

KYA322 Semiconductors

Ross Turner
rjturner@utas.edu.au

1

Energy Bands

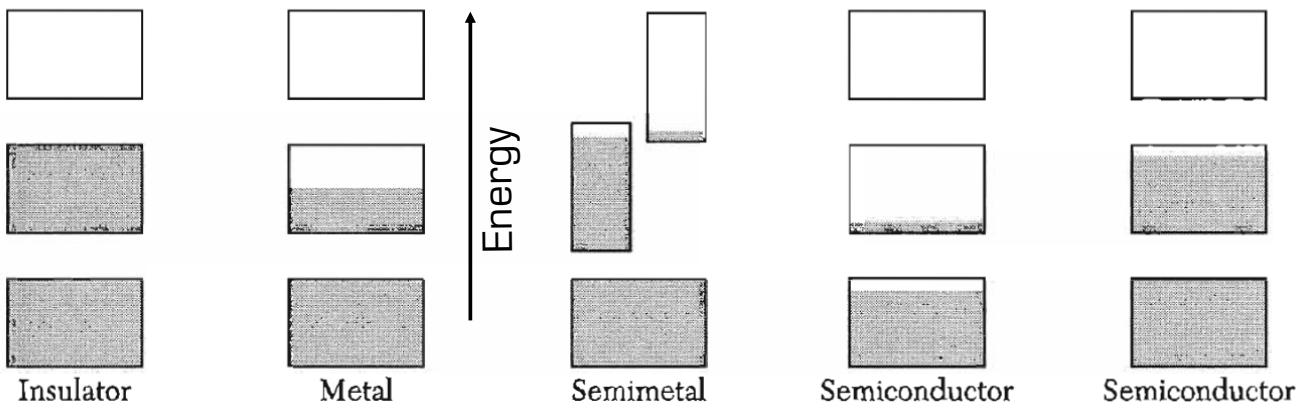
The free electron model did a reasonable job of calculating some properties of metals, but it doesn't do a great job of explaining the differences between conductors and insulators, or the existence of semiconductors.

E.g., a conductor can have resistivity as low as $10^{-10} \Omega \cdot \text{cm}$, but an insulator can be as high as $10^{22} \Omega \cdot \text{cm}$. This is an enormously high range that the free electron model doesn't account for.

The most general treatment uses the Schrödinger equation for electrons in the presence of a strictly periodic lattice of atomic potentials.

Model the electrons as moving independently of each other in an average field that is determined by the symmetries of the crystal structure.

There are a few different approximate analytical techniques, and computational techniques based on them that add complexity.



Schematic of electron band occupancy for various types of solids, grouped by electrical properties. The vertical size of the boxes indicates the allowed energy regions. Shading indicates the parts of the bands that contain electrons.

A semimetal has one band nearly filled and one nearly empty even at absolute zero, while pure semiconductor becomes an insulator at 0 K; higher bands are partly filled due to thermal excitation.

Other materials can become semiconductors because of electron deficiency due to impurities (right-hand column).

Energy Bands

The electrons in solids group themselves into allowed energy bands, separated by *band gaps* within which no wavelike solutions to the Schrödinger equation exist.

The forbidden regions arise from interactions between the electrons and the lattice structure of the crystal.

Qualitatively, this picture clearly explains the difference between insulators and conductors:

- If all allowed regions are either completely full or completely empty, then there are no mobile electrons that can respond to an applied E field.
- If one or more bands are between ~10-90% filled, or ~10-90% empty, the material behaves like a metal.
- If a band is very close to completely filled or completely empty, the material behaves as a semiconductor

"Nearly Free Electron" Model

In this approximation, the lattice potential is treated as a small, time-independent periodic perturbation to the free electron case.

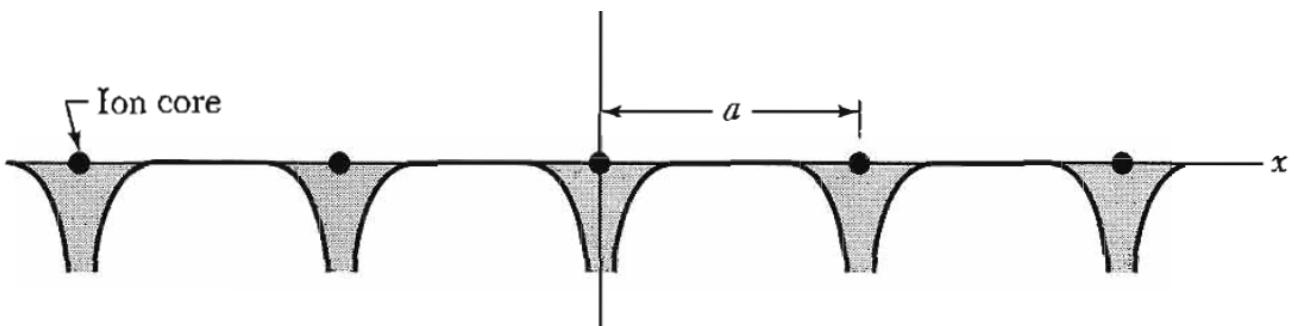
Bragg reflections from the atom cores in the lattice cause the energy gaps in solids.

In a 1-dimensional solid with crystal spacing a , the first band gaps occur for wavevectors $\pm\pi/a$, corresponding to the edge of the first Brillouin zone. Successive band gaps occur at $k = \pm n\pi/a$.

The wavefunctions at $k = \pm n\pi/a$ are standing waves. These are made of equal parts rightward and leftward travelling waves, superposed on each other. There are two solutions, with opposite parity:

$$\begin{aligned}\psi_+(x) &= 2 \cos(\pi x/a) \\ \psi_-(x) &= 2i \sin(\pi x/a)\end{aligned}$$

