

Non-radiative transitions in semiconductors

A M Stoneham

Theoretical Physics Division, AERE Harwell, Oxon OX11 0RA, UK

Abstract

Non-radiative transitions affect many aspects of semiconductor performance. Normally they reduce device efficiency by suppressing luminescence, creating defects, reducing carrier lifetimes, or enhancing diffusion during operation. The present review surveys both the theoretical and practical understanding of non-radiative transitions. It includes general theoretical results and the associated ideas, with the emphasis on phonon-induced and defect Auger processes. Most of the purely formal aspects are omitted, but the points of principle where uncertainties remain are discussed. The review also covers the relation between basic theoretical studies and practical applied work on device degradation. This includes a description of the atomic processes involved in the more important mechanism of device deterioration and the theoretical understanding of the mechanism of these underlying processes. Finally, there is a survey of models proposed for 'killer' centres.

This review was received in June 1981

Contents

	Page
1. Introduction	1253
1.1. Energy surfaces: the configuration-coordinate diagram	1254
1.2. Promoting and accepting modes	1255
1.3. Thermodynamics of non-radiative transitions	1257
2. Non-radiative transition rates	1258
2.1. Transitions between two bound states	1258
2.2. Free-bound transitions	1267
2.3. Effects of perturbations	1271
3. Cooling transitions	1272
3.1. Transitions within a single level	1272
3.2. Transitions between two states	1274
3.3. The Dexter-Klick-Russell criterion: luminescence or not?	1275
3.4. Evolution of the excited state	1275
4. Special cases of non-radiative transitions	1276
4.1. Recombination-enhanced processes	1277
4.2. Persistent photoconductivity	1279
5. Auger processes	1281
5.1. Defect-free Auger processes	1281
5.2. Impurity-associated Auger processes	1281
5.3. Laser annealing	1284
6. Theory and practice and non-radiative transitions	1285
6.1. Non-radiative transitions in devices	1285
6.2. The fitting of defect parameters	1286
6.3. Relation to the theory of electronic structure	1287
6.4. Killer centres	1290
References	1291

1. Introduction

When a solid is stimulated into an excited state, it may recover by several alternative processes. Some, the radiative processes, lead to the emission of light. Others simply redistribute the absorbed energy among the lattice vibrations or among electronic excitations. These, the non-radiative transitions which ultimately degrade the absorbed energy to heat, are the subject of this review.

Non-radiative transitions manifest themselves in many ways: suppressed luminescence, reduced carrier lifetimes, the creation of intrinsic defects, enhanced diffusion, and so on. Most of these diminish the efficiency of semiconductor devices. In a few cases only are the non-radiative transitions beneficial, for example when a semiconductor is to be used as a bolometer or when one wishes to isolate regions electrically or optically. However, in all cases the control of non-radiative transitions is potentially beneficial. Many devices work at well below their theoretical efficiencies, and there are clear advantages in reducing power consumption and the associated difficulties of over-heating of components. Yet to control non-radiative transitions one needs either a large investment in materials preparation in the hope that the problem can be circumvented or an understanding of the processes so that these effects can be minimised more easily. It is with the understanding of non-radiative transitions that theory can be of especial help. Much of the formal theory is exceptionally complex, and there remain some problems of principle. Despite this, there are a number of results which are sufficiently general and accurate to be of practical value. The present review concentrates on these in the context of III-V semiconductors. Wider discussions of multiphonon transitions are given in several other reviews (Jortner and Mukamel 1975, Stoneham 1975 chap 14, 1977a, Mott 1978, Toyozawa 1978, 1980, Newmark and Kosai 1981) and above all in the fine recent book by Englman (1980). The Auger processes, in which energy is transferred to electronic excitations, have also been reviewed recently (Robbins and Landsberg 1980), and these workers cover in depth areas I shall only discuss briefly.

When planning this review, I did consider giving a systematic analysis of the centres seen. This rapidly became unacceptable. Whilst huge amounts of data were available (see, for example, Grimmeiss 1978), much relevant to this review was obtained uncritically, sometimes inconsistently, and often for ill-defined defects. In consequence, I have elected to concentrate on some of the areas where points of principle arise and, in addition, to avoid areas previously reviewed. The present paper is thus complementary to my own previous reviews and to Englman's book.

Even though I shall concentrate on semiconductors, it would be misleading not to mention some other areas with related phenomena. The four fields in which non-radiative transitions are especially well understood are (i) transition metals (including rare-earth ions) in oxide hosts, where much work has been stimulated by the need for phosphors, (ii) alkali halides, where both reorientation and radiolysis have been the subject of detailed studies, (iii) molecules and molecular crystals, and (iv) chemical reaction studies. It is notable that, in all these cases, progress has been rapid because the centres involved and the states involved have been successfully identified. In semiconductors, sadly, this is not usually the case, with consequences that should become clear.

This introduction concludes by discussing a number of general features common to many aspects of non-radiative transitions.

1.1. Energy surfaces: the configuration-coordinate diagram

One of the central concepts is that of an energy surface. The related idea of a configuration-coordinate diagram is widely used and abused. As a simple way to describe complex phenomena, the idea is magnificent. As a complete description, it is potentially deceptive. This subsection supplements the discussion in Stoneham (1975), with emphasis on special points we need later.

1.1.1. Normal modes and other coordinates: perfect crystal. We start from the perfect host lattice in its electronic ground state (assumed non-degenerate for present purposes). The motion of the nuclei is determined by a Hamiltonian of the form

$$H = \sum_i \frac{1}{2} M_i p_i^2 + \sum_{i,j} \frac{1}{2} x_i V_{ij} x_j + \varepsilon$$

where the vectorial and tensor subscripts are suppressed. The individual atoms are labelled i , and their momenta p_i and displacements x_i ; the harmonic interactions are described by the force-constant matrix V_{ij} . Provided we ignore the higher-order terms ε , one can take linear combinations of the x_i to form *normal modes* q_α so as to eliminate cross-terms proportional to $(q_\alpha q_\beta)$ with $\alpha \neq \beta$:

$$H \approx \sum_\alpha \frac{1}{2} \mu_\alpha \dot{q}_\alpha^2 + \sum_\alpha \frac{1}{2} K_\alpha q_\alpha^2 + \varepsilon.$$

The new effective masses are μ_α and the new (diagonal) force constants are K_α .

Note a given combination of displacements $\sum_i a_i x_i$ is *not* a normal mode in general, and motion in such a general coordinate cannot be regarded as motion of an independent oscillator.

1.1.2. Defect terms for energy surfaces. When there is a defect present in its ground state we can write the extra terms in the energy as (a) corrections to the existing masses, force constants and higher-order terms, which alter the normal modes ($q_\alpha \rightarrow \tilde{q}_\alpha$), (b) an additional constant term, important in the formation energy but not relevant dynamically, and (c) an important linear term:

$$h = -\sum_\alpha f_\alpha q_\alpha = -\sum_\alpha \tilde{f}_\alpha \tilde{q}_\alpha$$

where the f_α are the generalised forces on modes q_α . The effect of the linear term is to displace the equilibrium positions of the nuclei. The linear term does not affect the eigenvalues or eigenvectors of the harmonic Hamiltonian:

$$\frac{1}{2} \mu \dot{q}^2 + \frac{1}{2} k q^2 - f q = \frac{1}{2} \mu (\dot{q} + \delta)^2 + \frac{1}{2} k (q + \delta)^2 - \frac{1}{2} f^2/k$$

where $\delta = f/k$ is the displacement and $f^2/2k$ is the *relaxation energy*.

The configuration-coordinate picture attempts to represent all these terms, especially the linear ones, as well as possible within the framework of a *single* coordinate. It is quite common for the one-coordinate picture to fail (see §1.2). Even when a one-coordinate description works well, it may do so only for a finite range of temperatures or phenomena. One can imagine choosing a particular configuration coordinate in various ways:

- (i) maximise relaxation energy;
- (ii) optimum optical lineshape;
- (iii) optimal low-temperature transition probabilities;
- (iv) optimal high-temperature transition probabilities;

and various more exotic options. One has no assurance that one chosen effective frequency will represent all features well, nor even that the configuration coordinate will itself be sensibly constant from case to case. There is one obvious conclusion which I shall stress in several places: if one does not know rather a lot about the atomic nature of a defect, there is no point in going beyond rather primitive descriptions of non-radiative transitions in making fits to experimental data.

1.2. Promoting and accepting modes

We shall use later an expression for the non-radiative transition rate in the form

$$W = \frac{2\pi}{\hbar} |M|^2 G(\omega)$$

where M is a transition matrix element and $G(\omega)$ is a generalisation of the usual $\delta(\omega)$ expressing energy conservation. It can be seen that the lattice vibrations may enter in two different ways. First, the energy-conservation factor will balance the electronic energy with the vibrational energy into which it is converted. The modes involved here are the *accepting modes*; the configuration-coordinate diagram implicitly refers to just such a mode. Secondly, the matrix element inducing the transition may vary with lattice geometry. If so, the modes which affect the matrix element are the *promoting modes* (see figure 1).

Several remarks are essential. First, these 'modes' are strictly 'reaction coordinates', i.e. they are not necessarily the normal modes of either the ground or excited state of the system. Secondly, a mode may be both promoting and accepting, neither promoting nor accepting, or either one: the two categories are not exclusive. Sometimes there are distinctions by symmetry. A good example is given by Huang and Rhys (1950). Consider the F centre in ionic crystals where, in essence, there is a 2p to 1s transition at a hydrogen-like defect. The totally symmetric (A_{1g}) modes are the accepting modes, reflecting the difference in radial charge density in the two states. The promoting modes, which must

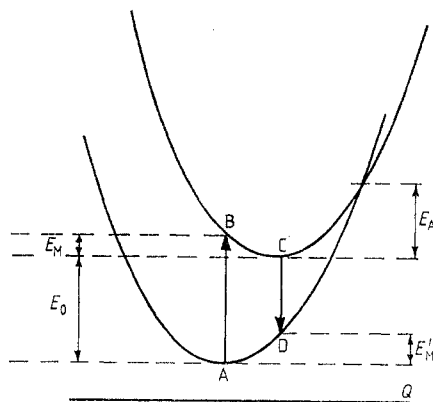


Figure 1. Notation for a simple two-level system. Q is the accepting mode. E_0 is the 'zero-phonon' energy. E_M and E_M' are the relaxation energies. The absorption energy is $E_0 + E_M$ and the emission energy $E_0 - E_M'$. The parameter Λ is $E_M(E_0 + E_M)^{-1}$.

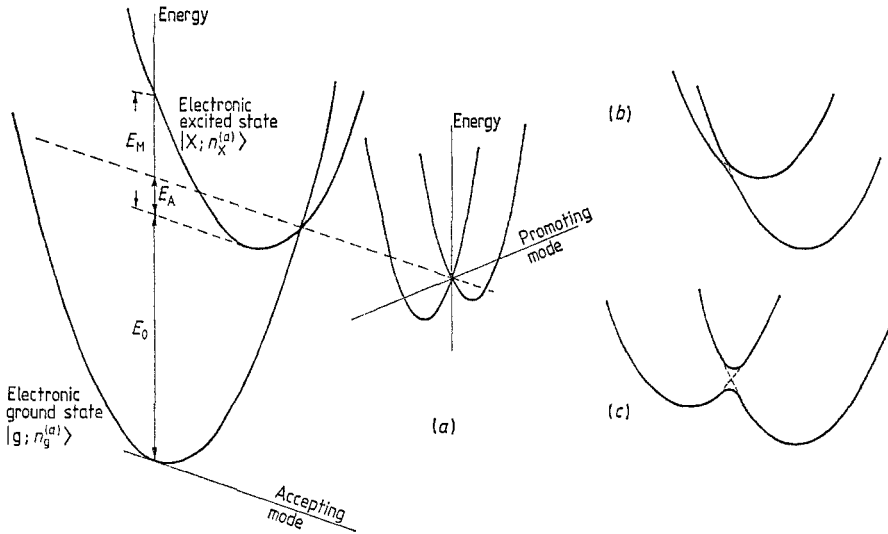


Figure 2. Examples of energy surfaces. (a) Energy surfaces for an idealised centre, without degeneracy, interacting with both a promoting and an accepting mode. (b) Avoided crossing. This corresponds roughly to (a), though here the accepting mode is also a promoting mode. (c) Avoided crossing. An alternative form to (b).

mix the $1s$ and $2p$ states, have the same symmetry as an electric field (T_{1u}) at the site. Here the accepting and promoting modes are quite distinct, with opposite parity. Figure 2(a) shows a generalisation of the configuration-coordinate diagram for a case like this, and emphasises one of the circumstances in which one-coordinate models are inadequate. This may be contrasted with figures 2(b) and (c), where there is an avoided crossing resulting from the accepting mode acting partly as a promoting mode. Thirdly, the matrix element does *not* need to be linear in the promoting mode. Figure 3 shows a variety of possibilities, all leading to characteristically different temperature dependences.

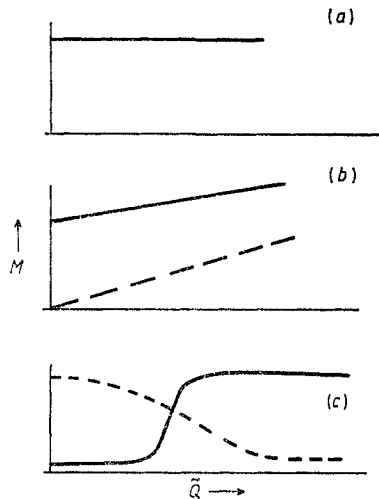
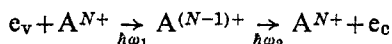


Figure 3. Dependence of matrix element on geometry. Here \tilde{Q} is the promoting mode. Various characteristic forms are shown, illustrating both 'linear' (b) and 'threshold' (c) deviations from the Condon approximation (a).

1.3. Thermodynamics of non-radiative transitions

There are several slight but important problems in assigning a specific 'level in the gap' to a particular defect state. These may be outlined as follows.

(i) *Thermal versus optical transitions*: the Stokes shift. Broad-band optical transitions occur essentially at constant geometry (the Franck–Condon principle), whereas thermal transitions take place with essentially the minimum energy. The difference (figure 1) is the *relaxation energy* and can easily be several tenths of an eV. Note in particular that this sequence of optical transitions, whose sole effect is to take an electron in the valence band (e_v) to the conduction band (e_c), requires more energy ($\hbar\omega_1 + \hbar\omega_2$) than the gap E_G :



because of the relaxation energy in each state of the reaction. Here A^{N+} and $A^{(N-1)+}$ are the charge states of the defect species A involved.

(ii) *Defect energies*. Theoretical defect energies and experimental ones need not correspond. First, there is a distinction between one-electron levels and the (many-electron) level which may be needed. One-electron levels of deep centres and 'energies in the gap' are not related simply. Secondly, most theories predict thermodynamic internal energies U . Such theoretical values cannot be compared directly with the free energies G actually observed at constant pressure and finite temperature.

(iii) *Thermodynamic quantities*. In suitable cases, experiments can give both enthalpies H and entropies S . One example is when both the thermal emission rate e_t and capture cross section σ_t can be measured over a range of temperatures. The principle of detailed balance shows that the two are related by

$$e_t = \sigma_t v_{th} N_{eff} \exp(-\Delta G/kT)$$

where $v_{th} \equiv \sqrt{3kT/m^*}$ is the thermal velocity and N_{eff} is the effective density of states. Rearranging:

$$\Delta G = kT \ln(\sigma_t v_{th} N_{eff}/e_t).$$

If $\Delta G \equiv \Delta H - T\Delta S$ is plotted as a function of T , the low-temperature limit is ΔH and the slope is $-\Delta S$.

The entropy contains two main contributions. One is configurational, essentially involving the degeneracy in our context. The second is vibrational. As has been noted (e.g. Lowther 1980) the two are not distinct when there is a Jahn–Teller effect. Moreover, the vibronic entropy may be distinctly temperature-dependent. There is a simple rule (which is essentially the point of van Vechten (1975a, b, see also Heine and Henry 1978)) that the vibrational entropy contribution is only large when the transition involves a substantial change in charge density, e.g. a compact to a diffuse state. As a final point, we note that the entropy changes at constant pressure and at constant volume differ by a leading term $\Delta S_p - \Delta S_v \equiv [(\text{thermal expansion coefficient})/(\text{isothermal bulk modulus})] \times (\text{change in volume at constant pressure})$. This may be quite large.

(iv) *Apparent activation energies*. Quantities like emission rates and capture cross sections often have a temperature dependence dominated by an Arrhenius factor of the form $\exp(-\varepsilon/kT)$. However, there is usually a slowly varying pre-exponential factor too. If the true dependence is given by the function $f \equiv T^n \exp(-\varepsilon/kT)$, then the effective activation energy $\varepsilon_{eff} \equiv kT^2 (df/dT)/f$ becomes

$$\varepsilon_{eff} = \varepsilon + nkT.$$

The correction term can be quite significant if ε is small.

(v) *Isolated defects.* Activation energies are measured ideally for isolated defects in a uniform host. Such ideals may not be realised. In particular, high concentrations of defects can introduce impurity bands associated with their excited states, and these may suggest artificially low ionisation energies. Some experiments exploit situations involving high electric fields or carrier density inhomogeneities, both of which can mislead the unwary.

