

IX

INTERACTION OF LIGHT WITH ELECTRONS IN SOLIDS

9.1. General outline. Classical theory

IN this chapter we return to the action of light waves, which was also discussed in Chapter III. There we confined ourselves to the case in which, both before and after the transition, the electronic system was in its ground state. We must now discuss the more general case in which there are electronic transitions. This includes all optical problems in metals as well as the photoelectric effect. It also includes the light absorption and emission in all solids at frequencies comparable with atomic absorption lines. The exceptional cases of coherent scattering and vibrational Raman effect have already been treated in Chapter III.

Thus in metals one important class of processes, dominant at low frequencies, i.e. usually in the infra-red and visible regions, is concerned only with the properties of the ordinary conduction band. There are then still two important regions according to whether the light frequency is smaller or larger than the collision frequency $1/\tau$. When $\omega\tau \ll 1$, the effect of the field is essentially the same as that of a static field, so that the ordinary static conductivity is adequate for describing the behaviour of the metal. This is certainly the case for waves in the radio-frequency region, and in the far infra-red. Since the collision times for ordinary metals at room temperature are in the neighbourhood of $3 \cdot 10^{-14}$ sec., the light wavelength for which the period equals the collision time is of the order of 10^{-3} cm.

In the opposite limit the collisions are negligible and we are dealing essentially with electrons in a perfect lattice.

We can survey the important processes of this class by following the classical theory due to Drude (1904). We assume each electron to have a mean collision time τ defined as in § 6.1. Then the mean component of the electron velocity is given by

$$\frac{\partial \mathbf{v}}{\partial t} = \frac{e}{m} \mathbf{E} - \frac{1}{\tau} \mathbf{v}, \quad (9.1)$$

where \mathbf{E} is the component of the applied field. If the applied field depends on time through a factor

$$e^{i\omega t}$$

(as usual the complex exponential serves mathematical convenience

and the real field is the real part of the final solution) it therefore follows that

$$\mathbf{v} = \frac{e\mathbf{E}}{m\{i\omega + (1/\tau)\}}. \quad (9.2)$$

Hence the current density due to all electrons is

$$\mathbf{j} = \frac{ne^2}{m\{i\omega + (1/\tau)\}} \mathbf{E} = \frac{\sigma_0}{1 + i\omega\tau} \mathbf{E}, \quad (9.3)$$

where σ_0 is the conductivity for constant field (cf. (6.16)). Rationalizing the denominator,

$$\mathbf{j} = \frac{\sigma_0}{1 + \omega^2\tau^2} (1 - i\omega\tau) \mathbf{E} = \frac{\sigma_0}{1 + \omega^2\tau^2} \left(\mathbf{E} - \tau \frac{\partial \mathbf{E}}{\partial t} \right). \quad (9.4)$$

A current which is proportional to the rate of change of the electric intensity, and hence out of phase with the electric field, is physically equivalent, not to a conduction current, but to the displacement current arising from a dielectric constant.

Hence the effective conductivity at frequency ω is

$$\sigma(\omega) = \frac{\sigma_0}{1 + \omega^2\tau^2}, \quad (9.5)$$

and the effective dielectric constant is

$$K(\omega) = 1 - \frac{4\pi\tau\sigma_0}{1 + \omega^2\tau^2}. \quad (9.6)$$

These relations show that, as long as $\omega\tau \ll 1$, the current is almost entirely in phase with the electric vector (as is evident from (9.4)), and therefore for most purposes the dielectric constant is unimportant. When $\omega\tau \gg 1$, most of the current is out of phase with \mathbf{E} , and the properties of an electromagnetic wave travelling through the medium are then mainly influenced by the dielectric constant, which, in this case, is approximately given by

$$K(\omega) = 1 - \frac{4\pi\sigma_0}{\omega^2\tau}.$$

This expression is independent of the collision time, since we introduced σ_0 merely as a convenient abbreviation for $ne^2\tau/m$, so that

$$K(\omega) = 1 - \frac{4\pi ne^2}{m\omega^2}. \quad (9.7)$$

For normal metals this quantity is negative even for frequencies considerably greater than the collision frequency. A negative dielectric constant means an imaginary refractive index, and therefore strongly damped waves.

