ABSTRACT
Diffusion on the nano/atomic scales in multilayers, thin films has many challenging features even if the role of structural defects can be neglected and ‘only’ the effects related to the nano/atomic scale raise. Different examples for diffusional nanoscale effects we discovered recently will be given in this paper. We illustrate that the continuum descriptions of the diffusion cannot be applied automatically on such short distances. They predict faster kinetics than the atomistic models and the interface shift is always proportional to the square-root of the time ($x \propto t^{1/2}$ ⇒ $x^2 \propto t$: parabolic kinetics). However, the kinetics can be even linear ($x \propto t$) on the nano/atomic scale. Furthermore, the continuum descriptions foretell infinitely fast kinetics as the time goes to zero ($v=dx/dt \propto 1/t^{1/2}$), which is a long standing paradox of the diffusion theory. A possible resolution of this paradox is also presented. Moreover, it will also be illustrated that an initially diffused interface can sharpen even in completely miscible systems.

Keywords: interface motion, interface sharpening, non-Fickian kinetics

INTRODUCTION
In nanostructured materials, where the density of grain- and interphase-boundaries is high, the diffusion and kinetics of surface segregation, i.e. the effective material flow is always influenced by the contributions of these boundaries (Beke, 2002). But even if the role of these boundaries can be neglected (single crystalline or amorphous materials) there are a lot of open questions concerning the diffusion on atomic scale, because for short diffusion distances the classical continuum approximations (Fick's laws) cannot describe correctly the atomic movements close to interfaces (Erdélyi, 1998; Erdélyi, 2002a). By computer simulations, we have recently shown deviations from Fick’s law (Erdélyi, 2002a; Erdélyi, 2004a) and that an initially diffused interface can sharpen even in completely miscible systems (Erdélyi, 2002b). We have successfully performed experiments, too, verifying these effects (Erdélyi, 2004b).

BASIC EQUATIONS
Continuum approach
The most basic equations to describe the diffusion are Fick’s equations. The first equation states that the atomic flux ($j$) is always antiparallel to the gradient of the concentration ($\rho$, atomic density) (in one dimension):

$$j = -D \frac{\partial \rho}{\partial x},$$

where $D$ is the diffusion coefficient. The combination of this equation with the conservation of material yields Fick’s second equation (in one dimension):
\[ \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} \left[ D \frac{\partial \rho}{\partial x} \right], \] (2)

where \( t \) is the time. Here it is important to emphasize that \( D \) in general composition dependent, thus cannot be factored out. As will be seen also in this paper, the composition dependence of \( D \) is very important and has very significant influence on the diffusion kinetics on the nano/atomic scales. It originates from the fact that usually the diffusion coefficient of an A atom different in an A and in a B matrixes. Thus, in a diffusion couple the \( D \) cannot be considered constant, especially if we take into account that the interface is not atomically sharp, i.e. there is a more or less intermixed region between the pure A and B matrixes. Consequently the diffusion coefficient changes continuously while e.g. an A atom diffuses from the pure A matrix into the pure B. This feature is also called diffusion asymmetry.

However Eq. (2) can describe only the time evolution of the composition profile, but does not include the stress effects, which can have important influence onto the diffusion especially on the nano/atomic scale.

Basically three types of stresses can be considered: (i) ‘built-in’ stress originating form the lattice mismatch at the interface, (ii) thermal stress due to the difference, between the thermal expansion coefficients of the layers (Pompe, 1998) and (iii) diffusional stress developing during interdiffusion because of the net volume transport caused by the difference of atomic currents of the constituents through the interface. In Stephenson's (Stephenson, 1988; Beke, 1998) model these types of stresses can be treated and stress relaxation is also considered. Stephenson - in one-dimensional, isotropic binary system - derived a set of coupled differential equations for the description of the resultant stress development and stress relaxation by viscous flow, convective transport (Kirkendall shift) and composition evaluation:

\[ \frac{Dp}{Dt} = - \frac{2E}{9(1-v)} \left[ \sum_{i=A}^{B} (\Omega_i \nabla j_i) + \frac{3}{4\eta} p \right] \] (3)