(vi) *Shifts relative to band edges.* It is frequently said from examination of the temperature or pressure dependence of some transition energy that 'a level L is pinned to band edge B', and some conclusion is then proposed about the wavefunction of the bound state corresponding to L. For very shallow donors and acceptors this has some merit, though the wavefunction may still have significant components from higher bands. For deep levels, or for isovalent impurities, the description is more misleading than helpful. Distant bands can be extremely important even if only one band is important; a simple examination of model systems (like the Slater-Koster model) shows the band edge alone does not define the level for a given perturbation.

2. Non-radiative transition rates

This section derives some of the important expressions for non-radiative transition rates. The underlying results are in three main categories. First, there are transition rates between two distinct electronic states. It is these which are used in the second category, in which the evolution and kinetics among an ensemble of states are considered. Thirdly there are the cooling transitions (described in more detail in §3) by which vibrational energy in the accepting mode is transferred to other vibrational degrees of freedom (dissipative relaxation). These three types of result, combined in different ways, lead to carrier capture rates, to questions of the existence or non-existence of luminescence, to interpretations of the effects of perturbations, and to assessments of the usefulness of the various possible approaches to electronic structure on which predictions might be based. On this last aspect, it is important to distinguish between those results which depend only on the energy surfaces $E_i(Q)$ and those which need electronic wavefunctions $\psi_i(r; Q)$ too. Many of the features of non-radiative transition theory are common to chemical reaction theory. A very clear and stimulating discussion of reaction paths on energy surfaces is given by Muller (1980). Figure 4 summarises some of the relationships.

2.1. Transitions between two bound states

2.1.1 *Basic definitions.* In the non-radiative transition, electronic energy is transformed into thermal energy. Clearly, if the initial and final electronic states are labelled i and f,

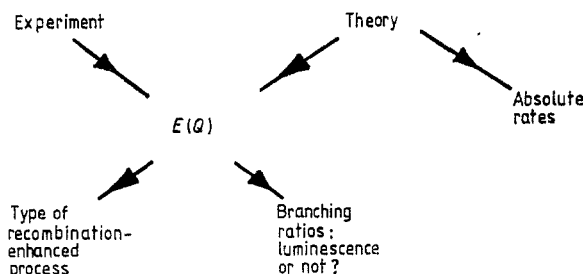


Figure 4. Status of energy surfaces $E(Q)$ in non-radiative transition studies.

and if the corresponding occupancies of the phonon modes are $\{n\}$ and $\{n'\}$, the transition rate involves probabilities of the form $w_{if}(\{n\}, \{n'\})$. These transition probabilities are not those usually observed in solids, where experiment normally keeps track only of the electronic state. The observed probabilities are of the form

$$W_{if} = \left\langle \sum_{\{n'\}} w_{if}(\{n\}, \{n'\}) \right\rangle_{\{n\}}$$

with an average over the initial phonon occupancies and a sum over the possible final vibrational states. The initial state need not correspond to thermal equilibrium. The lack of equilibrium can result from selective population of the initial states, or from transient non-equilibrium phonon distributions.

We now use two ideas mentioned in §1, namely potential energy surfaces $E_i(\mathbf{Q})$ and $E_f(\mathbf{Q})$, which will be described by a configuration-coordinate diagram, and a factorisation of the wavefunction into an electronic and a vibrational part. The transition probability will then take the form

$$(2\pi/\hbar) |(\text{electronic matrix element})|^2 \times (\text{lattice response component}).$$

The electronic component and the factorisation will be discussed in more detail later (§2.1.3). For present purposes we concentrate on the two energy surfaces shown in figure 1. This diagram defines several energies. The energies involved in low-temperature optical measurements are the energy $(E_0 + E_M)$ for an upward Franck–Condon transition, the corresponding downward transition energy $(E_0 - E_M)$ and the energy of the zero-phonon line, E_0 . The ‘crossover’ energy E_A above the upper minimum will prove important later. It is related to the other energies by

$$E_A = (E_0 - E_M)^2 / 4E_M$$

when $E_M = E_{M'}$. If we introduce an effective phonon energy $\hbar\omega$, we can now define three dimensionless parameters:

(i) $p \equiv E_0/\hbar\omega$, the number of accepting phonons corresponding to the zero-phonon line energy. Multiphonon processes have large values of p ; the *statistical limit* is $p \gg 1$. Transitions with $p = 1$ occur as components of cascade capture, and spin–lattice relaxation involves both $p = 1$ (direct) and $p = 2$ (Raman) processes.

(ii) $S_0 \equiv E_M/\hbar\omega$, the Huang–Rhys factor. The *strong coupling* limit is $S_0 \gg 1$ and the *weak coupling* limit is $S_0 \ll 1$. These regimes are defined by S_0 alone for optical spectra, where the lineshape function (§2.1.4) is critical. In such situations it is immaterial whether the upper minimum lies inside or outside the lower energy surface (figure 5). For non-radiative transitions, S_0/p is more critical.

(iii) $\Lambda \equiv E_M(E_0 + E_M)^{-1} = S_0(S_0 + p)^{-1}$, the ratio of the relaxation energy to the Franck–Condon absorption energy. Λ determines the relative values of E_M and E_A . For $\Lambda < \frac{1}{2}$, E_A exceeds E_M ; for $\Lambda = \frac{1}{2}$, E_A vanishes, the upper minimum coinciding with the point of intersection. For $\Lambda > \frac{1}{2}$ the upper minimum lies outside the lower curve (figure 5).

2.1.2. Partitioning: upper bounds to rates. When one considers optical transitions, there is a fairly clear division into (i) the electronic system which is to be excited or de-excited, (ii) the electromagnetic field, and (iii) the interaction between the electronic system and the field. There is rarely any difficulty in deciding between which states the transitions occur, simply because the energy transfer takes place between two *distinct* subsystems. When one considers non-radiative transitions, however, one is concerned with a closed system, and problems can arise if the states between which the transitions occur are

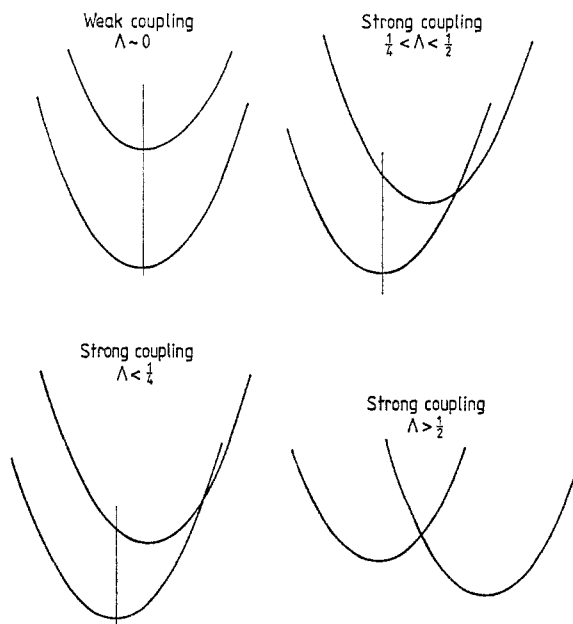


Figure 5. Four important cases from weak to strong coupling. All four are realised in practice. Λ is $S_0(S_0+p)^{-1}$.

chosen carelessly (see also Englman 1980, §8.3). As a trivial example, suppose one has a time-independent Hamiltonian $H=H_0+h$, with time-independent eigenstates ψ_{i0} of H_0 and ψ_i of H . Clearly, if the system starts at state ψ_i at $t=0$, it will remain there. However, if the time-dependent Schrödinger equation is solved with the system ψ_{i0} at time $t=0$, the component h of the full Hamiltonian will cause the state to change with time, possibly very rapidly, and in a way which depends on the partitioning of H into H_0 and h . This example makes it obvious that both the choice of basis and the boundary condition are important. In much the same way the usual second-order perturbation theory would give different results if the electronic states were for a static lattice than if they were Born–Oppenheimer states.

Since spurious overestimates of transition probabilities can arise, it is useful to look at upper bounds to rates. There are no rules quite so powerful as the f sum rule of optical transitions, but we may note several arguments.

(i) Certain transitions require a *lattice condition* to be satisfied. If Q is the reaction coordinate (usually the accepting mode) then the condition is normally either (a) Q has some critical value, $Q=Q_c$, to within some specified accuracy δQ , or (b) Q exceeds some threshold, $Q \geq Q_c$. The first class includes cases where transitions involve a rapid electronic motion without energy exchange to the sluggish lattice; the second includes the traditional ‘over an energy barrier’ transitions. The important point is that only *one* transition between two discrete states, if any, occurs each time Q satisfies the condition. Even if the electronic transition can occur *many* times whilst the condition is satisfied, the *net* effect of these many forward and reverse transitions is either one transition or no transitions. The net transition rate which results is thus:

(number of times per second the lattice condition is satisfied)

× (efficiency factor η , $0 \leq \eta \leq 1$).

It should be stressed that this does not hold, without changes, when a continuum of states is involved. However, for two states:

maximum rate = rate of occurrence of events satisfying the lattice conditions.

This rate is itself bounded by ω_{\max} , the most rapid lattice vibration frequency (the longitudinal optic frequency in most cases; a local mode frequency where appropriate). Thus these phonon-induced transitions occur at a rate less than ω_{\max} .

(ii) The same restriction also applies when one assumes a limit from either dephasing or energy relaxation towards a final state. The rates of change of Q (defined by, for example, $(dQ/dt)/\langle Q^2 \rangle^{1/2}$) or of dephasing (defined by the fractional dispersion of the phonons involved) are again bounded by ω_{\max} .

(iii) Whilst limits on Auger processes are less clear, parallel arguments suggest the plasma frequency as a bound.

These several arguments provide only general rules, and are not complete. Nevertheless, they do refute the view sometimes heard that theory A is better than theory B merely because it predicts a higher rate.

2.1.3. Electronic matrix elements. The electronic matrix element frequently causes feelings of discomfort. One reason is that, even when the matrix element is correctly defined, it may be beyond the power of many methods to calculate accurately. Another is that choices which are mainly matters of convenience are sometimes regarded as matters of principle. As noted in §1, there is plenty of freedom in choosing a basis, provided the remaining terms in the Hamiltonian are included to high enough order and provided the correct time-dependent property is calculated. This point is taken up in a recent paper by Morgan (1981).

(a) *Adiabatic approximations.* For non-degenerate systems (and here it is Jahn-Teller systems and the like, rather than systems with isolated close approaches of energy surfaces, which are excluded) four basis sets are commonly invoked:

- (i) crude adiabatic (static, Hartree, etc);
- (ii) Born-Oppenheimer adiabatic;
- (iii) Born-Huang adiabatic;
- (iv) diabatic.

There is a particularly lucid discussion of (i)-(iii) by Azumi and Matsuzaki (1977). Markham's (1956) penetrating paper concentrates on the same approaches. Both papers list earlier workers and their nomenclature, and Englman (1980, §8.1) gives a more recent survey. Morgan (1981) gives explicit calculations in several schemes. The diabatic choice is particularly common in molecular collision theory, and is sometimes invoked in non-radiative transitions when there is a crossing of energy surfaces. Sometimes the diabatic choice is implicit rather than explicit, notably in the use of the Landau-Zener expressions. There are substantial differences between the normal non-radiative transition situation and the molecular collision case so that, for practical purposes, diabatic representations are mainly of use for rather specific technical problems in which the crossover region is dominant.

From the many studies, several points emerge. First, it is always possible to transfer from one basis to another. This is usually done by perturbation expansion (see Azumi and Matsuzaki 1977) though choice of detail remains. Huang (1981) demonstrates elegantly the equivalence of the adiabatic and static forms and points to potential dangers. Secondly, one is frequently faced with a decision about the Condon approxima-

tion: is a matrix element of the form

$$\Omega_{12}(Q) \equiv \int d^3r \psi_1^*(r, Q) \Omega \psi_2(r, Q)$$

independent of Q ? The evidence in general terms is rather weak (see Stoneham 1975, §10.3), even for Ω representing the optical dipole operator. In the static approximation, the Condon approximation is always implied. This is not an advantage, as sometimes claimed; it simply means that the geometry dependence of the wavefunctions must be handled separately. Normally such dependence is handled by the Herzberg–Teller expansion (for details, see Azumi and Matsuzaki (1977)). Thirdly, some discussions of the relative merits of static or adiabatic approximations are framed in the form ‘is the best description of the individual initial and final states one in which the lattice distortion is fixed by the mean electronic charge distribution in that state (essentially the static approximation) or one in which the distortion follows the electronic motion?’. The question is important (for discussions, see Markham (1956), Buimistrov and Pekar (1957) and Stoneham (1975, p243)). However, it is almost entirely irrelevant for non-radiative transitions in which *off-diagonal* elements are needed and for which the ‘best’ description of one state may not be the ‘best’ for the other. In particular, any phenomenon which depends on level crossing or on a narrowly avoided crossing will usually involve distortions well away from the ‘average’ geometry of either state, and will involve rapid changes with geometry of electronic wavefunction.

(b) *Beyond the Condon approximation.* Here a number of schemes have been examined, such as those shown in figure 3. The standard Condon approximation assumes $\Omega_{12}(Q)$ independent of Q . One can easily envisage other cases. Thus, if the two states involve different sites, and if Q affects the potential barrier between these sites, then Ω_{12} will saturate once the barrier vanishes. In this case one might expect threshold behaviour. The ‘linear’ and ‘threshold’ cases may be acceptable approximations for different temperature regimes in a single system, depending on the amplitude of vibration in the promoting mode. Discussions are given by Kovarskii and Sinyavskii (1963), Sinyavskii and Kovarskii (1962, 1964, 1967), Flynn and Stoneham (1970), Ridley (1978a, b) and Goto *et al* (1980).

(c) *Specific examples of matrix elements.* We now give some examples of matrix elements. If the initial and final vibronic states are $|i\rangle$, $|f\rangle$ then the total matrix element is $\langle i|H-E|f\rangle$; the transition operator is that part of the Hamiltonian not diagonalised, less a term which corrects for non-orthogonalities of $|i\rangle$ and $|f\rangle$. The electronic matrix element, which is one factor only, is conveniently given as the total matrix element divided by the overlap of the vibrational components, i.e.

$$\langle i|H-E|f\rangle/\langle\chi_i|\chi_f\rangle = M_{if}.$$

Consider, for example, the case of Born–Oppenheimer adiabatic basis states. For these the total matrix element is

$$\sum_{\nu} \frac{1}{2M_{\nu}} \{ \langle\chi_i^*|\chi_f\rangle \langle\psi_i^*|P_{\nu}^2|\psi_f\rangle + 2\langle\chi_i^*|P_{\nu}|\chi_f\rangle \langle\psi_i^*|P_{\nu}|\psi_f\rangle \}$$

with momenta $P_{\nu} = -i\hbar\partial/\partial Q_{\nu}$ and masses M_{ν} . In this case the effective electronic matrix element is

$$\sum_{\nu} \left\{ \langle\psi_i|P_{\nu}^2/2M_{\nu}|\psi_f\rangle + \frac{\langle\chi_i^*|P_{\nu}|\chi_f\rangle}{M_{\nu}\langle\chi_i|\chi_f\rangle} \langle\psi_i^*|P_{\nu}|\psi_f\rangle \right\}.$$

Even without detailed evaluation, several features are clear. First, there are selection

rules. Some of these reflect point symmetry. Thus if Q_ν has odd parity, only ψ_i and ψ_f of opposite parity will have a finite second term, and only ψ_i and ψ_f of the same parity have a finite first term. Other selection rules derive from the harmonic oscillator selection rules: P_ν connects oscillator states differing by $n = \pm 2$. When the Condon approximation fails these selection rules are weakened with significant consequences (Ridley 1978a, b). Secondly, and indeed resulting from the oscillator matrix elements, there are differences in definitions which can confuse. Thus Brailsford and Chang (1970) and Freed and Jortner (1970) make rather similar approximations, but define empirical matrix elements C_{ν}^{BC} and C_{ν}^{FJ} , respectively, related thus:

$$|C_{\nu}^{BC}|^2 = \frac{\hbar\omega_\nu}{2} (n_\nu + \frac{1}{2} \pm \frac{1}{2}) |C_{\nu}^{FJ}|^2.$$

Thirdly, since the matrix element contains a sum of terms from the various promoting modes, there can be interference between their contributions (see, for example, Stoneham 1967, Sarai and Kakitani 1981).

Most authors start from expressions like the one just given for the effective electronic matrix element and evaluate the terms of their choice directly. In principle, this should present no special problems; in practice, calculations tend to be of low accuracy. This is why Huang and Rhys (1950) worked in terms of the electric field produced by longitudinal optic phonons for 2p-1s transitions at an F centre, and why Henry and Lang (1977) used a quasi-classical model. These two results lead to the following electronic matrix element:

Huang and Rhys:
$$|M_{HR}|^2 = \frac{8\pi}{3} \alpha \left(\frac{\hbar\omega}{p}\right)^2 (2n+1) i_p \bar{x}_{if}^2 / \Omega_c$$

where α is the polaron coupling constant, p is the net number of phonons emitted, r_p is the polaron radius $[(\hbar^2/2m^*)/\hbar\omega]^{1/2}$, \bar{x}_{if} is the matrix element $\langle \psi_{2p} | x | \psi_{1s} \rangle$ and Ω_c is the cell volume. The temperature-dependent factor comes from the promoting mode in this case:

Henry and Lang:
$$|M_{HL}|^2 = (\varepsilon^2 a^3 / \Omega_c) \equiv \left(\frac{\hbar^2}{2m^*}\right)^{3/2} \varepsilon^{1/2} / \Omega_c$$

where ε is the binding energy and $a \equiv (\hbar^2 \varepsilon / 2m^*)^{1/2}$ is the effective radius of the bound state. One notes that neither result shows a dramatic dependence of $|M|^2$ on binding energy, as sometimes suggested.

