Available online at www.sciencedirect.com
ScienceDirect
Acta Materialia 56 (2008) 1243–1256
Acta MATERIALIA
www.elsevier.com/locate/actamat

The Soret effect in diffusion in crystals

Robert J. Asaro [a,*], Diana Farkas [b], Yashashree Kulkarni [a]

[a] Department of Structural Engineering, University of California, San Diego, La Jolla, CA 92093, USA
[b] Department of Materials Science and Engineering, Virginia Tech, Blacksburg, VA 24061, USA

Received 10 November 2007; received in revised form 14 November 2007; accepted 15 November 2007
Available online 2 January 2008

Abstract

The Soret effect, as it occurs in the diffusion of solutes in crystals, is analyzed using the principle of microscopic reversibility whereby an approach is developed for interpreting and computing Q^*, the heat of transport. To compute the transport of energy during the diffusion jump process, and then Q^*, the processes of thermal activation to the transition state and the decay from the transition state are considered to be inverses; thus the net process is reduced to the analysis of the purely mechanical decay process. Molecular statics and dynamics are then suggested as the means to simulate the decay process and the case of carbon diffusion in body-centered cubic α-iron is so analyzed as an example. Our results show that, for the case $Q^* \approx -Q_m$, where Q_m is the activation energy for carbon diffusion, this is in agreement with experimental evidence reported in the literature. Thus both the sign and magnitude of Q^* are correctly predicted. Various cases, such as the diffusion of substitutional solutes and vacancies, are also considered within our approach along with implications for future study.

Keywords: Soret effect; Thermal diffusion; Heat of transport

1. Introduction

Under isothermal conditions, the diffusion of atoms in crystals is controlled by the height and local configuration of the saddle point, i.e. the transition state; it is not important to know the details of how the diffusing particles are activated to the saddle point. However, the thermal flux accompanying the particle flow is dependent on the details of the activation process (the Soret effect), and so is the particle flow that occurs in a temperature gradient. The two effects are connected by an Onsager symmetry relation [1–3].

Wirtz [4] extended the theory for isothermal particle diffusion to particle diffusion in a temperature gradient by assuming that, at any point along the activation path, the isothermal formula

$$\exp\left(\frac{-\Delta Q}{kT}\right) \tag{1}$$

could be used for the probability of activation, where ΔQ is the increment in activation energy at that point and T is the temperature of the plane through that point normal to the temperature gradient. In this formulation, the basic problem is to decide on the spatial distribution of the total activation energy, Q_m: what increments should Q_m be considered to be made of, and precisely where are they supplied? Or, stated differently, what is the appropriate temperature, T, to be used in the activation formula of Eq. (1) for each particular increment ΔQ? So far it appears that this problem has not been approached in a fundamental way. Rather, one has essentially fitted Wirtz's [4] description to account for experimental heats of transport. Thus Shewmon [5] rationalizes the negative heat of transport for interstitial C in α-Fe by describing the activation process as one of activating the final octahedral site in the diffusion jump to accommodate the C atom. Descriptions picturing the C atom to be activated in its original site sufficiently that it may jump into a neighboring site would, according to the Wirtz theory, imply a positive heat of transport, in disagreement with experiment. Similarly,

* Corresponding author. Tel.: +1 858 534 6888; fax: +1 858 534 6373.
E-mail address: rasaro@ucsd.edu (R.J. Asaro).

1359-6454/$34.00 © 2007 Acta Materialia Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.actamat.2007.11.019

Oriani [6,8] has described vacancy diffusion by an activation process involving half vacancies, which again gives fair agreement with some observations on the heat of transport. Oriani [6] also attempted to extend the theory to separate the heat of transport, Q^*, into a "lattice part", Q^*_L, and a "chemical part", Q^*_C, where the first part is associated with lattice distortions and the second part with short range chemical binding to the host atoms at particular sites favorable for bond formation, and argued that the chemical contribution would always be positive, i.e. $Q^*_C > 0$, and of the order of the chemical binding energy in the bonds broken in the diffusion process.[1]

The above theories and rationalizations of experiment are most interesting and informative about the details of the activation process when experimental results are known. Also, the contribution of very local chemical binding effects to the Soret effect can be said to be fairly well understood: clearly a bond can only be broken by an activation process centered on the bond, and when the chemical effect is short range, the position of activation is then quite well defined. But when it comes to the lattice part, which is less confined spatially, the present theories are quite lacking as they stand. They are not able to predict a priori, from known interatomic potentials, what should happen; they can, at best, be used to reconstruct the activation sequence when experimental heats of transport are known.

The basic problem is this: the theories are given in terms of activation processes, which are comparatively rare, and "improbable" processes, whose most probable development is hard to predict directly. One is said to think in terms of averages and highest probabilities; the most probable realization of some event that is in itself improbable is conceptually difficult.

However, using the theorem of time reversibility, the problem can be recast into such a form that one can more safely apply the sort of probability considerations one is used to. A statement of the theorem of time reversibility appropriate for the present purposes would run as follows: the most probable path of the improbable event of activation is the time inverse of the most probable path of decay.[2] This is exactly the reasoning behind the Onsager [1] theory for the famous reciprocal relations $D_{ij} = D_{ji}$ in linear transport equations expressed in terms of conjugate fluxes and forces. According to this principle, one may use the following procedure: the diffusing atom is placed at the saddle point, and the configuration is then slightly disturbed[3] so that it becomes unstable and decays by the emission of lattice vibrations into the surrounding crystal. The forces causing accelerations and emission of radiation will far exceed the forces due to thermal vibrations until the decay is essentially complete and the decay energy has been spread out and become submerged in the overall thermal energy of the lattice. Thus, to a good approximation, the dynamics of the decay can be studied by purely mechanical considerations, to decide such questions as which atoms are left in highly excited vibrations after the first part of the decay. The highly excited atoms can then be regarded as sources radiating phonons into the surrounding crystal. When the radiation is sufficiently symmetric, it does not give rise to net energy transport in any one direction; the "source" is the "center of mass" of the energy radiated.[4] When in this manner the mode of decay and the energy displacements in the decay have been determined, the mode of activation is directly obtained as the inverse process, by the principle of time reversibility. Combining the two processes, activation and decay, the complete energy displacement accompanying the diffusion jump is formed.

Various cases of diffusion will be discussed qualitatively from this point of view. The discussion supports Shewmon's [5] description of C in α-Fe and γ-Fe. The description of vacancy diffusion will be extended. The heats of transport in some systems for which measurements are not yet available will be discussed and predicted qualitatively. But, perhaps most important, since the discussion is based on the consideration of decay processes, the conclusions can largely be explored and revised by classical simulations on the atom movements in decay when interatomic forces are well enough known. Simulations conducted via either molecular dynamics or finite temperature quasi-continuum methods are among the possible approaches. In what follows we employ molecular dynamics simulations to describe and verify the procedure for the most interesting case of C diffusion in α-Fe.

A central theme of our discussion concerns the elucidation of the sign and magnitude of the heat of transport appearing (below) in Eq. (3) and the transport equations, Eqs. (7) and (8). For this we note the comments of Darken

[1] Oriani [6] also considered a third component, Q^*_v, ascribed to the vibrational energy residing in the local vibrational modes caused by the presence of the diffusing atom in the host lattice. Since such localized vibrational energy must be one part of the equilibrium enthalpy of solution, and since the equilibrium enthalpy of transport is subtracted from the total energy of transport to give the proper heat of transport, it appears that this term should not be counted in the heat of transport.

[2] Expressed this way, the principle of time reversibility has wide applicability in phenomena of central interest in chemistry and materials science, such as nucleation phenomena. For example, it follows that those subcritical droplets which do grow up to the critical size in a supersaturated vapor are, on average, cooler than the ambient vapor (Feder et al. [11]). Similarly, one could show that precipitation in a supersaturated solid solution occurs by an initial fluctuation increasing the concentration over some region and followed by the formation of a subcritical crystallite which grows by consuming the excess concentration surrounding it so that, at the critical size, the crystallite is surrounded by a uniform solution and finally growth occurs, leading to a depleted zone around the growing crystallite.

[3] According to standard diffusion theory, the diffusing atom has translational energy $\sim kT$ through the saddle point, and the host atoms are in vibration about the saddle point configuration. However, the precise nature of the disturbance bringing the configuration out of equilibrium is quite immaterial for the mode of subsequent motion.

[4] In some cases, asymmetries in the radiation may be important and the center of mass of the energy radiated will then depend on how fast the radiation loses its net momentum by collision with thermal phonons. This may be an important effect in vacancy diffusion, where the vacancy will throw a "shadow", as will be discussed later in this paper.

and Oriani [8]: "that the heat of transport is associated, and in principle calculable from, the manner in which, on the average, a statistical fluctuation concentrates energy about a small region, and from the manner in which the activation energy is dissipated into the lattice after the unit step of diffusion".

