ADVANCES IN PHYSICS, 1981, VOL. 30, NO. 5, 593–676

Luminescence and recombination in hydrogenated amorphous silicon

By R. A. STREET

Xerox Palo Alto Research Center, Palo Alto, California 94304, U.S.A.

[Received 30 June 1981]

ABSTRACT

Luminescence and related investigations of recombination in hydrogenated amorphous silicon prepared by glow discharge and sputtering are described. Emphasis is given to a detailed discussion of the various competing radiative and non-radiative recombination mechanisms. The evidence of the luminescence data is compared to other measurements that involve recombination processes. The experiments demonstrate that the dominant luminescence is a tunnelling transition between band tail electrons and holes. Other luminescence transitions involving defects and impurities, and the role of phonon interactions in the transition are discussed. Experiments in which luminescence is used as a probe of the properties of a-Si : H and related materials are also discussed.

CONTENTS

	PAGE
§1. INTRODUCTION.	594
1.1. Radiative and non-radiative transitions in crystalline semiconductors.	595
1.2. Preparation and properties of a-Si : H.	597
§2. UNDOPED a-Si : H.	599
2.1. A review of published data.	599
2.2. Phonon interactions.	602
2.3. Recombination mechanisms.	607
2.3.1. Low temperature decay.	609
2.3.2. The role of dangling bond defects.	612
2.3.3. Temperature dependence.	614
2.3.4. High excitation intensities.	619
2.3.5. Excitation energy effects and surface recombination.	624
2.3.6. Summary of the recombination processes.	627
2.4. Other luminescence properties.	629
2.4.1. Time-resolved luminescence spectra.	629
2.4.2. Electric field dependence and electroluminescence.	632
2.4.3. Pressure dependence.	632
2.4.4. Spin dependent luminescence.	633
2.5. Defect luminescence.	637
2.6. Above band gap luminescence.	642
§3. DOPED AND COMPENSATED a-Si : H.	643
3.1. Phosphorus and boron doping.	644
3.2. Compensated a-Si : H.	647
3.3. Other impurities.	651

0001-8732/81/3005 0593 $02·00 © 1981 Taylor & Francis Ltd

§4. Related materials issues. 652
4.1. Glow discharge and sputtered material 652
4.2. Bombardment damage, hydrogen evolution, annealing and rehydrogenation. 655
4.3. Fatigue effects. 660
4.4. Related materials. 662
§5. Related recombination processes. 665
5.1. Room temperature geminate recombination. 665
5.2. Photoconductivity. 667
5.3. Optically induced E.S.R. and absorption. 670
§6. Summary. 673
Acknowledgments. 673
References. 673

---

§1. Introduction

In 1976 the present author reviewed the luminescence of amorphous semiconductors (Street 1976). At that time the great majority of data had been collected on chalcogenide glasses. In these materials luminescence is observed at about half the band gap energy with a characteristically broad line which originates from a large electron–phonon interaction. The luminescence centre is a defect whose detailed origin is still much debated. In 1976 the study of luminescence in amorphous silicon was still in its infancy, following the first observation by Engemann and Fischer (1973). These measurements, and all the subsequent work, were made on material containing hydrogen, produced by glow discharge or reactive sputtering. Amorphous silicon made by sputtering or evaporation without hydrogen has no detectable luminescence and poor photoconductivity, precluding useful studies of recombination. On the other hand the hydrogenated material has excellent semiconducting properties including effective doping, and efficient luminescence and photoconductivity. The ability to control and vary the material properties has allowed a thorough investigation of the recombination processes so that at the present time the recombination is better understood than in the chalcogenide glasses.

This paper is mainly intended to summarize what is known about the recombination processes in a-Si : H. The areas of disagreement or uncertainty are given extra emphasis. An attempt is made to present a unified framework of recombination processes with which to interpret the many different related experiments. The emphasis is on luminescence data, but we try to relate the results to other experiments in which recombination is important. In §2 the results on undoped a-Si : H are described. Various experiments on the dominant luminescence band are used to develop the recombination model in §2.3 which is then applied to other experiments described in §2.4, and to doped material in §3. Section 4 describes experiments in which luminescence is used to gain information about materials questions. In §5 the recombination is discussed within a wider context of photoconductivity and other optically excited phenomena.

1.1. *Radiative and non-radiative transitions in crystalline semiconductors*

Before starting the discussion of a-Si : H, it may be helpful to summarize briefly what is known about luminescence in crystalline semiconductors. In this way, the recombination models for a-Si : H can be placed in a wider context. An enormous literature exists on luminescence in the cubic group IV, III–V and II–VI semiconductors, which are the closest crystalline analogies to a-Si : H, and no attempt is made at a complete description of this knowledge. To within a fair approximation, the luminescence of these semiconductors can be categorized as 'band edge' or 'deep', depending on whether or not the transition is within $\sim 0{\cdot}1$ eV of the band gap energy. The band edge transitions almost always involve either intrinsic states or donor or acceptor impurity levels for which the effective mass approximation is applied. The 'deep' recombination processes tend to involve defects, deep impurities such as oxygen, or transition metals. Of the two classes, the band edge transitions are by far the better understood. We shall describe briefly what is observed in each case and the general experimental conditions for these observations.

The low temperature band edge transitions tend to be of four types: free exciton, bound exciton, free carrier to neutral acceptor (or donor), or donor–acceptor pairs. This series also corresponds to the usual ordering in energy with the free exciton naturally having the highest energy. These transitions involve states which are not strongly localized, having Bohr radii ranging from 10 to 100 Å. The electron and hole therefore do not interact strongly with the lattice. Thus, the luminescence spectra generally consist of sharp zero-phonon transitions and phonon side bands whose intensity varies from case to case. Coupling to the lattice increases with the binding energy of the carriers (Hopfield 1959) and so is generally strongest for donor–acceptor pairs.

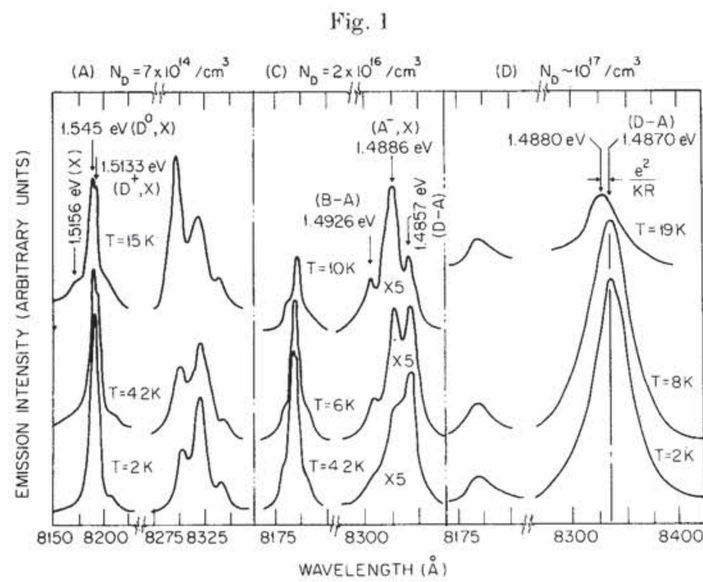
The efficiency of any luminescence transition is determined by the competing processes, either radiative or non-radiative. For example, the different band edge transitions compete with each other. At low temperatures, free excitons are only observed if the doping level is extremely low, otherwise the excitons are rapidly trapped and observed as bound excitons. Similarly, if there is a sufficient density of both donors and acceptors, pair bands will tend to dominate the band edge luminescence, because these have the lowest energy. It is, therefore, generally the case that as the purity of the material improves, the luminescence intensity moves to the higher energy transitions.