2.1.4. The lattice response component. The second important factor is a measure of how readily the lattice takes up energy when there is a perturbation. This factor appears in several contexts in solid-state physics, notably optical lineshapes, quantum diffusion and the present types of non-radiative transition. The response is most conveniently expressed in terms of a *lineshape function* $G(\omega)$ or its Fourier transform, the *characteristic function* $g(t)$. These appear in expressions for optical cross sections (see Stoneham 1975, §10.7) where, for example, $G(\omega)$ is the generalisation of the simple delta function $\delta(\omega)$ appropriate for an optical transition with photon energy $\hbar\omega$ when there is no coupling to the lattice:

$$G(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \exp(-i\omega t) g(t)$$

$$g(t) = \int_{-\infty}^{\infty} d\omega \exp(+i\omega t) G(\omega).$$

In non-radiative transitions $\hbar\omega$ is the energy which is *not* taken up by the accepting modes; it is zero for some important cases. When energy is partly dissipated via Auger or radiative mechanisms, or when the occupancy of the promoting modes changes, then ω is finite.

The lineshape function can be written in the form

$$G(\omega, T) = A v \sum_{\{n\}\{n'\}} \left| \int dQ \chi_{i(n)}(Q) \chi_{f(n')}(Q) \right|^2 \delta(\omega - \omega_{if(n)\{n'\}})$$

where the χ are the vibrational wavefunctions for the initial (i) and final (f) states with occupation numbers defined by $\{n\}$, $\{n'\}$. The standard simplifications then follow: the neglect of anharmonic terms, the assumption that the modes remain the same, i.e. that whilst they may change mean value and frequency, they are not mixed (invalid in Jahn-Teller systems and possibly many others), and the assumption that no frequencies change. In this limit the matrix element $G(\omega)$ factorises into a product of terms from each mode. The delta function, however, couples the modes through the $\omega_{if(n)\{n'\}}$ term, and it is here that the Fourier transform to $g(t)$ gives technical advantages. We shall not discuss this here (see, for example, Englman and Jortner 1970, Stoneham 1975, Robertson and Friedman 1976, Englman 1980), and shall simply quote some of the most important cases of $G(\omega, T)$ for $\omega=0$, thermal equilibrium in the initial state, and linear coupling.

We begin with results for decay from an excited state in the statistical limit $p \gg 1$. This is the commonest case encountered in non-radiative transitions in semiconductors. Results for the reverse, excitation, process can be obtained using detailed balance, with care.

(i) *Expressions valid at $T=0$.* In this limit the coupling strength is not important:

$$G(T=0) = \frac{1}{\sqrt{2\pi}} \frac{1}{\hbar\omega E_0} \exp \left\{ -S_0 - p \left[\ln \left(\frac{p}{S_0} \right) - 1 \right] \right\}.$$

Note that this can be rewritten so as to replace $p \ln(p/S_0)$ by $\gamma E_0/\hbar\omega$, where $\gamma \equiv \ln(p/S_0)$ varies only slowly with the various parameters. This is the source of the widely quoted result that the rate is proportional to $\exp(-\gamma E_0/\hbar\omega)$ at absolute zero temperature, a special case of the *energy gap rule*. The factor $\exp(-S_0)$ which appears can be regarded as a 'phonon dressing' factor for the transition matrix element of §2.1.3, i.e. $M \rightarrow M_{\text{eff}} = M \exp(-S_0/2)$; at high temperatures, however, the corresponding factor is absorbed into other parts of $G(\omega)$, so one must be careful to use M and M_{eff} consistently.

(ii) *Expressions valid for weak coupling.* When $S_0 \ll 1$ one can write

$$G(T) = G(T=0) [1 + n(T)]^p$$

where n is the occupation number. This will give a temperature dependence tending towards T^p at higher temperatures. A second result is a more general form of the *energy gap rule*. The argument can be seen rather generally. Suppose the rates of a p -phonon process and a $(p-1)$ phonon process are related by

$$W_p = \beta W_{p-1}$$

with β a small factor to be discussed. Suppose also that the lowest-order process always dominates, i.e. the smallest value of p and the largest $\hbar\omega$ dominate. Then the major process has $p = E_0/\hbar\omega_{\text{max}}$ and a rate $W_0 \beta^p$ (which cannot be taken seriously for small p) which can be rewritten as

$$W = W_0 \exp \left(\frac{E_0}{\hbar\omega_{\text{max}}} \ln \beta \right)$$

i.e. the rate decreases exponentially as the energy release E_0 (recall β is small, so $\ln \beta$ is negative).

(iii) Expressions valid for strong coupling, $S_0 \gg 1$.

$$G(T) = \frac{1}{\hbar\omega} \exp\left(\frac{E_0}{2kT} - S_0(2n+1)\right) I_{-p} [2S_0 \sqrt{n(n+1)}]$$

where $I_N(x)$ is a Bessel function of order N . For excitation to a higher state, the sign of the $(E_0/2kT)$ term in the exponential is changed. In suitable and quite general limits (see Englman and Jortner 1970, Ridley 1978a, b, Markvardt and Landsberg 1981) one finds the simpler and better-known form:

$$G(T) = \frac{1}{\sqrt{2\pi}} \frac{1}{E_M kT^*} \exp(-E_A/kT^*)$$

with $kT^* = \frac{1}{2}\hbar\omega \coth(\hbar\omega/2kT)$, so that kT^* and kT are identical at high temperatures. This last expression shows thermal activation. Ignoring the slow temperature dependence of the pre-factor, the activation energy is $E_A \equiv (E_0 - E_M)^2/4E_M$, i.e. just the energy of the crossover above the initial state (figure 1). As the coupling and the temperature increase, the low-temperature 'tunnelling' component (which involves the 'dressed' matrix element $M \exp(-S_0/2)$) gives way to the activated 'crossover' component. Note too that E_A contains terms linear and quadratic in E_0 : there is no longer a simple energy gap rule.

(iv) Various generalisations. In some of the generalisations it is hard to separate the matrix element part and the lattice response part. Readers who become bewildered when making direct comparisons should note first that most authors give the temperature dependences of the promoting and accepting mode contributions combined together (e.g. Huang and Rhys 1950, Englman and Jortner 1970). Secondly, violations of the Condon approximation lead to further changes (e.g. Ridley (1978a, b), whose results include as special cases the earlier ones of Kovarskii and Sinyavskii (1963)). Some of these changes come from altered energy denominators, others from altered selection rules. Thirdly, there are always the multifarious factors depending weakly on temperature which appear in capture cross sections and not in all expressions for non-radiative transitions; examples include the thermal velocity and Sommerfeld factor dependences (e.g. Bonch-Bruevich and Landsberg 1968, Pässler 1978).

The more direct generalisations include the following.

(a) Initial state not in thermal equilibrium. Brailsford and Chang (1970) developed the theory for a single initial vibronic state, and their work can be used for analysis of any defined initial distribution. We shall discuss what such a distribution might be in the next subsection.

(b) Changes in frequency in the transition. Here Lin and Eyring (1972) collect together results. Even here there is no mode mixing.

(c) Anharmonic terms, again with no mode mixing. Sturge (1973) and Mikami *et al* (1975) both discuss cases where anharmonic terms produce very profound effects. They use two entirely different methods: Sturge exploits Morse potentials instead of parabolic energy surfaces, whereas Mikami *et al* add cubic terms which 'tilt' the surfaces.

(d) Several of the standard formulae assume weak coupling to many modes (a distinction which is not relevant unless there is phonon dispersion). Sarai and Kakitani (1981) discuss the large displacement limit. They also note the existence of interference effects when there are several distinct promoting modes.

(e) Generalisations to discuss the consequences of phonon dispersion. These have been analysed in some detail by Weissman and Jortner (1978) and to a lesser extent in

papers such as those by Englman and Jortner (1970) and Pässler (1974a, b). For our present purposes one should note first that the effective frequency is temperature-dependent, and secondly that the effective frequency depends on the property of interest: it will be different for calculating the ground-state relaxation energy, the optical absorption and the non-radiative rate. The effective frequency is usually discussed in terms of moments, weighted by the mode Huang–Rhys factor, e.g. $\sum_i \omega_i^N S_{10}(An_1 + B) / \sum_i S_{10}(An_1 + B)$ where A and B are small integers (see, for example, Evangelou 1981). So far as experiment is concerned, it is often found that ω_{eff} is close to either the highest host frequency or to some local mode or local resonance frequency. However, one knows too that, in optical spectra, the dominant phonons may also be of lower energy, from near maxima in the density of states. We shall see later that other cases do occur, and it is *not* true that the same frequency will be found for every centre in a given host.

We now turn to the limit in which very little energy is transferred to the accepting modes, i.e. $p \simeq 0$. This covers diffusion, spin–lattice relaxation and various defect reorientation and tunnelling problems. A fuller discussion is given by Stoneham (1975, 1979). The important processes are these.

(v) One-phonon processes, in which a phonon of energy $\hbar\omega = E_0$ is absorbed or emitted. The rate is proportional to $(n+1)$ in emission, and so linear in T in the usual condition $kT \gg E_0$.

(vi) Two-phonon processes, in which the energy differences of two phonons $\hbar\omega_1 - \hbar\omega_2 = E_0$ combine to provide the transition energy. The phonons with energy $\sim kT$ dominate usually, giving a rate proportional to T^N , where N is typically 5–9. Note the different origin of the high power from case (ii). In some cases there may be a resonant interaction with an excited state, provided the excitation energy Δ is less than $\hbar\omega_{\text{max}}$. In such systems, the rate varies as $\exp(-\Delta/kT)$.

In both case (v) and the resonant instance of case (vi) the transitions involve only a narrow energy band of phonons. Non-equilibrium effects are well-documented, and include the *phonon bottleneck* and *phonon avalanche* situations. Two-phonon processes can also exhibit interference effects between the resonant and non-resonant contributions.

2.1.5. The Landau–Zener method. Many quasi-classical discussions of non-radiative transitions combine the idea of a classical motion over a potential energy surface with a branching each time surfaces meet, the branching ratio being given by the so-called Landau–Zener probability (Landau 1932a, b, Stueckelberg 1932, Zener 1932). Whilst this approach has some merits, it is neither generally applicable nor specifically suited to many of the situations in which it is used.

The original equations gave results for two (isolated) intersecting levels E_1 and E_2 with a matrix element ε_{12} independent of separation. If E_1 and E_2 vary linearly with spacing R , the branching probabilities are given in terms of P , where

$$P = \exp\left(-\frac{2\pi}{v} \frac{|\varepsilon_{12}|}{|\partial(E_1 - E_2)/\partial R|}\right)$$

with v the classical approach velocity. For very slow velocities, there is a high probability of following the adiabatic surface; for high velocities, the motion tends to be ballistic. Generalisations by Coulson and Zalewski (1962), Bykhovskii *et al* (1965) and Delos and Thorsen (1972) cover various points, including continua of states.

Even for molecules, there are open questions and problems (Bates 1960, Coulson and Zalewski 1962), some relating to the choice of basis. Some of these features of concern also extend to solids: it is not clear whether ε_{12} is independent of R , nor that E_1

and E_2 are adequately linear, nor even that the exponent in P is satisfactorily defined. Still less is it clear that the classical velocity v is sensibly constant at important parts of the trajectory such as classical turning points. Probably the main weakness in solid-state applications is the neglect of other degrees of freedom (specifically vibrational) and the implied neglect of dissipation other than through some primitive phenomenological model. One of the most useful applications is that of Henry and Lang (1977), who made the specific generalisation discussed in §2.2. The accuracy of the Landau-Zener approach in cooling transitions is less well established, though the results obtained (see §3) agree qualitatively with those obtained otherwise.

2.2. Free-bound transitions

The discussions in the earlier subsection considered transitions between only two discrete levels. There are at least three circumstances in which this does not suffice. First, even if there is just a single bound state, the conduction band from which an electron is captured comprises a continuum of states (Henry and Lang 1977, figure 6(a)). Secondly, for defects like shallow donors, there may exist a continuous range of states between the

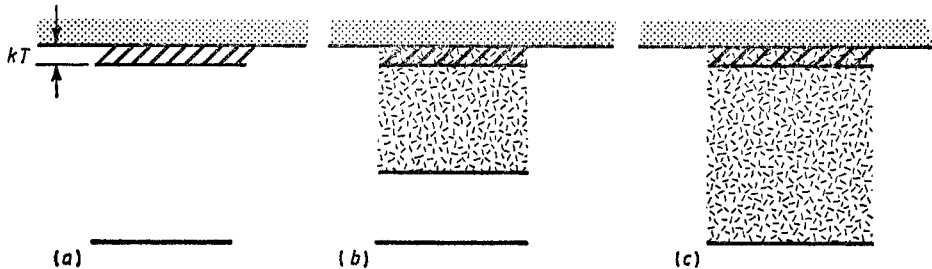


Figure 6. Cascade processes. The upper and lower shaded areas correspond, respectively, to the continuum of unbound states and the region of discrete but dense states among which cascades are possible. (a) Direct multiphonon, e.g. Henry and Lang (1977); (b) intermediate cascade, e.g. Rees *et al* (1980); (c) cascade, e.g. Lax (1960).

ground state and conduction band such that no splitting exceeds the maximum phonon energy. In such cases (Lax (1960) and many subsequent workers; figure 6(c)) cascade theory is appropriate. Thirdly, there may be a spectrum of closely spaced levels, yet with a gap between the lowest of these and the ground state. Here (Rees *et al* 1980, figure 6(b)) one expects a final multiphonon transition after what is akin to a cascade process. In all these three cases, which we now consider, one estimates a *capture cross section* rather than a state-to-state transition probability. Typical values are given in table 1.

2.2.1. Capture into an isolated level from a continuum. One important aspect is that the excited vibrational states with the electron in the ground (trapped) state $|g\{n\}\rangle$ are degenerate with conduction band states $|c\{n'\}\rangle$ with low vibrational excitation. In semi-classical descriptions the vibration excursions which bring the trapped state close to the free states are critical. Precise crossing is not essential (cf figure 6) but obviously there are differences in detail. One should not forget that transitions among the unbound states occur. T N Morgan (1979 private communication) has noted a useful description in terms of separate one-phonon continua, two-phonon continua, etc.

There is thus the possibility of transitions both into and out of the captured state

Table 1. Capture mechanisms.

Mechanism	Radiative	Auger	Phonon
Typical cross section	10^{-5} – 10^{-4} Å ²	Wide variation: 10^{-1} Å ² typical	10^{-2} – 10^4 Å ² , e.g. for electron capture: trap: 10^1 – 10^4 Å ² (Ge:Sb ⁺). Neutral centre: 10^{-1} – 10^1 Å ² (Ge:Ni ⁰). Repulsive: 10^{-8} – 10^{-5} Å ² (Ge:Mn ⁻)
Conditions for dominance	Depends on states and energies	High carrier and defect concentration ($N = 10^{17}$ cm ⁻³ typical for Ge)	Low defect concentration; strong coupling to lattice
Temperature dependence	Small dependence	Small dependence	Power law or exponential
Concentration dependence	None	Linear with concentration until screening is important†	None

† Since the role of screening is sometimes controversial, we note the careful discussion in appendix C of Burt (1981a) and the distinction between dynamic screening (usually small) and static screening (often large).

during the capture process. The actual rate at which the carriers reach the ground vibronic state is a balance between the rate of these capture and emission transitions and the cooling transitions discussed later. Normally the cooling transitions are assumed fast.

Sumi (1980a, b) and Henry and Lang (1977) both illustrate their analyses using a quasi-classical description and the Landau-Zener formula to describe the possible outcomes of the two types of transition. The specific aspect which is important in both the Henry-Lang and Sumi results is the Landau-Zener prediction of a branching ratio at an intersection and its dependence on the thermal velocity in the accepting mode and on the relative derivatives of the energies of the two states as functions of the accepting mode. The results are given for slightly different systems (figure 8). Sumi's results are expressed in the form

$$\text{emission rate} = (\text{emission rate, uncorrected}) \eta(\gamma)$$

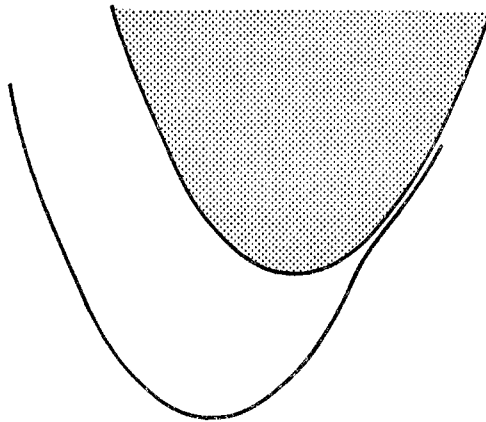


Figure 7. Avoided crossing with a continuum of states. This corresponds roughly to the discrete cases of figure 2.

