Theory of defect processes

Phenomenological rate equations, chemical rules and computations based on processes at the atomic and microstructural levels contribute to our understanding of corrosion, radiation damage and catalysis.

A. Marshall Stoneham

Most scientific studies of imperfect solids concentrate on the properties of individual isolated defects. Such properties include their electronic structure, as shown by optical and spin-resonance data, formation energies, and diffusion parameters. Yet these properties alone do not always determine those practical applications of solid-state physics that exploit the behavior of defects. The principal extra ingredients can often be described as defect processes—how defects interact with each other and how the imperfect lattice evolves. That is, the time-dependent behavior of the imperfect solid, not just its static properties, affects how it can be used.

The importance of defects and defect processes has been realized for many years; some of the early work on tarnishing,[1] for example, dates from the 1920's. The rapid development of experimental methods in the 1940's and 1950's moved the main thrust of defect studies to what one might call a "botany" of defects. This detailed classification and analysis of defects and their properties is now largely complete in some areas (intrinsic defects in alkali halides, shallow centers in semiconductors and transition-metal ions in oxides, for example), yet incomplete in others (as in the case of deep levels in semiconductors). This approach can be contrasted with the "technological" one, in which the behavior of real materials is described. The study of defect processes provides a link between the two approaches. Here I will try to give some idea of the state of the theory of defect processes (for the theory of defects, see reference 2), emphasizing mechanisms and transient phenomena. The examples I will discuss are cases where experiment alone is not sufficient for a full understanding of the process. Such cases are typically those with inconvenient time scales, either very fast or very slow, or else situations so complex or so dirty that experiment cannot separate the many components involved. In complex situations, another important role for theory is to provide a framework in which the phenomena can be understood. But the real test of a theory is the *quantitative* understanding of behavior. We are thus concerned with quantitative predictions of the physical variables that make a system behave in one way rather than another, and ultimately with the use of such a theory to control or modify processes in desirable ways. By contrast, other authors, Hans Haken,[3] for example, are more concerned with qualitative analogies between phenomena in systems far from equilibrium and their basis in the underlying differential equations. Such an approach leads to a different kind of understanding of the complex phenomena in crystals and other systems.

Table 1 shows a primitive classification of defect processes. It is superficial, in that widely different phenomena may be grouped together and some phenomena could fit several categories, but the classification clearly shows the range and variety of defect processes in solids. The four main categories depend on the degree to which a steady state is achieved. The processes have a unifying factor, for the critical steps are determined by atomic motion, and by defect populations and their evolution. It is these components which are now open to quantitative prediction and assessment.

Marshall Stoneham is in the theoretical physics division at the Harwell Atomic Energy Research Establishment in England.

Types of theory

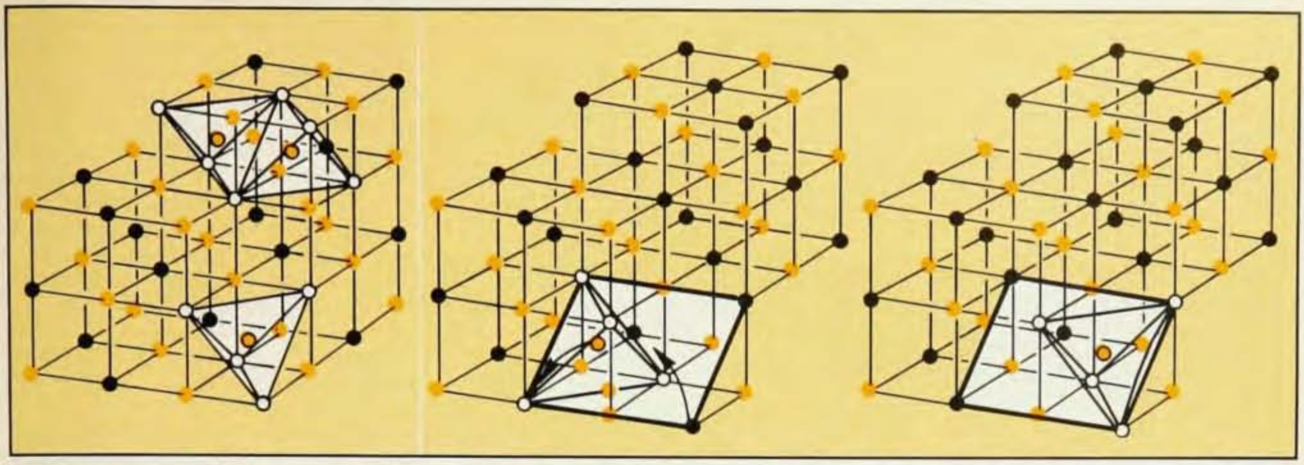
It is a tautology that one must pose the right question to get the right answer. Likewise, one must decide what type of theory is needed to establish a theory of defect processes.

For illustration, consider the oxidation of a simple metal exposed for long periods to oxygen gas. This is a classic example of corrosion, which itself includes an enormous diversity of phenomena from the simple oxidation of metals to the effects of corrosive liquids and the added complications of electrochemical and mechanical factors. Obviously, the study of corrosion is a very practical subject, emphasizing protection from corrosion—for example by avoiding faults of geometrical design, or by interfering with the corrosion mechanism by protective coatings, by electrical methods, or by manipulating sensitive components of the environment, or by identifying circumstances in which corrosion will be negligible. Thermodynamic data are an important guide, for they suggest what equilibrium state is relevant, though they do not describe how rapidly it is attained.

It is when one turns to examining the *rates* and the way in which they can be influenced that defect processes become important. We can illustrate the four kinds of theory by examples for oxidative corrosion.