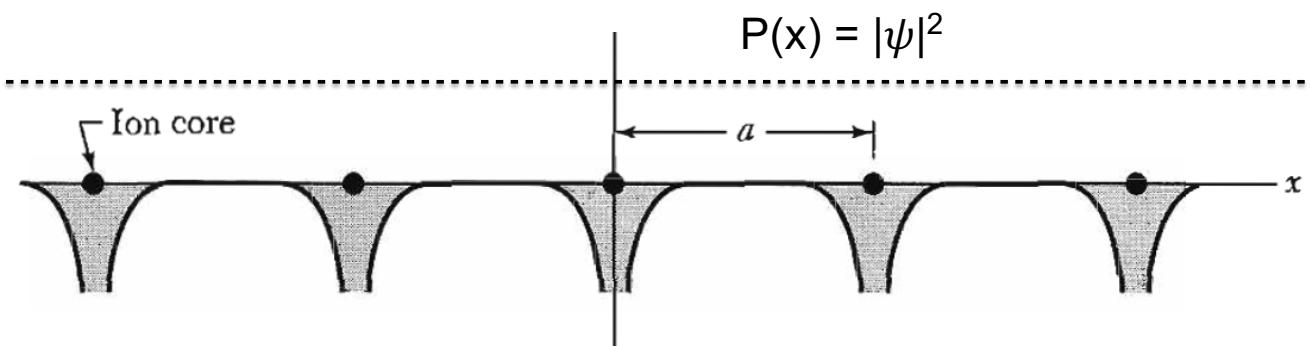
Boundary conditions



- The length L is identified as the size of the one-dimensional object that contains N atoms with an average spacing a .
- In this case we impose the boundary condition that $\psi(x) = \psi(x+L)$.
- In any finite system the allowed values of k are quantized. Here, k is restricted to be a solution of $k = 2\pi n/L$, where n is an integer.
- Successive allowed values of k are separated by $2\pi/L$, so the number of energy states between $-\pi/a$ and $+\pi/a$ is equal to $L/a = N$.

Probability Densities

The probability density for finding the electron at position x is $|\psi(x)|^2$, so the probability for the electron to be found in a given region between x and $x+dx$ is $\int \psi^* \psi dx$. A travelling wave, corresponding to a free electron, has equal probability of being found anywhere along the line, and thus constant charge density along the wave.



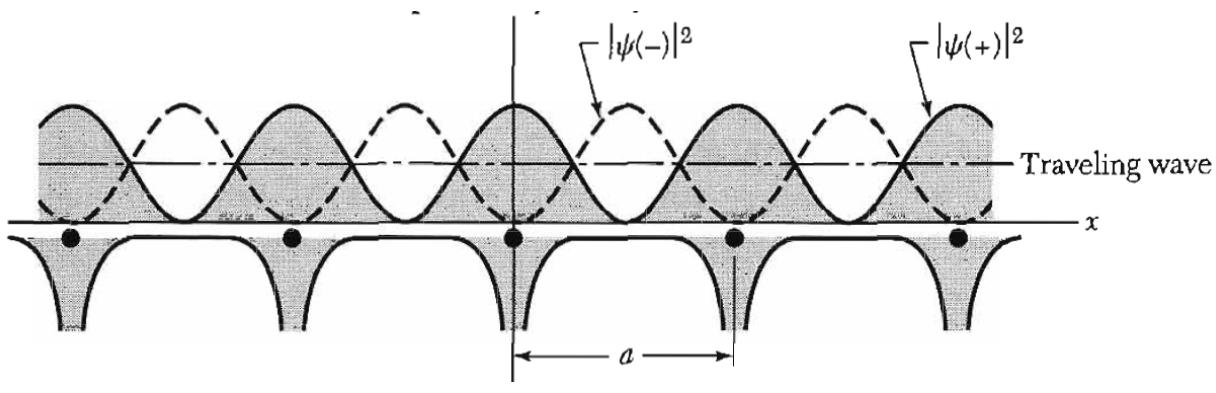
The Origin of Band Gaps

The positive and negative parity standing wave functions indicate different concentrations of electrons at specific locations in the lattice.

The solution ψ_+ gives charge density $\propto \cos^2(\pi x/a)$, which concentrates the negative charge density around $x = 0, a, 2a, 3a, \dots$, minimising the potential energy.

The solution ψ_- is proportional to $\sin^2(\pi x/a)$, which maximises the potential energy.

The two standing waves, with the same k , have quite different energies.



Band gaps in a 1-d lattice

Band gaps appear where the lattice spacing is a half-integer multiple of the electron wavelength.

E

Dashed line = free electron.

$-\pi/a$

π/a

Energy of an electron in a one-dimensional lattice as a function of wave-number k .

9

Second allowed band

First allowed band

$-\pi/a$

π/a

k

10

Size of the Band Gap

For electrons with energies corresponding to k very far from $\pm \pi/a$, the energy distribution is continuous, and very like the free electron case.

The two wavefunctions at the edge of the Brillouin zone with $k = \pi/a$ are $\psi_+ = \sqrt{2} \cos \pi x/a$ and $\psi_- = \sqrt{2} \sin \pi x/a$.

The periodic potential can be approximated as $U(x) = U_0 \cos 2\pi x/a$.

The band gap can be calculated using time-independent perturbation theory to calculate the expectation value of E_- and E_+ . Doing the integrals shows that the size of the band gap is equal to the depth of the periodic potential,
 $\langle E_- - E_+ \rangle = U_0$.

In turn, U_0 depends on how many electrons are taken from the individual atoms to create the sea of conduction electrons.

The Bloch Theorem

A crystal can be thought of as a periodic potential with which the electron wave functions interact, limiting the allowed wavefunctions to certain specific forms.

The phonon spectrum had a cutoff at the Brillouin zone boundary ($k = \pm \pi/a$), but there is no such cutoff for electron probability density waves.

The Bloch theorem says that:

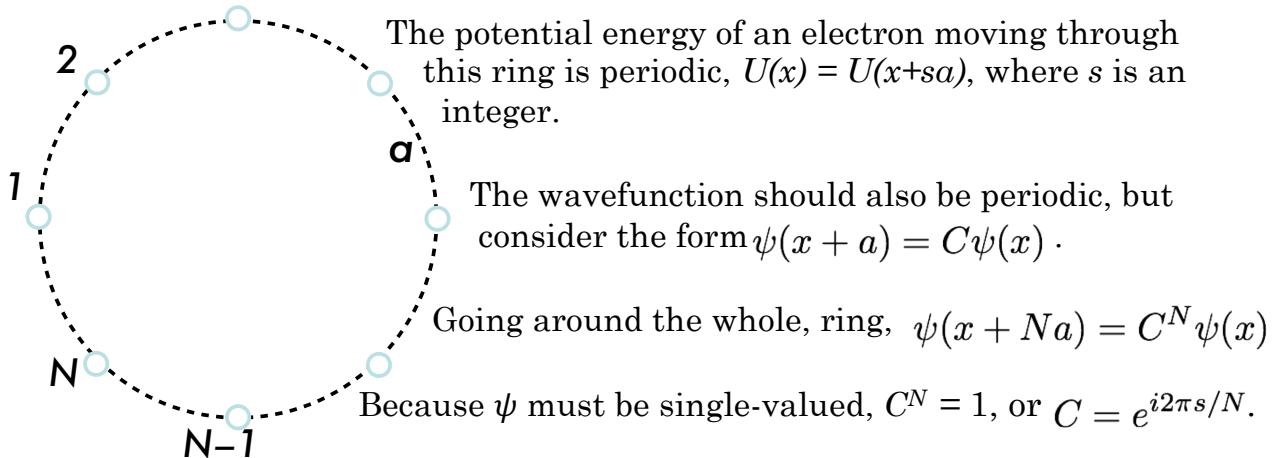
The eigenfunctions of the wave equation for a periodic potential are the product of a plane wave $e^{(k \cdot r)}$ times a function $u_k(r)$ with the periodicity of the crystal lattice.

$$\psi_k(r) = u_k(r) \exp(ik \cdot r)$$

A one-electron wave function of this form is known as a Bloch function, and it can be Fourier decomposed into a sum of travelling waves, and summed up into wave packets that describe the electron probability density distribution in the crystal lattice.