The dielectric constant changes sign for a critical frequency

$$\omega_1 = \sqrt{\left(\frac{4\pi ne^2}{m}\right)}. \quad (9.8)$$

According to Zener (1933) this result may be linked with the observation that the alkali metals become transparent in the ultra-violet, and the magnitude of the frequency at which this begins is in reasonable agreement with (9.8) if one assumes one electron per atom, and equates m with the mass of a free electron.

In deriving these results we have neglected the effect of transitions between different bands, which is to be discussed later. In addition we have taken free electrons instead of electrons in a periodic field; this is likely to alter the results by numerical factors in cases in which the energy surfaces are not nearly spheres, or in which the effective mass is very different from the free electron mass. The general nature of the results is not likely to be sensitive to this approximation.

A further assumption is, however, the existence of a definite collision time. The result can be extended, without change, to the case of isotropic motion with a collision probability depending arbitrarily on the angle of deflexion. In that case the considerations used at the end of § 6.1 apply, and since the perturbing force has the same angular dependence as that due to a static external field, the appropriate collision time is again τ_1 (cf. (6.21)) and therefore the same as in the ordinary conductivity.

However, when we are dealing with an anisotropic energy surface, and the collisions are not such as to distribute the electrons uniformly over the energy surface after each collision, then the problem takes the form of an integral equation, in which the term arising from the time dependence varies over the energy surface in a different way from that due to the collisions, and therefore the answers depend on the ratio between the two terms, i.e. on $\omega\tau$. It is to be expected that the solutions can no longer be represented in the form (9.5), (9.6) with a constant value of τ .

The problem becomes even more involved at low temperatures, but even there it will still be true that with increasing frequency the current will change from a mainly conductive to a mainly dielectric phase, and the frequencies for which the transition occurs will again depend on the collisions. We saw in § 6.7 that the frequency of collisions was proportional to T^3 , but that the mean angle of deflexion was proportional to T . Since collisions in which the electron is deflected by a small angle are unimportant also in the present problem, the effective collision

frequency for this purpose is again proportional to T^5 , as in the static conductivity. Hence one expects that (9.5), (9.6) should still be qualitatively correct, with the order of magnitude of τ given by the classical formula for the conductivity.

In experimental studies of the optical constants of metals one usually has to work with reflected light since the transmission is too small, except in the frequency region in which the dielectric constant (9.6) is positive but the transitions to higher bands have not yet started to play a part. In the case of reflection, one must be sure that the results are not influenced by the special properties of the surface. Apart from the purity and gross structure of the surface, which raise questions of experimental technique, one must remember that the electrons within a mean free path from the surface will behave differently from those inside the metal, since, in addition to ordinary collisions, they may be deflected by the surface itself. However, this effect is noticeable only over distances from the surface of the order of the distance an electron travels during the collision time (mean free path) or during a period of the light wave, whichever is the smaller. Since the electron velocity is very small compared to the light velocity, this distance is therefore much less than a wavelength, and it will in general be unimportant.

In using equation (9.2), it has been assumed that the field acting on each electron may be taken as the space average, which also appears in Maxwell's equations, and this is correct only if there is no correlation between the position of one electron and that of other electrons that contribute to the field acting on it. It is well known that in other circumstances, e.g. in calculating the dielectric constant of a non-conducting solid, in which each atom is in a highly symmetric position in relation to its neighbours, these correlations are important and give rise to the Lorentz-Lorenz factor.

If we are treating free electrons as strictly independent, there are, of course, no correlations. On the other hand, we know that both the Pauli principle and the electrostatic repulsion between the electrons tend to keep the electrons apart, so that each electron is, on the average, surrounded by a spherical region in which the chance of finding another electron is less than its mean over all space. It would appear plausible, therefore, that some part of the Lorentz-Lorenz correction should apply. On the other hand, Wilson (1936) concludes that no correction is necessary.

The existence of the periodic potential would seem to increase the correlations, and hence the case for a correction; certainly in the case of

strong binding, each electron is likely to be found near some atom, and any other electrons that might be nearby are most likely to be found in the neighbouring atoms, and are very unlikely to occupy the same atom.