Phenomenological theory: rate equations. This approach was established in the early work on corrosion,[1] by Gustav Tamann, Carl Wagner and his colleagues, Sir Nevill Mott and N. Cabrera. The basis for the rate equations rests on laws such as the diffusion equation, the law of mass action, and the conservation of charge. One example is the parabolic law central to the theory of oxidation. Suppose an oxide layer is growing on a metal at a rate limited by diffusion through the oxide, rather than by some surface or interfacial phenomenon. Then the concentration gradient falls off roughly as the

0031-9228/80/010034-09/$00.50



Defect clusters in metal oxides of the type $M_{1-x}O$. The diagram on the left shows the structure of a [4 : 1] cluster (below) (four vacancies and one interstitial) and a [6 : 2] cluster (above); the clusters are marked. The two diagrams on the right illustrates how a [4 : 1] cluster moves without dissociation when a lattice atom and an interstitial one make a transition through a saddle point. Black spots mark filled M^{2+} sites, solid color marks O^{2-} lattice sites. Vacant cation sites are shown as black circles and interstitial M^{3+} ions are filled circles. Figure 1

thickness of the oxide layer, and the diffusion equation shows that the thickness increases as the square root of the time. The same approach can be generalized to include surface-limited and space-charge influenced cases, though the resulting complexity may be impractical to handle.

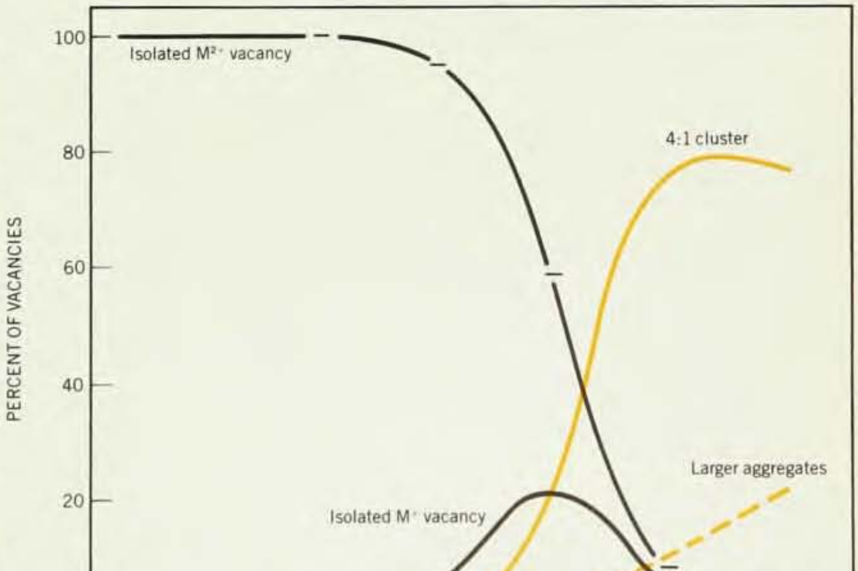
The main results of this approach, such as the parabolic law, are important because they provide a framework within which experimental results can be analyzed. Indeed, the parabolic law has become so much a part of practical corrosion studies that its origin in theory is sometimes forgotten.

Phenomenological theory: chemical. The basic ideas can be seen best by an example. Suppose the oxidation of a metal is limited by cation diffusion through the oxide, and suppose diffusion occurs by the motion of isolated cation vacancies. Any impurity that enhances the cation vacancy concentration should increase the oxidation rate. Chemical phenomenology can relate the impurity chemistry to the oxidation rate, and one can draw up systematic rules from this type of argument.

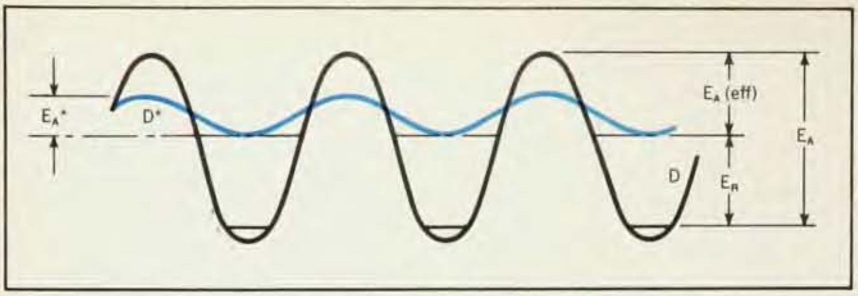
Disappointingly, the rules have rather limited success. The source of the difficulty is the correct identification of the diffusion mechanism, notably, precisely which species is diffusing and what role the impurities play. Both may be considerably more complicated than one would hope, even ignoring more subtle effects related to grain structure. The aim of the two remaining types of theory is to deal with just those features that limit the chemical phenomenological theories.

Atomistic theories involve quantitative microscopic calculations of the atomic aspects of corrosion. They attempt to answer questions such as, what defects are present in oxide layers? What are the modes of matter and charge transport? Such calculations are especially important because sensible assumptions about defects and their diffusion mechanisms can easily be wrong. Oxides are notorious in this respect. In TiO_{2-x}, for instance, a simple and sensible assumption might invoke titanium interstitials or oxygen vacancies. Yet neither is important, except possibly at extremely low non-stoichiometry (that is, when x is very small) or in association with impurities; the oxygen deficiency, it turns out, is achieved mainly through shear planes, as seen in electron microscopy. In many other cases, the nature of the dominant defect is very hard to deduce without accurate atomistic calculations. This is especially true of the complicated and diverse effects which involve chemical impurities. The shear planes in TiO_{2-x} also illustrate the important role that defect aggregates play in the behavior of substances.

Clearly, if the main defects, their populations, and their interactions with impurities are known, one can repeat the chemical theory arguments with much



Cluster formation in $Mn_{1-x}O$. We plot the fraction of the total of vacancies in each of the types of cluster as a function of the cation deficiency (or deviation from stoichiometry) x. Note the dominance of [4 : 1] clusters (solid color) for deficiencies larger than one per thousand. For very small values of x, M^{2+} vacancies (black dashed line) predominate; when the deficiency reaches several percent larger aggregates become important. (From reference 10) Figure 2



Recombination-enhanced diffusion. The sinusoidal curve represents the crystal potential; atoms diffuse by hopping from trough to trough. In local heating the activation energy, E_A, required for diffusion is effectively reduced to E_A(eff) by the energy E_R available from recombination. In local excitation the crystal is excited to the state D* (colored curve—its position with respect to E_R is arbitrary) with a reduced activation energy E_A*. Figure 3

greater certainty. One extra limitation remains, however, so that a further type of theory is needed.