2. Kinetic development of the phenomenological equations

Consider an atom diffusing from A to B over a free energy barrier, whose saddle point is at C (Fig. 1a). For simplicity, let the plane O–O through C be a reflection plane, so that decay from C to B can freely considered as the inverse process of activation from A to C.

Place the diffusing atom at the saddle point and observe it falling down to B. Suppose that the host atoms around the saddle point are left behind in a highly excited configuration; that is, they are not able to transfer all the distortional energy to the diffusing atom by accelerating it as it leaves. Some energy Q_1 is then radiated out from the saddle point region. Now suppose it is radiated out symmetrically. The diffusing atom has accelerated down to B and causes a highly excited state about B, and an energy Q_2 is radiated out from B (see Fig. 1b). Clearly

$$Q_m = Q_1 + Q_2 \tag{2}$$

when the two processes complete the relaxation.

With the decay process occurring as just described, it follows from time reversibility that the activation from A to C occurs by first an activation of Q_2 at A, exciting the initial site, and that next the host atoms about the saddle point are excited by Q_1, sufficient to let the excited diffusing atom get through (see Fig. 2). By a fluctuation, thermal energy in the surrounding crystal converges on A and excites the configuration about the initial site A by Q_2; this is the time inverse process of radiation of energy Q_2 from the excited destination site. The activation at C could be described similarly.

Clearly, the net effect of energy absorption and radiation for one atomic jump from A to B is that an energy Q_2 is absorbed from the surrounding lattice at A and given back to the lattice at B, the total effect of which is displacing the "center of gravity" of an amount of energy Q_2 in the surrounding lattice by a distance b, as depicted in Fig. 2. Thus, under isothermal conditions, a net particle current J_p would be accompanied by a heat current, J_Q, in the surrounding lattice

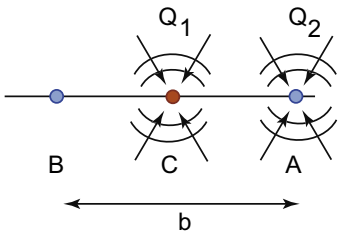


Fig. 2. Time reversed radiation pathway for the process of activation, the time inverse of Fig. 1b.

$$J_q = Q^* J_p \tag{3}$$

where $Q^* = Q_2$, Q^* is the heat of transport [9,10]. Only Q_2 contributes as obviously the two processes involving Q_1 at C cancel each other out and give no net energy transport into the surrounding medium.

We now derive the particle flux, following Wirtz [4] and Shewmon [5], for the case with prevailing concentration gradient dc/dx and temperature gradient, dT/dx (see Fig. 3). For jump frequencies we use the simple formula

$$v = v_0 \exp\left(\frac{-Q}{kT}\right) \tag{4}$$

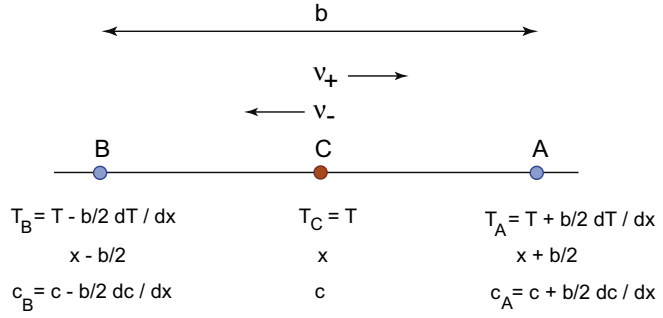


Fig. 3. Schematic of combined concentration and temperature gradients over the path A–B. Jump frequencies v_+ and v_- are indicated to the left and right, respectively.

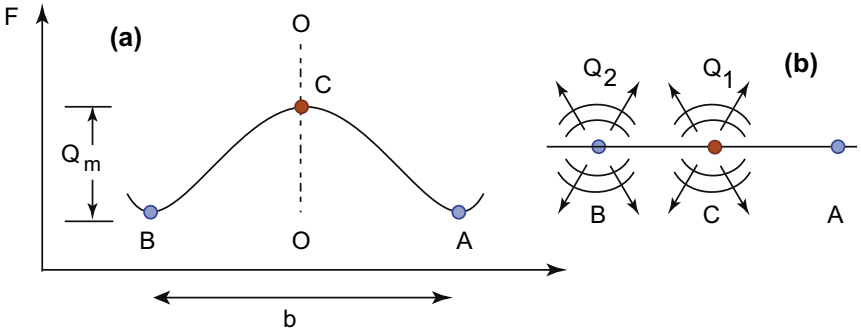


Fig. 1. Thermally activated pathway for an atom diffusing from point A to point B through saddle point at point C.

where T is the appropriate average reservoir temperature at the place from where the energy Q is absorbed. Jump frequencies to the right and left are denoted by v_+ and v_-, respectively. The particle current is

$$J_p = b(c_B v_+ - c_A v_-) \tag{5}$$

where

$$v_+ = v_0 \exp\left(\frac{-Q_2}{kT_B}\right) \exp\left(\frac{-Q_1}{kT_C}\right) \tag{6a}$$

$$v_- = v_0 \exp\left(\frac{-Q_2}{kT_A}\right) \exp\left(\frac{-Q_1}{kT_C}\right) \tag{6b}$$

An attempted jump to the right occurs with a frequency $v_0 \exp(-Q_2/kT_B)$, but only a fraction of times, $\exp(-Q_1/kT_C)$, is the saddle point prepared to let the particle through. v_- is described similarly.

Linearizing Eqs. (5) and (6) and using the above result $Q^* = Q_2$, we obtain

$$J_p = -D\frac{dc}{dx} - D\frac{Q^*}{kT^2} c \frac{dT}{dx} \tag{7}$$

where $D = v_0 b^2 \exp(-Q_m/kT)$.

Inserting Eq. (7) – after deleting the second term in it, since we want here the mass flux under isothermal conditions – into Eq. (3) and then adding an independent lattice thermal conduction term, $-\kappa dT/dx$, we obtain for the thermal flux

$$J_q = -DQ^* \frac{dc}{dx} - \kappa \frac{dT}{dx} \tag{8}$$

Eqs. (7) and (8) have the correct reciprocal relationship (when one recalls the proper definition of the "forces" to be used) and agree in form with the general phenomenological equations for transport of matter and heat. This shows that indeed the consequences of time reversibility have been correctly incorporated in the model on which the kinetic description was based. It is worthwhile to demonstrate this consistency, and thus firmly establish a connection to phenomenological thermodynamic theory. Consistency with phenomenological theory is also demonstrated in Appendix A, where additional insight is provided into the form of the flux equations, i.e. Eqs. (7) and (8).

Consider a system composed of two compartments separated by a porous and diathermal wall of thickness d. The compartments exchange heat and matter, but are otherwise isolated; this implies a fixed volume so that no external work is exerted on them. The fluxes of matter, J_p, and heat, J_q, are expressed in terms of conjugate forces via the phenomenological linear relations

$$J_p = L_{pp}\delta(-\mu/T) + L_{pq}\delta(1/T) \tag{9a}$$

$$J_q = L_{qp}\delta(-\mu/T) + L_{qq}\delta(1/T) \tag{9b}$$

where the δS have the meaning that positive jumps in their arguments promote positive conjugate fluxes. Noting that the first law of thermodynamics, for the case of either compartment, reads as

$$dU = TdS - pdV + \mu dn \tag{10}$$

we have

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{V,n} \quad \text{and} \quad \frac{-\mu}{T} = \left(\frac{\partial S}{\partial n}\right)_{V,U} \tag{11}$$

which provides the conjugate force vs. flux pairs used in the flux relations. Thus

$$\delta(-\mu/T) = -d\frac{\partial(\mu/T)}{\partial x} \quad \text{and} \quad \delta(1/T) = -d\frac{\partial(-1/T)}{\partial x} \tag{12}$$

The flux laws of Eq. (9) thereby become

$$J_p = -D_{pp}\frac{\partial(\mu/T)}{\partial x} - D_{pq}\frac{\partial(-1/T)}{\partial x} \tag{13a}$$

$$J_q = -D_{qp}\frac{\partial(\mu/T)}{\partial x} - D_{qq}\frac{\partial(-1/T)}{\partial x} \tag{13b}$$

where $D_{ij} = d \times L_{ij}$.

To make contact with the flux laws derived as Eqs. (7) and (8), let

$$\mu = \mu_0 + kT \ln c \tag{14}$$

as in a perfect solution; μ_0 here is a reference chemical potential.[5] When using the expression for μ in the flux laws, however, we set $\mu_0 = 0$ since we note that an energy flux $J_p\mu_0$ is inherent with a particle current and we are interested in only the "extra" energy current. Thus in Eq. (13) we set $\mu = kT\ln c$. Later, in the Appendix A, we re-derive the flux vs. force (i.e. gradient) relations of Eqs. (7) and (8), starting with a standard statement of these kinetic laws and introducing the concept of a *reduced energy flux* as defined by Howard and Lidiard [7] following, e.g. deGroot and Mazur [3]. With $D_{ij} = D_{ji}$, as follows for Onsager's reciprocity relations, we confirm that our flux relations given in Eqs. (7) and (8) indeed conform to the Onsager reciprocal structure.