\[ \nabla \nabla = - \sum_{i=A}^{B} (\Omega_i \nabla j_i) - \frac{3(1-2v)}{E} \frac{Dp}{Dt} \] (4)

\[ \frac{Dc}{Dt} = - \frac{1}{\rho} \left[ (1-c) \nabla j_A - c \nabla j_B \right] \] (5)

where \( P \) is the pressure, \( E \) is the Young modulus, \( \Omega_i \) the atomic volume of the constituents \( i = A, B \), \( \eta \) is the share viscosity, \( v \) is the Kirkendall velocity, \( c \) is the atomic fraction of material \( A \), \( \rho \) is the average volume density and \( j_i \) the atomic flux, which is given by (Beke, 1998)

\[ \dot{j}_i = - \rho D_i \nabla c_i - \rho L_i c_i \nabla p, \quad i = A, B \] (6)

here \( D_i \) is the intrinsic diffusion coefficient, \( R \) is the molar gas constant, \( T \) is the absolute temperature and \( \Theta \) is the thermodynamic factor.
Discrete approach

The model used in our calculations is based on Martin's deterministic kinetic equations (Martin, 1990), where time derivatives of atomic fractions of A atoms in the \(i\)-th atomic layer perpendicular to the \(x\)-axis can be given by:

\[
dc_i/dt = J_{i-1,i} - J_{i,i+1}.
\] (7)

Here \(J_{i,i+1}\) is the net flux of A atoms from plane \(i\) to \((i+1)\) per lattice site:

\[
J_{i,i+1} = z_v \left[ c_i (1-c_{i+1}) \Gamma_{i,i+1} - c_{i+1} (1-c_i) \Gamma_{i+1,i} \right],
\] (8)

where \(\Gamma_{i,i+1}\) is the probability per unit time that an A atom in layer \(i\) exchanges its position with a B atom in the layer \((i+1)\) and \(z_v\) is the vertical coordination number. For the jump frequency an Arrhenius-type temperature dependence is assumed:

\[
\Gamma_{i,i+1} = \nu \exp\left(-E_{i,i+1}/kT\right),
\] (10)

where \(\nu\) denotes the attempt frequency, \(k\) and \(T\) are the Boltzmann's constant and the temperature, respectively, \(E_{i,i+1}\) is the activation energy barrier of an exchange between an A atom being in layer \(i\) and a B atom being in layer \((i+1)\), which can be chosen in different ways (but satisfying the steady-state condition) as was shown by Martin (Martin, 1990). As we have shown (Erdélyi, 2004c) one of the most adequate choices is the following:

\[
E_{i,i+1} = \hat{E}^0 + \alpha_i + \epsilon_i
\] and \(E_{i+1,i} = \hat{E}^0 - \alpha_i - \epsilon_i\), (11)

where \(\hat{E}^0\) is a composition independent term (the composition independent part of the activation energy) containing the saddle point energy, and:

\[
\alpha_i = [z_l (c_{i-1} + c_{i+1} + c_i + c_{i+2}) + z_t (c_i + c_{i+1})]M
\] (12)

\[
\epsilon_i = [z_l (c_{i-1} + c_{i+1} - c_i - c_{i+2}) + z_t (c_i - c_{i+1})]V
\] (13)

Here \(z_l\) is the lateral coordination number, \(M=(V_{AA}+V_{BB})/2\) and \(V=V_{AB}+V_{AA}+V_{BB}/2\), where \(V_{XY}\) are the nearest neighbour pair interaction energies between species \(X\) and \(Y\) (\(X,Y=A\) or \(B\)) in the framework of a pair-interaction model. However, as was shown also in (Erdélyi, 2004c) \(M\) in \(\alpha_i\) determines the strength of the composition dependence of the transition rates, whereas \(V\) in \(\epsilon_i\) is the regular solid solution parameter (Beke, 2002), proportional to the heat of mixing and measures the phase separating \((V>0)\) or the ordering \((V<0)\) tendency. Thus they \(M\) and \(V\) can be deduced either from more sophisticated models or experiments, too [e.g. \(M\) can be determined from the composition dependence of \(D\): \(M=mkT/2Z\), where \(Z=2z_l+z_t\) and \(D(\delta)=D(0)\exp(m\delta)\). For example \(m\) is about to 10 and 16 in the Ni-Cu and Mo-V systems respectively (Erdélyi, 2004a; Erdélyi, 2002b; Erdélyi, 2004b), which means e.g. that the diffusion is about 4-5 orders of magnitude faster in the Cu matrix than in the Ni.]
VALIDITY LIMIT OF THE CONTINUUM APPROACH