Evolution of microstructure. Films of oxide on metals, and indeed almost all corrosion layers, contain microstructure: voids, grain boundaries, cracks, layers of different chemical or physical nature, and so on. This microstructure both determines the important transport processes and is determined by them. The theory of the evolution of the microstructure is the link between atomistic processes and the phenomenological rate-equation approaches. In the phenomenological approach one postulates a specific diffusion process to predict, for example, an oxidation rate. In the atomistic–microstructural approach one attempts to predict the macroscopic behavior in terms of known chemical and physical properties of the initial system. The approach thus parallels the aims of the phenomenological chemical one, which it contains as a limiting case. Just as nature can be complicated, so any general theory will be complex. However, most of the present theories of the microstructural evolution are relatively primitive. There are many special cases, in which results on specific systems are explained *a posteriori*. There are phenomenological theories too, which model a particular component—the gradual blockage of fast-diffusion paths, for example. Unfortunately, there are very few attempts to predict what microstructure should develop from a given initial metallurgical state of a sample, its environment and its geometry.

Within these various types of theory, there are various theoretical tools. Especially powerful are some of the large computer codes that calculate the electronic structure, lattice geometry and energy and associated observable properties. Members of my group at Harwell, for example, have developed a very flexible suite of programs, including codes such as Michael Norgett's HADES and Tony Harker's MOSES, for such calculations. These codes have been used extensively in a wide range of applications, in many countries, and with collaborators from many institutes. Large computers are not the only tool, for analytical approaches are very valuable. Having the tools is only part of the story too. Just as devising good experiments is hard, so is the identification of soluble problems that aid our understanding of defect processes.

Defects and defect motion

Atomistic calculations make predictions that are qualitative—such as what the main types of defects are and how they move—and also quantitative—the characteristic energy of defects, for example. Because the motions of many interacting ions are involved, the calculations use advanced computer techniques for any realistic representation of real solids. In general one assumes an interionic potential and works with a rigid-ion model or with the shell model or one of its relations.

Depending on the type of problem involved, one can use one of two general schemes for the calculations: static simulation or molecular dynamics. In static simulations one finds the geometry and energy of a perfect or imperfect crystal in equilibrium—it may be a constrained equilibrium like the saddle point of a classical diffusive lump. In molecular dynamics one follows the motion of the ions by integration of the classical equations of motion, obtaining series of snapshots from which the probabilities of various processes can be deduced. Both the static simulation and molecular dynamics routines have been implemented by workers in many countries. A good example of a general-purpose code for static simulations is the HADES code mentioned earlier.

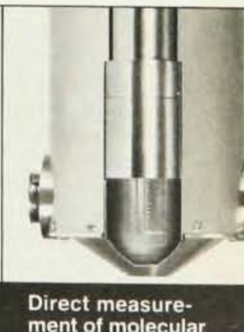
The special advantages of static simulation include flexibility and economy. Finding an equilibrium geometry for a defect lattice can be made extremely efficient. Thus if one wants to consider a truly isolated defect, one can indeed handle a large region of surrounding crystal, properly matched to an elastic dielectric continuum. If one wants to discuss defect clustering, one can handle aggregates of, say, ten or twenty defects, though obviously extreme cases will always be time-consuming. One can also treat dislocations. The simulations let one study activation energies for classical diffusion jumps and for small-polaron motion and allow one to predict charge-transfer energies.

What then are the disadvantages? First, the method is a static one, so that, in essence, it predicts internal energies, whereas experiment usually works with free energies. (Although it is, in principle, possible to obtain the free energy from the internal energy and the entropy, it is extremely difficult to calculate the entropy; only recently has anyone successfully calculated an entropy.) As a consequence, the method does not allow one to find absolute rates easily, although their temperature dependence is often accurately given. A second disadvantage is that one must choose the likely defects or processes by physical intuition, so it is always possible to miss out on an important contribution.

Molecular dynamics has complementary strengths and weaknesses. The method is far less economical, and one usually adopts periodic boundary conditions with a rather small repeating unit to reduce the computing effort. Complex defects and line defects cause severe problems for the codes, and treating the motion of small polarons (that is, polarons that are localized to within roughly a unit cell) requires significant changes in the standard procedure. Molecular dynamics, however, also has important virtues. It gives rates (probabilities per unit time)—though it may not easily give activation energies for rarer processes. One can use it to treat the liquid state as well as the solid state. And it automatically includes any processes that occur with

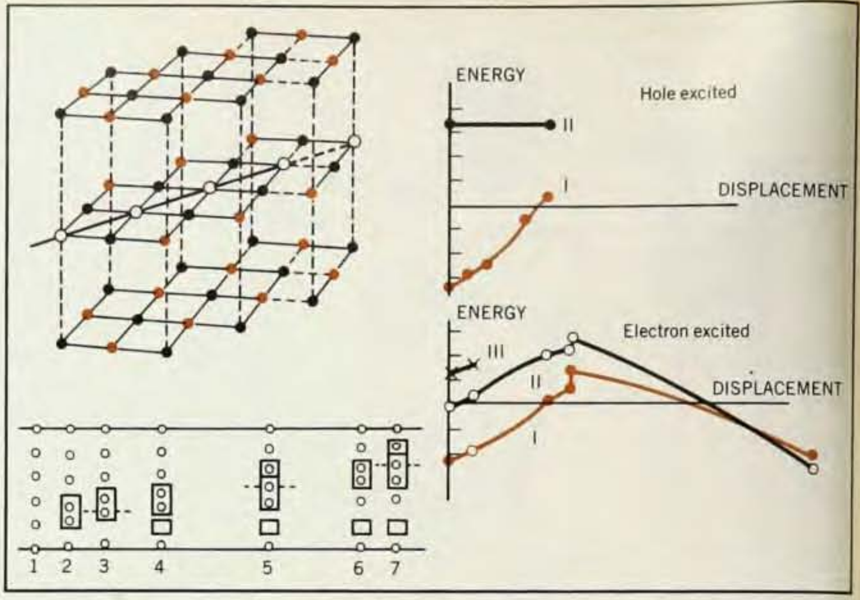
Table 1 Types of Solid State Process

Type	Main Features	Examples
Steady state, non-equilibrium	Transport, electronic or ionic	Forced diffusion in a field gradient
Quasi-steady state, non-equilibrium	Metastable states, frozen-in order or disorder	Void lattice and related formation of Glasses (conventional or metallic)
Non-steady state: Evolving system	Continued development with time	Solid-state reactions; Segregation
Non-steady state: Dynamic case	Rapid changes, without approach to thermal equilibrium	Early stages of radiation damage; Collision sequences



reasonable probability.[4] With static methods the working assumption is that the most likely process has the lowest energy barrier. While this is often so, there can be cases where processes with higher activation energies have higher rates, and molecular dynamics can identify them.