For the sake of simplicity, it was assumed that an excited center decays by emitting radiation symmetrically with respect to the origin of the excited center. In this approximation

$$Q^* \leqslant Q_m \tag{15}$$

the equality corresponding to $Q_1 = 0$ and $Q_2 = Q_m$ and the energy Q_2 displaced the same distance as the diffusing atom in one atomic jump. Oriani [6] asserts that the inequality of Eq. (15) should hold.

However, the assumption of symmetric radiation is not generally valid. Consider the radiation of Q_1, after the diffusing atom has come down on the left hand side, at B. The presence of the diffusing atom to the left of the excited saddle point atom introduces an asymmetry factor; if the diffusing atom facilitates the transmission of energy, the

[5] It can readily be confirmed that what is shown for the case of a perfect solution is also true for an ideal solution for which $\mu = \mu_0 + kT\ln(\gamma c)$ so long as the activity coefficient, $\gamma \neq \gamma(c, T)$.

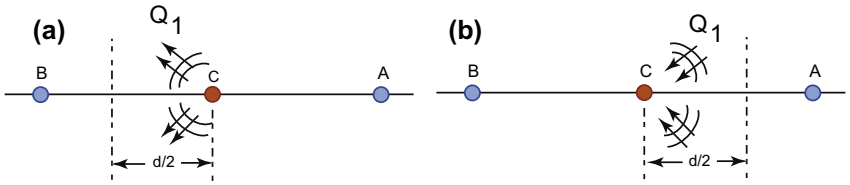


Fig. 4. Schematic of combined concentration and temperature gradients over the path A–B. Jump frequencies v_+ and v_- are indicated to the left and right, respectively.

radiation will have directionality to the left, and the "center of mass" of radiated energy comes to rest some distance, $d/2$, to the left of the saddle point, the distance depending on the interaction with the thermal phonons (see Fig. 4a). It follows by time reversibility that the activation process should be considered as absorption of energy from the right, taken from a region $d/2$ to the right of the saddle point, from a reservoir with the mean temperature of the plane $d/2$ to the right of the saddle point (see Fig. 4b). Treating this model by the same methods as before, transport equations of the correct general form again result, but with

$$Q^* = Q_2 + \frac{d}{b}Q_1 \tag{16}$$

It is seen that if $d > b$, then $Q^* > Q_2$. This demonstrates that no general significance should be attached to the inequality of Eq. (15); although it is usually obeyed, cases not agreeing with the inequality may be discovered.

With strong asymmetries in the radiation, a temperature dependence in Q^* would be expected, since the heat displacement due to asymmetry, say d in Eq. (16), would be temperature dependent. As d may be either positive or negative, depending on the sense of the asymmetry, the total Q^* may either increase or decrease with increasing temperature.

In the remaining discussion asymmetries in radiation will be ignored as a first approximation, except in the case of vacancy diffusion, where it appears that the asymmetry may be significant.

3. Model considerations for decay

Consider first that the host saddle point atoms are two balls held together by springs with a spring constant α, and that the saddle point configuration consists of the diffusing atom inserted in unstable equilibrium between the two balls pushed aside (see Fig. 5). Consider that the two host atoms can only move vertically, that they are stabilized horizontally by atoms not shown in the figure. It can be shown that the diffusing atom will be in contact with the host atoms until the host atoms come in contact as indicated in Fig. 5c. The potential energy of the saddle point configuration is $2 \times 1/2\alpha r^2$; thus

$$\frac{1}{2}mv^2 + MV^2 = \alpha r^2 \tag{17}$$

From the fact that contact exists until the host atoms are in contact there is the geometrical condition

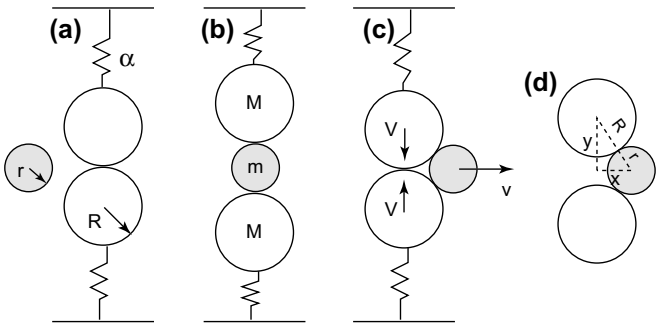


Fig. 5. Model scheme for the transfer of kinetic energy from host to a diffusing atom. In (d) note that $x^2 + y^2 = (R + r)^2$ and thus $x\mathrm{d}x/\mathrm{d}t + y\mathrm{d}y/\mathrm{d}t = 0$. With $V = \mathrm{d}(-y)/\mathrm{d}t$ and $v = \mathrm{d}x/\mathrm{d}t$, this results in the relation $V/v = x/y$ which is the basis for Eq. (18) below.

$$RV = (2Rr + r^2)^{1/2}v \tag{18}$$

Thus, the fraction of the total saddle point energy carried away by the diffusing atom's kinetic energy is

$$\frac{\Delta E}{Q_{\mathrm{act}}} = \frac{mR^2}{2M(r^2 + 2Rr) + mR^2} \tag{19}$$

or, in the range of most reasonable parameters,

$$\frac{\Delta E}{Q_{\mathrm{act}}} \approx \frac{mR}{4Mr} \tag{20}$$

with M and R typical of Fe atoms and m and r of C atoms, say, this is a negligible fraction. With $R = r$ and $M = m$, typical of self-diffusion, the ratio is still only $\sim 1/7$. One can conclude, then, that in this model the saddle point atoms would be left in a highly agitated state; that they will recoil, recompress and vibrate until the energy has radiated via neighbors into the surrounding crystal. In a first approximation, the energy can be said to be radiated from the point O, the saddle point.

In actual cases, since more than two atoms usually make up the immediate saddle point environment, the transfer of kinetic energy to the diffusing atom will be even less than estimated in the above model. If the diffusing particle is charged, and the forces are not mainly contact repulsion but due to internal electric fields, say, the situation is entirely different. In such a case a major fraction of the saddle point energy may be transferred to the diffusing ion during decay. The above model considerations are pertinent when the major part of the saddle point energy can be considered as strain energy.

The above considerations show quite conclusively that only a small fraction of the saddle point energy can be carried as kinetic energy by the diffusing atom. However, the situation may still be radically different from the case described above. The above model would apply directly when there is no severe hindrance for the diffusing atom to reach its final position; that it may just "roll" quietly into the final position after leaving the saddle point atoms. The other extreme case, which will now be discussed, is that the diffusing atom cannot move out of the saddle point without pushing other atoms aside (see Fig. 6). Again, the diffusing atom itself will carry negligible energy but will transmit energy from the saddle point atoms to other atoms; thus the excess energy liberated on decay need not be radiated out from the original saddle point atoms alone. Consider the model depicted in Fig. 6, which is self-explanatory. Again host atoms are considered to have no horizonal motion. As the original saddle point configuration shown in Fig. 6a relaxes and the original saddle point atoms relax, the two other atoms are pushed out; the configuration then goes through the symmetrical configuration sketched in Fig. 6b with velocities as indicated by the arrows, and then reaches the configuration sketched in Fig. 6c, or more precisely a point somewhat below that configuration because of some energy loss. The process then reverses and the configuration returns to near the original saddle point configuration, then reverses again, etc. The system is an oscillator composed of all five atoms, and energy is radiated out symmetrically with respect to the center of the oscillator, O, which is also the final position of the diffusing atom, sketched in Fig. 6b, with all atoms at rest.

The coupling of the oscillator to neighboring atoms will make the oscillator strongly damped with a decay time of, say, 2–3 cycles. Thus there will be some asymmetry in radiation, involving roughly the energy initially radiated out in about 1/4 cycle since the oscillator starts out from an asymmetric configuration, as shown in Fig. 6a. However, we shall assume that to a good approximation all the energy can be considered to be radiated out from O.

It follows from time reversibility that activation should be regarded as a gradual excitation of the oscillator to higher and higher amplitudes, energy being drawn symmetrically toward it O from the surrounding crystal. The details of the activation process of this model could be treated by the theory of stochastic processes in the manner of Seeger and Donth's treatment of activation of dislocation oscillators [12]; see also Jossang et al. [13].