Restricted Proof of the Bloch Theorem

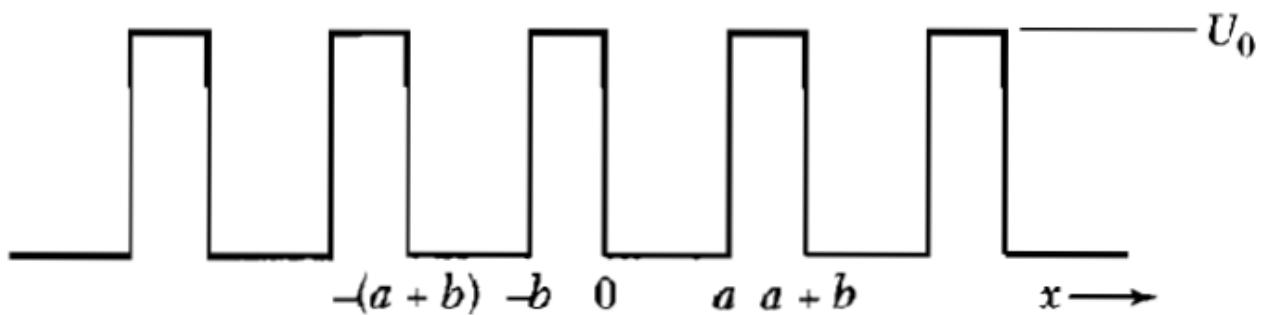
When the wavefunctions are non-degenerate (there are no other ψ_k that have the same energy), the Bloch theorem can be proved by considering a ring of lattice points of circumference Na .



So the full wavefunction can be any periodic function $u(x)$ with period a , multiplied by C , or $\psi(x) = u_k(x) \exp(i2\pi sx/Na)$, which is Bloch's result.

The Kronig-Penney Model

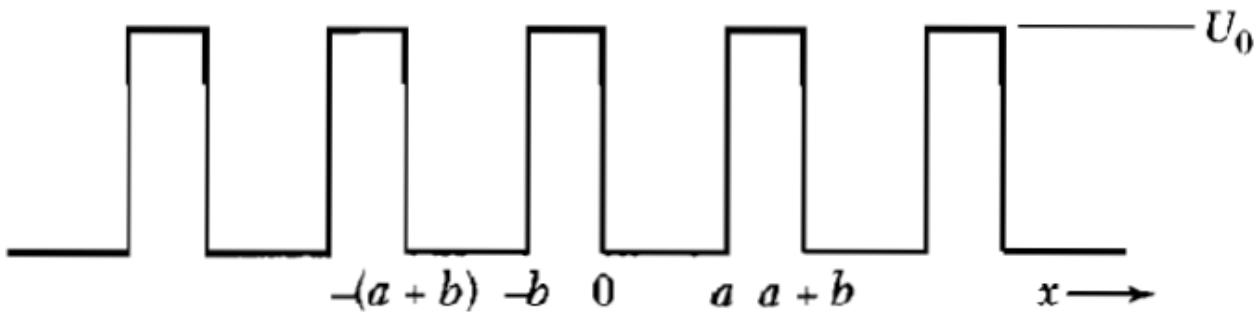
A **1-dimensional** periodic array of square well potentials can be solved for a wavefunction made of elementary functions.



In the region from $0 < x \leq a$, the potential $U = 0$, and from $a < x \leq b$, $U = U_0$.

The Schrödinger equation $-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + U(x)\psi = \epsilon\psi$ has the general solution where $U = 0$, $\psi = Ae^{iKx} + Be^{-iKx}$, $\epsilon = \frac{\hbar^2 K^2}{2m}$

The Kronig-Penney Model

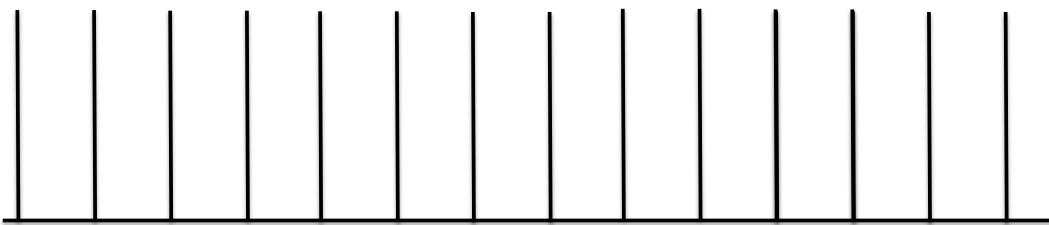


Where $U = U_0$, the solutions are $\psi = Ce^{Qx} + De^{-Qx}$, where $U_0 - \epsilon = \frac{\hbar^2 Q^2}{2m}$

The *Bloch theorem* constrains the solutions, so that ψ in the region from $-b < x \leq 0$ and from $a < x \leq a+b$ are related by $\psi(a \rightarrow a+b) = \psi(-b \rightarrow 0)e^{ik(a+b)}$

The constants A, B, C, D are chosen so ψ and $\partial\psi/\partial x$ are continuous, giving four equations and the condition that the determinant vanish. The full solution is $[(Q^2 - K^2)/2QK] \sinh Qb \sin Ka + \cosh Qb \cos Ka = \cos(k(a+b))$
but for $b \rightarrow 0$ and $U_0 \rightarrow \infty$, $Q^2 ba/2 \equiv P$, and $(P/Ka) \sin Ka + \cos Ka = \cos ka$

The Kronig-Penney Model

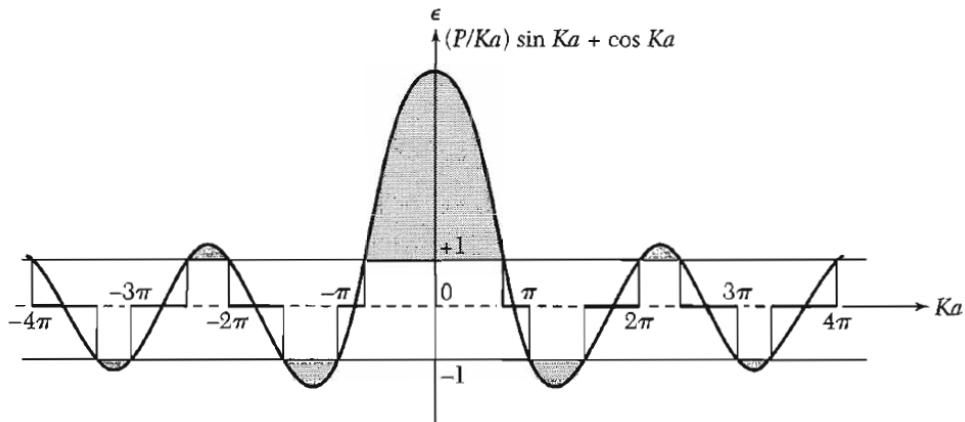


In the limit as the potential barriers get infinitely high and infinitely narrow, the solutions take on the form $(P/Ka) \sin Ka + \cos Ka = \cos ka$.