The static approach has a wide and varied history of successful use. Among the cases where processes have been analysed, rather than just the formation and motion energies of simple intrinsic defects predicted, are Richard Catlow's discussions[5] of the incorporation of oxygen into alkaline earth fluorides by hydrolysis, and his analysis of the defect behavior of nonstoichiometric UO_2 fuels. The work by Catlow, Kaz Diller and Michael Norgett on defects and defect aggregates in alkali halides[6] has been important in understanding the complicated development of radiation damage. In a quite different direction, the static modelling of the mechanism of fracture in silicon[7] has shown a number of features—such as lattice trapping—that are not present in simpler treatments. Fast ionic conduction has been a fruitful area for both static and dynamic methods. In the fluorites, for example, the static treatments[8] complement the dynamic methods[9] most effectively.



Photochemical damage in an alkali halide. The main action takes place along a (110) close-packed row of anions (indicated by the colored stripe in the diagram); the sequence of events along this row is shown below. The boxes indicate, first, the formation of a self-trapped exciton (2) and its subsequent motion through the crystal, leading to the production of an F center (the empty box in 4) and an H center (the extra filled box in 7). The graphs at right show the energy surface of the crystal as a function of an atomic displacement for several different states of excitation labelled by roman numerals. The colored curves show the energy of the recombined state; these differ because the geometry of the crystal depends on the state of the hole. Figure 4

Transition-metal oxides

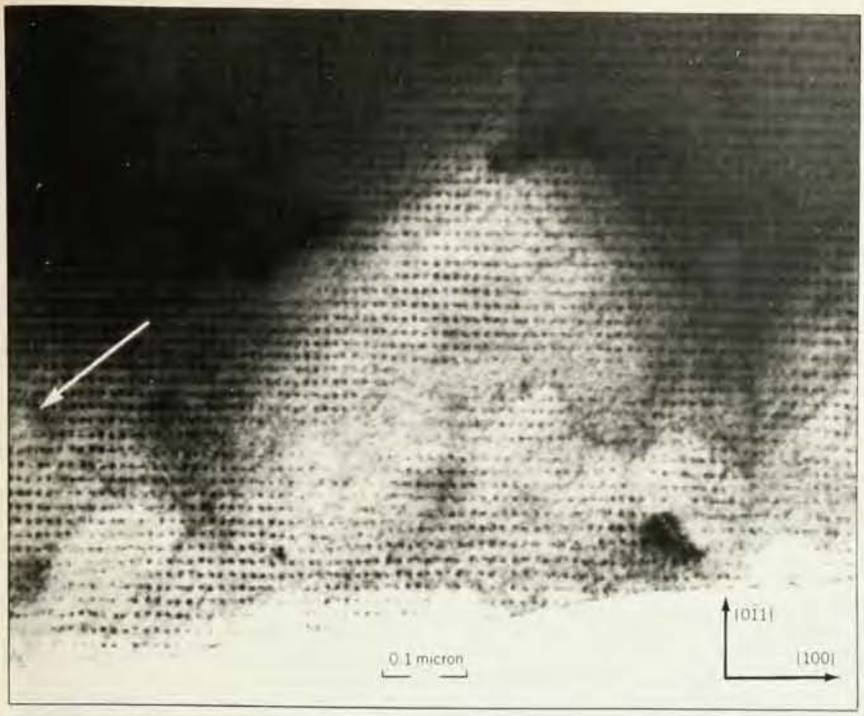
Oxidation of metals is one of the simplest cases of corrosion. Yet even when the parabolic law shows that mass transport is the rate-determining process, the behavior may be too complicated to analyze with the whole present range of experimental techniques. This is partly because of the effects of microstructure, and partly because of the subtle and varied effects of trace impurities. Theory is able to combine with experiment to give a clearer and surer picture of the main mechanisms underlying the enormous amount of experimental data.

At Harwell we have performed[10] calculations based on a well-tried model of ionic crystals, widely tested in alkali and alkaline earth halides, in alkaline earth oxides, and in a range of other oxides like Al_2O_3, TiO_{2-x}, and UO_{2+x}. The basic model combines the shell model, which gives a good description of the lattice dynamics and dielectric properties, with empirical models of short-range potentials. The computations do not explicitly include covalency, though some terms are implicit in the empirical potentials; the remaining covalency and crystal-field terms are small and can be estimated separately. Past experience with these calculations suggests they can be at least as reliable as the best experiments in most cases. We have used the calculations to predict the mechanism of charge transport, the principal defects, and their mode of motion in the cubic oxides, $M_{1-x}O$, of manganese, iron, chromium or nickel. These oxides have a deficiency of cations (rather than an excess of oxygen as, for example, in the case of UO_{2+x}), so that the form of cation vacancies and their charge compensation is thus of central importance.

A first question concerns the mechanism of charge transport. Do the free holes move as "large polarons," with a long mean free path and spread through large volumes of the crystal—like the holes in GaAs—or do they move as "small polarons," in which the hole is effectively localized to within a unit cell by the lattice distortion it causes—effectively producing an M^{3+} cation instead of M^{2+}—and diffuses randomly by hopping from cell to cell? To answer these questions one needs to know the relative importance of the polarization and distortion produced when an M^{3+} ion replaces M^{2+} and the competing reduction in kinetic energy achieved in a large-polaron band. The result can be expressed in the form that a large polaron is expected when the carrier's bandwidth exceeds a specific value. While there is no general concensus about the experimental results, they are consistent with the trend predicted: charge transport in MnO probably involves small polarons, and that in CoO and NiO is a close balance marginally favoring large polarons. (See table 2.)