Three mechanisms of non-radiative recombination are found to be important. First, there is thermal quenching in which the carriers are excited to the band edge from where they can diffuse to deep non-radiative centres. Thermal quenching of the higher energy transitions generally occurs at the lowest temperature because these have the lowest binding energies. The second non-radiative mechanism is surface recombination. Carriers in a high quality semiconductor can easily have diffusion lengths greater than the absorption depth of light used to excite luminescence. Diffusion of carriers to the surface and recombination through surface states can therefore be a very effective recombination process. A third mechanism is Auger recombination, which is particularly significant for bound excitons because they comprise three particles localized at the same site. Often the Auger rate can be orders of magnitude larger than the radiative rate.

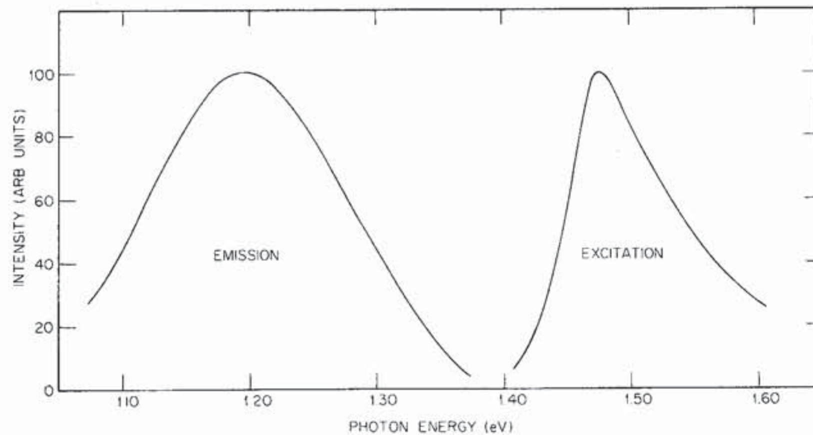
Deep luminescence in the cubic semiconductors is much less clearly understood than is the band edge luminescence. One problem is that it is often very difficult to

identify the localized state involved in a particular transition. There is also an increased tendency for these lower energy transitions to be non-radiative because the multiphonon recombination rate increases rapidly as the transition energy decreases. The larger binding energy of the states giving 'deep' luminescence also tends to increase the electron–lattice coupling although this is not true for transition metal impurities, which are largely decoupled from the lattice. However in other cases, the increased coupling is seen as a broadening of the luminescence spectrum which then becomes dominated by phonon side bands. Figure 1 shows two examples of luminescence spectra of GaAs, showing band edge and deep transitions.

Fig. 1



(*a*)



(*b*)

Typical low temperature luminescence spectra of GaAs showing (*a*) the sharp band edge transitions and (*b*) the broad deep transitions. Note the difference in the energy scale (Williams and Bebb 1972).

This brief discussion is intended to emphasize those features of crystalline luminescence which are relevant to the studies of amorphous materials. For example, a-Si : H has localized states at both band edges, which is analogous to a crystal with donors and acceptors, and indeed luminescence properties similar to those of donor–acceptor pair bands are found. Thermal quenching and Auger recombination effects are also observed, and even surface recombination, although as might be expected, the short diffusion lengths reduce the importance of this effect. In this way, we shall see that recombination in a-Si : H falls within the framework of our knowledge of crystalline semiconductors. However the phenomena that are discussed are adapted to the particular circumstances of amorphous semiconductors. For example, all of the detailed information related to band structure that is obtained from the Zeeman splitting of exciton transitions is entirely lost in an amorphous semiconductor because of the broadening due to disorder. On the other hand, the process of carrier thermalization and diffusion in a-Si : H seems to contain a richer variety of phenomena than is generally seen in crystals.

1.2. *Preparation and properties of a-Si : H*

The properties of a-Si : H have been the subject of several recent reviews (Knights and Lucovsky 1980, Fritzsche 1981, Paul and Anderson 1981) and so a detailed description is not given here. Hydrogen is the key ingredient that gives the amorphous silicon its good semiconducting properties. The hydrogen (or an equivalent element such as fluorine) terminates the majority of dangling Si bonds, and unlike the free Si dangling bond, the states associated with the Si–H bond are not in the forbidden gap. Thus, the density of deep localized states is dramatically reduced compared to unhydrogenated a-Si.

The two most common methods of preparing a-Si : H are glow discharge decomposition and reactive sputtering. In the glow discharge method, SiH_4 or some other appropriate gas (e.g., Si_2H_6 or SiF_4), dissociates in an electric discharge and deposits on a substrate. The dissociation is incomplete so that hydrogen remains in the films. In reactive sputtering, a silicon target is sputtered by an inert gas, usually Ar, in the presence of a small partial pressure of hydrogen. With either technique, the films generally contain 5–25 at.% hydrogen, and all of the properties of the material are sensitive to the hydrogen content and to the details of deposition. The dependence of the luminescence on preparation conditions is described in §4.1.

Although the material is amorphous, the atomic structure is in fact rather complex. a-Si : H has a tendency to grow into a columnar structure (Knights and Lujan 1979). The columns are oriented in the direction of growth and have a characteristic diameter of about 100 Å. The material between the columns is of low density containing a higher fraction of hydrogen than the interior of the columns. This structure is believed to be a consequence of the nucleation and growth process in which the growing surface is highly hydrogenated. The columnar structure can be inhibited in glow discharge material by growing films from pure SiH_4 and at low discharge power. Such films have the best semiconductor properties and the lowest defect density. As discussed in §4.1. the luminescence data give virtually no indication of the columnar structure.

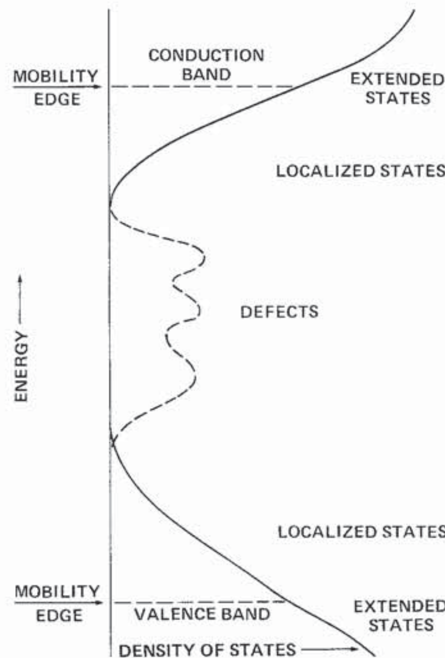
Hydrogen is bonded in SiH, SiH_2, SiH_3 and $(SiH_2)_n$ units, each of which has an identifiable infra-red absorption signature (Knights and Lucovsky 1980). The relative concentration of each structure is dependent on the growth conditions; for example, the material between the columns is predominantly $(SiH_2)_n$, and the

observation of this type of hydrogen bonding is generally a signal for material with poor electronic properties and a high defect density. Even when the columnar structure is absent, there is clear evidence for small hydrogenated voids in the material which results in a non-random distribution of hydrogen (Leadbetter *et al.* 1980).

