Chapter 8-1

Carrier concentrations representative of metals, semimetals, and semiconductors are shown in Fig. 1. Semiconductors are generally classified by their electrical resistivity at room temperature, with values in the range of 10^{-2} to 10^{9} ohm-cm, and strongly dependent on temperature. At absolute zero a pure, perfect crystal of most semiconductors will be an insulator, if we arbitrarily define an insulator as having a resistivity above 10^{14} ohm-cm.

Devices based on semiconductors include transistors, switches, diodes, photovoltaic cells, detectors, and thermistors. These may be used as single circuit elements or as components of integrated circuits. We discuss in this chapter the central physical features of the classical semiconductor crystals, particularly silicon, germanium, and gallium arsenide.

Some useful nomenclature: the semiconductor compounds of chemical formula AB, where A is a trivalent element and B is a pentavalent element, are called III-V (three-five) compounds. Examples are indium antimonide and gallium arsenide. Where A is divalent and B is hexavalent, the compound is called a II-VI compound; examples are zinc sulfide and cadmium sulfide. Silicon and germanium are sometimes called diamond-type semiconductors, because they have the crystal structure of diamond. Diamond itself is more an insulator rather than a semiconductor. Silicon carbide SiC is a IV-IV compound.

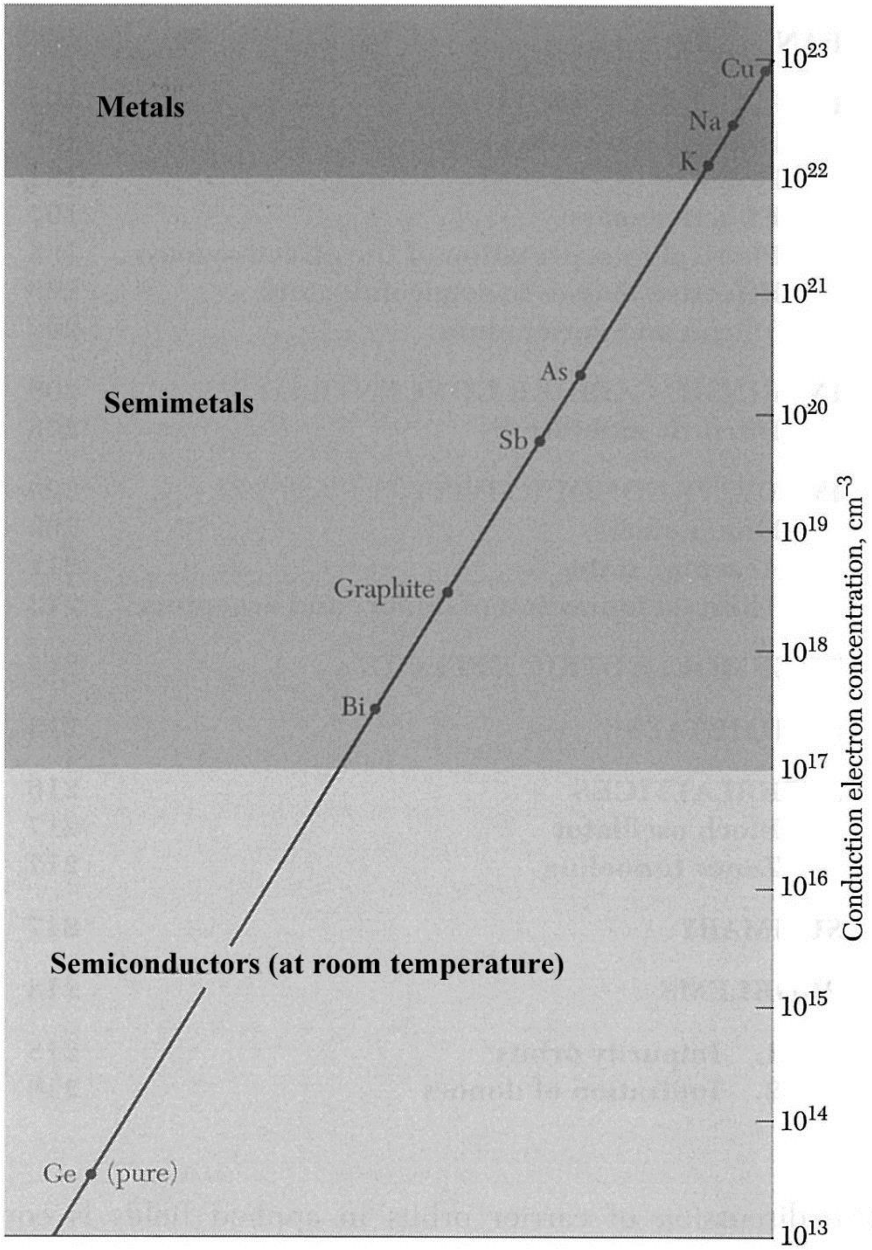


Figure 1 Carrier concentrations for metals, semimetals, and semiconductors. The semiconductor range may be extended upward by increasing the impurity concentration, and the range can be extended downward to merge eventually with the insulator range.