4. Interstitials in body-centered cubic (bcc) lattices

Consider first the case of interstitials occupying octahedral equilibrium positions, and with a saddle point at the tetrahedral site. Except for some questions regarding the saddle point, which will be commented on later, C and N in α-Fe belong to this class. A schematic of this situation is given in Fig. 7. At equilibrium, the interstitial, indicated as "X", extends the "dumbbell" of host atoms with the octahedral site's center labeled O. The dumbbell length with no extension is marked with the tick marks (see Fig. 7a). Now move the interstitial to the saddle point; dumbbell O contracts by x, as indicated in Fig. 7b, whereas dumbbell O' extends by y. Since T is a saddle point, the energy released in dumbbell O must be less than the energy put into dumbbell O'; otherwise T would have been the equilibrium position. As dumbbell O is severely stretched originally, more energy per unit length of contraction is released in dumbbell O than is put into dumbbell O' per unit length of expansion. It follows that

$$y \gg x \tag{21}$$

This is also reasonable geometrically for the larger size interstitials which prefer O sites. Of course, displacement components other than those depicted in this model should strictly speaking be included. However, the model includes the most essential features and is, in fact, verified later via our molecular static computations discussed in Section 7.1.

Now let the activated state collapse, the interstitial moving back to O. From energy considerations, it follows that the interstitial must be confined to the interval $T - T'$ in the resulting motion. Thus, the amplitude of vibration of the central dumbbell cannot be more than $\sim x$, while the amplitude of dumbbell O' must be $\geqslant y$ since it will "overshoot" contact due to the kinetic energy released during collapse unless that energy has been transferred to dumbbell O'' extending it. This is the crucial question: will dumbbell

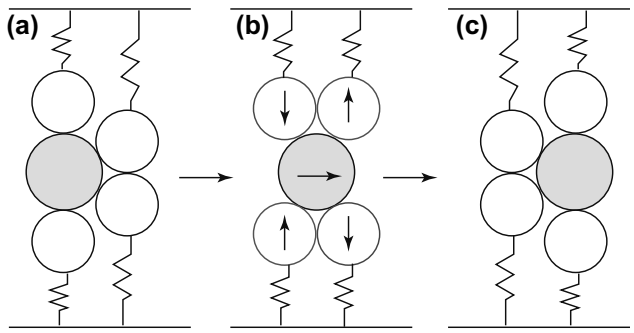


Fig. 6. Model for a diffusing atom transmitting energy to host atoms.

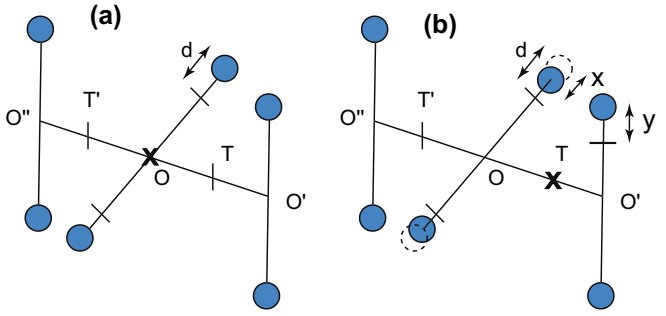


Fig. 7. (a) Model for a diffusing atom occupying an octahedral site in a bcc lattice. (b) Diffusing atom at a tetrahedral saddle point position.

O'' extend as dumbbell O' contracts so that the two dumbbells make up the ends of an oscillator symmetric about O, similar to Fig. 6, or can the contraction of dumbbell O' only cause relatively small extensions in dumbbell O'', so that the amplitudes of vibration in dumbbell O'' will be much smaller than the amplitudes of vibration in dumbbell O'? Geometrical considerations indicate that the motion of two dumbbells as far separated as O' and O'' cannot be effectively coupled by an interstitial to make up an oscillation like the one depicted in Fig. 6. It can be concluded, then, that the amplitudes of vibration during decay will be much smaller in dumbbell O' than in both dumbbells O and O''. The interstitial will roughly oscillate in the interval $O - T$ while dumbbell O' is executing violent vibrations and radiating energy. Practically the entire relaxation energy is radiated from dumbbell O'!

It might be thought that the host atoms in dumbbells O and O' would interfere when dumbbell O' is in motion, and that energy would be transferred directly in dumbbell O by such contact and not only via the interstitial. In the unstrained lattice the atoms in two neighboring crossing dumbbells are in contact! However, inspecting Fig. 7a and b, one sees that when O' is most contracted the interstitial must be near O, extending that dumbbell; thus the atoms in the two dumbbells cannot get into each others way.

Thus, it seems quite well established that when the interstitial decays from T to O, practically the entire relaxation energy is radiated from O'. It follows by time reversibility that when the interstitial is activated from O to T, practically the entire activation energy is supplied to dumbbell O', preparing the final site for the interstitial in diffusion to O' in the description of Shewmon (see Fig. 8a). Shewmon's [5] description of diffusion of C in α-Fe in terms of activated processes is substantiated in detail.

Schematically, the diffusion process can be pictured as follows in a way that is self-explanatory. The heat of transport is obviously

$$Q^* = -Q_m \tag{22}$$

for the diffusive jump $O - T - O'$.

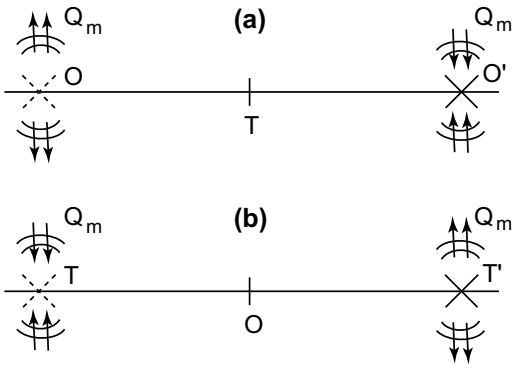


Fig. 8. (a) Diffusive jump of an interstitial atom from octahedral site O to O' via a saddle point at a tetrahedral site T. (b) Diffusive jump of a small interstitial occupying a tetrahedral site at rest and transiting an octahedral saddle point site. The jump is from $T - O - T'$.

In the case of interstitials relatively smaller than C or N in α-Fe, the equilibrium sites are the tetrahedral sites, i.e. the T sites. Suppose the saddle point for diffusion are the O sites, as described in Fig. 8b. In this case, except possibly for very small interstitials, it is reasonable to assume that the interstitial couples two neighboring crossing dumbbells effectively to make up a symmetric oscillator similar to the case described in Fig. 6, because in this case the two dumbbells in question are twice as close as in the previous case. It follows that the diffusion jump $T - O - T'$ schematically can be pictured as in Fig. 8b. The heat of transport is

$$Q^* = Q_m \tag{23}$$

A deviation from this result would be expected to be in the direction of a smaller heat of transport than just estimated, due to less effective coupling between neighboring dumbbells by the interstitial than assumed.

In the preceding, two extreme cases have been considered: octahedral equilibrium sites with saddle points at T sites and tetrahedral equilibrium sites with saddle points at $O-$ sites. It is well known that the equilibrium sites of C and N in α-Fe are O sites, but there have been suggestions that the T sites are metastable equilibrium sites and not the saddle points for diffusion (see, e.g. McCellan et al. [14] or Beshers [15]); ab initio calculations such as those of Jiang and Carter [21], however, indicate that the tetrahedral sites are indeed true saddle points. Nonetheless, it is worthwhile to consider the implications as the analysis is useful for discussing the later case of C interstitials in face-centered cubic (fcc) Fe. Alternatively, for somewhat smaller interstitials with stable T sites, the O sites may be sites of unstable equilibrium. Consider the case below, with stable O sites and metastable T sites. Reconsidering the preceding arguments for application to the present case, one concludes that decay from the saddle point S to O should be discussed the same way as decay from T to O was discussed and that the energy is radiated at O', and that similarly the metastable T site is associated with a symmetric oscillator. Thus, the diffusion from O to O' is described as in Fig. 9a, and

$$Q^* = -Q_m \tag{24}$$

as before, since the processes involving Q_T cancel.

Next consider that the T sites are stable and the O sites metastable (see Fig. 9b). Discussing along similar lines, one concludes that

$$Q^* = Q_m - 2Q_0 \tag{25}$$

since the transport distance of Q_0 is twice the interstitial jump distance. This is different from Fig. 9a; the processes involving Q_0 do not cancel.