A second question concerns the effect of impurities. How, for example is the motion of holes affected when the oxide is doped with alkali atoms (lithium or sodium), producing $M_{1-x}O{:}A$? We have studied both the binding of a hole to the impurity and the hopping of the hole around the impurity. Experimental data are more complete for the impurity systems. They indicate that the trapped hole is localized on one of the cations adjacent to the substitutional alkali impurity. The fact that the predictions are in good general accord with experiment is a formidable test, for the energies are only a few tenths of an electron volt, as compared with Madelung energies (essentially the electrostatic energy of the ions in the crystal) and polarization energies of a few tens of electron volts.

It is agreed that diffusion in these oxides is dominated by the motion of cation vacancies. It is far less clear what charge state or degree of aggregation is involved. Just as assimilation of anion vacancies into shear planes modifies the behavior of TiO_{2-x} dramatically, so clustering has important effects in other oxides. One can identify these effects qualitatively. Clustering changes both the defect concentrations and the fractions of mobile and immobile imperfections. Further, there may be cooperative motions that allow clusters to move with a lower activation energy than their component point defects. There is good experimental evidence for clustering of defects in oxides. For $Fe_{1-x}O$ the type of cluster can be deduced because there are enough clusters to be seen by neutron scattering.[11] In other transition-metal oxides of the same type, $M_{1-x}O$, the effects of clustering show in the precise dependence of the



Void lattice in molybdenum. The voids are about 20 Å in diameter and 200 Å apart; they are stabilized in this array by elastic interactions. (The plane shown is the [011] crystal plane.) A dislocation in the void array is marked with an arrow. To produce the voids the crystal was irradiated at 870°C with 2-MeV nitrogen ions at a rate of $7 \times 10^{17}/\mathrm{cm}^2$. (From ref. 20.) Figure 5

oxygen partial pressure on x, the degree of non-stoichiometry, and in the apparent dependence of motion and formation energies on both x and temperature. Two of these clusters are shown in figure 1. The main basic cluster proves to be the [4:1] cluster (here an $[N:M]$ cluster comprises N vacancies and M interstitials, and consequently $N - M$ net vacancies). These [4:1] clusters have the important features of stability and mobility; they are structurally interesting too, because they can be used as building blocks for the spinel structure of $\mathrm{Fe_3O_4}$. The quantitative theory of mass transport works out well: It can correctly predict formation energies and binding energies. This is important, for theory can isolate the various mechanisms and assess their relative importance. Further, mass-action arguments give the relative concentrations of different defect species for a known non-stoichiometry. A striking feature of the results, shown clearly in figure 2, is the rapid change from almost all isolated vacancies to almost all [4:1] clusters. For Mn_{1-x} this occurs at values of x between 10^{-3} and 10^{-2}.

The key points are thus that the theory successfully predicts the transport and thermodynamic properties of transition metal oxides and that defect clusters are an important component in most cases. Obviously this affects how one analyzes diffusion and corrosion experiments. For oxidation the most important implications concern the effects of impurities. Contrary to the assumptions of the simple "chemical" theories, impurities do not merely affect defect populations. They may change the major defect type (charge state, cluster form, cluster size) and they may affect the mobilities of these different defect types. For example, if an impurity were to encourage vacancy and cluster formation, but interfered with the non-dissociative diffusion mechanism of the clusters, the observed effects would not be those given by the simple chemical theories. In so far as matter transport is the only important issue, the atomistic methods can handle just those problems that are beyond the scope of the "chemical" methods. The effect of impurities on the evolving oxide microstructure is another and more formidable issue.

High carrier densities

Many diverse phenomena are affected by high densities of free carriers. These phenomena often involve transient states with very short time scales, though longer-term developments also occur. They often have very practical implications because they may cause degradation: One does not want semiconductor devices or solar cells to deteriorate under just the conditions in which they are most useful. The various processes can be put into two broad classes, depending on whether the effect of the extra carriers is to modify the electronic or lattice vibrational components of the system.

Local heating: the carriers provide energy (from their recombination energy) to increase the vibrational energy in the reaction coordinate. The recombination energy E_R reduces the effective thermal activation energy from E_A to $E_A(\mathrm{eff}) = E_A - \alpha E_R$, where α is a number less than unity (see figure 3). In general, local heating effects should be rare in solids, though possibly common in small molecules.

It is central to the model that the reaction coordinate must correspond rather closely to a normal mode of the system within a relatively narrow limit. If it does not, two problems arise. One is that it is hard to excite efficiently a motion which is not a good approximation to a normal mode. The other is that it is hard for such a systematic motion to survive long enough to have its desired effect; the energy is not effectively localized (as measured by amplitudes of vibration of specific atoms) but is spread over the whole system in delocalized modes. This delocalization can occur even before a single diffusive jump has occurred. Clearly this is less of a problem in small molecules, where all modes are, of necessity, strongly localized. Indeed, one method of laser isotope-separation exploits this: The excitation energy from a laser is transferred in part directly to the reaction coordinate and in part indirectly, *via* other modes that excite the reaction mode through anharmonic interactions. Thus the molecules can dissociate, despite the anharmonic shifts of the higher vibrational states from resonance.

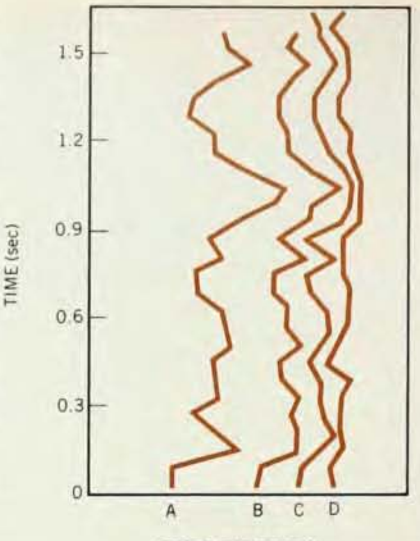
For practical purposes, the importance of a local heating mechanism depends on whether or not a local mode or resonance occurs. In simple cases, where interatomic forces are known to be changed greatly or where very light or very heavy masses are involved, one can use standard methods to show the localization occurs. There is good evidence for the mechanism in a defect involving hydrogen in SiC, for example.[12] In many other cases, however, one must be content with noting that, if the frequency associated with the reaction coordinate lies in the midst of the bulk modes, localization is probably negligible and local heating unimportant.

