solutions by using an additional piece of information. In the present paper, by using theoretical predictions of the heat of formation to estimate the slope of the Henry's law line, it was possible to select the correct solution. The scheme was successfully demonstrated on the CoNi system which was shown to obey Raoult's law. Excellent agreement was obtained between the predicted values and the data reported in the literature.

REFERENCES