A highly purified semiconductor exhibits intrinsic conductivity, as distinguished from the impurity conductivity of less pure specimens. In the **intrinsic temperature range** the electrical properties of a semiconductor are not essentially modified by impurities in the crystal. An electronic band scheme leading to intrinsic conductivity is indicated in Fig. 2. The conduction band is vacant at absolute zero and is separated by an energy gap E_g from the filled valence band.

The **band gap** is the difference in energy between the lowest point of the conduction band and the highest point of the valence band. The lowest point in the conduction band is called the **conduction band edge**; the highest point in the valence band is called the **valence band edge**.

As the temperature is increased, electrons are thermally excited from the valence band to the conduction band (Fig. 3). Both the electrons in the conduction band and the vacant orbitals or holes left behind in the valence band contribute to the electrical conductivity.

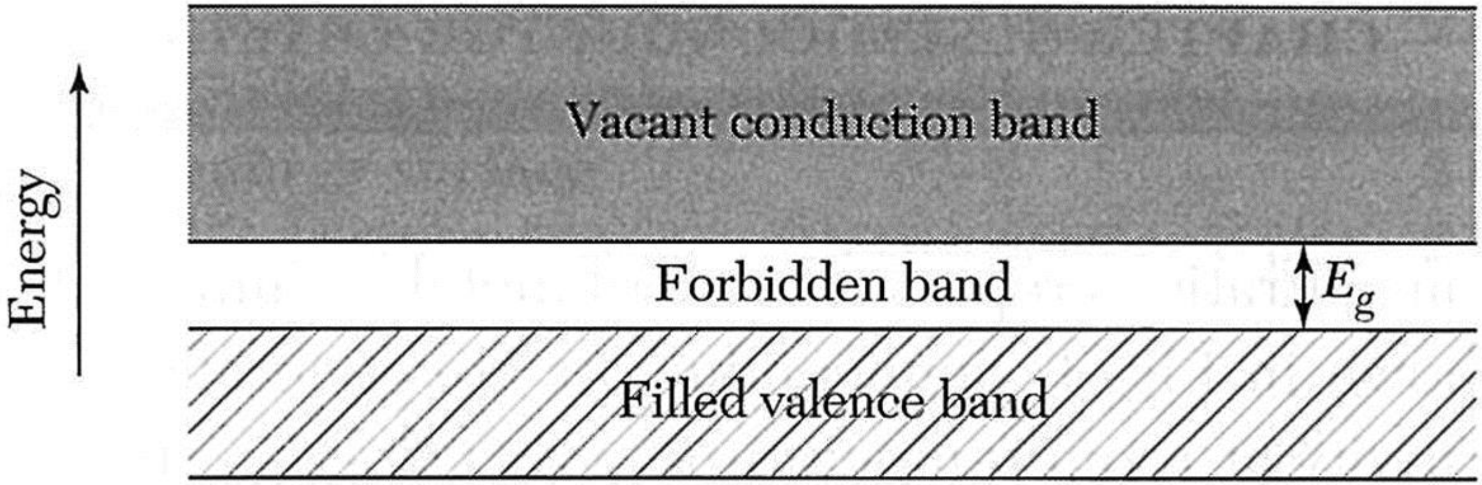
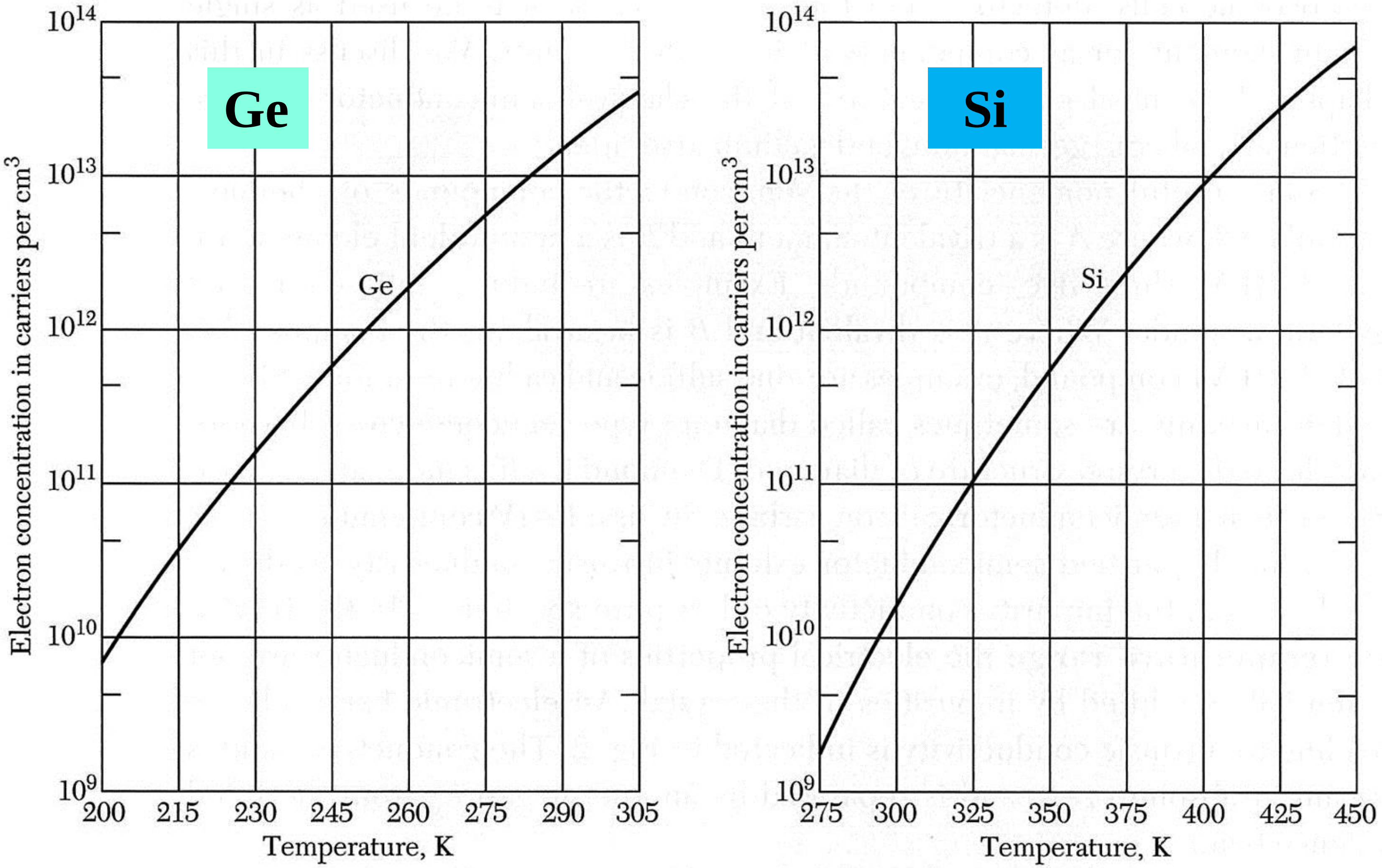