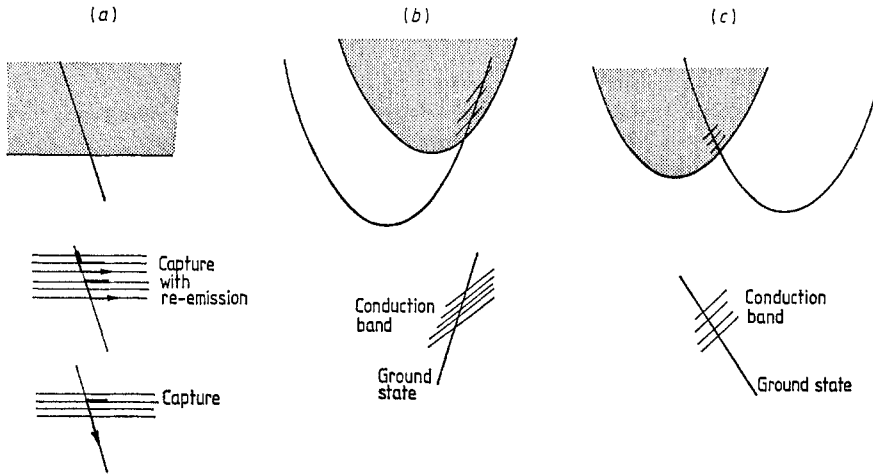


Figure 8. The three cases (Henry and Lang (a); Sumi I (b) and II (c)) in which are treated capture and re-emission involving continua. Both the standard configuration-coordinate and simplified forms are given.

where the emission rate can be related to the capture rate by detailed balance. Here η is a monotonic increasing function (figure 9) of the parameter γ given by

$$\gamma = \frac{4}{3} \left(\frac{BS_0^2}{\Delta^3} \right)^{1/2} \frac{kT}{\hbar\omega}$$

where $2B$ is the full bandwidth of the (tight-binding) band, and Δ is the Franck-Condon energy for emission. Note that the effects of re-emission become negligible as γ becomes large ($\eta \rightarrow 1$), e.g. at high temperatures, strong coupling or wide bandwidth. This is the opposite conclusion from Henry and Lang (1977) whose corresponding factor is their \bar{P}_{tt} (their equations 106 and 110) which has a slow contrary dependence. I have not resolved this difference.

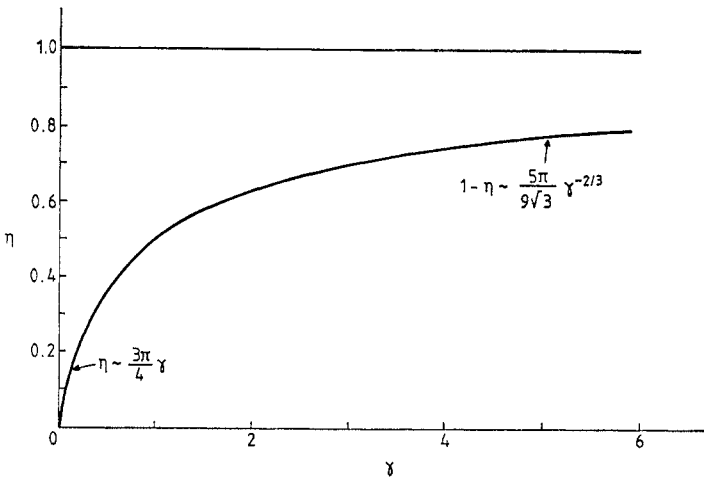


Figure 9. Sumi's results showing the variation of η (the pre-exponential factor in units (mean frequency/ 2π) with $\gamma \equiv (4/3) (S_0^2/\Delta^3)^{1/2} kT/\hbar\omega$.

2.2.2. *Capture at a deep level via an excited state.* In a formal sense, this is one of the simpler cases, for the final multiphonon step may indeed be between non-degenerate levels, and the cascade sequence may be less critical than for shallow defects. Discussions for Si:Se and Si:S are given by Rees *et al* (1980) and Gibb *et al* (1977), who note that capture will usually be described by two time constants. In particular, capacitance transients monitor the change in charge state, not necessarily the occupancy of the real ground state. Discussions of rate equations generally and of other relevant points are given by Landsberg (1956), Stoneham (1975, §14.5) and Pickin (1978); further features are covered in §2.2.3.

2.2.3. *Cascade processes.* The cascade hypothesis asserts that carriers are captured into some highly excited state, and that they then continue to lose energy primarily by a cascade of one-phonon transitions between the bound states. Since the energy levels are separated by less than the phonon energy, we are normally considering shallow centres. Since shallow centres normally exhibit weak coupling, all the individual transitions can be estimated effectively by perturbation theory.

Calculations of cascade capture rates involve two main components. The first is the initial capture cross section $\sigma_{\alpha}(E)$ for an electron of energy E into one of several possible excited states $|\alpha\rangle$. The second is the 'sticking probability' P_{α} , for which the usual definition is the probability that a carrier in state $|\alpha\rangle$ will reach the ground state before escaping from the trap. There is a substantial literature on both components, the most convenient summaries being Bonch-Bruevich and Landsberg (1968), Stoneham (1975, §14.4) and Abakumov *et al* (1978). The last of these discusses developments which resolve some of the outstanding difficulties in the absolute values of the rates. Given these several surveys, I shall merely discuss some specific issues here.

The first point concerns the effects of net charge on the initial capture step. Clearly one expects differences between the processes $X^{+} + e \rightarrow X^{0} + e \rightarrow X^{-}$ and $X^{-} + e \rightarrow X^{2-}$ because of the difference in Coulomb interactions in the initial state. This can have striking consequences. The Coulombic contribution gives the so-called Sommerfeld factor, whose effect is to give a net increase or decrease in electron density near the centre relative to the free-electron case. For electrons of momentum p in the limit that kT is much less than the effective Rydberg (defined as $Z^2 m^* e^4 / \epsilon^2 \hbar^{-2}$, irrespective of the sign of the charge Z). Bonch-Bruevich and Landsberg give these expressions for the Sommerfeld factors:

attractive	\tilde{p}
neutral	1
repulsive	$\tilde{p} \exp(-\tilde{p})$

where \tilde{p} is p/p_Z with $p_Z = 2\pi\hbar^2 (\epsilon\eta/|Z| m^* e^2)^{-1}$.

Pässler's (1977) analysis leads to related factors in the multiphonon cross section expressed in terms of S , p and the binding energy ϵ . If we write

$$P = \frac{\hbar\omega}{kT} \left[\frac{1}{2} + \frac{kT}{\epsilon} \sinh^{-1} \left(\frac{p}{S} \sinh \frac{\epsilon}{2kT} \right) \right]$$

then the factors which appear in the rate are these:

attractive	$4P^{1/2}/\sqrt{\pi}$
neutral	1
repulsive	$(8/\sqrt{3})P^{2/3} \exp(-3P^{1/3})$.

For repulsive centres one confirms that higher energies and temperatures favour capture. It is known for optical transitions (see, for example, Stoneham 1975, p337 *et seq*; also Ridley 1980) that there can be significant corrections even for neutral centres because of the defect effects on the free-electron wavefunction. Sadly, we must still agree with the comment of Bonch-Bruевич and Landsberg: 'this factor may usually be neglected, at least at the present state of the subject'.

The sticking probability has been the subject of changes in view. These changes are of two types. First, the formal definition does not correspond to experiment. In experiment, capture can occur and contribute effectively to many observable properties without the ground state being involved. This feature is analysed by Belezny and Andor (1978). Secondly, Abakumov and Yassievich (1976) have corrected some of the earlier derivations of the sticking probability, and this leads to a significant increase in P_α . The source of the error is that one cannot always replace averages over energy $F(E)$ by the same function of the mean energy $F(\langle E \rangle)$. In particular, acoustic phonon emission continues to play a significant role when the carrier kinetic energy falls below mv_s^2 , with v_s the velocity of sound.

2.3. Effects of perturbations

In most practical devices non-radiative transitions occur at defects which are neither isolated nor in regions free from applied fields. The applied fields are either electric fields, stress fields or magnetic fields.

The effects of nearby defects are of three main types. One is that selection rules are relaxed (see, for example, Venzl and Fischer 1981). A second is that there may be impurity banding effects, so that the nature of the excited states is altered. This type of effect occurs at rather low concentrations: the effective Bohr radii of donor bound states with principal quantum number $n \sim 7$ in Si cannot be ignored for concentrations above 10^{14} cm^{-3} . The third effect is that, for shallow isoelectronic centres, even the modest fields of nearby impurities may prevent capture, and hence inhibit radiative recombination. This has been discussed in the context of dislocation in GaP by Tasker and Stoneham (1977). Magnetic fields have rather special effects, notably through the spin dependence of capture cross sections. Stress fields (apart from implied effects from the proximity of dislocations) appear relatively unimportant, except perhaps in guiding the forced diffusion of defects. Electric fields are of major importance, of course, since they underlie device operation, and there have been several studies of effects on capture and emission rates. The effects on shallow defects are reviewed by Bonch-Bruевич and Landsberg (1968), who comment on two specific aspects. One is that the field affects the velocity distribution of free carriers, increasing their mean velocity. This lowers the capture rate at attractive centres (whether Coulombic or short-range binding), where slowly moving carriers are most readily captured, but enhances the capture by repulsive centres where a barrier is to be overcome. The second effect is that reverse processes (emission) are enhanced. One contribution is the Poole-Frenkel effect, where the ionisation energy along the field direction is reduced; other contributions can appear from the very different phenomena associated with impact ionisation and autoionisation. Electric-field effects are of special importance when the energy changes induced (of order $E|e| \times \text{orbital diameter}$) exceed the energy level separations since the orbital diameter increases as the binding decreases. Shallow defects are especially susceptible.

Electric-field effects for deep levels, in which multiphonon mechanisms are critical, have been the subject of a series of studies by Korol (1977) and Makram-Ebeid and his

colleagues. Here both quasi-classical methods (Pons and Makram-Ebeid 1979) and the Bardeen tunnelling transition formalism (Makram-Ebeid and Lannoo 1981) have been used to find the effects of electric fields on emission rates, i.e. a multiphonon equivalent of the Poole-Frenkel phenomenon. Experimentally (see also Makram-Ebeid 1980a, b, Makram-Ebeid *et al* 1981), the results have consequences for both the conventional (DLTS) and optically-refilled (ODLTS) forms of deep-level transient spectroscopy. The key result is this: the emission rate can be expressed in the form:

$$e = e_{n_0} + \sum_{p=-\infty} W_p(S, p_1, \hbar\omega) \Gamma(\Delta_p | E|)$$

in which e_{n_0} is the value in the absence of the field:

$$W_p = \exp \left[\frac{p\hbar\omega}{2kT} - S \coth \left(\frac{\hbar\omega}{2kT} \right) \right] I_p \left(S \operatorname{cosech} \frac{\hbar\omega}{2kT} \right)$$

corresponding to the lattice factor in Huang-Rhys theory, and $\Gamma(\Delta_p)$ is the elastic tunnel transition rate from a localised state at depth Δ_p from the conduction band at the site of the trap; $\Delta_p \equiv E_c - E_T + p\hbar\omega$. In Korol's (1977) calculation, and in later calculations by Pons (1979), the dependence of Γ on a field of magnitude $|E|$ takes the form

$$\Gamma(\Delta_p | E|) = a \left(\frac{\Delta_p}{h} \right) \frac{\exp(-K)}{K}$$

$$K = b \frac{2m^*}{\hbar |E|} (\Delta_p)^{3/2}$$

with a, b being constants of the order of unity, taking values $a = \frac{9}{8}$ and $b = \frac{4}{3}$ in the Pons and Makram-Ebeid analysis. The net effect is that the emission rate increases rapidly with field and with temperature. One can devise parallels with the usual Poole-Frenkel result, but the strong phonon coupling leads to significant differences. We shall discuss some of the results of experiments in electric fields later.

3. Cooling transitions

It is a common assumption that a non-radiative transition takes place from a distribution over vibrational states in thermal equilibrium. In this section I consider transitions far from thermal equilibrium. These arise commonly in two cases: when a carrier is captured into a highly excited vibronic state, and when a carrier is optically excited into a highly excited vibronic state in a Franck-Condon transition. The main question discussed is the route by which a lower electronic state is reached. Just as in electronic transitions, we may define strong and weak damping. In weak damping, the case we shall usually consider, each vibronic state survives for several lattice vibrations. With strong damping (e.g. Seitz 1940) the oscillations may fail even to reach the classical turning point of the starting state before settling into the lowest state (e.g. moving directly from B to C in figure 10 without reaching B'). I shall not describe hot luminescence, i.e. radiative transitions directly from higher vibronic levels. Clearly, this only occurs with relatively weak damping; a brief survey is given in Stoneham (1975, §12.2.6).

3.1. Transitions within a single level

Suppose the defect is prepared in an excited vibronic state (whether a specific state or a mixture) of an isolated electronic level interacting only with an accepting mode Q . If

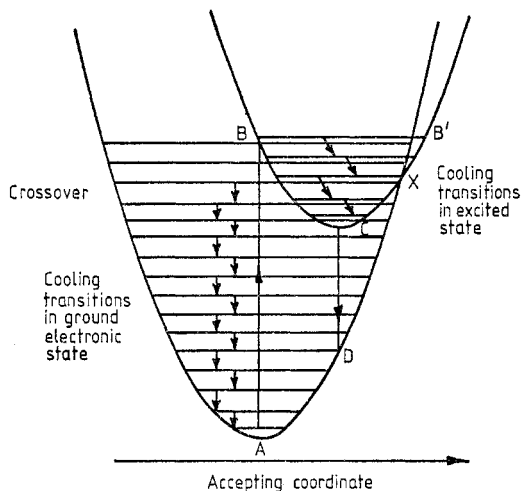


Figure 10. Optical absorption ($A \rightarrow B$), luminescence ($C \rightarrow D$) and various cooling transitions between vibronic levels. The crossover is labelled X. Note the accepting coordinate is not usually a normal mode.

this system is left to evolve at very low temperatures, there is a widespread view that it will end in the lowest vibronic state. As defined, however, no such thing will occur: there may be some complicated vibrational motion but, without a means of dissipating energy, there will be no evolution towards the lowest state. How, then, do cooling transitions occur? Obviously, anharmonic interactions with other modes q_i , through terms proportional to Qq_iq_j or Q^2q_i , will always occur. Another important mechanism arises because (as remarked in §1.1) the accepting mode is usually a reaction coordinate, not a normal mode. This distinction is important, and it is useful to elaborate. In a harmonic lattice the potential energy is a quadratic function $\frac{1}{2} \sum_{r,s} x_r V_{rs} x_s$ of the positions x . This form contains cross-terms $r \neq s$. The normal mode transformation takes linear combinations $q_\alpha = \sum_r a_{\alpha r} x_r$ so that the potential energy separates, giving $\frac{1}{2} \sum K_\alpha q_\alpha^2$. The importance of the transformation is that the harmonic system can now be treated as a collection of dynamically independent oscillators. Suppose a defect is introduced which does not change the modes (e.g. linear coupling only). The accepting 'mode' is chosen to maximise the relaxation energy, a point implicit in most discussions. But this choice will give, in general, $Q = \sum_\alpha A_\alpha q_\alpha$, a sum of contributions from several modes. If we write Q in the form $A_0 q_0 + \sum' A_\beta q_\beta$, where $A_0 q_0$ is the dominant component (there is some limited freedom here, since degenerate modes can be mixed freely), then the potential energy can be rewritten in the form $\frac{1}{2} K_0 Q_0^2 + \sum' B_\beta Q_0 q_\beta + \dots$. It is the terms like the second which allow dephasing and cooling transitions, i.e. loss of energy from motion in the accepting 'mode'.

We note one simple and acceptably general result. The cooling transition from the n th to the $(n-1)$ th vibrational state will be induced by either an anharmonic term or a term coming from the extent to which Q is not a normal mode. In both cases the operator involved has the form $f(q_i, \dots) Q$, linear in Q . The transition probability will be proportional to $f^2 |\langle n | Q | n-1 \rangle|^2$, i.e. to $f^2 n$. The cooling rate is roughly proportional to the degree of excitation (n) in the accepting 'mode'.

A second obvious point is this. When one is really dealing with an accepting mode which is a normal mode (e.g. a local mode associated with a light atom, like hydrogen),

the anharmonic terms should dominate and cooling could be slow. If, however, the accepting mode is built from a strong admixture of many modes (e.g. when it has a frequency near the peak of the density of states of phonons in the perfect solid), cooling will be very rapid, and sustained systematic motion in the accepting 'mode' will not occur.

3.2. Transitions between two states

We now consider the intersecting energy surfaces and assume the defect has been excited to a state above this crossover (whether the crossing is avoided or not is immaterial here). As cooling occurs, will the system end on the upper or lower energy surface?

An outline of the solution can be seen without difficulty (A M Stoneham 1977 unpublished work, Stoneham and Bartram 1978, Bartram and Stoneham, to be published). The first important idea is this: within a narrow energy range about the crossover, the vibronic states from both surfaces will be heavily mixed. There will be some states with roughly equal contributions from upper and lower surfaces, and the others will be equally balanced between those principally derived from one or other of the surfaces. The second important idea is that it is the transition out of this crossover band of states which decides whether the system ends on the upper or lower surface: the branching ratio is determined by what happens near the crossover. If we combine the ideas with the result just derived (in which the cooling rate is proportional to the phonon occupancy) we conclude:

$$\frac{\text{probability of ending on lower surface}}{\text{probability of ending on upper surface}} \simeq D(n_{\text{LX}}/n_{\text{UX}})$$

where D is a degeneracy factor and n_{LX} , n_{UX} are the degrees of vibrational excitation at the crossover for the lower and upper surfaces, respectively. If the effective frequencies are the same, the branching ratio is simply in the ratio of the energies E_{XA} , E_{XC} of figure 10.