The electronic properties are also very sensitive to the deposition conditions (see, for example, Mott and Davis (1979) for a general discussion of the electronic properties of amorphous semiconductors). The dangling bond density as measured by electron spin resonance varies from 10^{19} cm^{-3} down to about 10^{15} cm^{-3}. As the defect density decreases, the transport changes from hopping at the Fermi energy to conduction at the band edge. Low defect density material has a room temperature resistivity $\sim 10^{12}\,\Omega$ cm and is close to being intrinsic. Electrons are the more mobile carrier by two to three orders of magnitude at room temperature, and so dominate the conduction processes. Transport of both electrons and holes is believed to be shallow trap-limited with an intrinsic mobility of about 10 cm^2/V s. Doping both n-type and p-type can be achieved by adding the appropriate gases, usually PH_3 or B_2H_6, to the deposition system. The conductivity is largest for n-type material and reaches $10^{-1}\,\Omega^{-1}$ cm^{-1} with an activation energy of $\sim 0{\cdot}15$ eV, indicating that the Fermi energy has moved up to the band tails.

The band gap of a-Si : H is 1·6–1·8 eV and is also sensitive to the hydrogen content and the details of the deposition. The absorption edge is exponential, in common with most amorphous semiconductors, but the origin of the exponential edges has never been satisfactorily resolved. A schematic energy level diagram for an

Fig. 2



Schematic density of states diagram for an amorphous semiconductor showing the extended states separated from the localized band tail states by the mobility edge and defect states in the gap.

amorphous semiconductor is shown in fig. 2. The main features are the conduction and valence bands of delocalized states, the localized band tails arising from disorder in the ideal network, and defect states. Each of these features will be seen to influence the luminescence properties of the material.

§2. Undoped a-Si : H

2.1. *A review of published data*

Luminescence spectra of hydrogenated amorphous silicon have been reported by about 15 groups. The energy of the various peaks observed are given in table 1 which summarizes a selection of the measurements. At first sight the results seem confusing because peak energies are scattered in the range 0·7–1·5 eV without any obvious connection. However, there are various consistent features. There is almost universal agreement that in samples with a high luminescence efficiency, usually

Table 1. Some examples of the observations of luminescence in a-Si : H. The luminescence energies include our identification of the peaks as discussed in the text. The material deposition system is noted (in the case of glow discharge, it is either capacitative or inductive) as is the temperature of the measurement.

Group	Material type	Luminescence peaks	Temperature
Marburg (F.R.G.)	Glow discharge (ind.)	0·93, 1·13, 1·23 (A)	77
		0·7 (B), 1·3 (A)	77
Xerox (U.S.A.)	Glow discharge (cap.)	1·1 (C), 1·3–1·4 (A)	10
		0·9 (B)	10
Sheffield (U.K.)	Dundee (G.D.-ind.)	1·3 (A), 0·9 (B)	10
	Sputtered	1·5 (A)	10
RCA (U.S.A.)	Glow discharge	1·27 (A)	10
	Rehydrogenation	1·1–1·3 (A)	10
IBM (U.S.A.)	Sputtered on crystal Si	1·15 (A)	10
Tokyo (Japan)	Glow discharge	0·89, 1·10, 1·25 (A) 1·44	77
		1·33, 1·43	4·2
		1·2 (A)	2
Harvard (U.S.A.)	Sputtered	0·95 (B), 1·3 (A)	77
		1·25–1·35 (A)	77
Bell (U.S.A.)	Glow discharge	1·35 (A)	2
	Glow discharge on crystal Si	1·15 (A)	2
		2·3 (D)	
Brookhaven (U.S.A.)	Glow discharge	1·3 (A), 1·05	50
Lausanne (Switzerland)	Dundee (G.D.-ind.)	1·28 (A), 1·38	77

deposited at a substrate temperature between 200 and 400°C, there is a single broad luminescence band. The peak energy at low temperature (~ 10 K) is usually between 1·25 and 1·4 eV and varies over this range as the deposition conditions are changed. The band is featureless with a width of about 0·3 eV. In table 1 we indicate (label A) those cases where this peak can be identified. The variation in peak position is not surprising since it is known that the band gap is sensitive to the hydrogen concentration, and it is also reasonable to expect that the band tail widths, for example, depend on the details of the sample structure. By far the majority of experiments have been concerned with this luminescence peak, and its properties are therefore the dominant topic of this review.

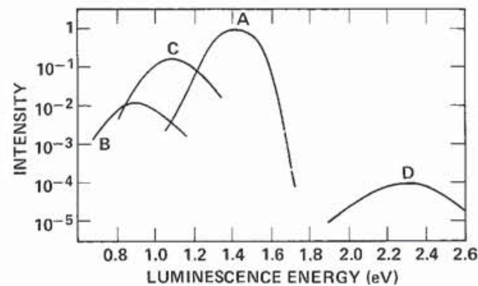
The second most readily identifiable and well characterized transition is one at 0·8–0·9 eV (labelled B in table 1). This is most easily observed in a-Si : H doped with either boron or phosphorus, but is also seen in some undoped material. It is usually associated with a high defect density and is observed best when peak A has a weak intensity. The efficiency of peak B is never high, and has not been observed to exceed $\sim 1\%$ of the maximum efficiency of peak A. Apart from its low energy, a distinguishing feature of peak B is its relatively weak temperature dependence. Typically the luminescence intensity drops by a factor 10–100 at room temperature, compared with a drop of $\geqslant 10^4$ for peak A. It is worth emphasizing that the different temperature dependence is important for the identification of the transition. It is often difficult to identify one of the transitions in a-Si : H from the luminescence energy alone because of the large variations in peak position that can occur, depending on the preparation method.

Another luminescence transition (C) near 1·1 eV has been identified as associated with oxygen impurities. This is, in fact, the only band so far observed that can be attributed to impurities. The oxygen can either be introduced during deposition, or by post-deposition oxidation. This transition is also most easily distinguished by its thermal quenching which occurs at a lower temperature than peak A. Because the luminescence has only been observed as a low energy shoulder to peak A, it has only been measured in a restricted temperature range ~ 10–100 K. The transition has also not been widely reported.

Finally a luminescence peak at higher energy (labelled D in table 1) has been found. There is a very weak transition at $\sim 2{\cdot}3$ eV with an intensity about 10^{-4} of peak A. The most striking feature of the luminescence is that it occurs above the band gap energy, and so it has been associated with thermalizing carriers. The recombination has a short decay time, which distinguishes it from the other transitions, and in addition the temperature dependence is very weak, with the intensity only dropping by about a factor of 3 at room temperature. Recently, it has been suggested that this luminescence band in fact originates from surface contamination and is not an intrinsic property of a-Si : H.