The structure of this function leads to forbidden regions, where Ka is not between 0 and ± 1 .

Remember that K is the wavevector related to the energy of the wavefunction ψ , while k is the wavevector associated with the periodicity of the lattice (the Bloch function).

The Kronig-Penney Model



Where $Ka = (2m\epsilon/\hbar^2)^{1/2} a$ is outside the range 0 to ± 1 , there are no travelling wave (Bloch) solutions to the wave equation, giving rise to energy band gaps in which no electron energy levels can exist (shaded regions).

The plot here is for $P = 3\pi/2$ and the parameter $P = Q^2ba/2$

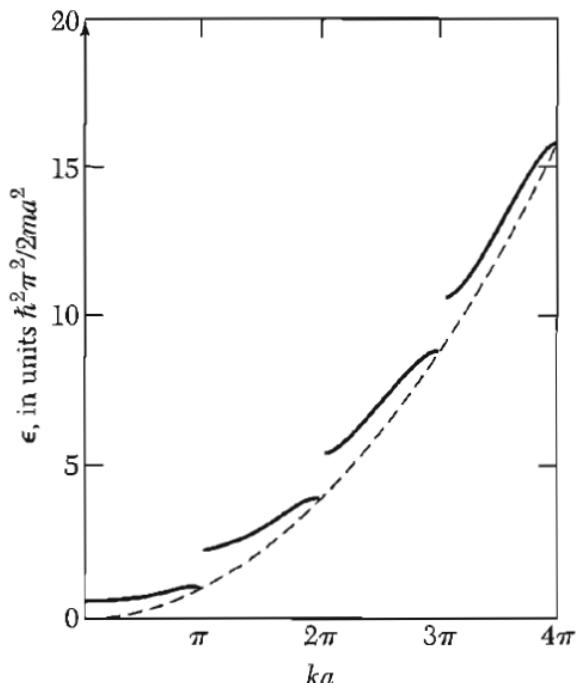
The Kronig-Penney model

This plot shows the electron energy vs. wavenumber for the Kronig-Penney potential, plotted for $P = 3\pi/2$.

Energy gaps appear where $ka = \pi, 2\pi, 3\pi, \dots$

The Kronig-Penney energy gaps and wavefunctions are a special case of the more general “nearly free electron model”.

Specifically it only applies in 1-dimension, but it's instructive to examine because of the insight into how the band gaps come about.



Electron in a Periodic Potential

To solve the Schrödinger equation in general for a three-dimensional case will produce a massive system of matrix equations with periodic components in both the Bloch wavefunction and the mean field potential.

If $U(x)$ is the potential energy of an electron in a lattice with constant a , $U(x)$ is periodic such that $U(x) = U(x+a)$, the potential can be expanded in a Fourier series of the reciprocal lattice vectors G .

$$U(x) = \sum_G U_G e^{iGx}$$

In a real crystal, the magnitude of U_G tends to decrease rapidly with increasing G . If $U(x)$ is a real function then we can write

$$U(x) = \sum_{G>0} U_G (e^{iGx} + e^{-iGx}) = 2 \sum_{G>0} U_G \cos Gx$$

Bloch functions

The eigenfunctions (orbitals) that represent electrons with energy eigenvalue ϵ are solutions of the Schrödinger equation

$$\left(\frac{p^2}{2m} + \sum_G U_G e^{iGx} \right) \psi = \epsilon \psi$$

where this is given in the *one-electron approximation*: that is, $\psi(x)$ represents the probability density distribution of a single electron in the potential of the ion cores of the lattice *plus* the average potential of all the other conduction electrons.

The solutions ψ can also be written as the Fourier series in all the wavevector components permitted by the boundary conditions,

$$\psi = \sum_k C_k e^{ikx}$$

This is unlike the phonon case because the electromagnetic field exists at all points and not just the position of the ion cores.

Crystal Momentum of an Electron

There are a few different ways to describe the importance of the wavevector k used to label the Bloch functions.

- If the lattice potential vanishes (or is very small compared to the electron energy, the Bloch function k becomes identical to the wavevector of the free electron with $\psi_k(r) = e^{ik \cdot r}$.
- Under a translation vector T that takes the electron from r to $r+T$, $\exp(ik \cdot r)$ is the phase factor that multiplies the Bloch function.
- The quantity $\hbar k$ is called the *crystal momentum* of the electron and enters into the conservation laws that govern collision processes between electrons and phonons: $k + q = k' + G$.

The Empty Lattice Approximation

Actual band structures in crystals are usually displayed as plots of energy vs. wavevector in the first Brillouin zone (any vector outside the first zone is reduced to its equivalent by subtracting the appropriate lattice vector).

This is a pretty straightforward procedure - a reciprocal lattice vector G must be found such that the true “free” electron wavevector $k = k' + G$.

$$\begin{aligned} \text{The electron energy can be written as } \epsilon(k) &= (\hbar^2/2m)(k + G)^2 \\ &= (\hbar^2/2m) [(k_x + G_x)^2 + (k_y + G_y)^2 + (k_z + G_z)^2] \end{aligned}$$

In the *empty lattice approximation*, the potential is described as strictly periodic, and weak. In this approximation neither the lattice translation nor electron interactions affect the electron energy.