This result proves to be important. It does, however, avoid several issues, and some of these have been discussed elsewhere (Dexter and Fowler 1967, Bartram and Stoneham 1975, 1979, Stoneham 1977, Stoneham and Bartram 1978, Nasu and Kayanuma 1978, Kayanuma and Nasu 1978, Kayanuma 1979, Kusunoki 1979, Leung and Song 1980). First, it assumes cooling is slow enough that the system can indeed exhibit several vibrations in each vibronic state. Seitz (1940), for example, considers a case of strong damping in which, essentially, the system 'slides' down the upper curve without necessarily approximating the classical crossover. Secondly, the picture ignores the behaviour before the crossover band is reached; in fact, there are transitions from one surface to the other in the early stages (see, for example, Leung and Song 1980), and there are also constraints depending on the way the system is prepared. Thirdly, the precise mechanism of mixing is important. Stoneham and Bartram (1978) treat a separate promoting mode explicitly (as in figure 2(a)) whereas Kayanuma and Nasu (1978), for example, replace the promoting mode by a constant transfer matrix element. Finally, we should return to the ever-popular Landau-Zener picture (§2.1.5). This considers two energy surfaces with the system approaching the intersection with a thermal velocity. The relative probabilities of remaining on the same surface or of transferring to the other are then given in terms of the velocity and the energy surface geometries. To a certain level this prescription parallels the ideas given: the critical stage occurs when the system has cooled so the classical turning point and the crossover coincide. However, there are unsatisfactory features for solids (which are not important in the traditional molecular

applications): one is not dealing with well-defined isolated surfaces, nor does one represent the cooling properly, though it is possible (e.g. Kusunoki 1979) to include a simplified phenomenological damping.

3.3. The Dexter–Klick–Russell criterion: luminescence or not?

One important result follows directly from the assumption underlying the results in the last subsection, that all the important transitions from one surface to another take place via the band of states at the crossover. This means that if, at low temperatures, optical excitation is to states *below* the crossover, the system will merely cool to the lowest vibrational state of the upper surface, and then luminesce in appropriate cases. If excitation is to states *above* the crossover, then the branching ratio will ensure the system usually cools directly, non-radiatively, to the lowest state of the lower surface. Thus, whether luminescence occurs or not depends on the two configuration-coordinate curves and the precise state into which excitation occurs. Like all simple descriptions, it has obvious limits even for isolated defects. Weak damping has already been mentioned. Other channels can short-circuit the routes discussed. Cases with $S \ll 1$ and p small, as in spin–lattice relaxation or cascade processes, will have direct few-phonon non-radiative processes and only weak radiative transitions.

Dexter *et al* (1955) put this idea in a useful form by assuming the initial optical absorption to involve a Franck–Condon transition (accepting mode coordinate fixed) with energy E_{abs} . For luminescence to occur in a Franck–Condon transition at energy E_{em} , assuming equal vibrational frequencies in ground and excited states, this condition must hold:

$$\Lambda' \equiv (E_{\text{abs}} - E_{\text{em}})/2E_{\text{abs}} \leq \frac{1}{4}.$$

Twenty years later, when far more data were available for a comprehensive check, Bartram and Stoneham (1975) found a rather better criterion to be

$$\Lambda \equiv (\text{relaxation energy } S\hbar\omega)/E_{\text{abs}} \leq \frac{1}{4}.$$

The two conditions would be identical if the energy surfaces were exactly those assumed. However, the Bartram–Stoneham form has two advantages. First, it only requires data from optical absorption to predict whether luminescence occurs, since S and $\hbar\omega$ can be estimated from the temperature dependence of the linewidth. Secondly, it emphasises correctly that only the energy surfaces near and above the crossover matter. Any special effects in the relaxed excited state (e.g. Jahn–Teller distortions) are irrelevant, though they will affect E_{em} and hence Λ' .

The condition $\Lambda \leq \frac{1}{4}$ divides defects into two classes very well, notably the substantial list of F-centre data cited by Bartram and Stoneham. Even for the F centres, there are special cases (LiCl is one (Takiyama 1978)) which appear to be exceptions, though the position is not completely clear. More interesting is the range $\frac{1}{4} \leq \Lambda \leq \frac{1}{2}$, since this provides a test of the rule given in the last subsection. Including the correct degeneracy factor (Bartram and Stoneham 1979) the luminescence efficiency should not exceed $\frac{1}{2}(1 - 2\Lambda)^2$, and this is indeed consistent with observation in magnitude with values reported for F centres in NaI and NaBr. Figure 11 illustrates the variation of the luminescence efficiency with Λ .

3.4. Evolution of the excited state

The dynamics of evolution of the excited state is an area in which there are very few

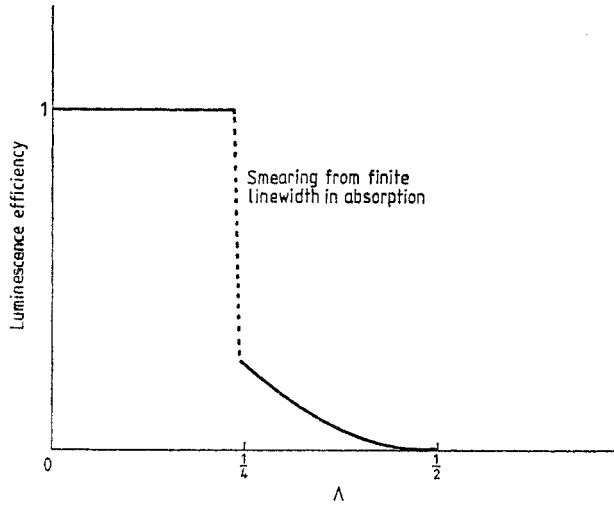


Figure 11. Efficiency of luminescence as a function of $\Lambda \equiv S_0 (S_0 + p)^{-1}$ in a simple two-level system.

calculations. One reason is that the system may be far from equilibrium, and many standard results do not apply directly (see Stoneham 1980). A second reason is that the way in which the system is excited is important: it may well make a difference if a short pulse of intense optical excitation is used instead of prolonged low-intensity monochromatic light (see, for example, Stoneham and Bartram 1978). The issue is still less obvious when excitation is the result of carrier recombination. A third reason is that our knowledge of the energy surfaces in these regions may be minimal. Indeed, for many of the important systems showing recombination-enhanced motion, the nature of the defect may not be known, let alone the energy surface. One can hope for a full theoretical analysis only in special cases where the atomic structure is known in detail. A promising system is GaP:(Zn, O) (Feenstra and McGill 1981) where recombination-induced reactions have been studied.

Given these circumstances, one is forced to resort to simple analytic calculations (e.g. those in this subsection, or the analogies with chemical reaction theory exploited by Weeks *et al* (1975)), or to use numerical methods. Englman and his colleagues have used this second approach (Barnett and Englman 1970, Englman and Barnett 1970, Englman and Ranfagni 1980, Englman 1981) and have been able to model many of the phenomena of interest: dephasing, competition between 'horizontal' (tunnelling) and 'vertical' (cooling) transitions. Engleman's recent exploitation of a wavepacket description within the WKB method in the context of complex energy surfaces may well have application in semiconductor systems, though even here a phenomenological treatment of dissipation is involved.

4. Special cases of non-radiative transitions

This section discusses some important classes of non-radiative transitions which embody some of the principles discussed in the last two sections. Non-radiative transitions have many effects beyond carrier capture and competition with luminescence, and the dramatic phenomena of recombination-enhanced processes and persistent photoconductivity

show this clearly. A further example, following the same principles, is that of energy transfer (see, for example, Knox 1968, Dexter *et al* 1969, Soules and Duke 1971). Again one finds a lattice factor precisely parallel to the function $G(E)$ of §2. Whilst there are other systems—especially organic—in which energy transfer is more important, one should not ignore the possibility of transfer from one species to another in semiconductors.

4.1. Recombination-enhanced processes

For the purposes of this subsection, I shall regard any process enhanced by raising carrier densities as recombination-enhanced, whether the carrier densities are raised optically, by ionising radiation, by injection, or by some indirect mechanisms. Such processes have two main characteristics.

(i) The processes are athermal, or show a relatively weak temperature dependence. As a result, accelerated aging experiments at high temperatures may give unduly optimistic ideas of lifetime. Thus Bellamy and Kimerling (1978) noted the degradation of GaAs/Pt Schottky barrier structures for use in IMPATT devices was controlled in accelerated aging tests by an interface reaction with a 1.6 eV activation energy; in the lower temperature avalanche conditions relevant in operation, the rate-controlling step was recombination-enhanced diffusion with an activation energy 0.3 eV.

(ii) The processes occur during device operation, and hardly ever 'on the shelf'. The excitation rate is important, not merely the total dose. Characteristic degradation effects of recombination enhancement are seen in injection-mode devices, where minority carrier injection under forward bias leads to dramatic effects not seen under reverse bias, nor on heating.

Kimerling (1978) has given an excellent review of these processes. Specific aspects are surveyed by Dean and Choyke (1977), Stoneham *et al* (1978) and Stoneham (1979, 1980).

The consequences of recombination enhancement are very varied, and not confined to semiconductors. The underlying processes are usually defect production (possibly transient), dislocation climb, enhanced diffusion rates, and phenomena which involve these processes. Defect production is documented especially clearly in alkali halides (Itoh 1976, Stoneham 1979), though there is some evidence too for SiO₂. The importance of recombination-enhanced dislocation climb was recognised rapidly through the dark-line defect mechanism: dislocation networks acting as non-radiative recombination centres were seen to grow on the active regions of GaAs heterostructure lasers and GaP light-emitting diodes. One of the earliest noted and most puzzling cases of enhanced diffusion is the elusive interstitial in Si and Ge, where the motion is enhanced by ionisation created by the same bombardment as produces the defect.

It is possible to group all these various phenomena into three main categories.

(i) Local heating models. Here the recombination energy is converted into vibrational energy in the mode corresponding to the reaction coordinate for the motion of interest. A combination of this recombination energy and of thermal energy is used in passing over the potential barrier to motion. Since any losses are positive, the enhanced and unenhanced activation energies can differ at most by the recombination energy.

(ii) Local excitation models. Here the recombination energy is transmitted to the electronic degrees of freedom. Thus the defect may be excited into a state with a lower activation energy for motion. Here the enhanced activation energy may be reduced from the unenhanced one by more than the recombination energy, though in such cases there must be a selection rule to ensure that diffusion is via the excited electronic state only when enhancement occurs.

(iii) Bourgoin–Corbett model (1972). The moving defect occupies different sites in different charge states. Successive alternate capture of electrons and holes take the defect from one charge state to another, and hence from one site to another. Since this process is limited by capture, and capture rates tend to fall with rising temperature, one may even find a negative activation energy here.

This nomenclature is not universally used; however, descriptions like ‘hot spot’ or ‘phonon kick’ imply specific features of the precise reaction which are often not known or

Table 2. Recombination-enhanced processes. In this table I have retained the assignments of the original authors and included only those cases where a careful analysis has been attempted. E_T =activation energy of thermal anneal, E_{RE} =activation energy of injection anneal and E_R =energy release by capture of a free carrier.

System	Mechanism proposed	Comments
e ⁻ damage centres in GaAs 0.18 eV 0.41 eV	Local heating	Not charge-state effect $E_T=1.75$ eV, $E_R=1.2$ eV, $E_{RE}=0.98$ eV $E_T=1.4$, $E_R=1.09$, $E_{RE}=0.34$
e ⁻ damage centres in GaP 0.14 eV 0.23 eV 0.32 eV 0.48 eV 0.62 eV 0.74 eV	Local heating	Not charge-state effect 1.7 2.15 1.3 2.1 1.3 2.0 1.7 1.8 1.3 0.62* 0.62 2.1 1.5 0.83
GaP: implanted H isotopes	Not identified	Strong non-monotonic isotope dependence of anneal of (intrinsic) damage centres
GaP: (Zn, O)		Effective dissociation of well-defined centre
Si:Al	Local heating	Not charge-state effect
Si:Bi	Bourgoin–Corbett, but involving +, 0 and – states	Different charge states monitored by EPR
SiC:H Si	Local heating	Light impurity atom ensures local modes occur
KCl:exciton	Local excitation	Production of neutral vacancy and interstitial via identifiable excited state. Theory of energy surfaces exists
Diamond self-interstitial	Local excitation	Probably not local heating, nor Bourgoin–Corbett. Mechanism may also hold for Si, Ge. Theory of energy surfaces exists

of only limited application. In looking at the broader aspects of these many processes, the useful classification is simply whether the recombination energy is channelled into the electronic or vibrational degrees of freedom. The main questions addressed in this subsection are the factors which affect the rates and the ways in which one can decide which mechanism is operating. (For a summary, see table 2.)

In deciding which mechanism is important, a good starting point is the discussion of interstitial Al in silicon by Troxell *et al* (1979). Their arguments can be put into two broad classes (see Stoneham 1980). One group requires *consistency*; if these arguments are violated, the mechanism can be ruled out. The second group of arguments hinge on

what is *reasonable*; in principle, other experiments or detailed calculations could change the status of a mechanism. The 'reasonable' arguments mainly concern efficiencies of complex processes. The consistency arguments are more varied: does the mechanism leave the system stable when there is no enhancement? Do the energies balance correctly: is the recombination energy sufficient? If there is saturation, does it occur where one would expect from known capture cross sections?

The main features of the local heating model are given by Weeks *et al* (1975) and Kimerling (1978); Markvardt (1980) and Markvardt and Landsberg (1981) have re-examined the multiphonon aspects. The Weeks *et al* result gives a reaction rate which contains an efficiency η , which describes the fraction of recombination events which result in a successful reaction. This is written in the form

$$\eta = \left(\frac{k_D}{k_L} \right) \left| \frac{E_T - E_R}{E_T} \right|^{S-1}.$$

Here the first factor (k_D/k_L) is the ratio of the rate of energy flow within the 'defect molecule' to the rate of loss to the surrounding lattice. The second factor expresses the probability that energy distributed among S molecular modes can find its way to the reaction coordinate (here we intentionally avoid the Weeks *et al* phrase 'critical reactive mode', since the reaction coordinate is not normally a normal mode). The language here is that of chemical reaction theory. We can also express the same physical ideas in more conventional solid-state terms. This is important, for a solid-state defect is rarely an isolated unit weakly interacting with its host. Further, the notions of energy flow within the defect molecule and to the environment are not concepts which can be applied in all cases. In the terms of §§2 and 3 we must ask three questions. First, which states are populated initially? Here we need to know both the vibronic states involved and whether there are phase relations which correspond to a systematic motion in some reaction coordinate. Obviously it is important how the state is populated. For example, there will be differences between population via Franck-Condon transitions and via a cooling transition from some mixed state. Secondly, if there is a systematic motion, is it in the right direction? This is most easily seen in the Franck-Condon case. If the mode Huang-Rhys factors S_{n0} are written in terms of the F_n , the differences in defect forces between the initial and final states, as $S_{n0} \equiv \frac{1}{2} F_n^+ (1/\hbar\omega_n^3) F_n$, then the 'phonon kick' is along $T \equiv \sum_n F_n (1/\frac{1}{2} m\omega_n^2) e_n$, where e_n is the corresponding eigenvector. If the reaction coordinate is written $R \equiv \sum_n A_n e_n$, then $R \cdot T$ is a measure of the degree to which the motion is in the right direction, and $(R \cdot T)^2$ measures the energy transfer. Thirdly, if the systematic motion is excited in the right direction, will it persist long enough for the reaction to occur by a combination of thermal and recombination-induced motion? Clearly this involves the issues discussed in §3; equally clearly, there are very few systems for which reasonable *a priori* predictions could be made.

4.2. Persistent photoconductivity

Normally, photoconductivity persists for only a short time after optical excitation. However, in many compound semiconductors photoconductivity can persist at a high level for hours, sometimes days, after the excitation has been removed. It is essential that the sample remains at low temperatures. Indeed, the conductivity persists until the sample is heated above a characteristic temperature. The phenomenon has been observed in both III-V and II-VI hosts: $Al_x Ga_{1-x}As$ (Nelson 1977), $GaAs_{1-x}P_x$ (Craford *et al* 1968), CdS (Wright *et al* 1968), $Cd_{1-x}Zn_xTe$ (Burkey *et al* 1976), $CdTe$ (Lorentz *et al*

1964, MacMillan 1972, Iseler *et al* 1972), including alloys with changes of anion and of cation. By far the most fully studied systems have been GaAs or $\text{Al}_x\text{Ga}_{1-x}\text{As}$.

The implications are most easily seen by noting that, if simple capture of the carriers were involved, the capture cross section would be less than $10^{-14} \text{ \AA}^2 (10^{-30} \text{ cm}^2)$, many orders of magnitude below that of any well-characterised centre in other semiconductors. Moreover, there is a pronounced temperature dependence at surprisingly low temperatures: even at 100 K, thermal activation with an energy of a few tenths of a volt is seen.