These four peaks are the only transitions that are well enough characterized to be identifiable. The approximate spectra of each are shown in fig. 3. In fact these bands account for most of the reported observations, and some of the exceptions can be explained away, as discussed below. Little can be said of the remaining reported luminescence peaks except that they require further confirmation and characterization. It is perhaps pertinent to point out that in the measurement of broad luminescence bands it is very easy for the spectral dependence of the detection system to introduce spurious features into the luminescence spectrum. Care is needed to normalize the data properly.

Fig. 3

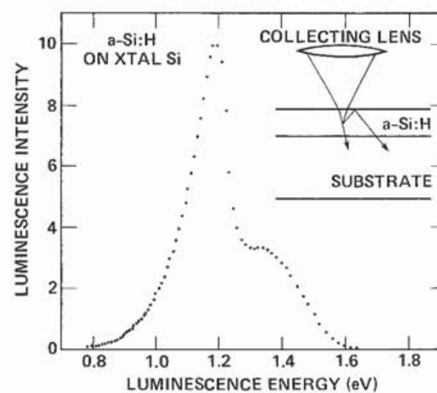


Composite luminescence spectra of a-Si : H showing the four well characterized peaks discussed in the text.

There is one case in which a different spectrum is observed, which is worth discussing in some detail. Figure 4 shows a highly asymmetrical spectrum with a peak at 1·15 eV and a shoulder at higher energy. This spectrum has been reported on three different occasions, and in each case from a-Si : H deposited on a crystalline silicon substrate (e.g. Brodsky *et al.* 1977). In fact the spectrum is the usual peak A except that the intensity decreases abruptly above the band gap of the crystalline silicon substrate. The substrate has a large effect even though the luminescence is observed in back-scattering from the illuminated surface.

The explanation of this surprising effect is shown in fig. 4. The luminescence is emitted isotropically inside the sample. Because of the high refractive index (~ 3) only a cone of half angle $\theta \sim 20°$ is emitted from the surface. This corresponds to about 5% of the luminescence. Most of the remainder is totally internally reflected. If the substrate is absorbing, then all of this light is immediately lost. However, if the substrate is transmitting, and the substrate holder has a reflecting surface, then all this light will eventually be emitted from the sample. Thus the absorbing substrate

Fig. 4



Luminescence spectrum of a-Si : H deposited on crystalline Si showing the effect of the substrate band gap. The insert is a schematic diagram of the emission.

can potentially reduce the luminescence intensity by a factor of 20. In practice the decrease is about a factor of 5, the remainder presumably either being absorbed in the substrate holder, or channelled to the perimeter of the sample. This effect shows that care must be taken in the choice of a substrate to avoid such spurious results.

Some similar considerations also relate to the deposition of samples on ground glass substrates. This procedure is routinely applied to get rid of interference effects in the spectrum. However, roughened substrates also give a higher detected luminescence intensity. We believe that the reason is that ground surface mitigates the effects of total internal reflection, so that a larger fraction of the luminescence is emitted directly from the sample without reflections. In practice the gain of intensity is about a factor of 3.

2.2. *Phonon interactions*

It is now generally accepted that the photoluminescence of chalcogenide glasses has a large Stokes shift. The Stokes shift arises from a strong electron–phonon interaction such that the atomic environment of the recombination centre is distorted when the electron occupancy changes. Under these circumstances the energy released by an electronic transition is given up as a photon in conjunction with several phonons. In the ideal case, the luminescence spectrum consists of a zero-phonon line at the electronic energy of the transition, and a series of side bands at lower energy, separated from the zero-phonon line by integral units of the phonon energy. In practice, for moderately strong coupling, the side bands merge to form a Gaussian band, and the zero-phonon line is often too weak to be observed. Its position, however, can be determined from the spectrum if the phonon energies are known. It is important to determine the strength of the phonon interaction in any luminescence band in order to identify the electronic energy of the transition. Generally if the phonon side bands cannot be observed explicitly in the spectrum, then the best test for phonon coupling is to observe the absorption process which is the inverse of the luminescence transition. This absorption is also broadened by the phonon coupling, but the side bands occur at the high energy side of the zero-phonon energy.

Phonon interactions are observed in virtually all luminescence processes. The strength of the effect often depends on the degree of localization of the states involved because the distortion of a centre increases with the electron density at the centre. For example in crystalline GaAs, bound excitations have very weak coupling because the binding energy is only ~ 5 meV, and the wavefunction has a large radius. However, transitions involving vacancy-complexes have a zero-phonon binding energy of $\sim 0{\cdot}1$ eV and much stronger phonon coupling, giving luminescence line widths of about 0·15–0·2 eV (see fig. 1).

In a-Si : H experiments have been reported which claim to identify the electron–phonon interaction (Street 1978). These results have been criticized by Paesler and Paul (1980), and, in a later paper, Collins *et al.* (1980 a) give arguments to show that their data can be explained without appealing to a Stokes shift. This author believes that the criticism is not valid. In view of the importance of the issue it seems valuable to discuss the arguments in some detail.

The initial experiments were based on the property that phonon coupling shifts the luminescence and absorption bands respectively to lower and higher energy from the zero-phonon energy. The absorption and luminescence of the same states can be

related quantitatively by the detailed balance relation between the radiative lifetime τ and the absorption coefficient α,

$$\tau^{-1} = (8\pi n^2/\lambda^2)(g_1/g_2)N^{-1}\int \alpha \, d\nu, \tag{1}$$

where n is the refractive index, λ is the wavelength, g_1 and g_2 are the degeneracies of the ground and excited states, N is the density of states and ν is the frequency (Nelson *et al.* 1966). The objective is to calculate the absorption coefficient that corresponds to luminescence. The absorption is then compared with the measured absorption. If the absorption at the luminescence energy is at least as large as the calculated value, then no Stokes shift is required. On the other hand if it is substantially less than the calculation, the only explanation is that the absorption lies at some higher energy due to the phonon coupling.

An immediate problem is that we know neither N nor τ in eqn. (1). However, the answer can be obtained without an explicit knowledge of these quantities. We assume that the sample is optically excited at a rate G absorbed photons/cm^3, and also that this excitation rate results in N_{occ} occupied recombination centres where $N_{occ} < N$. The rate of photon emission due to radiative recombination is given on the one hand by N_{occ}/τ, and on the other hand by $y_L G$ where y_L is, by definition, the internal quantum efficiency. Thus

$$y_L G = N_{occ}/\tau. \tag{2}$$

This expression has been criticized by Paesler and Paul (1980) who argue that the density of states refers to that at the excitation energy, which is typically well up into the bands, and not to the recombination centres. The above discussion of eqn. (2) is intended to show that this is not the case. Eqn. (2) is obtained by considering only the recombination centres. It is not necessary to known what the excitation energy is, or any details of the thermalization process by which carriers reach the recombination centres. In the original discussion (Street 1978), the y_L was omitted from eqn. (2) since the material was assumed to have unity quantum efficiency. Paesler and Paul (1980) in fact agree that the derivation is correct in that case. Here we shall complete the argument including the quantum efficiency explicitly to show how it affects the final result.