Figure 2 Band scheme for intrinsic conductivity in a semiconductor. At 0 K the conductivity is zero because all states in the valence band are filled and all states in the conduction band are vacant. As the temperature is increased, electrons are thermally excited from the valence band to the conduction band, where they become mobile. Such carriers are called "intrinsic."



(a) (b)

Figure 3 Intrinsic electron concentration as a function of temperature for (a) germanium and (b) silicon. Under intrinsic conditions the hole concentration is equal to the electron concentration. The intrinsic concentration at a given temperature is higher in Ge than in Si because the energy gap is narrower in Ge (0.66 eV) than in Si (1.11 eV). (After W. C. Dunlap.)

BAND GAP

The intrinsic conductivity and intrinsic carrier concentrations are largely controlled by E_g/k_BT, the ratio of the band gap to the temperature. When this ratio is large, the concentration of intrinsic carriers will be low and the conductivity will be low. Band gaps of representative semiconductors are given in Table 1. The best values of the band gap are obtained by optical absorption.

In a **direct absorption process** the threshold of continuous optical absorption at frequency ω_g measures the band gap $E_g = \hbar\omega_g$ as shown in Figs. 4a and 5a. A photon is absorbed by the crystal with the creation of an electron and a hole.

In the **indirect absorption process** in Figs. 4b and 5b the minimum energy gap of the band structure involves electrons and holes separated by a substantial wavevector $\mathbf{k}_c$. Here a direct photon transition at the energy of the minimum gap cannot satisfy the requirement of conservation of wavevector, because photon wavevectors are negligible at the energy range of interest. But if a phonon of wavevector **K** and frequency Ω is created in the process, then we can have

$$\boxed{\mathbf{k}(\text{photon}) = \mathbf{k}_c + \mathbf{K} \cong \mathbf{0} \;; \qquad \hbar\omega = E_g + \hbar\Omega \;,} \qquad \hbar\Omega << E_g$$

as required by the conservation laws. The phonon energy $\hbar\Omega$ will generally be much less than E_g: a phonon even of high wavevector is an easily accessible source of crystal momentum because the phonon energies are characteristically small (~0.01 to 0.03 eV) in comparison with the energy gap. If the temperature is high enough that the necessary phonon is already thermally excited in the crystal, it is possible also to have a photon absorption process in which the phonon is absorbed.