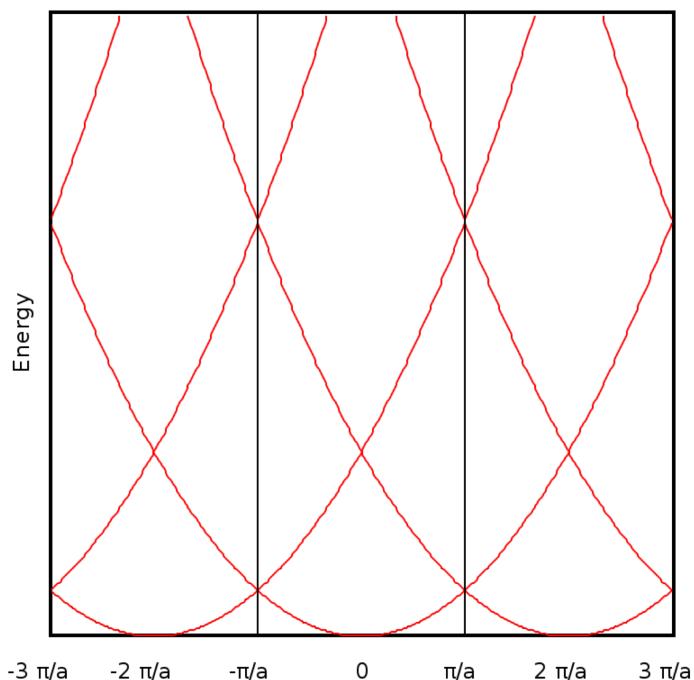
Although the potential is weak, the extent of the crystal is large, so eventually free electrons are scattered and undergo Bragg reflections from the crystal planes.

The Empty Lattice Approximation

Free electrons have a very simple dispersion relation where the energy is quadratic in the wavevector k .

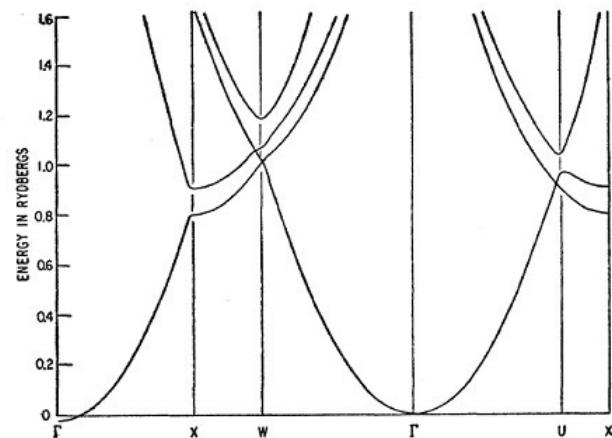
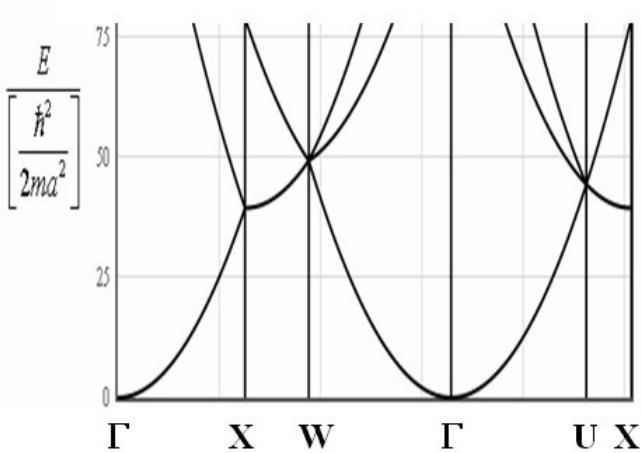
The figure shows the electron band structure for an empty lattice.

Note that the size of the forbidden band gaps has collapsed to zero, because perturbing potential approaches zero.



The Empty Lattice Approximation

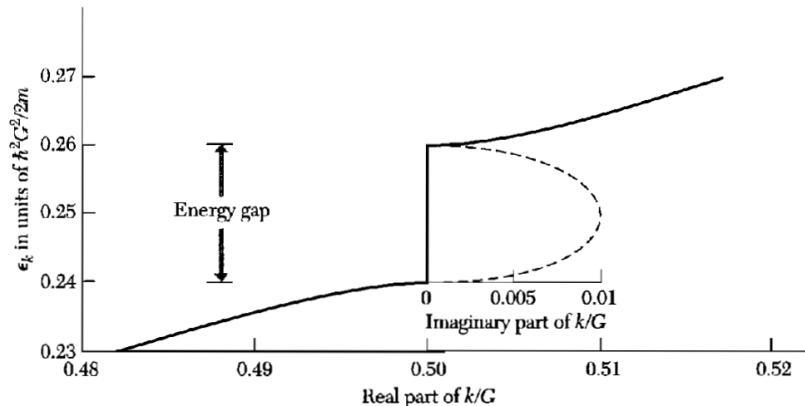
In real crystals, the dispersion relations look like paraboloids centred around each reciprocal lattice point, except they bend to hit the Brillouin zone boundaries at 90° (so the group velocity goes to zero there).



Left: Empty lattice approximation for an fcc crystal. Right: band structure for aluminium (an fcc metal). Figure from Harrison (1960), Phys. Rev., vol. 118, p. 1182. Note that band gaps exist in the real metal.