Finally, one should discuss the direct $T - T$ jump, possibly important for the smallest interstitials with T sites as equilibrium sites. The strain energy in the equilibrium sites will be small. In diffusing, the interstitial must go through a portal made up of three atoms midway between the two T

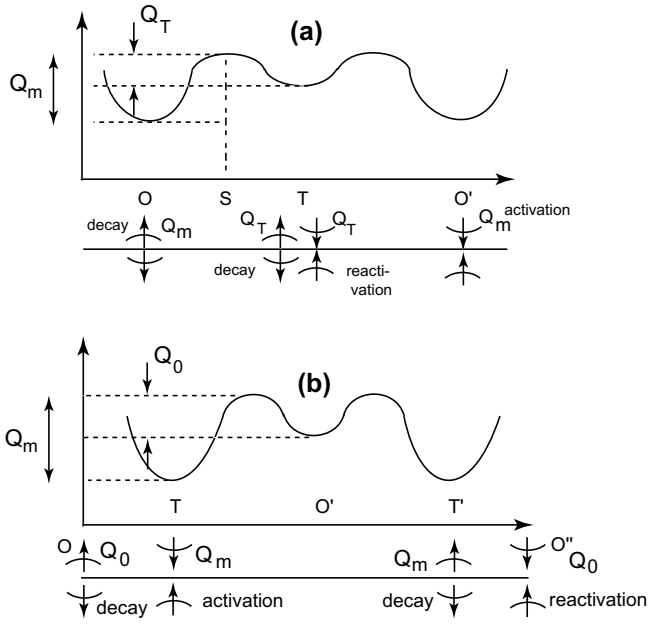


Fig. 9. (a) Diffusing atom with a stable O site and a metastable T site. Diffusive step is from $O - T - O'$. (b) Diffusing atom with stable T sites and metastable O sites. Diffusive step is $T - O' - T$.

sites, as described by Beshers, and the saddle point energy will consist of the strain energy opening up the portal. One would think that this case is best illustrated by a scenario involving co-located energy absorption and radiation and that $Q^* \sim 0$, i.e. that Q^* would be a small fraction of Q_m. This case is, however, sufficiently unsymmetrical that detailed considerations are difficult without, say, the benefit of reliable atomistic simulation.

The smaller interstitials usually prefer the smaller tetrahedral sites in fcc crystals, possibly to achieve better contact for chemical bonding. It is not clear what the diffusion path will be in this case, and quite certainly it is asymmetrical and complex. No attempt at predicting the heat of transport for this case will be made.

5. Interstitials in fcc crystals

The largest interstitial sites in fcc crystals are the two equivalent families of octahedral sites $<1/2, 0, 0>$ or $<1/2, 1/2, 1/2>$ and C and N in γ-Fe occupy these sites. The octahedral site in the fcc lattice is far larger than the octahedral site in the bcc lattice and thus typically the interstitial will fit within those sites with far less distortional energy than for the octahedral sites in bcc crystals, which fact is evidenced by the much larger solubility of C and N in γ-Fe than in α-Fe. In these particular cases we expect a high increase in the distortional energy as the interstitial moves into the saddle point configuration, pushing host atoms aside, and that the interstitial in the saddle point configuration has free access to its final position, which it will then reach, leaving the saddle point atoms behind in a highly activated state. The diffusion process is illustrated atomistically in Fig. 10. One must expect that the model of

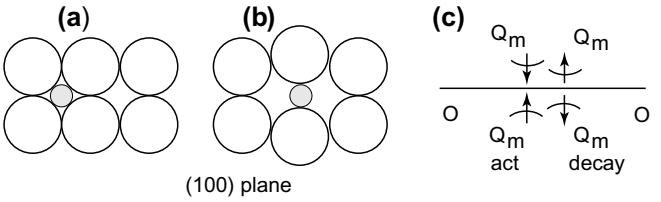


Fig. 10. (a) Diffusing atom with a stable T site and a metastable T site, and making a direct jump to another T site. (b) The saddle point configuration and (c) the energy radiated out in the jump process.

Fig. 5 applies and schematically the heat of transport can be illustrated as in Fig. 10c. In this case, clearly $Q^* \sim 0$. In terms of activation processes, the jump occurs by thermal activation of the saddle point atoms midway between the initial and the final position, which is exactly the description given by Shewmon [18].

The smaller interstitials usually prefer the smaller tetrahedral sites in fcc crystals, possibly to achieve better contact for chemical bonding. It is not clear what the diffusion path will be in this case, and quite certainly it is asymmetrical and complex. No attempt at predicting the heat of transport for this case will be made.

A most interesting special case is H in Pd. The hydrogen atom actually goes into the larger octahedral sites, contracting the lattice away to achieve contact with the neighboring host atoms. This behavior can only be explained as being due to strong attractive chemical forces. The reason for octahedral occupancy might then be that bonding to six neighbors is possible, as compared with four neighbors in the tetrahedral sites, and that factor outweighs the larger strain energy associated with the octahedral site's lattice contraction. There will be some strain energy in the saddle point, which, as in the preceding cases, give no contribution to Q^*. However, the chemical binding and the strain energy in the equilibrium states will both contribute positively to the heat of transport. For example, consider the interstitial in the saddle point. As it moves down towards the final site, it contracts the lattice and achieves better chemical binding, and a total energy $\sim Q_m - Q_{strain}$ (saddle point) is released around the final site, from out which it will be radiated. Assuming that the strain energy in the saddle point is the smaller part of the energy for motion, it follows that $Q^* \approx Q_m$. This result is also reasonable from the following point of view: with the interstitial in the original site, an activation of the saddle point atoms to open up a passage will hardly induce the interstitial to leave the initial site. The atoms surrounding the interstitial in the initial site must be activated and pulled out to make the interstitial less tightly bound to that site, and if this is the most important part of the process of motion, $Q^* \approx Q_m$ applies.

6. Vacancies in fcc crystals

Consider a diffusing vacancy, as illustrated in Fig. 11a and b. The strain energy in the saddle point configuration is symmetrical about the plane $p - p$ midway between the

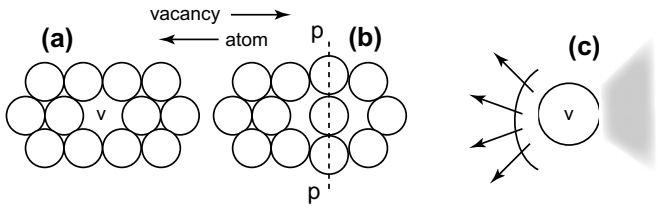


Fig. 11. (a, b) Vacancy diffusion in fcc crystals; the vacancy is marked as "v". The vacancy is moving to the right whereas the atom moves to the left. Note that the vacancy casts a "shadow" to its right so that energy is radiated toward the direction of the atom motion. This causes the heat of transport to be positive with respect to atomic motion.

initial position O and final position O'. Since, even for large diffusing atoms, only little kinetic energy is transferred to that atom and brought to the final site, one might think that the model of Fig. 6 applies. However, it is seen that, with the vacancy in the final position, the excess energy left behind is predominantly on the left-hand side of the vacancy and will thus be radiated directionally out to the left: the presence of the vacancy means that couplings that may transmit energy towards the right are missing, in other words the vacancy throws a shadow. It follows that in this case the saddle point strain energy, despite being released midway between the initial and final positions, will contribute positively to the heat of transport because it is radiated asymmetrically (positive with respect to the direction of the atom jump, negatively with respect to the direction of the vacancy jump). Further, if the jumping atom gains chemical binding energy jumping into the final site, this is another positive contribution to the heat of transport, and also for this contribution, the shadow effect will be important, making the contribution to the heat of transport larger than it otherwise would be.

It is concluded that the heat of transport for vacancies in fcc metals should always be positive. Also, since the energy displacement in asymmetric radiation must depend on the phonon–phonon interaction, one might expect a more pronounced temperature dependence in the heat of transport than in more symmetric cases. The heat of transport would be expected to decrease with increasing temperature. More measurements would be most illuminating.

7. Energy transport in the C–α-Fe

Molecular statics and dynamics were used to verify the assumptions made in the above and to provide quantitative estimates for the essential quantities, in particular for the heat of transport. The example of C in bcc α-Fe was chosen owing to the fact that its heat of transport is quite large and negative, both features making this example particularly interesting. For this purpose the GULP code was used [19] with the recently developed many-bodied Fe–Fe/Fe–C potential of Lau et al. [20]. This potential has been fit, via comparisons to the predictions of ab initio simulations [21,22], to the formation energies of C defect structures, including C interstitials. The potentials predict the most stable interstitial positions, as well as providing an accurate estimate of the activation energy of diffusion. Similar calculations were performed using the C–Fe interatomic potential developed by Ruda et al. [23]. For our molecular dynamics simulations, a modified version of this potential proposed by Ruda et al. [24] was used, which makes the octahedral site more stable relative to the tetrahedral site as is known from experiment and ab initio calculations [21]. For large simulation blocks this potential predicts an energy difference between the tetrahedral and octahedral sites of $\Delta E = E_{\text{tet}} - E_{\text{oct}} = 0.272$ eV. Thus the octahedral site is indeed forecast to be the more stable, although the difference in energy is less than predicted by ab initio calculations or experiment [21]. Aside from these modifications to the Ruda et al. potential, all other properties remain similar to those predicted from its original form.