CRYSTAL WITH DIRECT GAP **(T = 0)** CRYSTAL WITH INDIRECT GAP

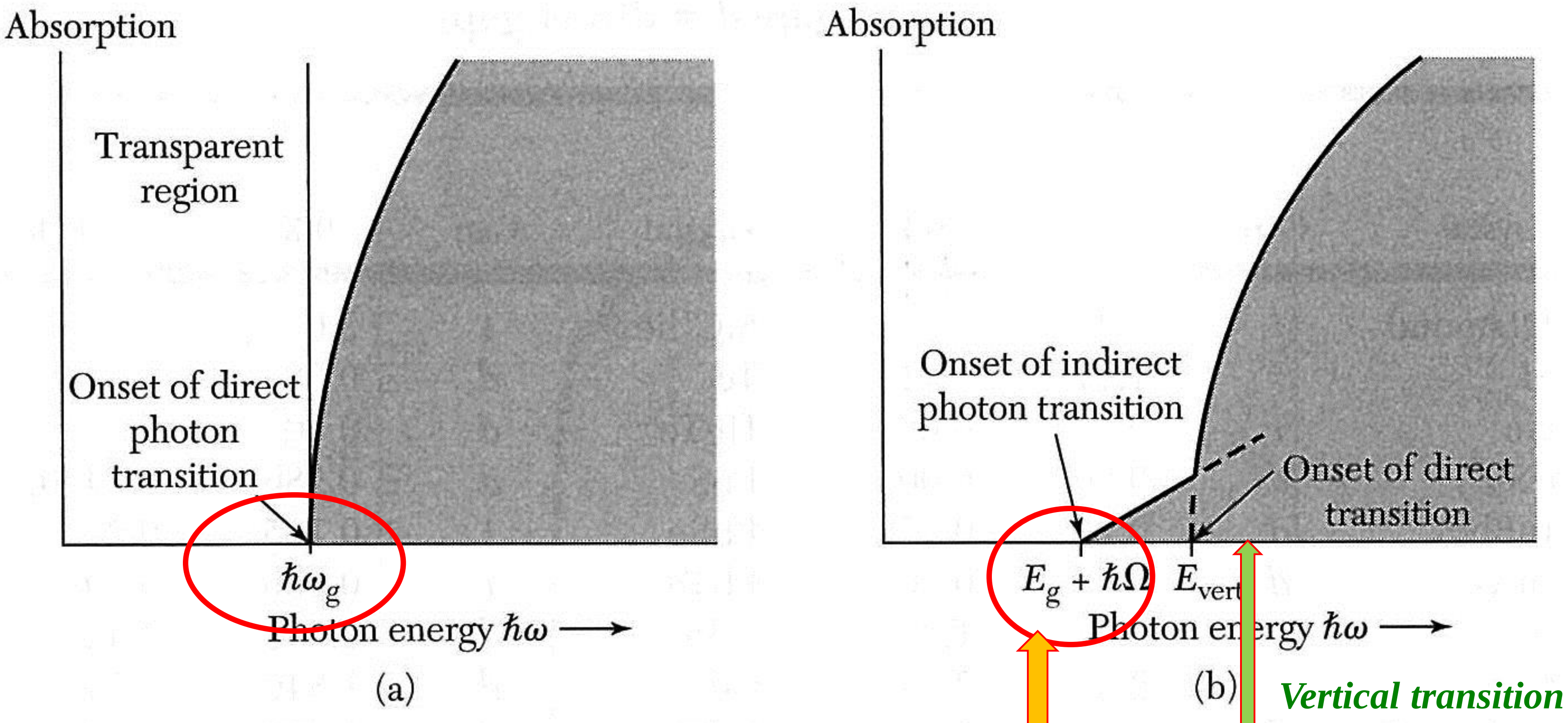


Figure 4 Optical absorption in pure insulators at absolute zero. In (a) the threshold determines the energy gap as $E_g = \hbar\omega_g$. In (b) the optical absorption is weaker near the threshold: at $\hbar\omega = E_g + \hbar\Omega$ a photon is absorbed with the creation of three particles: a free electron, a free hole, and a phonon of energy $\hbar\Omega$. In (b) the energy E_{vert} marks the threshold for the creation of a free electron and a free hole, with no phonon involved. Such a transition is called vertical; it is similar to the direct transition in (a). These plots do not show absorption lines that sometimes are seen lying just to the low energy side of the threshold. Such lines are due to the creation of a bound electron-hole pair, called an exciton.