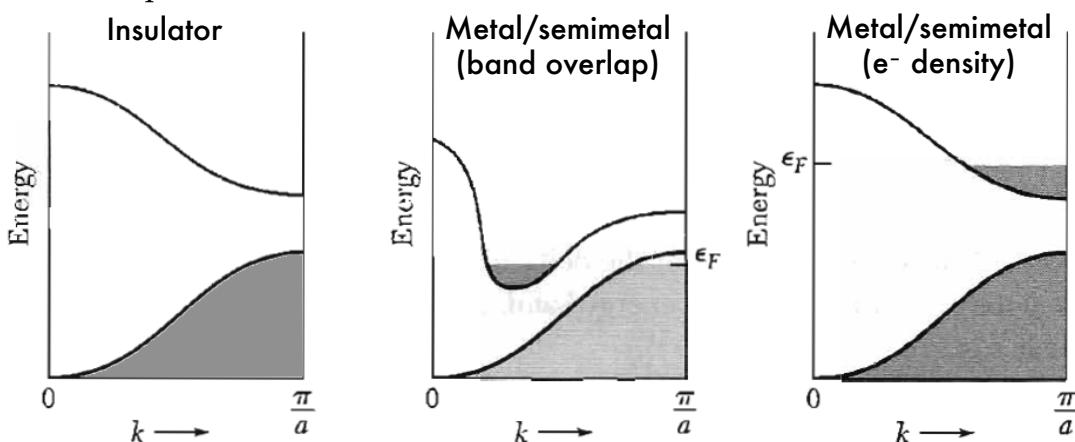
Summary of Band Theory

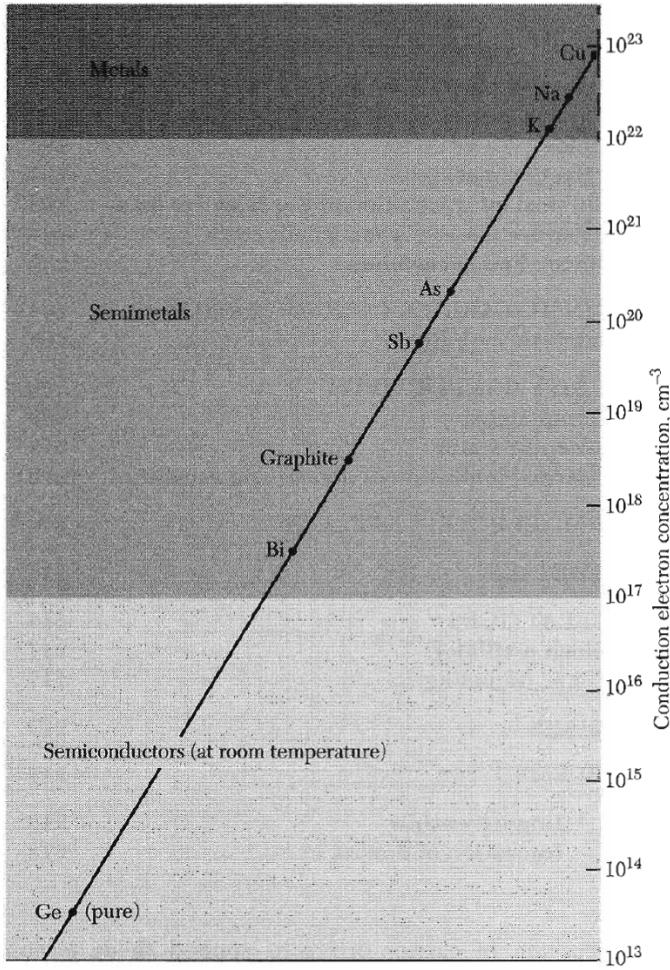
- Solutions of the wave equation in a periodic lattice are of the Bloch form $\psi_k(r) = u_k(r) \exp(ik \cdot r)$, where $u_k(r)$ is invariant under a lattice translation.
- There are energy ranges for which no (real) Bloch wave solutions of the wave equation exist. These forbidden regions have *complex k*. Note that in a 3-d, unbounded crystal, k must be real in order to keep the amplitude finite.
- Along a surface or an edge, complex wave vectors can exist (and along the edges of the Brillouin zone, are guaranteed to exist).



Summary of Band Theory

- Far from the energy gap boundaries, the wave functions may be approximated by plane waves (free electrons). Closer to the zone boundaries at $|G|/2$, wave functions may be approximated in many cases as the sum of a couple of plane waves., e.g., $\psi_k(x) \approx C_k e^{ikx} + C_{k-G} e^{i(k-G)x}$
- The number of orbitals in a band is $2N$, where N is the number of primitive cells in the specimen.





Conductivity & Band Theory

Density of charge carriers at room temperature for metals, semimetals, and pure semiconductors, at room temperature.

Semiconductors have strongly temperature dependent electrical conductivity and tend to act as insulators at $T = 0$.

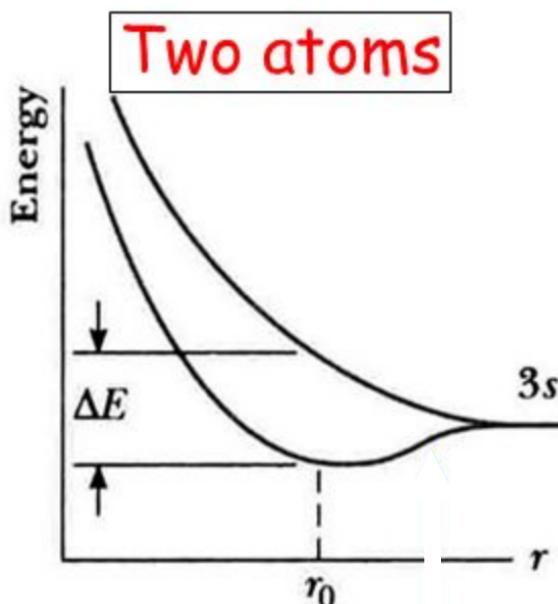
Transistors, diodes, detectors, photovoltaic cells, thermistors, etc. are frequently built around semiconductors.

Formation of Bands - Qualitative

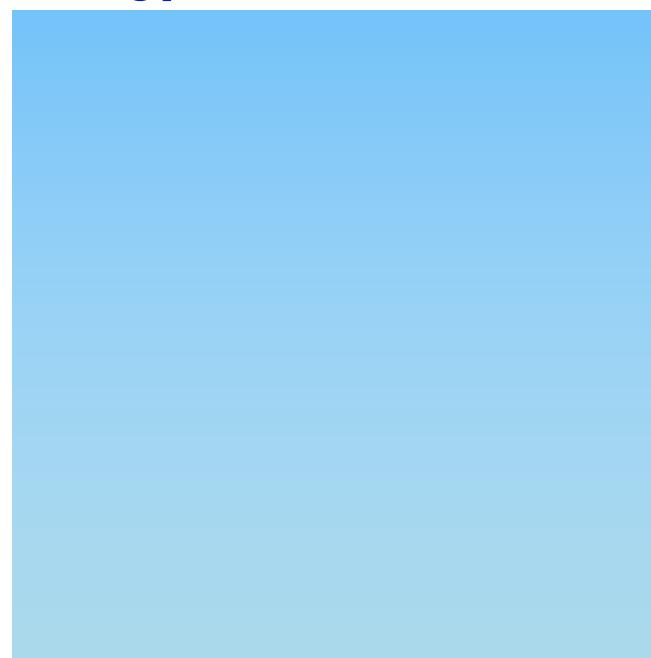
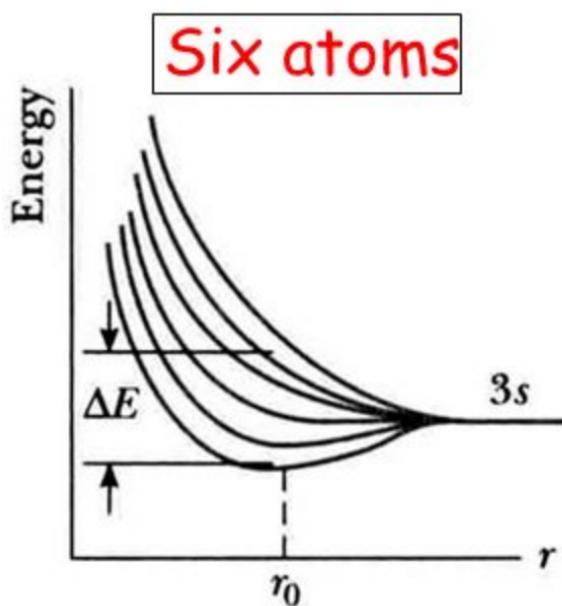
When two identical atoms are infinitely far apart, their electron orbitals do not affect each other. Taking for example the $3s$ orbital, the electrons in one sodium atom will have exactly the same energy with respect to the nucleus as any other well-separated sodium atom.