Since N_{occ} cannot exceed N,

$$y_L G \leqslant N/\tau, \tag{3}$$

and so from eqn. (1),

$$y_L G \leqslant (8\pi n^2/\lambda^2)(g_1/g_2)\int \alpha \, d\nu. \tag{4}$$

This is now an inequality that relates the absorption coefficient directly to measurable quantities. The next step is to obtain the maximum value of $Y_L G$ which is done by increasing G and observing the luminescence intensity. At some value of G we expect the luminescence to saturate. Some non-radiative recombination process takes over and any further increase of G is offset by a decrease in Y_L so that $Y_L G$ does not increase. This behaviour was observed (Street 1978) and from this result a lower

limit on the value of the absorption corresponding to the luminescence transition was obtained by using the luminescence line shape for the integral in eqn. (4). The final result gives

$$\alpha_{max} \geqslant 3 \times 10^2 y_L \text{ cm}^{-1} \tag{5}$$

where α_{max} is the value calculated at the peak of the luminescence spectrum.

We now come to the second part of the problem which is to measure the actual absorption coefficient in order to compare it with the calculated value. There are various complications associated with this experiment. In general it is not possible to make samples sufficiently thick for direct transmission measurements of the absorption, and so other techniques have to be used. A potentially more difficult problem is that it is necessary to be sure that the absorption that is measured in fact is related to the luminescence. It is quite possible that there is an extrinsic absorption process present which is totally unrelated to the luminescence transition, in which case a comparison with the calculation would be meaningless.

In the initial experiment the luminescence excitation spectrum was used as the technique to obtain the absorption. This experiment has the major advantage in that it immediately discriminates against the unwanted unrelated absorption processes which do not give luminescence. The excitation spectrum measures the luminescence intensity I as a function of the excitation photon energy,

$$I(\hbar\omega) = I_0 y_L (1 - \exp(-\alpha d)). \tag{6}$$

I_0 is a constant (generally unknown) related to the geometry of the detection system and sample parameters such as reflectivity. In two limiting cases eqn. (6) reduces to

$$I/I_0 = y_L, \qquad \alpha d \gg 1. \tag{7}$$

$$I/I_0 = y_L \alpha d, \quad \alpha d \ll 1. \tag{8}$$

The excitation spectrum is shown in fig. 5, and both regions can be easily identified. Since the thickness d is known, α can be readily deduced and is shown in the ordinate of fig. 5. Recent data has confirmed these measurements, and extended them to even lower energy (Bishop *et al.* 1981).

Obtained in this way the absorption coefficient is measured to be less than 10^{-1} cm^{-3} at the luminescence peak. To bring this value into agreement with eqn. (5) requires that y_L be no greater than 3×10^{-4}. However, most estimates find that in the type of samples used in this experiment the efficiency is at least 10^{-1} and possibly approaches unity. The enormous discrepancy between the measured and calculated values led to the deduction of a Stokes shift.

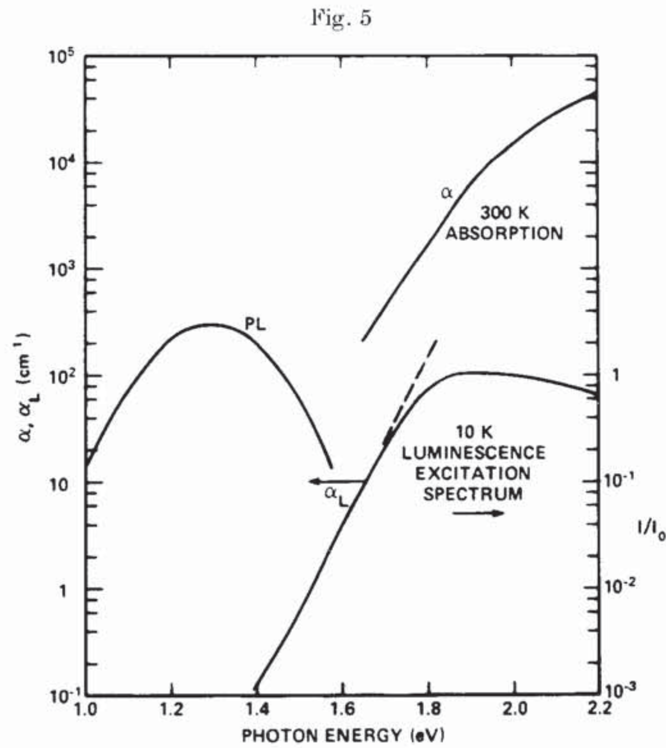
An alternative method of measuring the absorption coefficient at low energy is from photoconductivity, which is given by

$$\sigma_{ph} = G\tau e\mu(1 - \exp(-\alpha d)). \tag{9}$$

where G is the incident photon flux, and a reflectivity correction has been ignored. α can be obtained in a similar way as for the luminescence excitation spectrum. A complication of this method is that the $\mu\tau$ product can vary with energy. One solution to this problem is to use the distinction between primary and secondary photoconductivity (Crandall 1980 b). A second solution is to correct explicitly for the changes in $\mu\tau$ (Moddel *et al.* 1980). Both methods have been used.

Several groups have used photoconductivity to evaluate α, although not all have taken care to correct for the changes in $\mu\tau$. In all cases an exponential absorption

Fig. 5



The luminescence excitation spectrum of a-Si : H showing the derived absorption coefficient α_L compared to the room temperature optical absorption. The luminescence spectrum is adjusted to the absorption strength predicted by detailed balance (Street 1978).

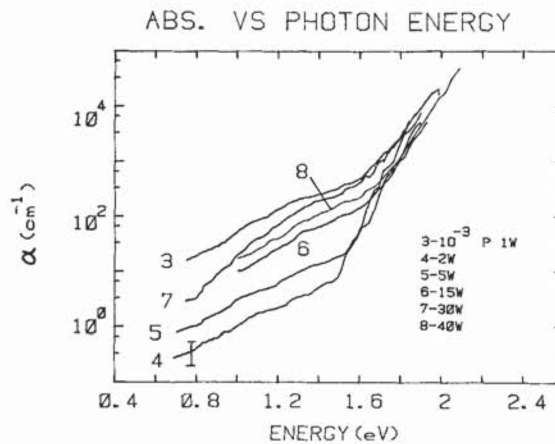
edge is observed above $\sim 1{\cdot}6$ eV which is clearly a continuation of the fundamental absorption edge observed in standard transmission measurements at higher energy. Below 1·6 eV a broad absorption band is often observed which has the appearance of an extrinsic band. The absorption coefficient at the peak of this band ($\sim 1{\cdot}3$ eV) varies from 50 cm^{-1} down to 0·5 cm^{-1} and below. Both Pankove *et al.* (1980) and Collins *et al.* (1980 a) point out the similarity in energy and shape of the extrinsic absorption band and the luminescence peak. This comparison leads them to identify the two spectra as arising from the same transition, and so they deduce that the Stokes shift is absent.

We believe that the bulk of the evidence indicates that this conclusion is incorrect. Our explanation of the extrinsic band is that it is due to defects (probably dangling bonds) and that it is not related to the 1·3–1·4 eV luminescence. The evidence is the following.