Direct Transition

Indirect Transition

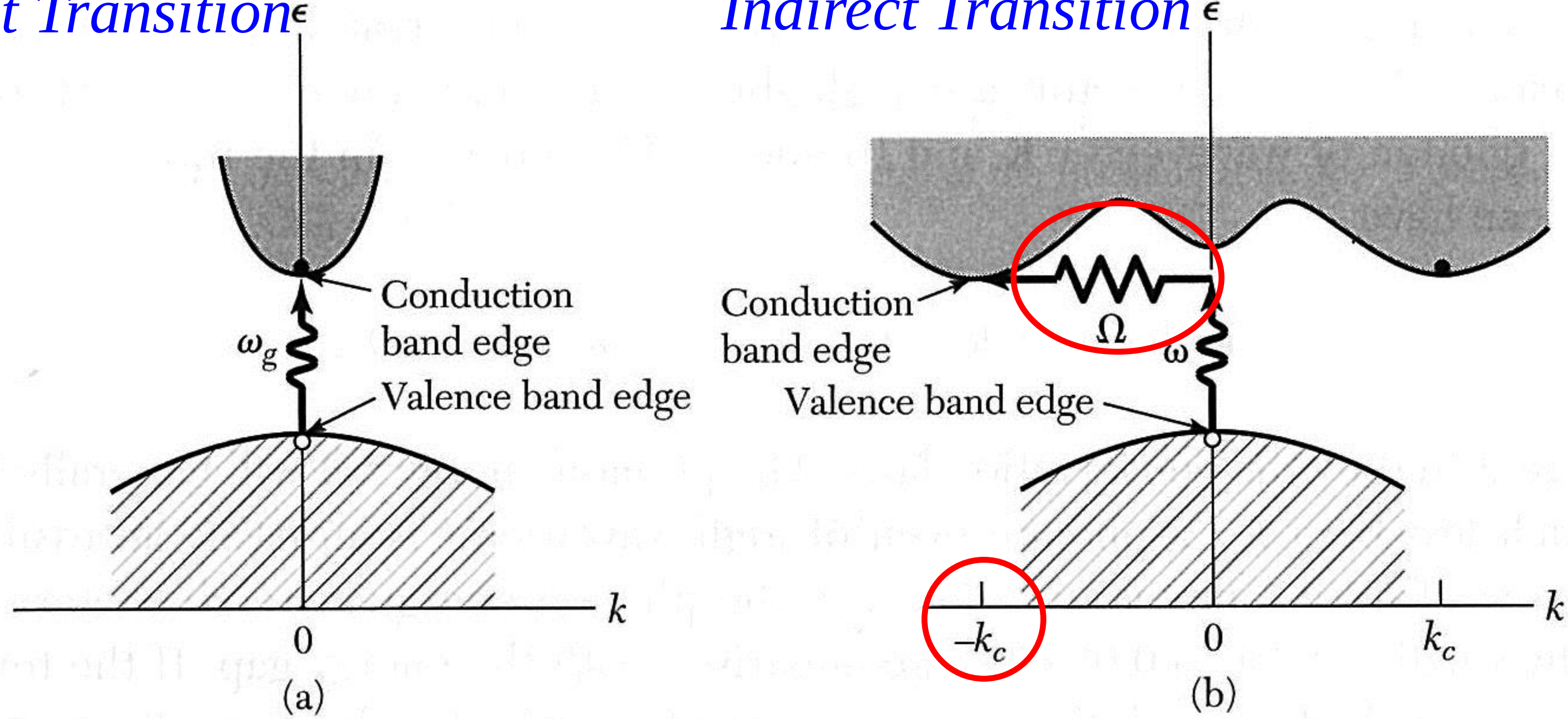


Figure 5 In (a) the lowest point of the conduction band occurs at the same value of **k** as the highest point of the valence band. A direct optical transition is drawn vertically with no significant change of **k**, because the absorbed photon has a very small wavevector. The threshold frequency ω_g for absorption by the direct transition determines the energy gap $E_g = \hbar\omega_g$. The indirect transition in (b) involves both a photon and a phonon because the band edges of the conduction and valence bands are widely separated in **k** space. The threshold energy for the indirect process in (b) is greater than the true band gap. The absorption threshold for the indirect transition between the band edges is at $\hbar\omega = E_g + \hbar\Omega$, where Ω is the frequency of an emitted *phonon* of wavevector $\mathrm{K} \cong -\mathbf{k}_g$. At higher temperatures phonons are already present; if a phonon is absorbed along with a photon, the threshold energy is $\hbar\omega = E_g - \hbar\Omega$. *Note:* The figure shows only the threshold transitions. Transitions occur generally between almost all points of the two bands for which the wavevectors and energy can be conserved.