Some potential explanations can be dismissed readily from the systematics of cases seen. There is no sign that the behaviour is correlated with a repulsive Coulomb interaction which hinders capture; indeed, known repulsive centres show much smaller effects from the repulsion. Even the donor or acceptor nature of the capture centre does not seem critical, except perhaps at the lowest temperatures. Likewise, from the range of hosts one can doubt any explanation based on some peculiarity of the band structure. Spin-orbit couplings are too large to rely on spin selection rules. It is likely that any

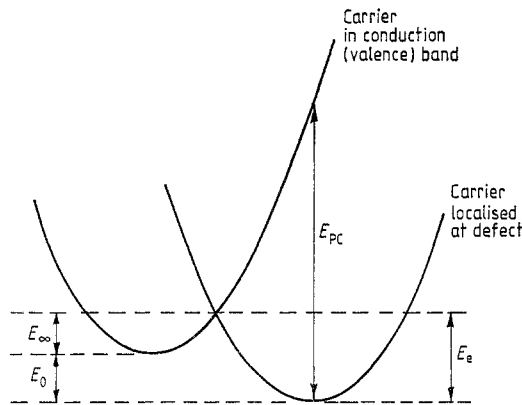


Figure 12. Energy surfaces for the large Stokes shift model of persistent photoconductivity. The energies involved are E_∞ from free-carrier capture, E_0 from Hall data, E_{PC} from the photoconductivity threshold, and E_e from thermal emission.

common defect energy-level structure must have some simple generic feature, rather than some fortuitous complex structure. Two models remain.

(i) The strong vibronic coupling model (Lang and Logan 1977, Lang *et al* 1979). Here the assumed configuration-coordinate model includes two important states (figure 12): a free-carrier state, with only small relaxation, into which the carrier is excited, and a strongly relaxed state with the carrier localised at the defect. The large lattice relaxation is the ultimate reason why the capture is inhibited. Similar energy diagrams have also been used to explain long-lived spin resonance in defect-excited states in diamond (Loubser and van Wyk 1978).

(ii) The spatial separation model (Queisser and Theodoru 1979, Theodoru and Queisser 1980) in which a macroscopic potential barrier separates the photogenerated electrons and holes. This barrier could be a junction, or an interface between different compounds.

Both models have strong support in specific cases. How often situations favour one mechanism or the other does not seem to be agreed yet.

5. Auger processes

In Auger processes, most (possibly all) the energy released in the non-radiative transition is taken up as the kinetic energy of the free carriers. From this we can conclude at once that (a) Auger processes become increasingly favoured as carrier concentrations rise, (b) they will have a modest temperature dependence in most cases, since carrier concentrations change less than vibration amplitudes with temperature, and (c) Auger processes will compete more effectively with multiphonon processes at low temperatures, as the multiphonon contribution becomes small.

Given the multitude of possible Auger processes (see, for example, Landsberg and Robbins 1978) and the several reviews available (Stoneham 1975, §14.4.3, Robbins 1980b) it would seem superfluous to attempt a comprehensive survey. Rather I have singled out a small number of features which warrant attention. Other specific examples occur later, including the possible role of colloids as killer centres (§6.4).

5.1. Defect-free Auger processes

Even though this review is mainly concerned with impurity-induced non-radiative transitions, some preliminary remarks about Auger recombination in perfect crystals may be useful. Here energy and momentum conservation are strong constraints. A typical process might involve three particles, either two holes and one electron, or two electrons and one hole. The transition probability involves an integral over all final-state wavevectors and energies, and an average over initial-state wavevectors and energies. The integrand contains four main factors. Two parallel those in multiphonon transitions, namely a matrix element and an energy-conservation factor. The other two factors ensure momentum conservation and incorporate the various statistical aspects as combinations of Fermi-Dirac occupation numbers.

The main features can be seen from quite general arguments (e.g. Haug 1978). These give the standard dependences on carrier concentration one would expect from mass-action arguments (e.g. $\sim n^2p$ for two electrons and a hole in a non-degenerate semiconductor), though there are complications at high excitation levels. The temperature dependence comes from the statistics factor in the absence of special selection rules, and so any temperature effects should often be modest. As usual, absolute rates cause some problems, including screening. For fuller discussions of matrix elements, see Antonic and Landsberg (1963), Dzwig (1979), Lochmann and Haug (1980) and Burt (1981).

A final point concerns phonon-assisted Auger transitions. Usually the phonon acts to conserve momentum: it is analogous to a promoting phonon, not an accepting phonon. Since only one phonon is involved, the temperature dependence is weak. This process is distinct from the Auger multiphonon process we discuss later.

5.2. Impurity-associated Auger processes

The impurity-associated Auger processes differ in three main respects. First, the momentum conservation is altered, for the translational symmetry has been destroyed. Secondly, the consequences of energy conservation are modified by the discreteness of the energy levels of the defect. Finally, even the free-particle wavefunctions are modified by the presence of the defect: plane-wave band functions need not suffice.

The Auger processes can be classified in several ways. One is given in table 3, which

Table 3. Impurity-associated Auger processes.

Numbers of carriers involved		Lifetime dependence (concentrations N of defects, n, p of free carriers)	Examples†
Free	Bound		
(a) 2	1	Np (electron) Nn (hole), Nn	$h + e + \boxed{e} \rightarrow \boxed{} + e$ $e + e + \boxed{} \rightarrow \boxed{e} + e$ Both seen for donors in Ge (Koenig <i>et al</i> 1962). Other mechanisms listed
(b) 1	2	N	$h + \boxed{ee} \rightarrow e + \boxed{}$ No cases demonstrated; proposed for an He-type defect in GaP, possibly GaP:O by Jaros (1978) and Riddoch and Jaros (1980)
(c) 0	3 or more	No direct dependence on N, n, p , though population of the initial state will depend on carrier concentrations	$\boxed{e(eh)} \rightarrow \boxed{} + e$ Exciton at neutral donor in (a) GaP:S (Nelson <i>et al</i> 1966), (b) Si donors (Schmid 1977, 1978), (c) Cu ₂ O (Khas 1965), (d) alkali halides $\boxed{h(eh)} \rightarrow \boxed{} + h$ Exciton at neutral acceptor. (a) GaP:X (Dean <i>et al</i> 1971), (b) Si (Schmid 1977, 1978), (b) Si (Schmid 1977, 1978, Osborne and Smith 1977) $\boxed{e(eh)} \rightarrow \boxed{} + (eh) \text{ (Merz } et al \text{ 1969)}$
<i>Other Auger-type process</i>			
(d)	$e + \boxed{e^*} \rightarrow \boxed{e} + e^*$	Inelastic scatter (ZnS:Mn, Gordon and Allen 1980)	
(e)	$\boxed{e^*}_1 + \boxed{e}_2 \rightarrow \boxed{e}_1 + \boxed{e^*}_2$	Inter-defect energy transfer, possibly involving multiphonon transitions (many systems; see, for example, Knox 1968, Dexter <i>et al</i> 1969, Soules and Duke 1971)	
(f)	$\boxed{e^*h} \rightarrow \boxed{eh^*} + \hbar\omega$	Intra-defect energy transfer with multiphonon cooperation (Itoh <i>et al</i> 1980)	
(g)	$e + \boxed{} + [\oplus e] \rightarrow \boxed{e} + \oplus + e$ $e + \boxed{} + [\ominus h] \rightarrow \boxed{e} + \ominus + h$ $[\oplus ee] + [\ominus h] \rightarrow [\oplus e^*] + \ominus$	Inter-defect Auger processes for GaP:O (Dean and Henry 1968, Gal <i>et al</i> 1981)	

† For fuller details see Stoneham (1975, p541) or Landsberg *et al* (1964) or Landsberg and Robbins (1978).

covers the processes seen. Many other processes are postulated, and I have not listed those suggested but still to be observed in III-V or similar semiconductors. Clearly, in most cases the rate of reaction is given in the form (defect concentration) \times (some power of carrier concentrations). The powers ν are typically 1 or 0, so defect Auger processes can be important at much lower carrier concentrations than free-carrier processes.

The best-studied impurity Auger processes are those involving three bound carriers, as in an exciton bound to a neutral donor or acceptor. Here the striking feature is the dependence on donor (or acceptor) binding energy: the rate increases as the fourth power of the binding energy (Dean *et al* 1971). Of the E_B^4 dependence, roughly $E_B^{3/2}$ comes from the scaling of the bound-state wavefunction and $E_B^{5/2}$ from momentum conservation. Rough estimates of the absolute rates for GaP:S using results scaled from

Table 4.

Device	Phenomenon	Mechanism proposed	Reference (R = review)
cw lasers; lasers for optical communications	Rapid degradation (time scale of hours at room temperature)	Seeded dislocation	Hayashi (1981) (R)
	Gradual degradation (time scale of the order of years)	Non-seeded dislocation loops; climb by interstitial condensation, accelerated by stress	Hayashi (1981) (R)
Injection lasers (buried heterostructure)	Pulsations	Macroscopic regions or surfaces which are non-radiative rather than amplifying	Henry (1980a, b)
Double-heterostructure lasers	Large dark spots	Recombination at scribed surface. Minority carriers transported to groove by current flow in uncontacted samples containing p-n junctions	Henry and Logan (1977)
	Dark line defects (GaAs, GaP)	Growth of dislocation network. Climb in active region; dislocation gives non-radiative recombination	Kimerling (1978) (R), Petroff and Hartman (1974), Hutchinson and Dobson (1975), van Vechten (1975a, b)
	Catastrophic damage	Local melting from heat generated by recombination. Epitaxial recrystallisation leaves frozen-in defects	Henry <i>et al</i> (1979)
Light-emitting	Suppression of luminescence	(a) ZnO dissociation in GaP.	Dean and Choyke (1977) (R)
		(b) 0.75 eV killer centre in GaP	Dapkus and Henry (1976)
		(c) 0.92 eV killer centre in GaP	Wight (1977)
	Dark line defects. Abrupt degradation. Dislocation, surface and interface recombination	Three-dimensional networks of dislocation dipoles. Epitaxially recrystallised material made non-radiative. Generation of dislocation loops and other structures by glide at high temperatures limited by carrier diffusion	Petroff <i>et al</i> (1976), Ueda <i>et al</i> (1980), Wight (1977)
Tunnel diodes	Full in peak current with time in operation		
IMPATT devices (GaAs:Pt Schottky barrier)	Degradation	Interface reactions at high temperatures; recombination-enhanced reactions at low temperatures	Kimerling (1978) (R)

the internal conversion problem of nuclear physics (Nelson *et al* 1966) give unexpectedly good agreement for the ratio of the Auger to radiative recombination rates.

Since the bound-carrier energy levels are discrete, Auger transitions only involving bound carriers cannot occur as a rule without an extra energy reservoir being involved. Phonon cooperation is to guarantee energy, not momentum, conservation. There is thus a basic difference between phonon-assisted Auger processes with and without defect involvement. Rebsch (1979) has discussed the multiphonon enhancement for processes involving one bound and two free carriers. Another impurity case is energy transfer between defects, well-known for many years (Dexter *et al* 1969, Soules and Duke 1971) and involving theory closely resembling that of §2. Indeed, the main difference comes from the fact that each defect has accepting modes of its own, so the lineshape function $G(\omega, T)$ is replaced by a convolution of two lineshape functions. J W Allen (1981 private communication) has argued that energy transfer is necessary to explain consistently the optical lineshapes and the luminescence quenching of ZnSe:Mn and also some sample-to-sample variations observed. Another example of a phonon-assisted Auger process occurs in the self-trapped exciton in KCl, where it is important because it provides a new channel of decay. The process, which could have parallels in semiconductors, involves the transfer of excitation of the electron component of the exciton to the hole component (Itoh *et al* 1980).

Finally, we come to the question of transition matrix elements. For very shallow centres, effective mass theory and scaling are adequate. For deeper defects, this is not so. Bess (1958) and Jaros (1978) have both commented on the effects of correlation among the bound particles and the modification of the free-carrier wavefunctions by the defect. The calculation of Itoh *et al* incorporated a self-consistent molecular orbital treatment of the modulation of the electron-electron interaction by the lattice vibrations. Treatments at this level would seem essential for compact states, except in rare cases where one could draw on free-atom data.

5.3. Laser annealing

Three related methods of modifying the surface region of a solid, typically an ion-implanted semiconductor, are *pulsed laser anneal*, *continuous laser anneal* and *electron beam heating*. Of these, the commonest, the most controversial as to mechanism, and the most relevant here is pulsed laser anneal. A laser pulse of a few J cm^{-2} of above band-gap light produces effects broadly consistent with local melting and epitaxial resolidification. Bigger pulses cause damage, smaller pulses have little effect; the essence of effective application is to produce consistent anneals within the 'window' of operating conditions. For many purposes it is not important to know whether the 'molten' state resembles that of the normal thermal melt or whether the high excitation produces a distinct 'plasma-like' state.

In all three methods the energy input per atom is enormous. If 1.5 J cm^{-2} is actually absorbed in a $10 \mu\text{m}$ layer of Si, the energy input is around 5 eV per atom, four times larger than the low-temperature band gap and an order of magnitude higher than the latent heat of melting. Certainly some light will be reflected, and some may be re-radiated rapidly. Nevertheless, it is of some practical importance to know the spatial and temporal distribution of the energy which finally ends as atomic motion, since this affects the dynamics of recovery and the distribution of impurity species with depth.

The fullest analyses of the non-radiative transitions are those of Yoffa (1980) and Dumke (1980). Both note that Auger recombination is fast at high carrier densities.

Indeed, Yoffa observes that the electrons and holes rapidly reach quasi-equilibrium, with a quasi-Fermi level, at a temperature which is lowered by thermal excitation of plasmons. The extra degrees of freedom associated with plasmons are here an important energy sink, especially on a very short time scale, as the quasi-equilibrium is set up in roughly 10^{-14} s; the Auger processes give recombination on a time scale of a few picoseconds.

6. Theory and practice and non-radiative transitions

Links between calculations of (for example) the electronic structure of a defect and the observed deterioration of behaviour in a semiconductor device may be tenuous, and the details rarely established in depth. This section examines the relationship. To do this, three main features are examined. First, which are the semiconductor problems which appear to be determined by non-radiative transitions? Secondly, how does one characterise defects which cause non-radiative transitions? Thirdly, how can theory be used to decide how a particular defect will behave and its influence on semiconductor behaviour? It will be seen that theory is used in two main ways. One is phenomenological: can one model device behaviour in terms of specific carrier concentrations, mobilities and recombination rates (which may be influenced by defects)? Can one characterise uniquely specific defects by a few parameters like binding energy, relaxation energy, etc? The other use involves the quantitative calculations of defect electronic structure. It will become clear that the factors which influence non-radiative rates themselves impose limits on those theoretical methods likely to be useful in all but a few special cases.

6.1. Non-radiative transitions in devices

No doubt scientific curiosity has provoked some studies of non-radiative transitions. Yet the main driving force in this area has been and will continue to be demand for more efficient devices. In this subsection we note some of the specific device problems and the types of non-radiative transition postulated to explain them (see table 4). We remark that the defects responsible may appear in a device at any stage: in crystal growth, in device assembly, and during operation.

(a) *Processes limited by carrier transport.* In these an effective non-radiative mechanism is necessary, but its nature is not critical. Examples include the large dark spot phenomena in double-heterostructure lasers, as well as many of the usual dislocation and surface effects.

(b) *Macroscopic inhomogeneity.* Here the non-radiative regions giving rise to pulsation in injection lasers are an example. In principle, precipitates and inclusions may have much the same properties as non-radiative point defects, and a special case will be discussed in §6.4.

(c) *Point-defect production.* In certain cases dislocation climb occurs far too rapidly to be explained by thermal diffusion of those defects present in equilibrium. In these cases one must postulate a mechanism of recombination-induced defect production.

(d) *Recombination-enhanced motion.* This is manifest in several forms: in dislocation climb, in dislocation glide, and in various defect reactions (§4.1). I include here all effects of transient defects too.

(e) *Conventional non-radiative transitions.* These occur at point defects and line defects where they are, of course, essential, but not always rate-determining. This is (at least

partly) because careful materials preparation eliminates many of the obvious problem species.

Within the context of the present review, processes (c), (d) and (e) are the most important. Of these, the role of dislocations needs special comment (a recent summary of the issues is given by Petroff (1979)). Without going into details, we note first that dislocation climb involves the motion of two species (e.g. Ga and As). Secondly, since defects of these two species are unlikely to be present in equal supersaturation in as-grown samples, one has to postulate a defect reaction in which the one in excess can react with the dislocation jog so as to generate a vacancy–interstitial pair of the other. Climb then occurs by the absorption of the species in excess and the emission of one member of the vacancy–interstitial pair. Petroff and Kimerling (1976) (see also Petroff 1979) have proposed one model based on excess interstitial Ga and As Frenkel defect production, though some uncertainties remain. Indeed, both the climb mechanism and the precise origin and nature of the point defects are in doubt. However, it is clear that recombination-enhanced motion of point defects is an essential component.

6.2. The fitting of defect parameters

How does one characterise a defect, given data on capture cross sections as a function of temperature and possibly some extra information like optical absorption, or perhaps electric-field effects? For present purposes I assume there are no awkward gaps in the data, and that there is no problem with incomplete resolution of contributions from several defects.