9.2. Transitions between bands

In transitions of an electron from one band to another, we have again to apply the by now familiar conservation of the wave vector. Strictly speaking, this should involve an equation between the initial and final wave vectors \mathbf{k} and \mathbf{k}' of the electron and that of the photon, but, unless we are dealing with X-rays, the photon wave vector is negligibly small, and it is a sufficient approximation to assume $\mathbf{k}' = \mathbf{k}$. Thus an electron in a given state \mathbf{k} in the conduction band may make optical transitions only to a state with the same value of \mathbf{k} in another band. However, the absorption frequency depends on \mathbf{k} ; it is given by

$$\hbar\omega = E_{l'}(\mathbf{k}) - E_l(\mathbf{k}), \quad (9.9)$$

where l and l' refer to the two bands. The variation of energy with \mathbf{k} differs in different bands; for example, one band may have its lowest energy at $\mathbf{k} = 0$, and another its highest. There will, however, be a minimum value of (9.9), and therefore a minimum frequency below which this process cannot take place. However, this frequency is not easily interpreted since it depends on the nature of the energy functions for both bands.

Experimentally this absorption edge should be more pronounced if its frequency is higher than the frequency (9.8) at which the one-band part of the dielectric constant ceases to be negative, as in the case of the alkalis.

Formulae for the values of K and σ including the inter-band transitions can be found in the literature; since, however, little use is made of these results in practice, we shall not derive them here.

If we increase the light frequency, we shall be concerned with transitions of electrons from the conduction band to higher bands, and also with the transitions from the lower energy states into the conduction band or just above it. The first of these cases is not interesting, since it will show very little structure, because in this case the energy change of the electron is again given by (9.9) with both states in broad bands and the different upper bands overlapping more and more for higher energy.

On the other hand, the absorption of X-rays by the inner electrons gives very much clearer results, since we may regard the bands belonging to the X-ray levels of the atoms as having negligible width, so that in (9.9) $E_l(\mathbf{k})$ is practically constant. The frequency of the absorbed

radiation is therefore directly the energy of the final state, apart from a constant.

The strength of the absorption is proportional to the density of states in the final band, multiplied by a transition probability. This is proportional to the square of the matrix element

$$\int \psi_{\mathbf{k}'}^*(\epsilon \cdot \text{grad}) \psi_{\mathbf{k}} d^3r, \quad (9.10)$$

where ϵ is the polarization vector of the radiation. The integral is to be taken over the whole crystal. However, by the Bloch theorem the integrand is periodic, so that, apart from factors, it is sufficient to integrate over the basic cell. In each basic cell $\psi_{\mathbf{k}}$ is then essentially one of the inner atomic wave functions. Consider this matrix element first for $\mathbf{k} = 0$. Then we have seen from the Wigner-Seitz approximation (§ 5.2) that $\psi_{\mathbf{k}}$ is almost exactly isotropic. If therefore $\psi_{\mathbf{k}}$ belongs to a p -level, the two factors in (9.10) have the same angular dependence, and therefore the integral does not vanish. In that case it does not depend on \mathbf{k} very sensitively.

On the other hand, if the inner state is an s -state, then we have for $\mathbf{k} = 0$ an isotropic function multiplied by the gradient of another isotropic function, and the result vanishes upon integration over angles. In that case the matrix element vanishes for $\mathbf{k} = 0$; it must therefore be proportional to \mathbf{k} near $\mathbf{k} = 0$, and the transition probability proportional to \mathbf{k}^2 . Near the other end of the band, for large values of \mathbf{k} the wave function may have to be a maximum on certain faces of the polyhedron that encloses the unit cell, and vanish on others; it has therefore more nearly the symmetry of a p -wave and the previous statements are reversed.

We expect therefore that the probability is reasonably constant throughout a band, except that it may decrease proportionally with the distance from one or the other end of the band, depending on where the minimum energy lies, and depending on the inner state involved.

The best results by this method are obtained using fairly soft X-rays, since with higher energies (very deep initial level) the width of the conduction band is a small fraction of the total energy difference. In this way Skinner (1938) and others obtained curves which give a clear idea of the level density as a function of energy in many metals.

The same method may be used with emission of soft X-rays, with the difference that the absorption method shows up only those levels which are normally empty and the emission depends on the upper level being occupied. The two results should supplement each other and it is found

that the absorption curve has a sharp cut-off at its low-frequency end, and the emission curve extends from that frequency down.