In the **indirect absorption process** in Figs. 4b and 5b the minimum energy gap of the band structure involves electrons and holes separated by a substantial wavevector $\mathbf{k}_c$. Here a direct photon transition at the energy of the minimum gap cannot satisfy the requirement of conservation of wavevector, because photon wavevectors are negligible at the energy range of interest. But if a phonon of wavevector $\mathbf{K}$ and frequency Ω is created in the process, then we can have

$$\boxed{\mathbf{k}(\text{photon}) = \mathbf{k}_c + \mathbf{K} \cong \mathbf{0} \ ; \qquad \hbar\omega = E_g + \hbar\Omega \ ,} \qquad \hbar\Omega << E_g$$

as required by the conservation laws. The phonon energy $\hbar\Omega$ will generally be much less than E_g: a phonon even of high wavevector is an easily accessible source of crystal momentum because the phonon energies are characteristically small (~0.01 to 0.03 eV) in comparison with the energy gap. If the temperature is high enough that the necessary phonon is already thermally excited in the crystal, it is possible also to have a photon absorption process in which the phonon is absorbed.

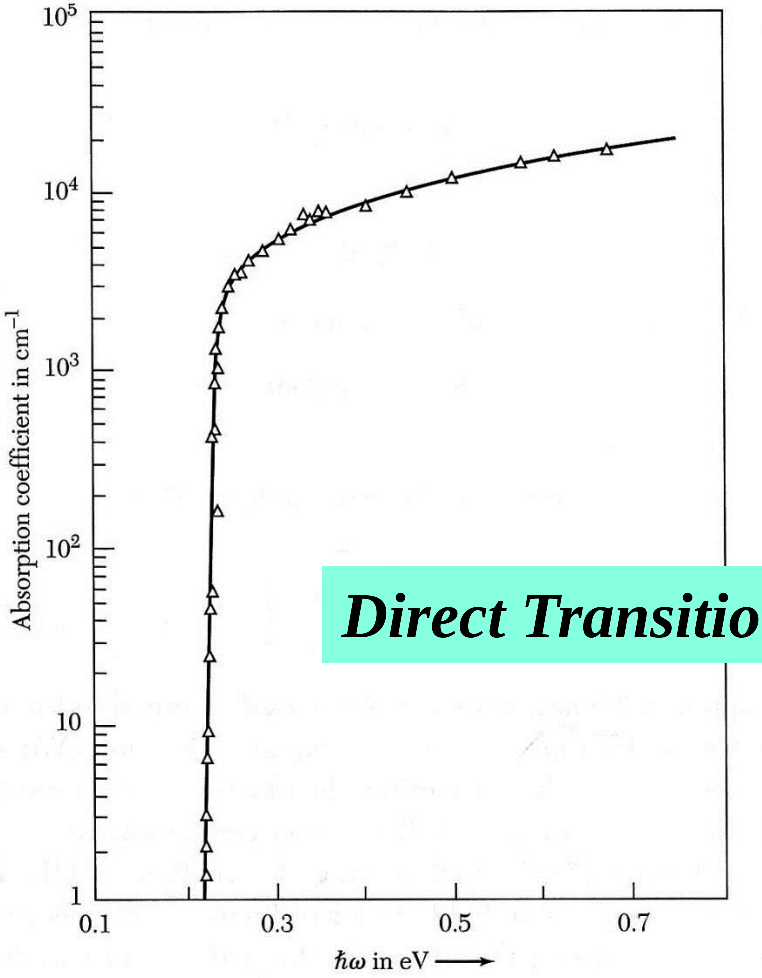
Table 1 Energy gap between the valence and conduction bands

(*i* = indirect gap; *d* = direct gap)

		E_g, eV				E_g, eV	
Crystal	Gap	0 K	300 K	Crystal	Gap	0 K	300 K
Diamond	*i*	5.4		SiC(hex)	*i*	3.0	—
Si	*i*	1.17	1.11	Tc	*d*	0.33	—
Ge	*i*	0.744	0.66	HgTe[a]	*d*	−0.30	
αSn	*d*	0.00	0.00	PbS	*d*	0.286	0.34–0.37
InSb	*d*	0.23	0.17	PbSe	*i*	0.165	0.27
InAs	*d*	0.43	0.36	PbTe	*i*	0.190	0.29
InP	*d*	1.42	1.27	CdS	*d*	2.582	2.42
GaP	*i*	2.32	2.25	CdSe	*d*	1.840	1.74
GaAs	*d*	1.52	1.43	CdTe	*d*	1.607	1.44
GaSb	*d*	0.81	0.68	SnTe	*d*	0.3	0.18
AlSb	*i*	1.65	1.6	Cu_2O	*d*	2.172	—

[a] HgTe is a semimetal; the bands overlap.