Ideally, one would end with some four parameters, or more in a more general model, namely an enthalpy H_i and entropy of ionisation, a relaxation energy, $S_0\hbar\omega$, and an effective phonon energy $\hbar\omega$. These parameters are then used in three ways. One is as a label, like ‘the electron trap at 0.4 eV below the conduction band’. A second is as a test of a particular atomic model: is the degeneracy right? Is the optical threshold consistent with a specific charge state? The third use is as a consistency check: are the trends among related cases consistent? Do the zero-phonon lines or phonon replicas show isotope effects? The trends might be those with ionisation energy, where quantum defect theory is the connecting link, or it might be a chemical trend, as in the discussions by Hjalmarson *et al* (1980).

The ideal situation is rarely, if ever achieved, even with the most careful fits to formulae like those in §2. Indeed, *in the circumstances which usually prevail, any detailed fit to a complicated form of $G(\omega, T)$ or $\sigma(T)$ will give a false illusion of accuracy.* This depressing conclusion arises from the fact that one is trying to describe a complicated incompletely specified system with a grossly oversimplified model. However, the conclusion focuses attention on the fact that it is *effective* parameters, such as an *effective* phonon energy, which are involved, and the implied effects on other parameters. The best one can hope for, if one has no established atomic model, is a set of parameters which describe adequately and consistently the range of experimental results. The fit may not be unique.

Why does this situation arise? The most important contributing reason is this: if one does not have an atomic model, one does not know what temperature dependence (or even which selection rule) to associate with the promoting mode. The factor $|M|^2$ in §2 could be independent of temperature, or linear, the two choices usually made. Yet one can identify circumstances in which it will show thermal activation or even the opposite trend with temperature (§2.1). A second problem is that the usual theory con-

cerns thermodynamic internal energies (zero temperature and constant volume rather than constant pressure). A third problem concerns the effective frequencies. Even if there is linear coupling alone, so long as there is interaction with a range $\Delta\omega$ of lattice modes, the effective frequency will itself depend on temperature. The issue becomes still more complicated when the effective frequencies change between the two states involved (as will usually be the case for deep levels because of the change in charge density) or when there is substantial anharmonicity. There are also many other minor corrections. Auger processes could provide another channel at low temperatures (e.g. Robbins 1980a). Thermal expansion of the host is usually unimportant, but should not be forgotten. Changes in concentrations of other shallow defects may lead to impurity banding and shifts of ionisation energies. Dynamic Jahn–Teller effects can lead to a change of the symmetries of the important vibronic states over modest temperature ranges. Nor must one forget the Sommerfeld factors which depend on the defect charge state, which is often unknown.

These arguments apply to defects whose atomic nature is uncertain. In special cases there may be enough information to over-determine the parameters of the model, and hence to have a consistency check too; an example is the neutral vacancy in diamond (Stoneham 1977) where the Jahn–Teller effect can be analysed in some detail. The most-analysed results where doubts about atomic nature remain are for GaAs, and here there has been controversy recently. The situation can be understood from the two tables. The first (table 5) assembles results on the B centre in GaAs. This gives one of the only two deep levels present in undoped GaAs grown by liquid-phase epitaxy. The level lies at 0.71 eV above the valence band and is not associated with some of the transition metals (Cu, Fe, Mn, Cr) first suspected. The capture cross section is given in figure 3 of Henry and Lang (1977).

Table 6 summarises some of the results for two defects in GaAs. One is the E3 level at 0.31 eV below the conduction band, formed by bombardment by electrons with MeV energies in layers of GaAs grown by vapour-phase epitaxy (Lang and Kimerling 1974). The E3 level may be associated with an intrinsic vacancy, both Ga and As vacancies having their supporters. The other is the EL2 centre, possibly associated with oxygen (though some samples with more EL2 defects than total oxygen impurities have been reported), and a native defect in bulk GaAs layers as well as those from vapour-phase epitaxy. The ensemble of the data show a helpful, if incomplete, consistency. The data in tables 5 and 6 also show an important point: the effective frequency depends on the defect as well as on the host. One cannot simply guess the frequency from the phonon dispersion curves of the host. For comparison, the L0 phonon energies are 37 meV in GaAs and 49.5 meV in GaP; the T0 energies are 33 meV in GaAs and 45 meV in GaP.

A third table is in order too, covering two defects where there is some information about the atomic nature (table 7). One is the ZnO centre in GaP, the other a centre associated with Cr in GaAs which acts predominantly as a hole trap. There is some doubt about details of the GaAs:Cr centre, e.g. whether Jahn–Teller terms matter, or which charge state is involved. One notices a further different effective frequency.

6.3. Relation to the theory of electronic structure

We now turn to what calculations of electronic structure can contribute to understanding non-radiative transitions. I shall discuss the case of point defects, though most of the arguments carry through unaltered for line defects.

It is worth remarking that, in many important systems, it is unlikely that either theory

Table 5. Analysis of the B centre in GaAs.

Author	Method	$\hbar\omega$ (meV)	S_0	p	S_0/p	Comments
Ridley (1978a, b)	Fit broad trend. Higher temperature values of $\sigma(T)$ important	37 (chosen)	1	~ 20	0.05	Trend of $\sigma(T)$ at lowest temperatures less satisfactory
Burt (1979)	Fit broad $\sigma(T)$ at lower temperatures	37 (chosen)	~ 3	~ 20	0.15	Exploits large p to get S_0 approximately from $\sigma(T=0)$. No assurance that σ accurate at high T without further refinement
Pässler (1980)	Fit fuller expression over wide temperature range	22.3	7.7	32.3	0.24	If Auger processes are as important as Robbins suggests, these parameters are not consistent with observed values
Robbins (1980a)	Revision for Auger processes	~ 18	~ 12	~ 40	0.3	Some dependence on assumed electron concentration
Markvardt (1981)	Semiclassical 1D model using detailed T dependence from experiment	~ 45	Not given	~ 16	Not given	Predicts energy difference of upper and lower states versus Q . Needs accurate $\sigma(T)$
Burt (1981a, b)	Graphical giving S_0/p against p	~ 18	~ 17	~ 40	~ 0.3	Some dependence on assumed electron concentration

or experiment alone can give a complete picture. Experiment alone cannot always probe, let alone characterise fully, the transient states which may be involved for very short times only, and theory alone will always be limited as to the range of states and geometries which can be studied in depth. However, the two together may give a rather

Table 6. E3 and EL2 defects in GaAs.

Level	Experiment	References	S_0	$S_0\hbar\omega$ (meV)	$\hbar\omega$ (meV)
E3	Capacitive transients	a	7.5 ± 2.4	75 ± 25	10.5 ± 1.5
	Cross section versus temperature	a, b	—	95 ± 10	—
	Differential capacitive transient	c	9 ± 1	100 ± 15	11 ± 1
EL2	Differential capacitive transient	c	6.5 ± 1.5	125 ± 15	20 ± 5
	Cross section versus temperature	c, d	5.5 ± 4	115 ± 50	20 ± 3
	photoionisation	e	—	120	—

a, Pons and Makram-Ebeid (1980).

b, Henry and Lang (1977).

c, Makram-Ebeid (1980a, b).

d, Mitoneau *et al* (1979).

e, Bois *et al* (1979).

Table 7. Impurity centres in GaAs and GaP.

Centre	Method	Reference	S_0	$S_0\hbar\omega$ (meV)	$\hbar\omega$ (meV)
GaP: (Zn, O)	Luminescence	a	10.2	19.0	18.6
	Differential capacitative transient	b	11.1	206	18.6 (chosen)
	Photoluminescence	c			6.0
GaAs:Cr	Cross Section versus temperature	c		170	28
	Optical absorption	d		180	40 ± 5
	Electric-field effects on emission rate	b		195 ± 15	35 ± 10

a, Henry and Lang (1977).

b, Makram-Ebeid (1980a, b).

c, Makram-Ebeid *et al* (1981).

d, Hennel *et al* (1980).

e, Feenstra and McGill (1981).

complete picture of the nature of non-radiative transitions and of the critical steps. This is the case, for example, in the area of photochemical damage in alkali halides.

From previous sections (e.g. figure 4) it is obvious that the energy surfaces $E_i(\mathbf{Q})$ comes from one of the two critical components. Thus, in simple multiphonon transitions $E_i(\mathbf{Q})$ determine the parameters such as S , p and $\hbar\omega_{\text{eff}}$. In cooling transitions the branching ratio depends on the precise shapes, especially near any crossing point. One of the models of persistent photoconductivity depends directly on the relationships of the several energy surfaces. Obviously the lowest-energy geometries for the different charge states must correspond to different defect sites if the Bourgoin–Corbett mechanism is to operate. For a local heating mechanism to be efficient, the effective frequency should not be close to the main peak in the phonon density of states of the host. For a local excitation model, one should be able to identify the important excited state.

Energy surfaces can be obtained either by direct calculation or by the analysis of experimental data. Indeed, in §6.2 the fits of capture data to S , p and $\hbar\omega_{\text{eff}}$ could be regarded as determining an energy surface, in just the way optical spectra are used to derive configuration-coordinate diagrams. Unfortunately, the procedure has only limited scope, for the contribution of promoting modes or of Jahn–Teller terms may be hidden. In special cases one can get rather detailed information from optical spectra, one example being the so-called GR1 centre (neutral vacancy) in diamond, where fine structure in the optical spectrum permits a reasonably complete analysis of the energy surface in this Jahn–Teller system (Stoneham 1977b). Direct calculations of energy surfaces are relatively rare, and may still over-simplify greatly the description of the vibrational degrees of freedom. One must exclude here papers which call the sum of one-electron eigenvalues a ‘total’ energy, since this double-counts electron–electron interactions and omits nuclear–nuclear interactions. Despite some valuable qualitative insights (e.g. Corbett *et al* 1973), the approximations have consequences which are often dramatic (e.g. Larkins 1971), especially for geometries such as the bond-centred interstitial (see the comments of Mainwood *et al* (1978)). Only a few cases of energy surfaces seem to have been studied self-consistently, and these few are mainly intrinsic defects of the simplest kind. There is very little overlap with systems in which non-radiative transitions are observed, the study of self-interstitials being closest (Mainwood *et al* 1978). This gap is not yet another

sign of the theorists' ivory tower, but reflects both the real difficulty of accurate calculation and the fact that the atomic natures of many key defects are still not known. Moreover, one specific technical difficulty has barely been touched: in non-radiative transitions, one is almost inevitably involved in energy surfaces in excited states. Problems arise of convergence, of accuracy, and in some cases even of identifying the important state among many contenders. Further, some approaches are strictly applicable only to the ground state, the density functional approach being one instance. Even where a method has adequate generality, the choice of basis may remain a problem. It is hardly surprising that, at the present time, a semi-empirical method (CNDO) has been the one most widely applied to the problems discussed here or to parallel situations in more ionic systems.

The second of the two critical components in non-radiative rates concerns the transition matrix element. For multiphonon transitions this inevitably leads to the question of the promoting mode, and to whether it is or is not the same as the accepting mode. For Auger processes at deep centres one has the particular complication that the wavefunctions of all the localised electrons change when one electron changes its state.

The general picture to emerge from §6.2 was that much theory had been aimed at phenomenology for simple models, and that in many cases these were in forms which have no major advantages in direct application. The discussion in the present subsection has equally unsatisfactory conclusions: there have been very few attempts to calculate energy surfaces or matrix elements, or indeed to provide a framework within which the complex components of non-radiative transitions can be handled.

6.4. Killer centres

A persistent problem is the identification of 'killer centres', those imperfections which cause rapid recombination even at low concentrations. For present purposes we shall assume transport of the carriers to the defect centre is not rate-determining. What models have been proposed for 'killers'?

(a) *Defects with many closely spaced bound states.* In these systems the cascade mechanism is possible (energy separations \lesssim phonon energies). Shallow donors and acceptors exhibit high carrier capture cross sections; indeed, any Coulomb-attractive centre will have close highly excited states which may prove effective. It is widely believed that transition-metal impurities are effective killer centres but, whilst some examples are known (e.g. Ni in GaP), there are wide variations in effectiveness from host to host.

(b) *Defects with favourable vibrational properties.* For fast multiphonon processes one has two demands which could conflict: a high-frequency accepting mode, so that few phonons need be emitted, and a large-amplitude motion in the promoting mode. One can envisage centres where both occur, e.g. 'split-interstitial centres' (Stoneham 1977) where there is both a high-frequency local mode and a low-frequency resonance. These types of defect are best known in metals (e.g. Dederichs *et al* 1972); examples have been seen in semiconductors, including the di-carbon centre in silicon (Watkins and Brower 1976), but the concentrations are so low that studies of non-radiative transitions have not been attempted.

(c) *Defects with strong electron-lattice coupling.* 'Strong' is used here in the sense of §1, i.e. a large change in displacement is involved. Simple vacancy centres have large state-dependent Jahn-Teller energies (see, for example, Watkins 1975) and so would be candidates, either in isolation or associated with impurities. It is tempting to argue dislocations will be effective because, in a restricted sense, one can relate the rebonding at

them to that which occurs in vacancies. However, this appears to be naive. Certainly decorated dislocations are effective in causing recombination (Blumtritt *et al* 1979). For clean dislocations there is also a dependence on the precise degree of dissociation, Burgers vector and chemical state (Ourmazd and Booker 1979, Petroff *et al* 1980). Moreover, a dislocation may merely concentrate non-radiative centres present in the bulk, and make no other contribution itself (Queisser 1980). Further, in special cases the dislocation may be critical yet only indirectly involved, as when the electric field from a charged dislocation inhibits capture into shallow levels (Tasker and Stoneham 1977).

(d) *Rapid Auger processes.* These are mediated by the electron-electron interaction. One of the important factors therefore is the localisation of the captured carriers; another is the effect of an attractive potential of the defect on the free carriers (e.g. the Sommerfeld factor). Thus deep levels have an advantage, and this is confirmed in detailed work. Specific examples proposed in this context are the antisite-vacancy complexes like $V_{\text{Ga}}-\text{P}_{\text{Ga}}^{2+}V_{\text{Ga}}^-$ or $V_{\text{P}}+\text{Ga}_{\text{P}}^{2-}V_{\text{P}}^+$ discussed by van Vechten (1975a, b) and the He-like defects discussed by Jaros (1978). The examples discussed by van Vechten could exhibit multiphonon transitions, as in (c) above.

A final example is of a rather different nature. In GaP there is a killer centre, the so-called 0.75 eV defect (Wight 1977), which has a large capture cross section and can cause recombination at a formidable rate without saturation. It correlates with dislocations, but carrier diffusion limits mean it cannot be a dislocation itself. Nor apparently is the defect a transition metal, as doping experiments show. Also the defect is not generated by standard radiation damage. In GaP grown by slow-growth liquid-phase epitaxy the killer centre is missing, though it appears as soon as these GaP samples are used in devices. One possible explanation is that the killer centre is a metallic precipitate, perhaps a Ga colloid. The recombination energy of carriers is transferred initially into energy of the conduction electrons, and energy from this reservoir is then removed by standard thermal conduction (A M Stoneham 1977, unpublished work). Provided the colloids can indeed mimic a 0.75 eV level, the model has distinct advantages. One can see how it might suppress luminescence (indeed, Glauber *et al* (1969) note that colloids in ionic crystals quench radiative recombination); one can see how the defects might correlate with dislocations (which would act as sinks for excess P) without being a dislocation; one can see why the centres might emerge on use as a device; one can understand the lack of influence of transition metals; and more detailed discussion shows how hard direct observation of the defect responsible would be. However, since colloids have been studied widely in other systems, there exists a basis of knowledge from which one might identify ways to minimise colloid formation and hence (if the model is correct) control this killer centre.