Calculation procedure

Since the discrete model is not able to describe the Kirkendall shift and the stress effects, we have chosen both the intrinsic diffusivities and the atomic volumes to be equal to each other ($D_A=D_B$ and $\Omega_A=\Omega_B$). Moreover, for simplicity’s sake, we have restricted the calculations onto completely miscible systems, i.e. when $V=0$.

In the framework of the continuum approach, a finite difference method was used to calculate the evolution of the concentration profile of a multilayer.

In the discrete model, a numerical method was used to solve the set of coupled equations (7-13) to obtain the time dependence of the concentrations in different layers.

In both cases we started from an initial square profile, with the $x$ axis parallel to the direction of diffusion. Furthermore, since the concentration profile in a multilayer is periodical and symmetrical during its evolution, only a half-bilayer has been considered during the numerical calculations. (Erdélyi, 1998)

Results

It was already shown in the middle of the seventies (Greer, 1985) by analytical calculations that considering composition independent diffusion coefficients the discrete and continuum approach give the same time evolution of a sinusoidal composition profile only if the modulation length is at least 6-10 times larger than the interatomic distance in the direction of diffusion: $\Lambda>6-10d$.

We have reproduced this result by our computer assisted comparison. However, we have also shown that in case of composition dependent diffusion coefficients (diffusion asymmetry), the two models results in the same time evolution of the composition profiles only for larger modulation lengths. The shifts of the critical $\Lambda$ (above which the two approach can be considered equivalent) in the function of the strength of the diffusion asymmetry $m'=\lg e$ ($m'$ gives in orders of magnitude the difference between the diffusion coefficients in the pure $A$ and $B$ matrices) is shown in Figure 1.

![Figure 1 - Values for the critical modulation length, $\Lambda_c$, above which the continuum model is valid, in the function of strength of the diffusion asymmetry (Erdélyi, 1998)](image)

As can be seen the range of the validity of the continuum model shifts strongly: the shift depends on the strength of the concentration dependence of $D$, and in many real multilayer systems with typical modulation length of few nanometer, it can break down.
TIME EVOLUTION OF THE COMPOSITION PROFILE IN CASE OF DIFFUSION ASYMMETRY

Computer simulations

Both from the continuum and discrete models we obtained that for strong composition dependent D, the interface between the A and B matrixes (Mo and V) remains abrupt and shifts until the component in which the diffusion is faster is not consumed even in completely miscible systems. (see Figure 2) It can be seen that the diffusion is very asymmetric: the Mo atoms diffuse very fast into the V matrix and are distributed homogeneously, whereas the V atoms practically cannot penetrate into the Mo matrix, since the diffusion is more than 7 orders of magnitude faster in the V than in the Mo. (Erdélyi, 1998)

We note here that in this case the difference between the discrete and continuum approaches has been manifested in the speed of the interdiffusion process (the kinetics is faster in the continuum model). (Erdélyi, 1998)

![Figure 2 - Concentration distributions at different times for the discrete (a) and continuum models (b).](image)

Experiments

This asymmetric behaviour is illustrated in Figure 3, which was obtained from measurement of the time evolution of the composition profile in amorphous Si/Ge system by Auger depth profiling technique. (Csik, 2001) This system is also completely miscible system and as usually in the nature, the diffusion is asymmetrical ($m'=4-5$, i.e. the diffusion is 4-5 times faster in the Ge than in the Si). As can be seen the Si atoms diffuse into the Ge and are homogenised there. As a consequence, the interface is shifted towards the Si resulting the thinning of the Si layers and thickening of the initial Ge layers of which Si content increases.
NON-FICKIAN INTERFACE SHIFT

As have been seen the interfaces remain sharp and shift even in completely miscible systems, which provide the possibility to speak about and investigate experimentally the kinetics of the interfaces in large variety of binary systems. Basically two types of systems can be investigated: i) completely miscible systems, where the interfaces remain abrupt during the diffusion process due to the strong diffusion asymmetry; and ii) phase separating systems in which the solubility is limited, where the interfaces remains sharp for chemical reasons.