The fact that these results agree, at any rate qualitatively, with the predictions of the simple theory is not at all a matter of course. As always, our model neglects the interaction between the electrons, except on the average. Now if an inner electron has been removed from one of the atoms, this must cause the potential distribution in that atom to be rather different from its neighbours. We shall discuss later a very similar effect in non-metals, where it will be found to be important. It would therefore be quite possible that in the neighbourhood of this particular atom the extra attractive potential would lead to an energy level for a conduction electron which is below the lower edge of the ordinary conduction band, and must then be a discrete level. If this were possible, one would expect to find, on the low-frequency side of the emission band, some discrete lines, or rather narrow peaks, since collisions may broaden these lines. This behaviour is not found in practice.

9.3. Photoelectric effect

A special case of light absorption is the photoelectric effect, in which the electron is able to leave the metal. If we apply to this (9.9), we find that the threshold for the effect is not given by the well-known equation

$$\hbar\omega = W, \quad (9.11)$$

but has a higher frequency. This can be seen immediately in the following way. To realize the threshold value (9.11) the transition must lead from a state on the border of the Fermi distribution ($E = \eta$) to a free electron of velocity zero. Assuming, for simplicity, that the free surface is the plane $x = 0$, then the translational symmetry of the crystal in directions parallel to this surface ensures that the components k_y and k_z of the wave vector are unchanged as a particle passes through the surface. If its velocity is negligible outside, it must have had $k_y = k_z = 0$ inside. If the wave vector is conserved in the optical transitions, as assumed in (9.9), then the initial state must also have had $k_y = k_z = 0$. Now in each band the equation

$$E_i(k_x, 0, 0) = \eta \quad (9.12)$$

has only two solutions, belonging to equal and opposite values of k_x , of which one will have the electron travelling away from the free surface. Hence the final energy $E_f(\mathbf{k})$ is also a well-defined quantity for each. Of these values none will in general coincide with $W + \eta$, i.e. with the energy of an electron at rest outside the metal.

This argument ignores the importance of the metal surface for the

motion of the electrons. The statement that in the optical transition the wave vector remains unchanged was derived using the translational symmetry of the crystal. Near a surface this translational symmetry in the direction normal to the surface is evidently lacking. Using the analogy between wave vector and momentum, we may say that the balance of momentum and energy which is impossible for a photon just above the threshold (9.11) can be restored by the surface forces, which may also take up momentum and transfer it to the crystal as a whole.

This effect is evidently of the greatest importance for electrons in a thin surface layer. In order to have the wave vector uncertain by an amount δk , the electron must be within a distance of about $2\pi/\delta k$ from the surface. One can therefore see that, unless the work function is very small, or one of the levels in the higher bands reached from (9.12) is nearly in the right place, the surface effect concerns a layer only a few atoms thick.

Since even in a strongly absorbing metal the light penetrates to a depth of 10^3 to 10^4 atomic distances, it is clear that, for the optical properties in general, the surface may be neglected. For the photoelectric effect it is of particular importance because electrons which have made the transition near the surface have a much greater chance of escaping than those starting from greater depth. If an electron has absorbed the phonon well inside the metal, it may make a collision before it reaches the surface. For this purpose the collisions with lattice waves or impurities are unimportant, since we know them to be almost elastic. They may deflect some electrons from the surface which would otherwise have escaped, but they also deflect some electrons towards the surface which previously travelled at an angle to it. However, the electron may make inelastic collisions with other electrons, and one such collision is in general enough to prevent its escape. Collisions between electrons were discussed in § 6.5 and found to be very infrequent, but this was for electrons which formed part of the Fermi distribution, and therefore had enough energy only for collision with a selected few other electrons in the border region.

In the present problem we are concerned with highly excited electrons, which may interchange energy with almost all the electrons in the Fermi distribution. In that case, the dimensional reasoning used in § 6.5 shows the mean free path to be comparable to the atomic distance, though probably larger numerically.

The following picture emerges: at frequencies just above the threshold

(9.11), only the 'surface photoelectric effect' is possible. At a higher frequency, which depends on the particular properties of the energy bands of the metal, a volume effect becomes possible in addition. This will in general be more intense than the surface effect, though not nearly by the same factor as one would estimate in the absence of inelastic collisions.

No detailed analysis of the volume effect seems to be available. The surface effect exists also in the absence of the atomic potential, and calculations for free electrons in the neighbourhood of a discontinuous or gradual potential rise near the surface give reasonable answers.

It has been pointed out by Mitchell (1934, 1936) that it is essential to take into account the change in the light wave near the surface due to the optical properties of the medium. This makes the problem somewhat too complicated for a detailed presentation here.