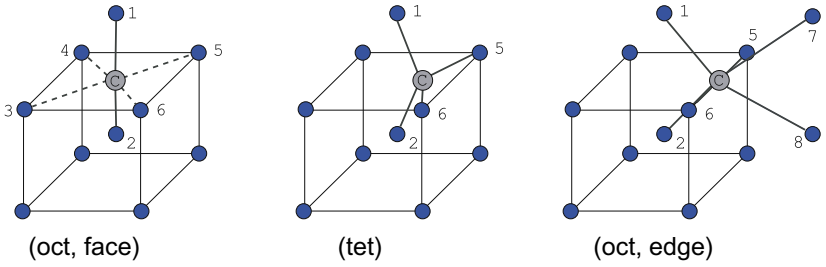
7.1. Molecular statics

Atomic positions were first calculated. Of immediate interest were the relative positions of Fe atoms comprising the "dumbells" described in Fig. 7 about the C interstitial in the stable octahedral and saddle point tetrahedral sites. Specifically of interest were the estimates of the relaxation distances referred to as x and y in Eq. (21), and in the discussion surrounding it. The atom separations are listed in the table (d) in Fig. 12. Inspection of the table in Fig. 12 shows that, for all the potentials used, indeed $y > x$. The ab initio calculations of Jiang and Carter yield $y = 0.278$ Å and $x = 0.099$ Å, respectively, whereas the GULP potential yielded $y = 0.342$ Å and $x = 0.097$ Å. Thus our arguments surrounding Fig. 7 are given vital substantiation. These results, which we argue underscore the sign and magnitude of Q^* in the C–α-Fe system, are not transparent. For instance, despite the larger radius of the tetrahedral site (0.36 Å) as measured from its center to the nearest Fe atom as compared to the radius of the octahedral site (0.19 Å), the stable site for C is the octahedral. Carbon atoms as shown by ab initio calculations (e.g. [21]), and as has been known experimentally for some time, prefer the strained octahedral site since it requires that C (with a covalent radius of 0.77 Å) has only two nearest neighbors, as opposed to four in the tetrahedral sites.

7.2. Molecular dynamics and energy transport

As noted above, the potentials of Ruda et al. [23,24] were used to study the evolution of the heat produced immediately after the decay of an interstitial carbon atom from a tetrahedral transition site to a stable octahedral site. In the modified potential, the Fe–Fe interactions are described by the EAM potential developed by Mendelleev et al. [25]; the details are reported in Refs. [23,24].

Simulations were carried out using two supercell simulation blocks, each with one carbon atom and with 16 and 128 Fe atoms, respectively. Using the 17 atom (i.e. including the C atom) supercell, we find that $\Delta E = 0.176$ eV,



Present study (using GULP with potentials from Lau et al. Phys. Rev. Lett. 98:215501 (2007)			
	oct (face)	tet	oct (edge)
1-2	3.955 (Å)	3.761	2.862
5-6	3.076	3.761	3.643
C-5	2.06	2.17	1.822
Jiang & Carter (ab initio) Phys. Rev. B 67:214103 (2003)			
1-2	3.559	3.361 (3.20)	2.811 (2.81)
5-6	2.797	3.353 (3.203)	3.552 (3.46)
C-5	1.98	1.82 (1.80)	1.775 (1.73)
Ruda *et al.* Scripta Meter. 46:349-355 (2002)			
1-2	3.48	3.291	
5-6	2.903	3.291	
C-5	2.04	1.82	

Fig. 12. Octahedral (a, c) and tetrahedral (b) sites of C in α-Fe with Fe atoms labeled 1–6. The table below shows distances between atoms comprising what has been referred to as "dumbells" in the preceding discussion. Atom positions were computed via molecular statics. Note that in connection with Fig. 7 and Eq. (21), the relaxation distance x would equal the difference between the separation of atoms 1 and 2 with C in its octahedral site as in (a) and the separation of these atoms with C in its tetrahedral site as in (b). The relaxation distance y in Eq. (21) is the difference between the separations of atoms 5 and 6 with C in its tetrahedral vs. its octahedral sites. For all cases the relation $y > x$ is substantiated, and especially so for the atom positions computed using ab initio methods.

somewhat lower than found using larger supercells. For the 129 atom supercell, we find $\Delta E = 0.209$ eV. During the decay process, this activation energy is converted to kinetic energy, which, in turn, is reflected as a temperature rise. After a time the system equilibrates and we expect that the final temperature can then be estimated from $n3/2kT = \Delta E$, where the factor n reflects the number of atoms within the cell. Our simulations indicated the final temperature rises as 86 K for the smaller cell and 13 K for the larger, which are in essential agreement with the estimates provided by the above expression. Our goal, as noted above, was to monitor how this energy was spatially distributed during the decay process as envisioned in, for example, Figs. 1, 4 and 8.

Standard molecular dynamics were utilized for the simulations with an implementation that maintained constant total energy and volume. During the decay from the tetrahedral site to the octahedral site, the potential energy difference is converted to kinetic energy, which was monitored by following the velocities of all atoms in the vicinity of the activated complex of atoms. Periodic boundary conditions were used in all three dimensions and no scaling of the atom velocities was performed. Time steps of 1 fs were used. Atom positions and velocities were analyzed every 100 time steps, i.e. every 0.1 ps. The initial configurations of the supercells were obtained via a conjugate gradient energy minimization of the atomic positions, with simultaneous minimization of the energy with respect to the total volume of the cell. The initial temperature was set at $T = 0.1$ K. This means that all Fe atoms were, in effect, initially at rest. To initiate the decay process, the C atom was given an initial velocity of 0.34 nm/ps (34 m/s) directed toward the octahedral site.[6] With these initial conditions, the system was followed for the next 0.5 ps, which is the time it took the C atom to reach the octahedral site.

In the large simulation, the pattern of energy flow during the decay process is shown in Fig. 13a–e. During the first 0.5 ps what is seen is a pattern characterized by the atoms in the wake of the decaying C atom acquiring most of the increases in kinetic energy. This, in turn, implies that energy is transported backward as expected from our qualitative analysis of the heat of transport. In fact, in Section 4 we argued that for C in α-Fe we expected $Q^* \approx -Q_m$. The results shown in Fig. 13 largely substantiate this, albeit in qualitative form due to the approximate value of the predicted activation energy that follows from our potentials and the size of the supercell. As estimated from the values of kinetic energy at $t = 0.5$ ps, we find that approximately 70% of the activation energy has been transported backwards during the decay process. At longer times, of course,

[6] Strictly speaking the initial velocity of the C atom would translate to an initial temperature of approximately 3.69 K for the simulation block.

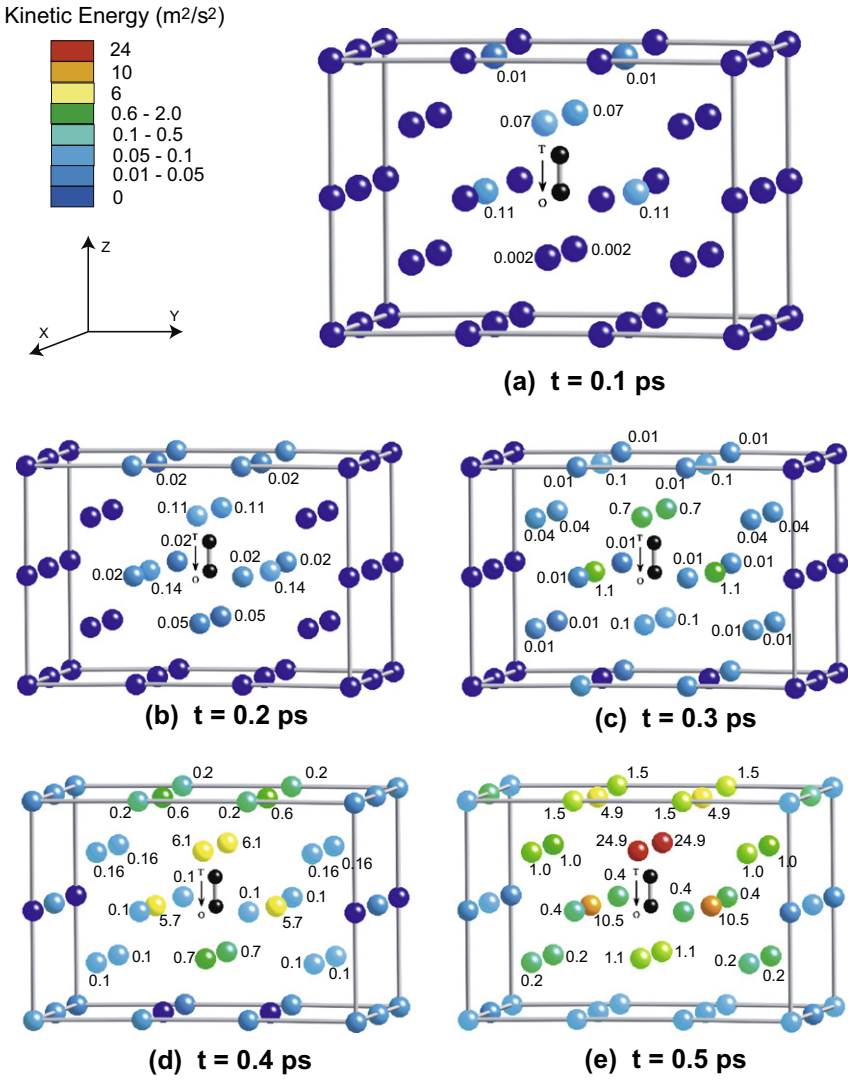


Fig. 13. Results for the distribution of kinetic energy during the decay process within a 128 atom simulation block. (a)–(e) show a time progression of the distribution of kinetic energy of atoms located in the immediate vicinity of the diffusing C atom. Among the atoms shown are those most intimately coordinated with the transition site and the more stable octahedral site.

the kinetic energy becomes more uniformly distributed, i.e. as the temperature equilibrates within the cell. In viewing Fig. 13a–e we remind the reader that full periodic boundary conditions were used, yet only the atoms most affected during the decay process are shown. Simulations using the 17 atom cell produced very similar results, indicating that nearly the entire activation energy was transported to the atoms in the wake of the decaying C atom, and hence we have not shown this for the sake of brevity. Thus, the simulations just presented appear to provide validation of the scenario presented in Figs. 7 and 8 for the C–α-Fe system. They also verify not only the fact, known from experimentation [5,8], that $Q^* < 0$ for this system but also that $|Q^*| \sim Q_m$ itself. The model considerations surrounding Figs. 7 and 8 are compelling but, given the nature of the atomic bonding in the C–α-Fe system, the patterns of energy flow cannot really be evaluated from the simple pictures of atomic bonding presented there.