It is known from Fick’s phenomenological laws that during annealing of a diffusion couple the displacement of a plane with constant composition (or an abrupt interface) is proportional to \( t^{1/2} \). However, we have shown first from computer simulations that this rule can be violated on the nanoscale either in completely (Erdélyi, 2002a) or restricted miscible systems (Erdélyi, 2004a). This is strongly related to the discrete character of the system on the nanoscale and to the highly neglected fact in the literature that the diffusion coefficients or mobilities depend on the composition.

Computer simulations

To investigate the interface shift, we have solved Eqs. (7-13) numerically (Erdélyi, 2002a; Erdélyi, 2004a). For the description of the interface shift, we started from the following initial condition: on the left- and right-hand side of the interface all the atomic planes were occupied only by \( A \) atoms as well as \( B \) atoms, respectively. In order to avoid ‘finite effects’, a continuous boundary condition was applied, i.e. when the composition of the atomic planes at the either ends of the sample were changed 10 pure \( A \) or \( B \) atomic planes were added to the sample. (Erdélyi, 2004a). The position of the interface was determined by the plane with the composition 0.5 (it can obviously lie between two atomic planes).

After determining this position \( (p) \), its logarithm versus the logarithm of the time \( (\log p \propto \log t) \) was plotted. Fitting a straight line to the data, its slope gave the power of the function describing the shift of the interface (called kinetic exponent and denoted by \( k_c \)). Obviously for parabolic interface shift \( k_c = 0.5 \). Thus if the kinetics is non-Fickian, \( k_c \neq 0.5 \) or the data do not fit on a straight line on the \( \log p \propto \log t \) plot.

Since we wanted to demonstrate the effects of the composition dependence of diffusion coefficients as well as the phase separation tendency on the kinetics of the interface shift, the parameters \( m' \) and \( V \) (or \( V/kT \)) were changed during the calculations.
Figure 4 shows the initial values of the kinetic exponent, $k_c$, (obtained by fitting to the interval corresponding to dissolution of the first 5 planes) versus $V/kT$ for different $m'$ values. It can be seen that $k_c$ is almost constant and, as it is expected, is very close to 0.5 for small $m'$ [the discussion of the small deviations from the 'pure' parabolic growth is given in (Erdélyi, 2004a)]. At the same time, the deviation from the square root kinetics increases with increasing $m'$ for a fixed value of $V/kT$.

The deviation from the parabolic law is a real "nano-effect", because after dissolving a certain number of layers (long time or macroscopic limit), the interface shift returns to the parabolic behaviour independently of the input parameters (see Figure 5).

We have shown very recently (Beke, 2006), that this transition can be understood from the analysis of the atomic currents in the different part of the sample. In principle three currents can be distinguished: i) $J_\alpha$ in matrix $A$, where the diffusion is very slow, ii) $J_I$ across the interface region and iii) $J_\beta$ in matrix $B$, where the diffusion is fast. (see Figure 6) However, $J_\alpha$ can be neglected, because practically there is no diffusion in matrix $A$. Moreover, at the beginning of the kinetics, when the composition gradient is very large, the flux in the $B$-rich phase ($\beta$ phase) is larger than across the interface ($J_I<J_\beta$). In this stage $J_I$ controls the diffusion. During the process $J_\beta$ becomes smaller and smaller because the tail of the composition profile in the $\beta$ phase grows more and more resulting in the decrease of the gradient of the composition. Although, $J_I$ also decreases with increasing time/number of layer dissolved, but $J_\beta$ decreases much faster. As a result, in a certain moment $J_\beta$ is the rate limiting. Thus the transition time or thickness must be deduced from the condition $J_I=J_\beta$ which has to be fulfilled at the interface.