Local excitation: the recombination energy (or corresponding excitation) is transferred to the electronic degrees of freedom, rather than the vibrational motion. The defect is excited electronically in some sense, altering the potential energy of the crystal geometry. The different potential energy surface in the excited state allows rapid motion. The mechanisms parallel the well-known cases in which light provides the excitation, as in the motion of self-trapped holes, in the charge-state dependence of defect diffusion rates and in catalytic processes where photoexcitation can change the intermediates.

Clearly, an important role for theory is to identify the class of processes involved. Given the energy of the crystal lattice as a function of the geometry—of the arrangement of its atoms—one can often predict mechanisms for any processes involving the motion of atoms, that is, for diffusion or diffusion-like processes. Even if one cannot make precise predictions one can at least eliminate some mechanisms. The calculations may be very complicated and one must be careful to identify the key features. One must include the effects of the displacement and polarization of more distant neighbors, and one may need to obtain the energies of one- and two-electron excited states. Usually there will be charge redistribution to include self-consistently. Nevertheless, these calculations are possible, and their range and detail is improving rapidly.

Sometimes the questions asked are much less complicated. For example, whether or not there is a local resonance is central to local-heating models, because the degree of localization affects the efficiency of the excitation, the persistence of the required motion, and the eventual thermalization of that motion. Calculations for the case of the self-interstitial in diamond[13] rule out local-heating models on just that basis. As another example, whether or not the defect has different sites in different charge states is central to the applicability of the Burgoin-Corbett model,[14] in which the defect moves by alternating changes in its charge state. The local-excitation models are normally much trickier, for one needs to compare energy surfaces (energy as a function of lattice geometry) for more than one electronic state. A good example of the way the local-excitation models work can be seen from studies of optically-produced damage in ionic crystals.

Optical excitation can produce vacancies and interstitials in alkali halides in a few picoseconds. The process involves the optical production of an exciton, the self-trapping of the exciton, and a non-radiative recombination giving damage.[15] There are analogies to be drawn between the alkali halide systems and others, not only the obvious alkaline-earth fluorides but also with systems as far removed as chalcogenide glasses. The phenomenon is a good subject for any theory of defect processes, partly because of the short time scale and the large number of different potential decay channels (so experiment alone cannot give the whole picture), and partly because experiments (notably by Milton Kabler and his colleagues at the Naval Research Laboratory) provide an impressive number of related measurements of the properties of the self-trapped exciton against which any theory can be tested.



Motion of iron particles over graphite. The particles, about 500 Å in diameter were observed in a carbon-monoxide atmosphere with an electron microscope. The largest motions shown represent excursions of about 1000 Å. The particles were pinned to a ledge on the graphite surface, so the motions are in one direction only. (From reference 21.) Figure 6

The self-trapped exciton can be regarded as an electron bound in the Coulomb field of a localized hole. In its ground state the hole is effectively immobilized, self-trapped by the formation of a molecular ion; such a hole, shared by two negative ions in the crystal, is called a V_k center. In KCl, for example, the V_k center corresponds to replacing $2Cl^-$ by Cl_2^-. The self-trapped exciton can recombine by several mechanisms

- non-radiative, giving damage (around 10^{-12} sec)
- spin-allowed radiative, or σ transition (around 10^{-8} sec)
- spin-forbidden radiative, or π transition (about 10^{-3} sec)

as well as a variety of non-radiative channels with a range of lifetimes. The spin-allowed transition is seen in only a few of the alkali halides. The theory of the self-trapped exciton has been concerned with its electronic structure, with the kinetics of the transitions among its levels and with the nature of the self-trapping and damage processes. Any theory that claimed to describe the damage process should certainly predict the various spectroscopic properties of the self-trapped exciton. In practice, the combination of several methods (including pseudopotential, unrestricted Hartree–Fock and semi-empirical molecular orbital techniques) gives excellent predictions of many observables. It is particularly gratifying that it was the theory which firmly established that the σ and π transitions involved distinct orbital states; that it was kinetics, rather than selection rules, which eliminated the σ transition in some crystals, and that the self-trapped exciton probably reorients as it evolves to its ground state. The quantitative predictions give confidence, for not only does theory agree with experiment, but also different theoretical methods agree with each other.

One of the problems of interest in the study of alkali halide crystals is the question of what electronic state is the precursor of damage. The initial damage consists of neutral centers: an F center, in which an electron is trapped at a negative-ion vacancy, and an H center, in which an extra anion is squeezed into a row of anions. There are two main contendors for the precursor: one with the hole in its lowest state and the electron in one of its lower states, and the other in which the hole is excited to a π-like state. Both types of state can arise in self-trapping and de-excitation, and indeed energy may be transferred between the electron and hole components of the exciton.