8. Discussion

8.1. Shewmon's results for C in α-Fe

The basic phenomenon studied by Shewmon [18] was concerned with the migration of iron carbides caused by thermal diffusion in a thermal gradient. This was part of his more general study of the important class of phenomena concerned with second-phase redistribution in alloys subject to thermal gradients. In Ref. [18], the equilibrium concentration of the diffusing specie (carbon in this case) is shown to increase monotonically with temperature. Thus, if a carbide were positioned within a temperature gradient, there would be a concentration gradient from the warmer to the cooler region which would, in itself, promote atomic carbon transport down the thermal gradient and lead to migration of the particle down the temperature gradient as well. Now, according to Eq. (7), for instance, with

$Q^* = 0$ or $Q^* > 0$, that is precisely what would be expected even taking into account the Soret effect of coupled thermal diffusion. Note that when $Q^* > 0$ the migration is larger than when $Q^* = 0$, which illustrates that a $Q^* > 0$ reinforces the effect of the concentration gradient. However, if $Q^* < 0$, as is the case for C diffusion in α-Fe, and if $|Q^*|$ is sufficiently large, the reverse trend is expected. Shewmon [18] in fact reported that trend experimentally, and his results qualitatively asserted that $Q^* < 0$ and was significant in magnitude. Wert [26] reported an activation barrier for C diffusion in α-Fe of approximately 0.87 eV, a value confirmed by the ab initio calculations of Jiang and Carter [21]. Based on these values, if $Q^* \approx -Q_m$ as suggested by our analysis, $Q^* \approx 5.24 \times 10^{23}$eV/mol (=−20 kcal/mol). Shewmon's later measurements [5] produced values closer to $Q^* \approx -24$ kcal/mol; reasons for this difference between his measured values and the documented value for Q_m are as yet unclear, but the experimental results are indeed consistent, as is the estimate we obtain from our theoretical reasoning and our molecular dynamics simulation.

Shewmon's method of measurement is of interest especially for possible future molecular dynamics simulation. According to Eq. (7), if a temperature gradient were imposed across an Fe–C alloy, a migration of C would produce a concentration gradient until such a point when

$$\frac{\mathrm{d}\ln c}{\mathrm{d}x} = -\frac{Q^*}{kT^2}\frac{\mathrm{d}T}{\mathrm{d}x} \tag{26}$$

at which $J_p \to 0$. This was the experimental approach that Shewmon [5] took. The method is amenable, however, to molecular dynamics simulation. For example, a temperature gradient may be imposed by continually adjusting the average kinetic energy at the ends of a relatively long simulation block that initially has a uniform concentration of C atoms. The simulation would then be continued until a fixed concentration difference of C develops and where $J_p \to 0$ and then where Eq. (26) is used to estimate Q^*. Such simulations are planned for future study. Our potentials provided a consistent picture for the process of C diffusion in α-Fe, but future simulations may be based on the more recently developed potentials developed by Lau et al. [20] which provide better consistency with the ab initio calculations of Jiang and Carter [21] and experimentally documented activation energies [26].

8.2. General comments on the use of the Onsager reciprocity theorem

Onsager's reciprocity relations [1], which herein take the form $D_{ij} = D_{ji}$, have their basis in the principle of time reversibility. The principle envisions fluxes in response to forces as expressed in, for example, Eq. (13). Fluctuations in energy and concentration also occur in systems in equilibrium, and a basis of Onsager's proof of reciprocity is to assume that the laws that govern the decay of such fluctuations are, in fact, those that govern macroscopic fluxes. Thus, we write on the microscopic level

$$\dot{x}_i = d_{ij}X_j \tag{27}$$

where $\dot{x}_i$ is a flux of mass when $i = p$ or energy when $i = q$, and X_j are the conjugate forces; the d_{ij} are the kinetic coefficients linking them. At the essence of the reciprocity proof is (i) that the ensemble average of x_iX_j is

$$\langle x_iX_j\rangle = -k\delta_{ij} \tag{28}$$

and (ii) that time reversibility implies that

$$\langle x_i(0)\dot{x}_j(0)\rangle = \langle \dot{x}_i(0)x_j(0)\rangle \tag{29}$$

where $x_i(0)$ and $\dot{x}_i(0)$ are the values of x_i and its rate at time $t = 0$ (see, e.g. Pathria [27]). Now if Eq. (27) is used in Eq. (29) and then Eq. (28) is invoked, we find that indeed $d_{ij} = d_{ji}$. The key assumption mentioned above is that Eq. (27), which applies to the decay of fluctuations, also applies to macroscopic processes (i.e. $d_{ij} = D_{ij}$) then leads to $D_{ij} = D_{ji}$. However, at equilibrium, when there is no net mass flow, the time average of energy density is fixed. Now consider the process whereby a diffusing atom jumps back and forth between a stable site and a transition state, establishing an equilibrium between the two states. Given that energy flows asymmetrically during both the activation and the decay processes, both microscopic reversibility and the fact that the time average of local energy density is fixed lead to the conclusion that the activation and decay processes are inverses with respect to energy flow. This is how we have used the principles involved to construct our approach as outlined in Section 1.

8.3. Implications for future study

We have presented a general approach for developing a qualitative interpretation of the heat of transport, Q^*, in crystals; the qualitative assessments thus gained include establishing the sign of Q^*. The approach then prescribes a clear methodology for computing, and thereby quantitatively establishing, the magnitude of Q^*; this methodology is based on an appropriate application of the principle of microscopic reversibility. Molecular dynamics (MD) offers a clear path for implementation of our method. For the case of C in α-Fe, our preliminary MD simulations indeed established that $Q^* \approx -Q_m$, as experiments have shown [8,5]. Oriani [6] summarizes experimental data on other interstitial solute systems, e.g. Q^*(C, γ-Fe) ≈ −2 kcal and Q^*(N, α-Fe) ≈ −18 kcal. Darken and Oriani [8] also discuss, albeit in a more qualitative fashion, thermal diffusion in a substitutional Cu–Au alloy; in this case they also report that $Q^* < 0$. All such cases are then candidates for future study via MD simulation. In fact, our methods define a relatively easy path for exploration, prior to experimental study, of other alloy systems where interatomic potentials are available with sufficient accuracy.

Acknowledgement

R.J.A. would like to thank Professor Jens Lothe, University of Oslo, for stimulating discussions and for his seminars given at Ohio State University during R.J.A.'s post-doctoral stay (1972/73), which inspired many of the ideas developed herein.

Appendix A. Derivation of the flux vs. force relations

The purpose here is to provide a derivation of the flux vs. force relations given in Eqs. (7) and (8) that makes specific link with the more familiar forms of the kinetic laws given in standard treatments, e.g. in deGroot and Mazur [3] or Howard and Lidiard [7], and that helps explain concepts such as the reduced energy flow discussed in connection with Eq. (14).