One interesting result is that the theory of free electrons near a plane surface predicts zero effect for a light wave which has its electric vector parallel to the surface (and hence always for normal incidence). The reason is that for free electrons the motion can be separated into the motion at right angles to the surface and that in a plane parallel to the surface. If the electric vector is parallel to the surface, the light wave does not influence the motion in the normal direction, and hence the wave vector in that direction cannot increase, as is required.

In fact, all metals do show an effect for any polarization or angle of incidence. This may be due to the fact that the electrons are not free, so that the motions in different directions are not separable, or to the effect of collisions with phonons, which are also capable of restoring the momentum balance. In addition, any deviation of the surface from the perfect plane which has been assumed, would tend to produce the same result.

9.4. Non-conducting crystals

The discussion of § 9.2 may also be applied to a case when the valency electrons just fill one band and the next one is empty. In that case the one-band transitions discussed in § 9.1 do not exist, and the substance is transparent up to an absorption edge in the visible or ultra-violet region. This absorption edge should correspond to a transition of an electron from the highest occupied band to the lowest empty band, usually in this case called the conduction band.

If the maximum energy in the lower band belongs to the same \mathbf{k} as the minimum energy in the higher band, then the absorption edge should

equal the width of the energy gap between the bands. It may well be that these two states belong to different values of k , and in that case the lowest absorption frequency compatible with (9.11) is larger than the width of the gap. However, as in the case of the photoelectric effect, there may be a small amount of absorption due to surface forces and other deviations which makes such transitions possible, and we must then expect the existence of a frequency range in which there is a weak absorption, which will also be sensitive to temperature and to the mechanical state of the crystal, followed by the strong absorption at a higher frequency.

The result of such an absorption process is one electron in the higher band, and one vacant place in the lower band. As long as we follow our usual approximation of neglecting the interaction between electrons, we should regard the 'hole' and the electron as moving independently. However, it is at once obvious that this is a much more doubtful assumption in this case than in the case of a metal. Indeed, the electron and the hole carry opposite electric charges and will therefore attract each other at all distances. In a metal, any excess charge will at once be neutralized by an adjustment in the density of conduction electrons, and hence two specific charges will influence each other only when very close together.

We must therefore expect not only states in which the electron in the upper band and the hole in the lower band separate and move independently, but also states in which they are bound together and revolve around each other in an orbit.

They may, in fact, remain in the same atom or group of atoms, in which case their description in terms of the band concept is not too convenient, and it is preferable to start from independent atoms rather than electrons moving singly in a periodic field. This means starting from the Heitler-London model of § 8.2, except that we now assume the ground state of each atom to be non-degenerate, and allow one or more atoms to be excited. The state in which one atom is excited then has an N -fold degeneracy, since the excitation of any one atom requires the same energy. This degeneracy is removed by the interaction between the atoms. The situation is precisely the same as for single electron states in the strong-binding approximation of § 4.2 or for the case of spin waves. The stationary states have again a well-defined wave vector and the energy is a function of this wave vector, so that we obtain a narrow band in place of each atomic excitation level. The width of this band is given by an interaction integral of the type

$$(\phi_0^*(1)\phi_1^*(2)W_{1,2}\phi_0(2)\phi_1(1)), \quad (9.13)$$

where ϕ_0, ϕ_1 are the wave functions for the ground state and an excited state of a complete atom, the arguments (1) and (2) refer to the coordinates of all electrons in atoms 1 and 2 respectively, $W_{1,2}$ is the interaction energy between all particles in atom 1 and those in atom 2, and the brackets again indicate integration over all coordinates. This expression is similar to the interaction integrals of § 4.2 or the exchange integral of § 8.2, except that it does not involve the bodily transfer of an electron from one atom to the other. It is therefore appreciable even where the atomic distance is so large that the wave functions do not overlap. If the atomic transition has a non-vanishing electric dipole moment associated with it, (9.13) decreases at large distances like the interaction between two dipoles, i.e. as the inverse cube of the distance between the atoms.

A state resulting from such an excitation travelling with a definite wave vector has been called an excitation wave or 'exciton' (Frenkel, 1936). It follows from our usual arguments that, since the crystal in the ground state has complete translational symmetry, the exciton formed by the absorption of a light wave in a perfect crystal must have a wave vector equal to that of the light wave, which for most purposes is negligible.