Optical absorption to determine the (direct) energy gap of InSb



Direct Transition at 0.23V

Figure 6 Optical absorption in pure indium antimonide, InSb. The transition is direct because both conduction and valence band edges are at the center of the Brillouin zone, $\mathbf{k} = 0$. Notice the sharp threshold. (After G. W. Gobeli and H. Y. Fan.)

The band gap may also be deduced from the temperature dependence of the conductivity or of the carrier concentration in the intrinsic range. The carrier concentration is obtained from measurements of the Hall voltage (Chapter 6), sometimes supplemented by conductivity measurements. Optical measurements determine whether the gap is direct or indirect. The band edges in Ge and in Si are connected by indirect transitions; the band edges in InSb and GaAs are connected by a direct transition (Fig. 6). The gap in αSn is direct and is exactly zero; HgTe and HgSe are semimetals and have negative gaps—the conduction and valence bands overlap.

EQUATIONS OF MOTION OF AN ELECTRON IN AN ENERGY BAND

We derive the equation of motion of an electron in an energy band. We look at the motion of a wave packet in an applied electric field. Suppose that the wave packet is made up of wavefunctions assembled near a particular wavevector k. The group velocity by definition is $v_g = d\omega/dk$. The frequency associated with a wavefunction of energy ϵ by quantum theory is $\omega = \epsilon/\hbar$, and so

$$v_g = \hbar^{-1}\, d\epsilon/dk \qquad \text{or} \qquad \boxed{\mathbf{v} = \hbar^{-1}\nabla_{\mathbf{k}}\epsilon(\mathbf{k})\ .} \tag{1}$$

The effects of the crystal on the electron motion are contained in the dispersion relation $\epsilon(\mathbf{k})$.

The work $\delta\epsilon$ done on the electron by the electric field E in the time interval δt is

$$\delta\epsilon = -eEv_g\,\delta t \ . \tag{2}$$

We observe that

$$\delta\epsilon = (d\epsilon/dk)\delta k = \hbar v_g\,\delta k \ , \tag{3}$$

using (1). On comparing (2) with (3) we have

$$\delta k = -(eE/\hbar)\delta t \ , \tag{4}$$

whence $\hbar dk/dt = -eE$.

We may write (4) in terms of the external force **F** as

$$\boxed{\hbar\frac{d\mathbf{k}}{dt} = \mathbf{F} \ .} \tag{5}$$

This is an important relation: in a crystal $\hbar d\mathbf{k}/dt$ is equal to the external force on the electron. In free space $d(m\mathbf{v})/dt$ is equal to the force. We have not overthrown Newton's second law of motion: the electron in the crystal is subject to forces from the crystal lattice as well as from external sources.

The force term in (5) also includes the electric field and the Lorentz force on an electron in a magnetic field, under ordinary conditions where the magnetic field is not so strong that it breaks down the band structure. Thus the equation of motion of an electron of group velocity $\mathbf{v}$ in a constant magnetic field $\mathbf{B}$ is

$$\text{(CGS)}\quad \hbar\frac{d\mathbf{k}}{dt} = -\frac{e}{c}\mathbf{v}\times\mathbf{B}\ ;\qquad\qquad \text{(SI)}\quad \hbar\frac{d\mathbf{k}}{dt} = -e\mathbf{v}\times\mathbf{B} \qquad (6)$$

where the right-hand side of each equation is the Lorentz force on the electron. With the group velocity $\mathbf{v} = \hbar^{-1}\mathrm{grad}_{\mathbf{k}}\epsilon$, the rate of change of the wavevector is

$$\text{(CGS)}\quad \boxed{\frac{d\mathbf{k}}{dt} = -\frac{e}{\hbar^2 c}\nabla_{\mathbf{k}}\epsilon\times\mathbf{B}\ ;}\qquad\qquad \text{(SI)}\quad \frac{d\mathbf{k}}{dt} = -\frac{e}{\hbar^2}\nabla_{\mathbf{k}}\epsilon\times\mathbf{B} \qquad (7)$$

where now both sides of the equation refer to the coordinates in $\mathbf{k}$ space.

We see from the vector cross-product in (7) that in a magnetic field an electron moves in $\mathbf{k}$ space in a direction normal to the direction of the gradient of the energy ϵ, so that **the electron moves on a surface of constant energy**. The value of the projection $k_{\mathbf{B}}$ of $\mathbf{k}$ on $\mathbf{B}$ is constant during the motion. The motion in $\mathbf{k}$ space is on a plane normal to the direction of $\mathbf{B}$, and the orbit is defined by the intersection of this plane with a surface of constant energy.