It is worth mentioning that from the analysis of the currents an atomistic explanation of the phenomenological interface transfer coefficient ($K$) can also be done, which has been missing in the reaction diffusion literature. Following the phenomenological definition of $K$, $J_I=K(c,c_e)$ ($c$ and $c_e$ denote the current and the equilibrium composition at the interface, respectively) and comparing this to the $J_I$ in the discrete model (Beke, 2006)

$$K = \nu_z \exp(-Q_k/kT)$$

with $Q_k = E_0 + z_\ell V + MZ = E_0 + z_\ell V + mkT/2$. In fact $K$ is proportional to the jump frequency from the $A$-rich phase to the $B$-rich one. This is different from the jump frequency in the $B$-rich phase (where the jump frequencies are larger at the same temperature) just because these frequencies depend on the composition. If there is an abrupt interface present at the very beginning of the intermixing, then the interface transfer controls the flux only until the gradients will be large enough to establish the diffusion flux $J_\beta$ larger than $J_I$. In fact the magnitude of the finite value of $J_I \cong K$ gives the permeability of the interface and it is determined by the $m$ and $V/kT$ parameters.
Experiments

The non-Fickian interface shift, i.e. the violation of Fick’s laws on the nano/atomic scales has been shown experimentally both in completely miscible (Erdélyi, 2002a) and phase separating (Erdélyi, 2004a) systems. In both cases we have follow the dissolution of a thin (10-20 atomic layers) epitaxially grown deposit into a single crystal substrate by AES or XPS, calculating the thinning of the deposit form the time evolution of the intensities of the measured substrate and deposit signals. For completely miscible system, we have used the Ni/Cu(111) system, and investigated the kinetics of the interface shift by AES. We obtained that the displacement of the interface (change of the thickness of the deposit) was practically proportional to the time.

In the Ni/Au(111) phase separating system, we have repeated the above experiment but by using XPS. Here, in very good agreement with the predicted value, \( k_c \) was obtained about 0.7.

Figure 4 - Kinetic exponent versus \( V/kT \) for different \( m' \) values. For small \( m' \) values there is a small maximum on the curves. Increasing \( m' \) at a fix \( V/kT \), \( k_c \) always increases, leading to a deviation from the parabolic law. (Erdélyi, 2004a)

Figure 5 - Change of \( k_c \) during dissolution (\( m'=7, V/kT=0.09 \)). The more layers are dissolved the closer the value of \( k_c \) is to 0.5.
Figure 6 - Scheme of the composition profile and the atomic fluxes in the linear ($t_1$, solid line) and parabolic ($t_2 >>> t_1$, dashed line) kinetic regimes. In the linear regime $J_\beta >> J_\alpha$, whereas in the parabolic one $J_\beta << J_\alpha$. The length of the arrows illustrates the intensity of the fluxes. Note that for large $|m|$ $J_\alpha$ is practically zero as indicated.

INTERFACE SHARPENING IN COMPLETELY MISCIBLE ALLOYS

Computer simulations

Using computer simulations, it was shown (Erdélyi, 2002a; Erdélyi, 2003) that on the nanoscale, for strongly composition-dependent diffusion coefficients, an initially diffuse A/B interface can become chemically abrupt even in ideal (either crystalline or amorphous) systems with complete mutual solubility.

Figure 7 - Composition distribution during intermixing in one period of a Mo/V multilayer calculated from an atomistic model, as in (Erdélyi, 2002a). The arrows represent schematically the 'flux distribution’, i.e. their lengths are proportional to the absolute value of the atomic flux.