Consider a multi-component system, with n diffusing species, and where the Latin indices i, j or k range over $1,\ldots,n$. Also let the index q denote the flux or force associated with energy flow. Later we introduce Greek indices, α, β, which, when involved in a summation, will range from $\alpha, \beta = 1,\ldots,n, q$. Begin by writing the flux vs. force relations as [3,7]

$$\begin{aligned} \hat{J}_k &= \hat{L}_{ki}\hat{X}_i + \hat{L}_{kq}\hat{X}_q, \quad k = 1,\ldots,n \\ \hat{J}_q &= \hat{L}_{qi}\hat{X}_i + \hat{L}_{qq}\hat{X}_q \end{aligned} \tag{A.1}$$

The summation convention over the indices i, j and k is employed and the "forces" are specified as

$$\hat{X}_i = -T\frac{\partial(\mu_i/T)}{\partial x}, \quad \hat{X}_q = -\frac{1}{T}\frac{\partial T}{\partial x} \tag{A.2}$$

Here J_q is the total energy flux as opposed to the reduced energy flux to be introduced just below. The forces are now defined so that, if Θ is the rate of entropy production per unit volume

$$T\Theta = \hat{J}_i\hat{X}_i + \hat{J}_q\hat{X}_q \tag{A.3}$$

The total energy flux includes the energy associated with the partial molar enthalpy, h_i, of the diffusing species i and thus we seek a transformation of variables of the form

$$J_\alpha = \xi_{\alpha\beta}\hat{J}_\beta \tag{A.4}$$

that renders

$$J_i = \hat{J}_i \quad \text{but} \quad J_q = \hat{J}_q - h_i\hat{J}_i \tag{A.5}$$

yet preserves the Onsager symmetry and the expression for entropy production. In tensor notation we write

$$\mathbf{J} = \boldsymbol{\xi}\cdot\hat{\mathbf{J}}, \quad \mathbf{J} = \mathbf{L}\cdot\mathbf{X}, \quad \mathbf{L} = \mathbf{L}^{\mathrm{T}}, \quad \mathbf{J}\cdot\mathbf{X} = \hat{\mathbf{J}}\cdot\hat{X} \tag{A.6}$$

This yields

$$X = \boldsymbol{\xi}^{-\mathrm{T}}\cdot\hat{X}, \quad L = \boldsymbol{\xi}\cdot\hat{L}\cdot\boldsymbol{\xi}^{\mathrm{T}} \tag{A.7}$$

The transformation we seek is such that the components of $\boldsymbol{\xi}$ are

$$\xi_{\alpha\beta} = \delta_{\alpha\beta} - \delta_{\alpha q}h_\beta \quad \text{with} \quad h_q \equiv 0 \tag{A.8}$$

Formally we find that

$$X_i = \hat{X}_i + h_i\hat{X}_q \tag{A.9}$$

and, since $\partial(\mu_i/T)/\partial T = -h_i/T^2$

$$X_i = -\frac{\partial\mu_i}{\partial x} \tag{A.10}$$

This means that the force involves the gradient of μ_i with respect to, say, concentration, but not temperature.

The energy flow defined by the second of Eq. (A.5) is a reduced flow in that what is removed from the total energy flow is the product of the enthalpy of the atomic species and their fluxes. This, in connection with the comments made with respect to an "excess flow" associated with the flux given by Eq. (8), means that the excess flow is equivalent to the reduced flow just computed via the transformation of Eq. (A.4) with (A.8). The transformation also leads to

$$X_q = \hat{X}_q \tag{A.11}$$

and the flux vs. force relations

$$\begin{aligned} J_k &= L_{ki}X_i + L_{kq}X_q \\ J_q &= L_{qi}X_i + L_{qq}X_q \end{aligned} \tag{A.12}$$

The heats of transport are formally introduced via

$$L_{iq} = L_{ij}Q_j^* = L_{qi} \tag{A.13}$$

the last equality due to the Onsager symmetry. When these are included into Eq. (A.12) we obtain

$$\begin{aligned} J_i &= L_{ij}(X_j + Q_j^*X_q) \\ J_q &= L_{qi}X_i + L_{qq}X_q \end{aligned} \tag{A.14}$$

where the Onsager symmetry is clear from the defining relation Eq. (A.13). Note that

$$L_{qi}X_i = L_{ik}Q_k^*X_i = L_{ki}X_iQ_k^* = J_k^cQ_k^* \tag{A.15}$$

where we mean by J_k^c the flux of specie k under strictly isothermal conditions, i.e. when $X_q = 0$. Eq. (A.14) may also be written as

$$\begin{aligned} J_i &= L_{ij}X_j + L_{ij}Q_j^*X_q \\ J_q &= Q_i^*J_i + \left(L_{qq} - L_{iq}Q_i^*\right)X_q \end{aligned} \tag{A.16}$$

This latter form more clearly reveals that when $X_q = 0$, $J_q = Q_i^*J_i$, which is how the heat of transport was originally introduced vis-à-vis Eq. (3). We may also note that consistency with Denbigh's discussion of the heat of transport [2] may be established by noting that he defines $Q^* = Q - \bar{H}$ for a single component system where Q is the total energy transported by mass flow and $\bar{H}$ is the partial molar enthalpy of the diffusing component. We have, in fact, created the same effect in our definition of a reduced heat flow via the second of Eq. (A.5). Similar relations are found in classical works on the thermodynamics of irreversible processes, e.g. Refs. [16,17,28], but where quite different approaches are used to arrive at the kinetic relations and where the reader is cautioned to understand whether the This latter form more

clearly reveals that when $X_q = 0, J_q = Q_i^* J_i$, which is how the heat of transport was originally introduced vis-à-vis Eq. (3). We may also note that consistency with Denbigh's discussion of the heat of transport [2] may be established by noting that he defines $Q^* = Q - \bar{H}$ for a single component system where Q is the total energy transported by mass flow and $\bar{H}$ is the partial molar enthalpy of the diffusing component. We have, in fact, created the same effect in our definition of a reduced heat flow via the second of Eq. (A.5). Similar relations are found in classical works on the thermodynamics of irreversible processes, e.g. Ref. [16,17,28], but where quite different approaches are used to arrive at the kinetic relations and where the reader is cautioned to understand whether the thermal fluxes are the full flux or the reduced flux considered here.

To make specific contact with Eqs. (7) and (8) we note that for the case of one diffusing specie $n = 1$, and we may associate $i \leftrightarrow p$ and identify $L_{pp} = Dc/kT$ and $\kappa = L_{qq}/T$; this, in fact, reproduces Eqs. (7) and (8).

References

[1] Onsager L. Phys Rev 1931;37:405–26;
Onsager L. Phys Rev 1938;38:2265–79.
[2] Denbigh KG. Thermodynamics of the steady state. New York: Methuen & Co; 1951.
[3] De Groot SR, Mazur P. Non-equilibrium thermodynamics. Amsterdam: North-Holland; 1962.
[4] Wirtz K. Phys Z 1943;44:221–33. There is an excellent discussion in English of Wirtz's approach given in Jost W. Diffusion. New York: Academic Press; 1952. p. 256–8. See also the short discussion of the Wirtz approach given by Denbigh [2, Section 2.6].
[5] Shewmon PG. Acta Metall 1960;8:605–11.
[6] Oriani RAJ. Chem Phys 1961;34:1773–7.
[7] Howard RE, Lidiard AB. Rep Prog Phys 1964;27:161–240.
[8] Darken LS, Oriani RA. Acta Metall 1954;2:841–7.
[9] Eastman EDJ. Am Chem Soc 1926;48(6):1482–93.
[10] Wagner C. Ann Phys 1929;3:629–49.
[11] Feder J, Russell KC, Lothe J, Pound GM. Adv Phys 1966;15:111–78.
[12] Seeger A, Donth H, Pfaff F. Faraday Soc 1957;23:19–25.
[13] Jossang T, Skylstad K, Lothe J. The relation between the structure and mechanical properties in metals, vol. II. London: Her Majesty's Stationery Office; 1963.
[14] McCellan RB, Rudee ML, Ishibachi T. Trans AIME 1965;233:1938–43.
[15] Beshers DN. Diffusion in body centered cubic metals. Metals Park, OH: ASM; 1965. p. 149.
[16] Weeks JD, Shuler KEJ. Chem Phys 1972;56:1883–9.
[17] Platten JKJ. Appl Mech 2006;73:5–15.
[18] Shewmon PG. Metall Trans 1958;45:642–7.
[19] Gale JD, Rohl AL. Mol Simul 2003;29:291.
[20] Lau TT, Forst CJ, Lin X, Gale JD, Yip S, Van Vliet KJ. Phys Rev Lett 2007;98:215501-1–4.
[21] Jiang DE, Carter EA. Phys Rev B 2003;67:214103-1–214103-11.
[22] Forst CJ, Slycke J, Van Vliet KJ, Yip S. Phys Rev Lett 2006;96:175501-1–4.
[23] Ruda M, Farkas D, Abriata J. Scripta Mater 2002;46:349–55.
[24] Ruda M, Farkas D, Garcia G. New EAM interatomic potentials for carbon interstitials in Fe, appropriate for the interaction of C with vacancies, surfaces, grain boundaries, self interstitials, and dislocations, submitted for publication.
[25] Mendelleev MI, Han SD, Srolovitz JG, Ackland J, Sun DY, Asta M. Philos Mag 2003;83:3977–94.
[26] Wert CA. Phys Rev 1950;79:601–5.
[27] Pathria RK. Statistical mechanics. New York: Elsevier; 1997.
[28] Haase R. Thermodynamics of irreversible processes. Menlo Park, CA: Addison-Wesley Publisher; 1969. p. 184–187.