At close separations, the electron wavefunctions overlap, and cannot be considered independently. Instead we have one wavefunction describing the entire system, and the energy states will split depending on whether the wavefunctions add or subtract.

They can't be equal because of the Pauli exclusion principle.



Continuous Energy Bands



The 3s level has now become a 3s band. N levels for N atoms.

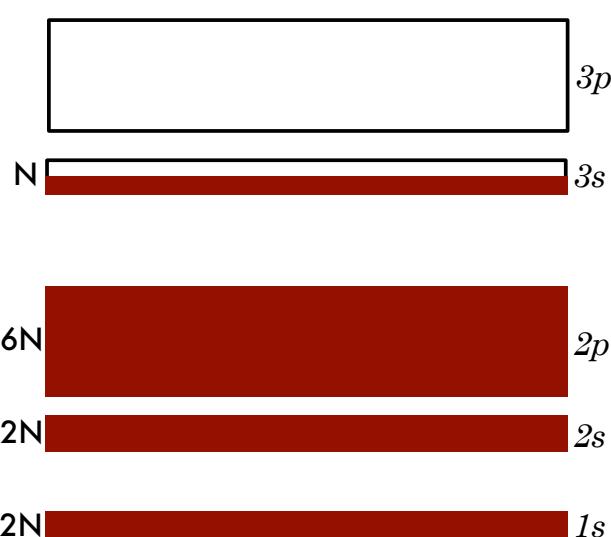
Band Capacity and Gaps

If a band is created by N atoms coming together, the capacity of the band can be taken from the degeneracy of the original (separated) energy levels, which can each hold $2(2l+1)$ electrons.

A picture of the bands of the ground state of sodium is shown at right.

The 3s band could accommodate $2N$ electrons, but is only half-filled, making sodium metal a good electrical conductor.

At $T > 0$, the 3s electrons could absorb a small amount of energy and be excited into the 3p state. This happens according to the Fermi-Dirac distribution function.



If the Fermi level corresponds to the top of an energy band, and there is a large gap between that band and the next, then the material acts as an insulator.

In many cases these properties could be approximately worked out just by considering the atomic properties of the elements in isolation. Other cases seem counter-intuitive, e.g., magnesium, which has a filled $3s$ shell yet is a good conductor.

The reason for this is that the separation of the atoms is so small that the $3s$ and $3p$ orbitals are merged together into a single band with a capacity of $8N$ electrons. However, only $2N$ of the available states are filled, and so Mg is a good electrical conductor.

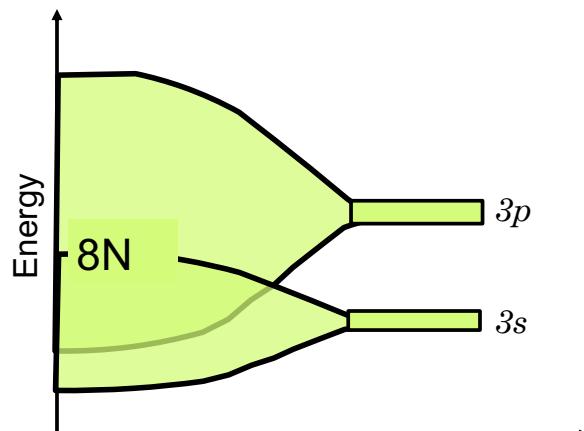


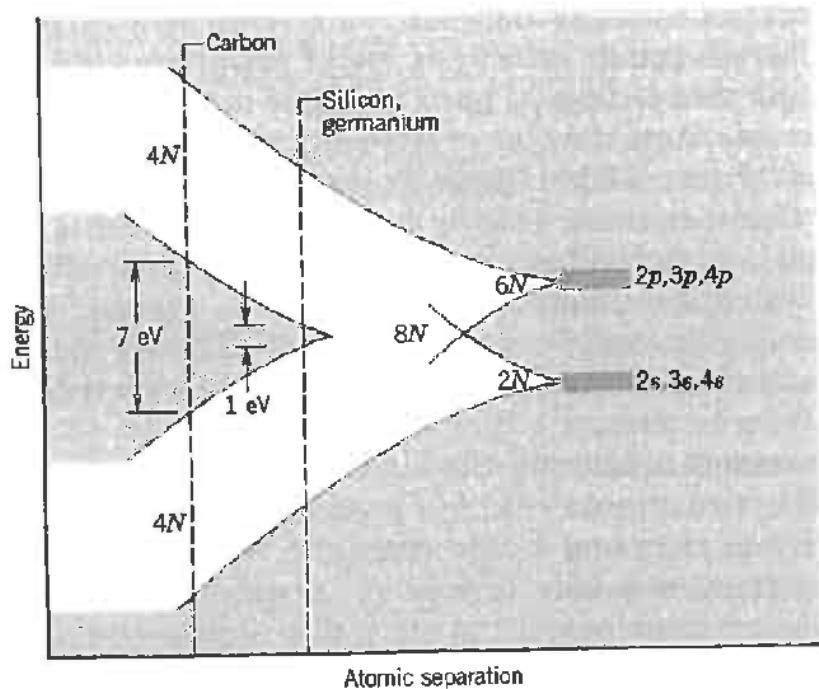
Image: sketch of band structure of Mg

Semiconductors

In carbon, there is mixing of the bands at moderate separations, as for magnesium.

For even closer separations, the mixed single band splits again, into two levels each with a capacity of $4N$.

Carbon is thus an insulator. Silicon and germanium have the same band structure, but the band gap is smaller, making them semiconductors.



Intrinsic Semiconductors

A semiconductor has a band gap of order 1 eV between the valence band and the conduction band. At $T = 0$ there is no thermal excitation across the band gap, and the material acts as an insulator. In contrast to a conductor, the conductivity increases with temperature.

At room temperature, the thermal energy kT is about 0.025 eV. The band gap is about 40 times the thermal energy, and so according to the Fermi-Dirac distribution, only about 1 in 10^9 electrons can contribute to the conduction.

The approximately exponential dependence of electron occupation on E_{gap}/kT accounts for the extreme variation in conductivity between conductors, semiconductors, and insulators for band gaps that range from 0 - 10 eV.

Holes

When an electron is excited across the band gap into the conduction band, it leaves behind a hole in the charge density of the valence band. The number of holes in the lower band is exactly equal to the number of electrons excited into the upper band.

If an electric field is applied to the semiconductor, the electrons in the conduction band will follow the force, producing an electric current.

The same is true of the electrons in the valence band, but there are very few available energy states for them to occupy. The net effect is that the vacant holes appear as quasiparticles that have positive charge, and travel in the opposite direction from the electrons in the conduction band.