References

- Abakumov VN, Perel VI and Yassievich IN 1978 *Sov. Phys.-Semicond.* **12** 1-18
 Abakumov VN and Yassievich IN 1976 *Sov. Phys.-JETP* **44** 345
 Antonic E and Landsberg PT 1963 *Proc. Phys. Soc.* **82** 337
 Azumi T and Matsuzaki K 1977 *Photochem. Photobiol.* **25** 315
 Barnett B and Englman R 1970 *J. Luminesc.* **3** 55
 Bartram RH and Stoneham AM 1975 *Solid St. Commun.* **17** 1593
 ——— 1979 *Bull. Am. Phys. Soc.*
 Bates DR 1960 *Proc. R. Soc. A* **257** 22
 Beleznyay F and Andor L 1978 *Solid St. Electron.* **21** 1305
 Bellamy WC and Kimerling LC 1978 *IEEE Trans. Electron Devices* **ED-25**

- Bess L 1958 *Phys. Rev.* **111** 129
- Blumtritt H, Gleichman R, Heydenreich J and Johansen H 1979 *Phys. Stat. Solidi a* **55** 611
- Bois D, Chantre A, Vincent G and Nouailhat A 1979 *Physics of Semiconductors 1978. Inst. Phys. Conf. Ser. No 43* ed B L H Wilson (Bristol: The Institute of Physics) p295
- Bonch-Bruевич VL and Landsberg EG 1968 *Phys. Stat. Solidi* **29** 9
- Bourgoin J and Corbett JW 1972 *Phys. Lett.* **38A** 135
- Brailsford AD and Chang TY 1970 *J. Chem. Phys.* **53** 3108
- Buimistrov VM and Pekar SI 1957 *Sov. Phys.-JETP* **5** 970
- Burkey BC, Khosla RRP, Fischer JR and Losee DL 1976 *J. Appl. Phys.* **47** 1095
- Burt MG 1979 *J. Phys. C: Solid St. Phys.* **12** 4827
- 1981a *J. Phys. C: Solid St. Phys.* **14** 3269–78
- 1981b *J. Phys. C: Solid St. Phys.* **14** L845
- Bykhovskii V, Nikitin EE and Ouchnnkova MY 1965 *Sov. Phys.-JETP* **20** 500
- Corbett JC, Bourgoin JC and Wiegel C 1973 *Radiation Damage and Defects in Semiconductors. Inst. Phys. Conf. Ser. No 16* ed J E Whitehouse (Bristol: The Institute of Physics) p1
- Coulson CA and Zalewski K 1962 *Proc. R. Soc. A* **268** 437
- Craford MG, Stillman GE, Rossi JA and Holonyak N 1968 *Phys. Rev.* **168** 867
- Dapkus PD and Henry CH 1976 *J. Appl. Phys.* **47** 4061
- Dean PJ and Choyke WJ 1977 *Adv. Phys.* **26** 1
- Dean PJ, Faulkner RA, Kimura S and Iligems M 1971 *Phys. Rev. B* **4** 1926
- Dean PJ and Henry CH 1968 *Phys. Rev.* **176** 928
- Dederichs PH, Lehmann C and Scholz A 1972 *Phys. Rev. Lett.* **31** 1130
- Delos JB and Thorsen WR 1972 *Phys. Rev. Lett.* **28** 647
- Dexter DL, Förster T and Knox RS 1969 *Phys. Stat. Solidi* **34** k159
- Dexter DL and Fowler WB 1967 *J. Chem. Phys.* **47** 1379
- Dexter DL, Klick CC and Russell G 1955 *Phys. Rev.* **110** 603
- Dixon RW and Joya WB 1979 *IEEE J. Quant. Electron.* **QE-15** 460
- Dumke WP 1980 *Phys. Lett.* **78A** 473
- Dzwig P 1979 *J. Phys. C: Solid St. Phys.* **12** 1809
- Englman R 1980 *Non-Radiative Transitions in Solids and Molecules* (Amsterdam: North-Holland)
- 1981 *Preprint*
- Englman R and Barnett B 1970 *J. Luminesc.* **3** 37
- Englman R and Jortner J 1970 *Molec. Phys.* **18** 145
- Englman R and Ranfagni A 1980 *Physica* **98B** 151, 161
- Evangelou SN 1981 *J. Phys. C: Solid St. Phys.* **14** 2117
- Feenstra RM and McGill TM 1981 *Phys. Rev. Lett.* **47** 925
- Flynn CP and Stoneham AM 1970 *Phys. Rev. B* **1** 3966
- Freed KF and Jortner J 1970 *J. Chem. Phys.* **52** 6272
- Gal M, Cavanett BC and Dean PJ 1981 *J. Phys. C: Solid St. Phys.* **14** 1507
- Gibb RP, Rees GJ, Thomas BW, Wilson BHL, Hamilton B, Wight DR and Mott NF 1978 *Phil. Mag.* **36** 1021
- Glaubermann A E, Lai Kim-t'en, Goldenberg A B and Golub SJ 1969 *Sov. Phys.-Solid St.* **2995**
- Goto H, Adachi Y and Ikoma T 1980 *Phys. Rev. B* **22** 782
- Grimmeiss HG 1978 *Ann. Rev. Mater. Sci.* **7** 341
- Haug A 1978 *Solid St. Electron.* **21** 1281
- 1979 *J. Luminesc.* **20** 173
- 1980 *Phys. Stat. Solidi b* **97** 481
- Haug A, Kerkhoff D and Lochmann W 1978 *Phys. Stat. Solidi b* **89** 357
- Hayashi I 1981 *Proc. 15th Int. Conf. on Physics of Semiconductors*
- Heine V and Henry CH 1975 *Phys. Rev. B* **11** 3795
- Hennel AM, Szuszkiewicz WS and Martinez G 1980 to be published
- Henry CH 1980a *Relaxation of Elementary Excitations* ed R Kubo and E Hanimera (Berlin: Springer-Verlag) p19
- 1980b *J. Appl. Phys.* **51** 3051
- Henry CH and Dapkus PD 1976 *J. Appl. Phys.* **47** 4067
- Henry CH and Lang DV 1977 *Phys. Rev. B* **15** 989
- Henry CH and Logan RA 1977 *J. Appl. Phys.* **48** 3902
- 1978 *J. Vacuum Sci. Technol.* **15** 1471
- Henry CH, Logan RA and Merritt FR 1978 *J. Appl. Phys.* **49** 3530

- Henry CH, Petroff PM, Logan RA and Merritt FR 1979 *J. Appl. Phys.* **50** 3721
- Hjalmarson HJ, Vogl P, Wolford D and Dow J 1980 *Phys. Rev. Lett.* **44** 810
- Holstein T 1978 *Phil. Mag.* **37** 49
- Huang K 1981 *Scientia Sinica* **24** 27
- Huang K and Rhys A 1950 *Proc. R. Soc. A* **204** 406–23
- Hutchinson PW and Dobson PS 1975 *Phil. Mag.* **32** 745
- Iseler GW, Kafalas JA, Strauss AJ, MacMillan HF and Bube RH 1972 *Solid St. Commun.* **10** 619
- Itoh N 1976 *J. Physique* **37** C7 27
- Itoh N, Stoneham AM and Harker AH 1980 *J. Phys. Soc. Japan* **49** 1364
- Jaros M 1978 *Solid St. Commun.* **25** 1071
- 1980 *Adv. Phys.* **29** 409
- Johnston WD 1974 *Appl. Phys. Lett.* **24** 494
- Jortner J and Mukamel S 1975 *MTP International Review of Science. Physical Chemistry Ser. 2* vol 1 *Theoretical Chemistry* ed A D Buckingham and C A Coulson (London: Butterworths) p327
- Kayanuma Y 1979 *J. Phys. Soc. Japan*
- Kayanuma Y and Nasu K 1978 *Solid St. Commun.* **27** 1371
- Khas Z 1964 *Czech J. Phys.* **15** 346
- 1965 *Czech J. Phys. B* **15** 568
- Kimerling L C 1978 *Solid St. Electron.* **21** 1391
- 1979 *Defects and Radiation Effects in Semiconductors 1978. Inst. Phys. Conf. Ser. No 46* ed J H Albany (Bristol: The Institute of Physics) p56
- Knox RS 1968 *Physica* **39** 361
- Koenig AJ, Brown RD and Schillinger W 1962 *Phys. Rev.* **128** 1668
- Korol EN 1977 *Sov. Phys.—Solid St.* **19** 1327
- Kovarskii TA 1962 *Sov. Phys.—Solid St.* **4** 1200
- Kovarskii TA and Sinyavskii EP 1963 *Sov. Phys.—Solid St.* **4** 2345
- Kubo R and Toyozawa Y 1955 *Prog. Theor. Phys.* **13** 160
- Kusunoki M 1979 *Phys. Rev. B* **20** 2512
- Landau L 1932a *Sov. Phys.* **1** 89
- 1932b *Z. Phys. Sowj.* **2** 1932
- Landsberg PT 1956 *Proc. Phys. Soc. B* **69** 1056
- Landsberg PT, Rhys-Roberts C and Lal P 1964 *Proc. Phys. Soc.* **84** 915
- Landsberg PT and Robbins DJ 1978 *Solid St. Electron.* **21** 1289
- Lang DV and Kimerling LC 1974 *Phys. Rev. Lett.* **83** 489
- Lang DV and Logan RA 1977 *Phys. Rev. Lett.* **39** 635
- 1979 *Physics of Semiconductors 1978. Inst. Phys. Conf. Ser. No 43* ed B H L Wilson (Bristol: The Institute of Physics) p433
- Lang DV, Logan RA and Jaros M 1979 *Phys. Rev. B* **19** 1015
- Langer JM 1980 *Proc. 15th Int. Conf. on Physics of Semiconductors*
- Larkins FP 1971 *J. Phys. C: Solid St. Phys.* **4** 3065
- Lax M 1960 *Phys. Rev.* **119** 1502
- Leung CH and Song KS 1980 *Solid St. Commun.* **33** 907
- Levine R and Jortner J 1976 *Molecular Energy Transfer* (New York: Wiley)
- Liu SH and Eyring H 1972 *Proc. Nat. Acad. Sci.* **69** 3192
- Lochmann W and Haug A 1980 *Solid St. Commun.* **35** 553
- Lorentz MR, Segall B and Woodbury HH 1964 *Phys. Rev.* **134** A751
- Loubser JHN and van Wyk JA 1978 *Rep. Prog. Phys.* **41** 1201
- Lowther JE 1980 *J. Phys. C: Solid St. Phys.* **13** 3665, 3681
- MacMillan HF 1972 *Unpublished thesis* Stanford University
- Mainwood AM 1978 *J. Phys. C: Solid St. Phys.* **11** 2703
- Mainwood AM, Stoneham AM and Larkins FP 1978 *Solid St. Electron* **21** 1431
- Makram-Ebeid S 1980a *Appl. Phys. Lett.* **37** 464
- 1980b *Boston Conf.* in press
- Makram-Ebeid S and Lannoo M 1981 to be published
- Makram-Ebeid S, Martin GM and Woodward DW 1981 *Proc. 15th Int. Conf. on Physics of Semiconductors* in press
- Markham JJ 1956 *Phys. Rev.* **103** 588
- Markvardt T 1980 *Solar Cells* **2**
- 1981 *J. Phys. C: Solid St. Phys.* **14** L345

- Markvardt T and Landsberg PT 1981 *Tbilise Conf. on the Radiation Physics of Semiconductors*
- Merz J L, Faulkner R A and Dean P J 1969 *Phys. Rev.* **188** 1228
- Mikami Y, Mizunoya K and Nakajima T 1975 *Chem. Phys. Lett.* **30** 373
- Mitoneau A, Mircea A, Martin GM and Pons D 1979 *Rev. Phys. Appl.* **14** 853
- Monemar B, Potemski R M, Small M B, Van Vechten J A and Woolhouse G R 1978 *Phys. Rev. Lett.* **41** 260
- Morgan T N 1981 *Preprint*
- Mott N F 1978 *Solid St. Electron.* **21** 1076
- Nannich Y 1977 *Jap. J. Appl. Phys.* **16** 2089
- Nasu K and Kayanuma Y 1978 *J. Phys. Soc. Japan* **45** 1341
- Nelson D F, Cuthbert J D, Dean P J and Thomas D G 1966 *Phys. Rev. Lett.* **17** 1262
- Nelson R J 1977 *Appl. Phys. Lett.* **31** 351
- Neumark G F 1973 *Phys. Rev.* **B 7** 3802
- Neumark G F, De Bitetto D J, Bhagarva R N and Harnock P M 1977 *Phys. Rev.* **B 15** 3147, 3156
- Neumark G F and Kasai K 1981 to be published
- Newman D H, Godfrey R F, Goodwin A R and Lovelace D F 1976 *Appl. Phys. Lett.* **29** 353
- O'Hara S 1977 *J. Phys. D: Appl. Phys.* **10** 409
- Osborne G C and Smith D L 1977 *Phys. Rev.* **B 16** 5426
- Ourmazd A and Booker G R 1979 *Phys. Stat. Solidi a* **55** 771
- Pässler R 1974a *Czech. J. Phys.* **B 24** 322
- 1974b *Phys. Stat. Solidi b* **65** 561
- 1975a *Czech. J. Phys.* **B 25** 219
- 1975b *Phys. Stat. Solidi b* **68** 69
- 1977 *Phys. Stat. Solidi b* **83** K55, K111, K163
- 1978 *Phys. Stat. Solidi b* **86** K39
- 1980 *J. Phys. C: Solid St. Phys.* **13** L901
- Petroff P M 1979 *J. Physique* **40 Colloq.** C6–201
- Petroff P M and Hartman R L 1974 *J. Appl. Phys.* **45** 3899
- 1975 *Appl. Phys. Lett.* **23** 469
- Petroff P M, Johnston W D and Hartman R L 1974 *Appl. Phys. Lett.* **25** 226
- Petroff P M and Kimerling L C 1976 *Appl. Phys. Lett.* **29** 461
- Petroff P M, Logan R A and Savage A 1980 *Phys. Rev. Lett.* **44** 287
- Petroff P M, Lorimer O G and Ralston J M 1976 *J. Appl. Phys.* **47** 1583
- Pickin W 1978 *Solid St. Electron.* **21** 1299
- Pons D 1979 *Thèse* Université de Paris VI
- Pons D and Makram-Ebeid S 1980 *J. Physique* to be published
- Queisser H J 1980 *Czech. J. Phys.* **B 30** 365
- Queisser H J and Theodoru D E 1979 *Phys. Rev. Lett.* **43** 401
- Rebsch J-T 1979 *Solid St. Commun.* **31** 377
- Rees G J, Grimmeis H G, Janzen E and Skarstam B 1980 *J. Phys. C: Solid St. Phys.* **13** 6157
- Riddoch F A and Jaros M 1980 *J. Phys. C: Solid St. Phys.* **13** 6181
- Ridley B K 1978a *J. Phys. C: Solid St. Phys.* **11** 2323
- 1978b *Solid St. Electron.* **21** 1319
- 1980 *J. Phys. C: Solid St. Phys.* **13** 2015
- Robbins D J 1980a *J. Phys. C: Solid St. Phys.* **13** L1073
- 1980b *Phys. Stat. Solidi b* **98** 11
- Robbins D J and Landsberg P T 1980 *J. Phys. C: Solid St. Phys.* **13** 2425
- Robertson N and Friedman L 1976 *Phil. Mag.* **33** 753
- Sarai A and Kakitani T 1981 *Chem. Phys. Lett.* **77** 427
- Schmid W 1977 *Phys. Stat. Solidi b* **84** 529
- 1978 *Solid St. Electron.* **21** 1285
- Seitz F 1940 *Modern Theory of Solids* (New York: McGraw-Hill)
- Sinyavskii E P and Kovarskii V A 1962 *Sov. Phys.—Solid St.* **4** 2345
- 1964 *Sov. Phys.—Solid St.* **6** 498
- 1967 *Sov. Phys.—Solid St.* **9** 1142
- Soules T and Duke C B 1971 *Phys. Rev.* **B 3** 263
- Steeple K, Dearnaley G and Stoneham A M 1980 *Appl. Phys. Lett.* **36** 981
- Stoneham A M 1967 *Phys. Stat. Solidi* **19** 787
- 1975 *Theory of Defects in Solids* (Oxford: Oxford University Press)

- 1977a *Phil. Mag.* **36** 983
— 1977b *Solid St. Commun.* **21** 339
— 1979 *Adv. Phys.* **28** 457
— 1980 *Solar Cells* **2** 43
Stoneham A M and Bartram R H 1978 *Solid St. Electron.* **21** 1325
Stoneham A M, Catlow C R A and Tasker P W T 1978 *Solid St. Electron. Devices* **2** S83
Stueckelberg E G C 1932 *Helv. Phys. Acta* **8** 369
Sturge M 1973 *Phys. Rev.* **B 8** 6
Sumi H 1980a *Proc. 15th Int. Conf. on Physics of Semiconductors*
— 1980b *J. Phys. Soc. Japan* **49** 1701
Takiyama K 1978 *J. Phys. Soc. Japan* **44** 1627
Tasker P W and Stoneham A M 1977 *J. Phys. C: Solid St. Phys.* **10** 5131–40
Theodoru D E and Queisser H J 1980 *Appl. Phys.* **23** 121
Toyozawa Y 1978 *Solid St. Electron.* **21** 1313
— 1980 *Relaxation of Elementary Excitations* (Berlin: Springer-Verlag) p3
Troxell J R, Chatterjee A P, Watkins G D and Kimerling L C 1979 *Phys. Rev. B* **19** 5336
Ueda O, Imai H, Fujiwara T, Yamakashi S, Sugawara T and Yamaoka T 1980 *J. Appl. Phys.* **51** 5316
van Vechten J H 1975a *J. Electron. Mater.* **4** 1159
— 1975b *J. Electrochem. Soc.* **122** 1556
Venzl G and Fischer S F 1981 *J. Chem. Phys.* **74** 1877, 1887
Vul, A Y, Golubev L V, Sharonova L V and Shmartsev Y V 1971 *Sov. Phys.-Semicond.* **4** 2017
Watkins C D 1975 *Lattice Defects in Semiconductors 1974. Inst. Phys. Conf. Ser. No 23* ed F A Huntley (Bristol: The Institute of Physics) p1
Watkins G D and Brower K L 1976 *Phys. Rev. Lett.* **36** 1329
Weeks J D, Tully J C and Kimerling L C 1975 *Phys. Rev. B* **12** 3286
Weissman Y and Jortner J 1978 *Phil. Mag.* **B 37** 1
Wight D 1977 *J. Phys. D: Appl. Phys.* **10** 431
Wright H C, Downey R J and Canning J R 1968 *J. Phys. D: Appl. Phys.* **1** 1893
Yoffa E J 1980 *Phys. Rev. B* **21** 2415
Zener C 1932 *Proc. R. Soc. A* **1** 696
Zgierski M Z and Pauer A P 1977 *Chem. Phys. Lett.* **45** 41