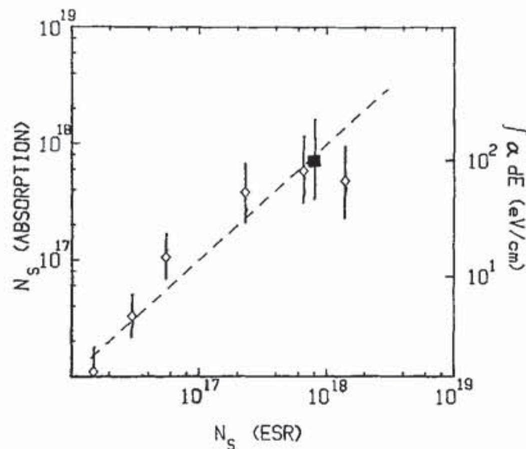
(1) There is substantial evidence that the strength of the extrinsic band correlates with the defect density but is inversely related to the luminescence efficiency. For example, there are several examples in which material with high solar cell efficiency has very weak extrinsic absorption (see, for example, Cody *et al.* 1980; and see fig. 6). This material is known to have a low defect density and high luminescence efficiency. Annealing studies of ion bombarded a-Si : H find that the extrinsic absorption decreases in strength as the

defects are annealed away (Stuke 1976), and a similar reduction with annealing is reported on sputtered material (von Roedern and Moddel 1980). Recently these trends have been confirmed by measurements using photothermal deflection spectroscopy (Jackson and Amer 1981). This technique allows a direct measure of the absorption coefficients down to about $10^{-1}\,cm^{-1}$. Figure 6 shows examples of the extrinsic absorption band in different samples and also shows that the integrated absorption strength is proportional to the measured spin density of the films. The proportionality constant also results in a reasonable value of the oscillator strength. These results are the clearest evidence that the extrinsic band is derived from transitions involving dangling bonds defect states.

Fig. 6



(a)



(b)

(a) The absorption tail as measured by photo-thermal deflection spectroscopy on various samples. (b) Plot of the strength of the absorption tail versus spin density (Jackson and Amer 1981).

(2) There is no evidence for the extrinsic band in the luminescence excitation spectrum. This experiment is only sensitive to absorption processes that yield luminescence. The inference is that if the extrinsic band is present, its absorption is unrelated to the 1·4 eV luminescence process.
(3) In heavily doped material there is a strong extrinsic absorption band (Knights 1976, Freeman and Paul 1979, Moustakas 1980). Such material has a large defect density, and the defect related luminescence band at $\sim$0·9 eV is observed. Excitation into the extrinsic band excites the 0·9 eV luminescence, but not the 1·3–1·4 eV peak (see § 2·5).

Based on the above discussion we believe that the evidence for a Stokes shift is extremely strong, and we shall adopt this result to discuss the consequences. An estimate of the size of the shift is obtained from fig. 5. About 0·4–0·5 eV separates the peak of the luminescence from an absorption of the magnitude calculated by detailed balance. The distortion energy of the recombination is approximately half this value or 0·2–0·25 eV. This in turn puts the zero-phonon energy at 1·5–1·6 eV. The band gap of a-Si : H is hard to determine because of the broad absorption edge, but room temperature values of 1·6–1·9 eV are often quoted, depending on how the gap is defined.

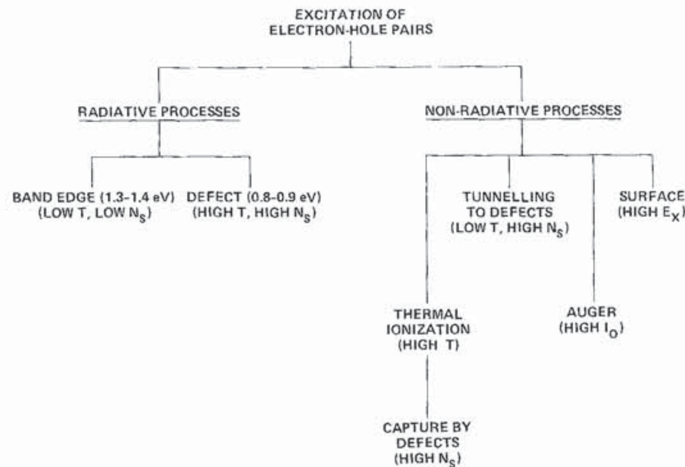
It is also worth pointing out some implications of the luminescence excitation spectrum. The observation that the excitation spectrum is apparently indistinguishable from the fundamental absorption edge, implies that the recombination centres are part of the conduction and valence bands. In fact we believe that the centres are band tail states as discussed in more detail in § 2.3.6. A more interesting implication concerns the origin of the exponential edge. Our results indicate that, at least up to an absorption coefficient of $\sim 3 \times 10^2\ \mathrm{cm}^{-1}$, the shape of the exponential edge is determined by the electron–phonon interaction. If this is so, then it is not possible to obtain the density of states directly from the shape of the absorption, as has been suggested by Crandall (1980 b). In addition if the phonon coupling effects remain important at higher absorption coefficients, the definition of the band gap energy becomes even more unclear.

2.3. *Recombination mechanisms*

In this section the development of a detailed recombination model for the 1·3–1·4 eV peak is reviewed. The model will then be used as the basis for the interpretation of other experiments. For the present we shall assume that the transition involves localized states at the conduction and valence band edges. It turns out to be convenient to describe the mechanisms before considering in more detail the microscopic origin of the states involved which is left to § 2.3.6. The reason is that the identification of those states depends largely on a good understanding of the recombination.

As usual in luminescence systems, there are several recombination mechanisms, both radiative and non-radiative, and the dominant mechanism depends both on material properties (such as defect density) and on the experimental conditions (e.g., excitation intensity, temperature, etc.). Figure 7 shows a chart of the recombination mechanisms that have been identified, and the experimental conditions under which each one prevails. There are four sources of non-radiative recombination—defects, thermal ionization, Auger, and surface recombination. The last two are relatively minor effects which can be avoided by using low intensity and weakly absorbed

Fig. 7



Schematic diagram showing the radiative and competing non-radiative processes, and the experimental conditions under which each dominates.

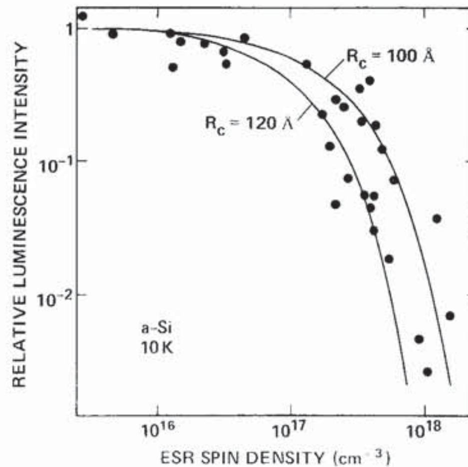
illumination and are discussed later. The first two are summarized here and then discussed in more detail below.