The predictions of the radiation-damage mechanism were made primarily to

Table 2 Polaron energies in transition-metal oxides

Crystal	$Mn_{1-x}O$	$Fe_{1-x}O$	$Co_{1-x}O$	$Ni_{1-x}O$
Minimum bandwidth, ϵ, for the large polaron to be stable (eV)	3.06	2.18	1.88	1.6
Type of polaron found experimentally	probably small			probably large
When an alkali atom is present				
Binding energy of polaron, E_B (eV)				
Theory	0.24		$0.52 - \Delta$[1]	$0.54 + \Delta$[1]
Experiment	0.2		0.38	0.31–0.43
Energy for hole E_H (eV) hopping around alkali				
Theory			0.24	0.19
Experiment			0.20, 0.25	Complex 0.05–0.16

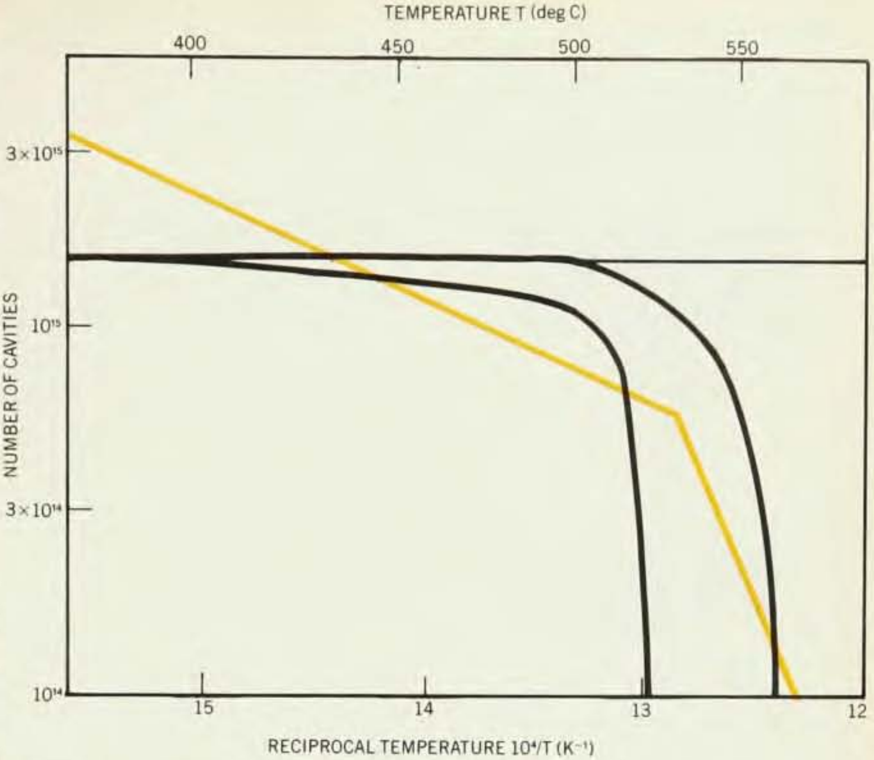
---

[1] Δ is the energy difference between large and small free polarons.

Adapted from reference 10.

answer this question.[16] Our calculations combined self-consistent molecular-orbital calculations with independent calculations of the lattice geometry by the sort of methods described earlier. The lattice distortion and electronic energies are interrelated, of course. However, here it is the hole that primarily determines the distortion, which one can calculate given potential energy curves for Cl_2^- in its various states; one must do this for the several different stages of events, starting with the self-trapped exciton and passing through the translational instability to the F and H centers. For each of these steps we calculated the electronic structure and total energy, using a 42-atom cluster for the early stages, and a 57-atom cluster for the later stages. The total energies define potential energy surfaces for the different stages of the electron and hole. These surfaces, in turn, determine the relative probabilities of different mechanisms. Figure 4 illustrates the results, showing both the sequence of events by which an exciton produces a pair of defects and the energy surfaces for the different states of the electron and the hole.

One important point is that the total energy for isolated F and H centers is higher than that of the self-trapped exciton in its lowest state. It is energetically forbidden for the exciton in its lowest state to give rise to damage, which is consistent with observation. Conversely, F and H centers have enough energy to recombine, producing the π-exciton luminescence. This has also been observed; it is not clear whether the barrier to recombination is overcome by tunnelling or by some other motion not discussed here. Still more important is that the potential energy surfaces strongly suggest that F and H center formation occurs when the hole is excited. When the hole is in its ground state, there is a barrier of over 2 eV against defect formation. This barrier persists even when the electron is excited. When the hole is excited, there appears to be no barrier at all: it should be able to produce defects athermally. The lifetime of this excited state need not be long, because one sees F centers produced within a few picoseconds of the initial optical excitation. It follows that if the hole is excited optically, starting from the lowest exciton state, then F and H centers are produced. Several workers, among them Richard Williams at the Naval Research Laboratory, and Noriaki Itoh and Masamitsu Hirai in Japan, have indeed seen F-center production after such optical excitation. Even if the results do not identify the excited state uniquely,[17] they do demonstrate a clear instance of a local excitation mechanism. Moreover, the present predictions give a consistent description of what is observed that agrees with current data.



Cavities in steel produced by irradiation as a function of temperature. The thin line shows the initial number of gas bubbles. The colored curve shows the results of variable-energy cyclotron experiments and the black curves show the results of the rate theory for two different values of the parameters. (From reference 19.) Figure 7

Microstructure and its evolution

Even though atomic processes ultimately determine solid-state behavior, it is not always convenient or constructive to discuss phenomena at this level alone. Many cases can be, and have been, discussed on a larger scale. Among these are the growth of grains in ceramics and "spinodal decomposition" in alloys (in which there is an instability with respect to the composition of the solid solution). These processes have many practical applications, from the control of grain size in ceramics for magnetic recording to alloy design. One finds four principal cases where discussion on the larger scale is particularly helpful.

The first case includes situations in which the microstructure actually determines which atomic processes are important. Corrosion is an obvious example. The grain structure determines whether diffusion occurs predominantly through the bulk or through the grain boundaries. Obviously, however, the atomic and microstructural aspects are interdependent, because the microstructure itself evolves through processes on the atomic scale.

The second important case covers situations in which several alternative extended defects may emerge by the capture or aggregation of point defects. One might have surfaces, dislocations, cavities and precipitates all competing for vacancies. We still need a theory that treats the evolution of the populations of these extended defects consistently. To a first approximation one may ignore the precise arrangement of the components, and the dominant interaction is that of competition.