Hence, no matter how wide the exciton band, the absorption spectrum of a perfect crystal should contain only a sharp line for each atomic absorption line. These discrete lines belong to frequencies below the beginning of the continuous absorption discussed before.

We see therefore that, as the result of the absorption of a photon, we may be concerned with three different states. (a) An electron is excited, but remains within the same atom; the excitation may then be passed along to other atoms in the form of an excitation wave, but each atom remains neutral throughout. (b) The electron leaves its atom, but remains in its neighbourhood, forming a kind of 'hydrogen atom' together with the positive hole which it left behind. This bound pair may again travel through the lattice together, though in a perfect lattice light absorption leads only to states in which the wave vector associated with the motion of the centre of gravity of the pair is zero. (c) The electron and the hole may separate and move independently with equal and opposite wave vectors.

The first state may not always be possible. For example, in an ionic crystal like NaCl, the least strongly bound electrons are those in the negative chlorine ion. This has no atomic excited states, and the result of a transition of such an electron is always its removal from the parent ion. Sodium ions do have excited states, but their excitation energy is

higher than that for removing an electron from the Cl^- , and therefore the discrete excitation levels overlap with a continuous spectrum and are broadened to a sufficient extent to lose their clear identity.

Another cause of broadening is the motion of the atoms in the lattice.

This point is best explained by reference to Fig. 16, which shows schematically the electronic energy of the ground state and the excited electron state as a function of some coordinate determining the positions of the atoms. For example, if we think of the transfer of an electron from a Cl^- ion to an adjacent Na^+ ion in NaCl , the abscissa of the figure might be the distance between these particular Cl and Na nuclei.

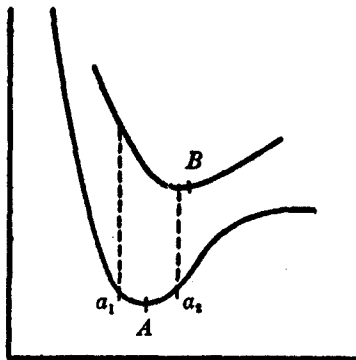


FIG. 16.

The point A , which represents the equilibrium position for the normal lattice, is no longer that for the excited state, since the attractive force between the two ions has been reduced. Because of thermal vibrations and zero-point motion, the initial state does not correspond exactly to A , but to a small spread of positions in its neighbourhood, indicated by $a_1 a_2$. It is evident that a transition which takes place when the atoms are in their extreme positions causes an energy change indicated by the vertical broken lines, and that the variation of this is much larger than the thermal energy itself, which is measured by the height of the lower curve at a_1 or a_2 above the minimum.

The use of this diagram is similar to the theory of molecular spectra. However, whereas in the problem of diatomic molecules there is only one coordinate, namely the distance of the two nuclei, the crystal has a very large number of degrees of freedom. Hence, for example, in the case of a molecule the curves of Fig. 16 would indicate that, after an absorption at or near A , the nuclei would not be in equilibrium and would be accelerated away from each other, the result being vibration with large amplitude, or dissociation, according to the shape of the curve.

In the crystal the vibration of one particular group of atoms must be strongly damped, since the energy will spread over neighbouring atoms, and the atoms should settle down near the new minimum B if such a point exists. The result is then a deformation of the lattice near the excited site, and then the excitation must be localized, and thus be a superposition of exciton states of all conceivable wave numbers. It is

therefore plausible that this description applies only if the interaction (9.13) is weak enough. If it is strong, we must expect that the states have the structure of excitation waves whose location is undetermined, so that there is no inducement for any particular set of atoms to adapt their positions to an excited state.

In exactly the same way an electron in the upper band, or a hole in the lower, may either be moving freely, or may attach itself to a definite site in the crystal, with a corresponding adjustment of the lattice near that place. After this, it would no longer be free to move since this would require either an increase in energy, if it moved to a place where the lattice structure was normal, or a movement of the heavy nuclei, if the distortion were to travel with the electron. The conditions for this to happen have not been studied very extensively. Empirically this phenomenon of electrons being trapped in a uniform lattice does not seem to be known; trapping of electrons always takes place at sites where the lattice is disturbed.

The process of emission of light is, of course, in principle just the inverse of absorption; however, the main question here is how the excited state is produced and what other processes compete with the light emission. These problems belong to the material of the next chapter.