The current in a semiconductor has two contributions: from the electrons in the upper band, and from the holes in the lower band. These will in general not be equal because the hole properties are determined by the totality of wavefunctions of electrons in the valence band. (Typically an individual hole acts as if it has larger mass than an individual electron).

Impurity Semiconductors

Impurities in a crystal can strongly influence the potential due to the lattice structure, thus changing the electrical conductivity.

Even for impurities at the level of 1 part in 10^6 , the impurities can dominate the electrons in the conduction band, which are only at the level of 1 in 10^9 for semiconductors. E.g., the addition of boron to silicon in the amount 1 part in 10^5 increases the conductivity by a factor of 10^3 at room temperature.

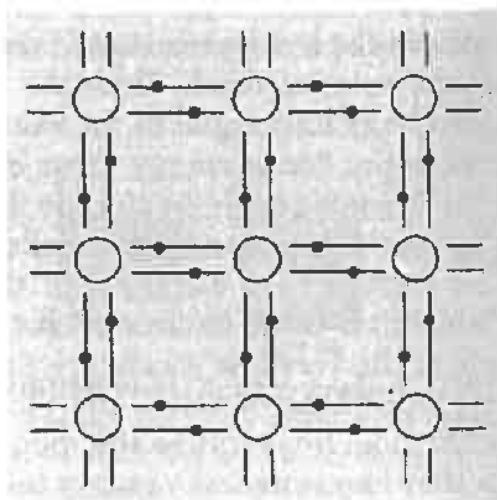
Thus the properties of semiconductors can be very carefully engineered by carefully controlling the amount and type of impurities introduced into the crystal (known as *doping*).

The impurities come in two basic types:

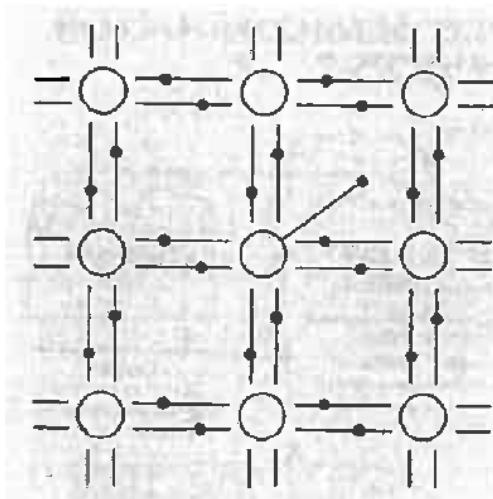
- additional electrons are added to the conduction band, or
- additional holes are added to the valence band.

Semiconductors: Bond Structure

In materials like silicon or germanium, there are four valence electrons in the hybrid orbitals.



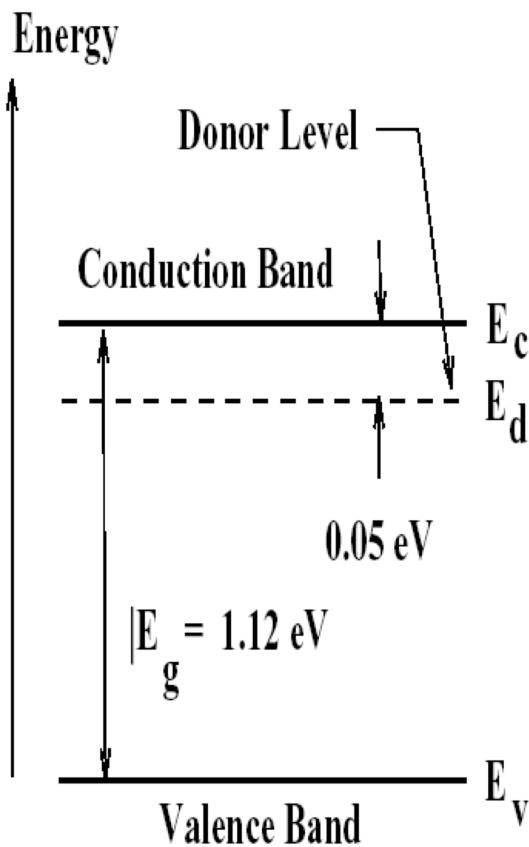
If a Ge or Si atom is replaced by a pentavalent atom, e.g., As or Sb, the fifth electron will not participate in bonding, but is relatively weakly bound to the impurity atom.



Donor States

The donor electrons sit just a fraction of an eV below the conduction band, and can easily be excited at room temperature.

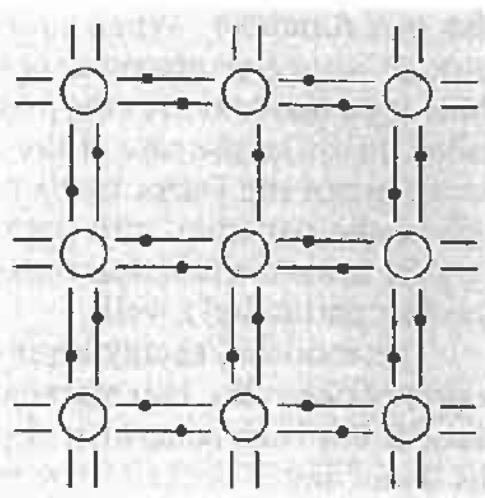
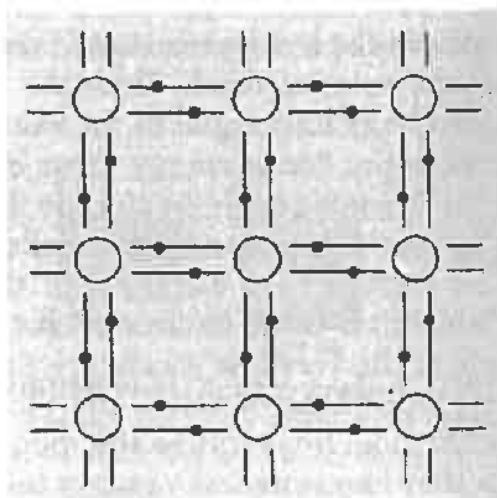
A semiconductor with this type of impurity is called an *n*-type semiconductor, because the contribution to the conductivity comes from an excess of electrons.



Semiconductors: Bond Structure

In materials like silicon or germanium, there are four valence electrons in the hybrid orbitals.

If a Ge or Si atom is replaced by a trivalent atom, e.g., B, Ga, or In, there will be a hole created, that can contribute to the conductivity.



Acceptor States

The hole attracts electrons from the upper levels, which is equivalent to saying the hole excited into a new state known as an acceptor state.

The acceptor states form discrete levels just above the valence band.

This type of semiconductor is called a *p*-type semiconductor, because the holes are the major contributor to the conductivity.