In the remaining cases there are important interactions beyond mere competition, almost always the elastic interaction, in which one component distorts the host and so affects the energy of another component. This interaction has a long range, and frequently overwhelms the purely electronic interactions so often discussed in the literature. The remaining two cases may also involve motion of macroscopic defects. One concerns the evolution of order, as in the development of the void lattice and many parallel cases. An example is shown in figure 5. The other involves the dynamic behavior of macroscopic components, including both motion and change of shape. Catalysts, for example, need not be the constant and steady base they are often supposed to be. Many experiments have shown dynamic behavior, in some cases clearly a consequence of a chemical reaction taking place on the catalyst. In the most dramatic examples, metal particles on non-metallic substrates may move at up to 1000 Å/sec, and show a most varied range of behavior (see, for example, figure 6). The motions themselves can have catalytic consequences. The particles do not have a constant area in their motion, nor need the number of any "critical" reaction sites be constant. The temperature of the particles may vary during their motion as their thermal contact with the substrate alters. In the longer term, degradation may occur because metal atoms are slowly dispersed or because the particles pick up poisons during their motion.

Important questions in the physics of nuclear reactor design concern the long-term behavior of the structural components. One does not want to wait tens of years to find out what can happen to the reactor components over that period. Simulation experiments have been a powerful alternative. One uses accelerators in short-term experiments to deduce the key parameters that determine the long-term behavior. The competition between different larger defects for the point defects produced directly by radiation damage is a key problem in relating the two time scales. A "Rate Theory" [18] has been developed specifically to handle this competition between voids, dislocation, and so on, for vacancies and interstitials. It goes beyond simpler approaches of chemical kinetics by considering each defect in a self-consistently chosen average environment and by letting the defect character evolve with size.

In irradiated steels one distinguishes between two types of cavity: the gas bubble, in which gas pressure is important, and the void, which may be regarded as empty. The distinction is never complete; in practice one talks of voids when the gas pressure is much less than that associated with surface tension; for bubbles the two pressures are comparable. The balance between the pressure and the surface tension affects the equilibrium concentration of vacancies. The surface tension attempts to reduce the cavity size, favoring emission of vacancies, whereas the gas pressure favors cavity growth.

When a steel is irradiated, both the swelling and the numbers of bubbles and voids are influenced by the initial cavity population and by the gas both initially present and produced during irradiation. Figure 7 gives an example of how the rate theory provides a successful model for this process[18] and shows that the observed void concentrations decrease with the temperature of irradiation. The rate of decrease changes rapidly at one point, giving a "knee" in the graph of the number of cavities versus reciprocal temperature—the "Arrhenius plot." The precise location of the knee depends on what gas is present and is successfully given by rate theory.

In this brief survey of defects and the solid-state processes they determine I have described only a few of the many important non-equilibrium phenomena which underly the technological use of solids and the theories that are necessary to understand them. These theories are not just empirical phenomenology, but have a real predictive power, both qualitative and quantitative. They thus affect the questions one may attempt to answer: Instead of being limited to those that are convenient to answer experimentally, one can explore those questions for which answers are needed.

* * *

This paper is based on an invited talk given at the 1978 Warwick Solid State Physics Meeting of the Institute of Physics. A fuller account of many of the points discussed has appeared in Advances in Physics **28**, 457 (1979).

References

1. G. Tamann, Z. Anorg. Chem. **111**, 78 (1920); C. Wagner, K. Grünewald, Z. Phys. Chem. B **40**, 455 (1938); N. F. Mott, N. Cabrera, Rep. Prog. Phys. **12**, 163 (1948).
2. A. M. Stoneham, *Theory of Defects in Solids*, Oxford U.P. (1975).
3. H. Haken, *Synergetics*, Springer, Berlin (1977).
4. C. H. Bennett, Thin Solid Films **25**, 65 (1975); A. de Fano, G. Jacucci, Phys. Rev. Lett. **39**, 951 (1977).
5. C. R. A. Catlow, Proc. R. Soc. **A353**, 533 (1977); C. R. A. Catlow, J. Phys. Chem. Sol. **38**, 1131 (1977).
6. C. R. A. Catlow, K. Diller, L. W. Hobbs, Phil. Mag., in press.
7. J. Sinclair, J. Phys. C **5**, L271 (1972), Phil. Mag. **31**, 647 (1975); J. Sinclair and B. R. Lawn, Proc. R. Soc. **A329**, 83 (1972).
8. C. R. A. Catlow, R. T. Harley, W. Hayes, J. Phys. C **10**, L559 (1977).
9. A. Rahman, J. Chem. Phys. **65**, 4845 (1976); M. Dixon, M. Gillan, J. Phys. C. **11**, L165 (1978).
10. C. R. A. Catlow, W. C. Mackrodt, M. J. Norgett, A. M. Stoneham, Phil. Mag. **35**, 177 (1977); Phil. Mag. **40**, 161 (1979).
11. C. R. A. Catlow, B. E. F. Fender, J. Phys. C **8**, 3267 (1975).
12. P. J. Dean, W. J. Choyke, Adv. Phys. **26**, 1 (1977).
13. A. Mainwood, F. P. Larkins, A. M. Stoneham, Solid-State Electron. **21**, 1431 (1978).
14. J. Bourgoin, J. W. Corbett, Phys. Lett. **38A**, 135 (1972).
15. D. Pooley, Proc. Phys. Soc. Lon. **87**, 245 (1966); H. Hersh, Phys. Rev. **148**, 298 (1966); N. Itoh, J. de Physique **37** (Supp.) C7-27 (1976).
16. N. Itoh, A. M. Stoneham, A. H. Harker, J. Phys. C **10**, 4197 (1977).
17. R. T. Williams, Semicond and Insulators **3**, 251 (1978).
18. S. D. Harkness, C.-Y. Li, Met. Trans. **2**, 1457 (1971); A. D. Brailsford, R. Bullough, J. Nucl. Mat. **44**, 121 (1972), J. Nucl. Mat. **69/70**, 434 (1976).
19. M. R. Hayns, AERE report R8806 (1977).
20. J. H. Evans, Nature **229**, 403 (1971), Rad. Eff. **10**, 55 (1971); the theory is reviewed by A. M. Stoneham, in AERE R7934 (R. S. Nelson, ed.) p. 319 (1974).
21. R. T. K. Baker, P. S. Harris, R. B. Thomas, Surf. Sci. **46**, 311 (1974). ☐