Figure 8 shows how the luminescence efficiency varies with E.S.R. spin density (Street *et al.* 1978 b). This data refers to the resonance at $g = 2{\cdot}0055$ which is commonly observed in amorphous silicon, and whose origin is now clearly established as a silicon dangling bond (Biegelsen 1980). The samples used for the data in fig. 8 are all glow discharge and cover a very wide range of deposition conditions. The luminescence intensity decreases very rapidly as the spin density N_s exceeds $\sim 10^{17}\,\mathrm{cm}^{-3}$. This data demonstrates that dangling bonds are non-radiative centres, and also explains why no luminescence is observed in non-hydrogenated a-Si, since N_s is typically $\sim 10^{19}\,\mathrm{cm}^{-3}$ in this material (Thomas *et al.* 1978). The fact that the luminescence has its maximum intensity in material with the lowest defect density implies a process which is in some way intrinsic, which is evidence for associating the transition with band tail states.

It is widely observed that the luminescence is quenched at temperatures above about 50 K (Engemann and Fischer 1973, Nashashibi *et al.* 1977 a, Tsang and Street 1979, Paesler and Paul 1980), and this effect is universally attributed to the ionization of carrier pairs. The thermal energy allows the carriers to diffuse apart and move through the material until some non-radiative event takes place. This is by far the dominant thermal quenching mechanism in semiconductor luminescence. The activation energy of the process will in general yield the binding energy of the carrier. Unfortunately the published data report values ranging from 0·07 eV (Paesler and Paul 1980) up to 0·23 eV (Tsang and Street 1979) so that no well defined values can be given. Nevertheless the small activation energy again supports the notion that band tail states are involved.

This outlines the two major non-radiative mechanisms. We now begin a more detailed discussion of the recombination, starting with the radiative process as observed by decay measurements.

Fig. 8



Plot of luminescence intensity versus spin density for a range of glow discharge deposited samples. Solid curves are a fit to eqn. (18) (Street *et al.* 1978 b).

2.3.1. *Low temperature decay*

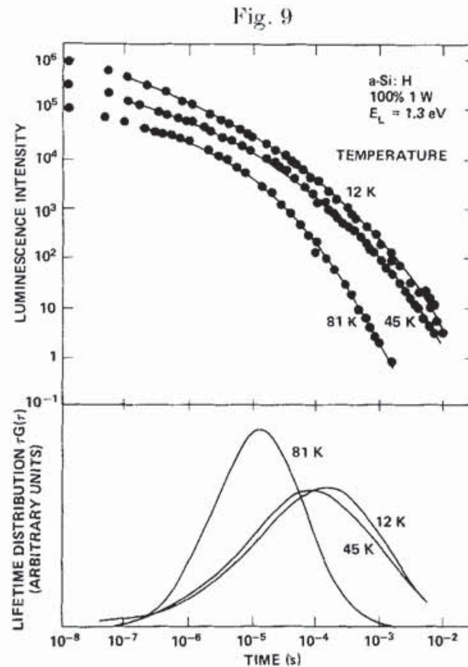
The initial measurements reported luminescence decay times of 10^{-8} s (Engemann and Fischer 1976). Subsequent studies found 3×10^{-5} s (Tsang and Street 1978) and more recently $\sim 10^{-3}$ s (Tsang and Street 1979). It was then recognized that these disparate results arise because the decay times in fact have a very broad distribution in time, and under these circumstances the apparent decay depends critically on the excitation pulse length. The complete distribution $G(\tau)$ of decay times can be found by using a range of excitation pulse lengths and an example is shown in fig. 9. The data is from a sample with low spin density ($\sim 3 \times 10^{15}$ cm^{-3}), and therefore the non-radiative processes are minimized. Here $\tau G(\tau)$ is plotted to account for the logarithmic time axis since

$$\tau G(\tau)\, d(\ln \tau) \equiv G(\tau)\, d\tau. \tag{10}$$

Even this data probably underestimates the decay times because of the excitation-intensity effects described in § 2.3.4. The mean decay time corresponds approximately to the peak in $\tau G(\tau)$, and can also be obtained by observing the decay after a very long excitation pulse at the lowest measurable intensity (see fig. 17). From this and other data (see, for example, Noolandi *et al.* 1980 a, b), the mean decay time in the low excitation limit for samples of high efficiency is found to be in the range 1–10 ms. No other groups have reported measurements to such long decay times. However, at shorter times the various reports find essentially identical data (Austin *et al.* 1979 a, Kurita *et al.* 1979, Shah and Di Giovanni 1981), indicating that there is a distribution $G(\tau)$ common to all.

The most striking feature of the decay is that the mean lifetime of $\sim 10^{-3}$ s is very long compared with the time of 10^{-8} s associated with an allowed transition of unity oscillator strength. There are several possible reasons for the long decay times; for example, the transition may be forbidden. The only possibility here is from the spin selection rules since orbital symmetry of the wavefunctions in a disordered system

Fig. 9



Examples of the luminescence decay, and the derived distribution of decay times (Tsang and Street 1979).

seems out of the question. However, spin dependent luminescence and related experiments discussed in §2.4.4 find that although the spin states do influence the recombination, the magnitude of the effect is much too small to explain the long decay times.

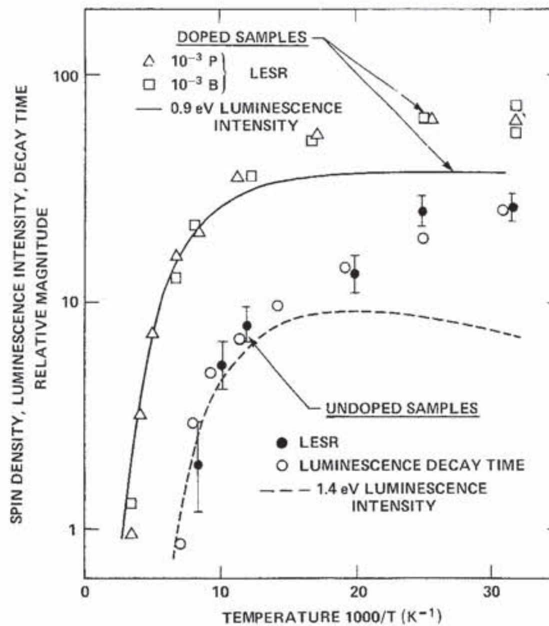
Kurita *et al.* (1979) have suggested that carriers may be trapped at sufficiently isolated sites so that the decay time is determined by the thermal release time from the trap rather than the radiative recombination time. This situation certainly occurs in many luminescence systems and is characterized by an activated decay time. However, fig. 10 shows the temperature dependence of the average decay time which increases only weakly as the temperature is reduced, and below 20 K there is no discernible change. This data indicates that although thermal activation may play a role above ~30 K, radiative recombination occurs even when the thermal energy tends to zero. The low temperature decay time is therefore deduced to be that of the radiative process.

Two other recombination models have been considered. The 'exciton' model (Engemann and Fischer 1976, 1977, Kivelson and Gelatt 1981) assumes that one carrier, for example the hole, occupies a localized band tail state. The electron is then bound to the hole by the Coulomb interaction in a state whose wavefunction is centred at the hole and made up predominately of conduction band extended states. This is in effect the equivalent of a bound exciton in a crystal. Kivelson and Gelatt (1981) have calculated the radiative lifetime τ_x of such an exciton, but it is convenient to use an earlier rough estimate by Mott (1977),

$$\tau_x = \tau_0 (R_h/R_e)^3. \tag{11}$$

R_h and R_e are the Bohr radii of the hole and electron, and $\tau_0 \sim 10^{-8}$ s.