This sharpening is surprising at first sight because the direction of diffusion is always opposite to the direction of the composition gradient: $j = -D \text{grad } c$. Indeed, for composition independent $D$, the composition profile will gradually decay and a flattening of the (sharp or diffuse) interface is expected. On the other hand, when the diffusion coefficient strongly depends on the local composition, the flux distribution can lead to a sharpening of the interface (see Figure 7). The sharpening can be qualitatively
predicted from the classical Fick I law although it is not able to provide correct kinetics on nanoscale. Since in ideal systems $D$ has a positive value, the direction of the flux is always opposite to the direction of the concentration gradient and, for concentration independent diffusion coefficients, this equation should lead to flattening of the interface. However, if $D=D(c)$, $j$ depends not only on the concentration gradient but also on the local composition of the sample. Figure 6 illustrates the ‘flux distribution’ at the interface in the initial state, when the film and the substrate are separated by a wide interface. As the concentration gradient is constant along the interface, regarding Fick’s first law, it is only $D$ on which the absolute value of the atomic flux depends. Therefore, in the case of concentration independent $D$ the atomic fluxes, independently of the position, are the same, whereas in case of $D=D(c)$ the ‘flux distribution’ follows the $D=D(c)$ function.

We have also shown using Eqs. (3-6) that sharpening takes place in case of stress effects intervening. (Erdélyi, 2003) Stress effects influences only the kinetics of the sharpening.

**Experiments**

We studied Mo/V multilayers (Erdélyi, 2004b). The structures (20 bilayers with a modulation length $\approx$ 5-6 nm) were produced by magnetron sputtering. The pure Mo and V layers were separated by a roughly 1.5 nm thick diffuse interface with a constant composition gradient. In order to follow the change of the composition profiles in situ during heat treatments, x-ray measurements were performed at the KMC2 beamline at the BESSY synchrotron. The samples were placed on a heater inside a hemispherical Be window under high-vacuum conditions. Measurements were carried out at temperatures that gradually increased from 293 K up to about 973 K in 10 steps. At each temperature consecutive symmetrical scans between 53° and 66° of the scattering angle $2\Theta$ were performed, measuring the scattering intensity around the (002) Bragg reflection of the Mo/V multilayer structure. The sample was kept at a certain temperature until no change in the diffraction pattern could be observed, and this required a minimum of two hours.

We found from the reconstruction of the diffraction patterns that the interface sharpened during the heat treatment. This conclusion is also supported by the analysis of changes of the intensity ratios during heat treatment. (Erdélyi, 2004b)

Finally, the fact that at a fixed temperature no more changes were observed after a certain time, i.e. a gradually increasing temperature was necessary, also gives a proof of sharpening. Since the interface is sharper, the Mo atoms are bound more strongly in the interface (the interface is more and more Mo-rich), consequently their diffusion into the V is slower, see also Figure 7. Thus, in order to counterbalance this effect, we had to increase the temperature slightly (the diffusivity has exponential dependence on both the temperature and the composition).

**CONCLUSIONS**

In the last 5-10 years, we have been intensively working on computer simulations of interface motion and transformation. With still continuously shrinking device structures, the question of how the microscopic laws may change on the nanoscale appears of utmost importance. This is not only interesting from fundamental point of view but also of practical interest. To plan and fabricate nanoscale devices need better understanding of the atomic scale processes. A possibly new behaviour could help to improve device properties or hinder its destruction.

It is known form Fick’s phenomenological laws that during annealing of a diffusion couple the displacement of a plane with constant composition (or an abrupt interface) is proportional to $t^{1/2}$. However, we have shown first from computer simulations that this rule can be violated on the nanoscale either in completely or restricted miscible systems. This is strongly related to the discrete character of the system on the nanoscale and to the highly neglected fact in the literature that the diffusion coefficients or mobilities depend on the composition.
Computer simulations also have shown that on the nanoscale, for strongly composition-dependent diffusion coefficients, diffuse interfaces can sharpen rather broaden in completely miscible binary systems during annealing. This sharpening is surprising at first sight, because the direction of diffusion is always opposite to the direction of the composition gradient. This phenomenon could provide a useful tool for the improvement of interfaces and offer a way to fabricate of e.g. better X-ray or neutron mirrors, microelectronic devices or multilayers with giant magnetic resistance. These phenomena predicted by computer simulations have been proved experimentally as well. This shows the efficiency and importance of the computer modelling in planning and fabrication of new devices.

ACKNOWLEDGEMENTS
This work was supported by OTKA grants No. F-043372, K61253. Z. Erdélyi acknowledges support from Bolyai János Foundation.

REFERENCES