Physical Derivation of $\hbar\dot{\mathbf{k}} = \mathbf{F}$

We consider the Bloch eigenfunction $\psi_{\mathbf{k}}$ belonging to the energy eigenvalue $\epsilon_{\mathbf{k}}$ and wavevector $\mathbf{k}$:

$$\psi_{\mathbf{k}} = \sum_{\mathbf{G}} C(\mathbf{k} + \mathbf{G}) \exp[i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{r}] \ . \tag{8}$$

The expectation value of the momentum of an electron in the Bloch state $\mathbf{k}$ is

$$\mathbf{P}_{\mathrm{el}} = (\mathbf{k}|-i\hbar\nabla|\mathbf{k}) = \sum_{\mathbf{G}} \hbar(\mathbf{k} + \mathbf{G})|C(\mathbf{k} + \mathbf{G})|^2 = \hbar(\mathbf{k} + \sum_{\mathbf{G}} \mathbf{G}|C(\mathbf{k} + \mathbf{G})|^2) \ , \tag{9}$$

using $\Sigma|C(\mathbf{k} + \mathbf{G})|^2 = 1$.

We examine the transfer of momentum between the electron and the lattice when the state $\mathbf{k}$ of the electron is changed to $\mathbf{k} + \Delta\mathbf{k}$ by the application of an external force. We imagine an insulating crystal electrostatically neutral except for a single electron in the state $\mathbf{k}$ of an otherwise empty band.

We suppose that a weak external force is applied for a time interval such that the total impulse given to the entire crystal system is $\mathbf{J} = \int \mathbf{F}\, dt$. If the conduction electron were free ($m^* = m$), the total momentum imparted to the crystal system by the impulse would appear in the change of momentum of the conduction electron:

Without interaction with the crystal potential

$$\mathbf{J} = \Delta\mathbf{p}_{\text{tot}} = \Delta\mathbf{p}_{\text{el}} = \hbar\Delta\mathbf{k} \ . \tag{10}$$

The neutral crystal suffers no net interaction with the electric field, either directly or indirectly through the free electron.

If the conduction electron interacts with the periodic potential of the crystal lattice, we must have

With interaction with the crystal potential

$$\mathbf{J} = \Delta\mathbf{p}_{\text{tot}} = \Delta\mathbf{p}_{\text{lat}} + \Delta\mathbf{p}_{\text{el}} \ . \tag{11}$$

From the result (9) for $\mathbf{p}_{\text{el}}$ we have

$$\Delta\mathbf{p}_{\text{el}} = \hbar\Delta\mathbf{k} + \sum_{\mathbf{G}} \hbar\mathbf{G}[(\nabla_{\mathbf{k}}|C(\mathbf{k}+\mathbf{G})|^2)\cdot\Delta\mathbf{k}] \ . \tag{12}$$

The change $\Delta\mathbf{p}_{\mathrm{lat}}$ in the lattice momentum resulting from the change of state of the electron may be derived by an elementary physical consideration. An electron is reflected by the lattice transfers momentum to the lattice. If an incident electron with plane wave component of momentum $\hbar\mathbf{k}$ is reflected with momentum $\hbar(\mathbf{k}+\mathbf{G})$, the lattice acquires the momentum $-\hbar\mathbf{G}$, as required by momentum conservation. The momentum transfer to the lattice when the state $\psi_{\mathbf{k}}$ goes over to $\psi_{\mathbf{k}+\Delta\mathbf{k}}$ is

$$\boxed{\Delta\mathbf{p}_{\mathrm{lat}} = -\hbar\sum_{\mathbf{G}} \mathbf{G}[(\nabla_{\mathbf{k}}|C(\mathbf{k}+\mathbf{G})|^2 \cdot \Delta\mathbf{k}] \ ,} \tag{13}$$

because the portion

$$\nabla_{\mathbf{k}}|C(\mathbf{k}+\mathbf{G})|^2 \cdot \Delta\mathbf{k} \qquad \textit{for a given } \boldsymbol{G} \tag{14}$$

of each individual component of the initial state is reflected during the state change $\Delta\mathbf{k}$.

Then, summation over all possible ***G***

The total momentum change is therefore

$$\Delta\mathbf{p}_{\mathrm{el}} + \Delta\mathbf{p}_{\mathrm{lat}} = \mathbf{J} = \hbar\Delta\mathbf{k} \ , \tag{15}$$

exactly as for free electrons, Eq. (10). Thus from the definition of $\mathbf{J}$, we have

$$\boxed{\hbar d\mathbf{k}/dt = \mathbf{F} \ ,} \tag{16}$$

derived in (5) by a different method. A rigorous derivation of (16) by an entirely different method is given in Appendix E.