Fig. 10



Temperature dependence of the average decay time showing only a weakly activated behaviour at low temperatures. The data are compared to the luminescence intensity of the band edge and defect transitions (Street and Biegelsen 1980 b).

The alternative 'pair' model (Tsang and Street 1979) is also taken directly from a known mechanism in crystalline semiconductors, that of radiative tunnelling, most commonly observed between donor–acceptor pairs (Thomas *et al.* 1965). In our case it is assumed that the electron and hole are localized in states separated by a distance R. The decay time τ_p is then given approximately by

$$\tau_p = \tau_0 \exp(2R/R_0), \tag{12}$$

where R_0 is the effective Bohr radius—whichever is larger of R_e and R_h.

There is no reason to doubt that excitons and pairs of the type described can exist in a-Si : H, or indeed in any other amorphous semiconductor. One expects the luminescence to be dominated by whichever is the more stable configuration, which is determined by the relative binding energies. Because the pair consists of a distant electron and hole, the Coulomb interaction will be weaker than for the exciton. On the other hand, the binding energy of the pair is increased relative to the exciton because both carriers are in localized states. Such considerations led Mott (1977) to the criterion that the stability of the two alternatives is determined by whether the Coulomb interaction of the exciton is larger or smaller than the band tail width. Unfortunately, the best estimates of both quantities are comparable (100–200 meV), so that it is not immediately obvious which configuration dominates. However, a decay time of 10^{-3} s requires an exciton with $R_e/R_h \sim 60$. We believe that this is well in excess of what is physically possible. On the other hand the pair model requires $R/R_0 \sim 6$ which is easily obtained. Long decay times, even exceeding 1 s are found for radiative tunnelling in crystalline semiconductors (Thomas *et al.* 1965). On the other

hand, we are not aware of any allowed excitonic transition with a lifetime approaching 10^{-3} s. We therefore believe that the decay data strongly established the pair model. In addition this model readily accounts for the distribution of decay times. From eqn. (12) it is seen that $\ln \tau \sim R/R_0$, and so the distribution shown in fig. 9 is simply that of R/R_0. For reasons described further in § 2.3.4, it is believed that after optical excitation the electron and hole diffuse a short distance apart and then recombine in a geminate process. Diffusion in three dimensions always leads to a broad distribution in R centred away from the origin, in qualitative agreement with the observed distribution.

Further specific evidence for the pair model comes from the time-resolved spectra described in § 2.4.1. It is also worth emphasizing that the pair model is a natural consequence of the presence of localized states at both band edges. A central feature of the model is that at low temperatures the carriers are rapidly trapped and become immobile. We shall see how the recombination mechanisms are modified at higher temperatures when carriers become more mobile.

2.3.2. *The role of dangling bond defects*

In this section we describe how dangling bonds act as non-radiative recombination centres. It is worth noting that the common model of carrier diffusion and capture by a trap predicts that the non-radiative recombination rate is inversely proportional to the trap density. Figure 8 shows that this relation is not obeyed. In fact one should not expect this model to work because the recombination mechanism described in the previous section involves carriers which are immobile after thermalization and can only recombine by tunnelling. For this reason, it is appropriate to apply the tunnelling model to the non-radiative recombination. The situation differs from that of the radiative process in two ways. Firstly, the position of the carrier is assumed to be uncorrelated with that of the defects. Secondly, the mechanism is assumed to be phonon (rather than photon) assisted tunnelling.

The standard expression for the non-radiative tunnelling rate P is

$$P = \nu_0 \exp(-2R/R_0), \tag{13}$$

where ν_0 is $\sim 10^{12}\,\mathrm{s}^{-1}$ and R is the distance to the defect state. From this expression, the luminescence efficiency can be calculated (Street *et al.* 1978 b, Tsang and Street 1978). For simplicity a single average value for the radiative decay time τ_R is assumed. The luminescence efficiency is then given by

$$y_L = \langle \tau_R^{-1}/(P + \tau_R^{-1}) \rangle, \tag{14}$$

where the average is over all possible configurations. The average is calculated assuming a random distribution of defects and electron–hole pairs so that the distribution $G(R)$ of nearest neighbour distances of a defect to a particular pair is (Reiss 1956)

$$G(R) = 4\pi R^2 N \exp(-\tfrac{4}{3}\pi R^3 N), \tag{15}$$

where N is the defect density. Next a critical transfer radius R_c is defined such that

$$\tau_R^{-1} = \nu_0 \exp(-2R_c/R_0) \tag{16}$$

or

$$R_c = R_0/2 \ln(\nu_0 \tau_R). \tag{17}$$

R_c is therefore the separation at which the radiative and non-radiative rates are equal. It is easy to see that because of the exponential dependence of P on R in eqn. (13), there is only a small range of R at which the two rates are comparable. Thus a reasonable approximation is that y_L is unity at $R > R_c$ and zero for $R < R_c$ so that

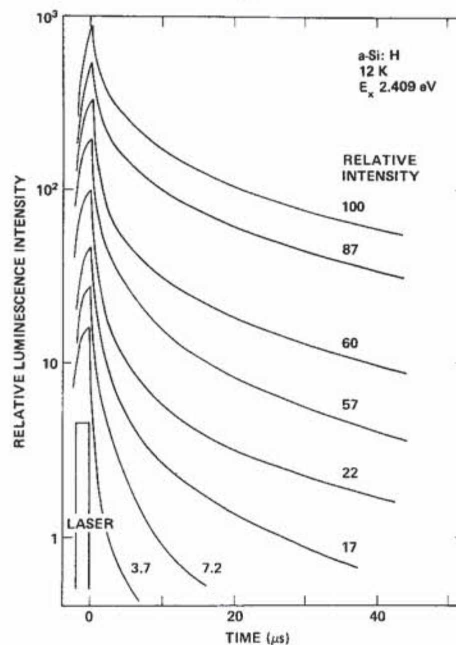
$$y_L = \int_{R_c}^{\infty} G(R) = \exp\left(-\tfrac{4}{3}\pi R_c^3 N\right). \tag{18}$$

Figure 8 shows that this simple expression reproduces the form of the data. A value of $R_c \sim 110$ Å is obtained and from that $R_0 \sim 12$ Å. (Street *et al.* 1978b).

There is an important distinction between this model and one involving diffusion and capture. In the tunnelling model some electron–hole pairs recombine non-radiatively and rapidly whilst others recombine radiatively. This is determined by whether the nearest defect is closer or further away than R_c. As the density of defects increases the fraction of pairs that recombine radiatively decreases. Thus, one expects to see fast and slow components separately in the luminescence decay. However, if diffusion occurs one expects to observe a single decay time that decreases as the non-radiative component becomes stronger. The reason for the difference is that in the first model there is a spatial average and in the second an average over time.

Figure 11 shows examples of the luminescence decay in a-Si : H samples with various luminescence intensities corresponding to different spin densities. As the

Fig. 11



Luminescence decay in samples of varying quantum efficiency, corresponding to different spin density (Tsang and Street 1979).