Junction analysis and temperature effects in semi-conductor heterojunctions
by Naresh Tandan

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in Electrical Engineering
Montana State University

Abstract:
A new classification for semiconductor heterojunctions has been formulated by considering the different mutual positions of conduction-band and valence-band edges.

To the nine different classes of semiconductor heterojunction thus obtained, effects of different work functions, different effective masses of carriers and types of semiconductors are incorporated in the classification. General expressions for the built-in voltages in thermal equilibrium have been obtained considering only nondegenerate semiconductors. Built-in voltage at the heterojunction is analyzed. The approximate distribution of carriers near the boundary plane of an abrupt n-p heterojunction in equilibrium is plotted. In the case of a p-n heterojunction, considering diffusion of impurities from one semiconductor to the other, a practical model is proposed and analyzed. The effect of temperature on built-in voltage leads to the conclusion that built-in voltage in a heterojunction can change its sign, in some cases twice, with the choice of an appropriate doping level. The total change of energy discontinuities (&ΔEc + ΔEv) with increasing temperature has been studied, and it is found that this change depends on an empirical constant and the 0°K Debye temperature of the two semiconductors. The total change of energy discontinuity can increase or decrease with the temperature. Intrinsic semiconductor-heterojunction devices are studied? band gaps of value less than 0.7 eV. are suggested if such devices are to be used at room temperature. Built-in voltage in the case of intrinsic semiconductor heterojunction devices varies slightly with the temperature. With -proper choice of effective masses, the built-in voltage remains essentially constant with the temperature. The advantages of intrinsic semiconductor devices are discussed. A three-dimensional energy-distance-momentum diagram is described; it suggests a change in the wave vector of a carrier, in transit from one semiconductor to other. Thus, the probability of transition of a carrier across the heterojunction is reduced.

JUNCTION ANALYSIS AND TEMPERATURE EFFECTS IN SEMICONDUCTOR HETEROJUNCTIONS

by

NARESH TANDAN

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

Approved:

Head, Major Department

Chairman, Examining Committee

Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

March, 1974

ACKNOWLEDGMENTS

I wish to express my deep appreciation to my advisor, Professor N. A. Shyne, for his thoughtful guidance and encouragement in this research. Thanks are also due to Professor J. P. Hanton for his useful suggestions and discussions in this research.

My thanks to the department head, Professor P. E. Uhlrich, for his encouragement and financial support throughout my stay at Montana State University.

My thanks to my parents without whose moral support and encouragement throughout all my graduate work I could not have completed this work. Finally, thanks to my wife, Anita, for typing the entire first draft.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Subscript "1" represents narrow-gap semiconductor

Subscript "2" represents broad-gap semiconductor

C	Capacitance in Farad.
E_c	Energy of conduction band gap in eV.
ΔE_c	Energy discontinuity in energy band diagram of a heterojunction (associated with electron affinity difference) in eV.
E_F	Fermi level, n and p semiconductor in eV.
E_{Fi}, E_{FI}	Intrinsic Fermi level in semiconductor in eV.
E_g	Energy band gap in eV.
E_v	Energy of valence-band edge in eV.
ΔE_v	Energy discontinuity in the valence-band energy diagram of a heterojunction in eV.
h	Planck's constant in Joule-sec.
$\hbar$	$\frac{h}{2\pi}$
J_n	Electron current density in A cm^{-2}.
J_p	Hole current density in A cm^{-2}.
K	Momentum wave vector.
k	Boltzmann's constant in eV° k^{-1}.
m_e^*	Electron effective mass Kg.
m_{de}^*	Density-of-state effective mass for electron in Kg.

m^*_{dh}	Density-of-state effective mass for holes in Kg.
m^*_{lh}	Light hole mass in Kg.
m^*_{hh}	Heavy hole mass in Kg.
m^*_l	Effective mass along the longitudinal ellipsoidal energy surface in Kg.
m^*_t	Effective mass along the transverse ellipsoidal energy surface in Kg.
m^*_h	Effective mass of hole.
N_a, N_A	Acceptor density in cm^{-3}.
N_C	Effective density of energy states at the conduction band edge in cm^{-3}.
N_d, N_D	Donor density in cm^{-3}.
N_V	Effective density of energy states at the valence-band edge in cm^{-3}.
n_o, n	Free electron density in the conduction band in cm^{-3}.
n_d	Unionized donors in cm^{-3}.
n_i	Intrinsic carrier concentration in a semiconductor cm^{-3}.
p_o, p	Hole density in cm^{-3}.
P_a	Unionized acceptors in cm^{-3}.
q	Charge of an electron, 1.6×10^{-19} conlomb.
T	Absolute temperature in °K.
V_{BHT}	Built-in voltage in heterojunction in volts.

V_{BHM}	Built-in voltage in homojunction in volts.
x	Distance in cm.
γ	Emitter injection efficiency.
ε	Dielectric constant of the semiconductor in F cm^{-1}.
μ_n	Mobility of electrons in cm^2 v^{-1} sec^{-1}.
μ_p	Mobility of holes in cm^2 v^{-1} sec^{-1}.
ϕ_m	Work function of a semiconductor in eV.
χ_e	Electron affinity of a semiconductor in eV.
χ_h	$(\chi_e + E_g)$ in eV.
$\psi(x)$	Electrostatic potential difference as a function of distance in volts.
σ	Conductivity in mhos.
θ	Debye temperature.

ABSTRACT

A new classification for semiconductor heterojunctions has been formulated by considering the different mutual positions of conduction-band and valence-band edges. To the nine different classes of semiconductor heterojunction thus obtained, effects of different work functions, different effective masses of carriers and types of semiconductors are incorporated in the classification. General expressions for the built-in voltages in thermal equilibrium have been obtained considering only nondegenerate semiconductors. Built-in voltage at the heterojunction is analyzed. The approximate distribution of carriers near the boundary plane of an abrupt n-p heterojunction in equilibrium is plotted. In the case of a p-n heterojunction, considering diffusion of impurities from one semiconductor to the other, a practical model is proposed and analyzed. The effect of temperature on built-in voltage leads to the conclusion that built-in voltage in a heterojunction can change its sign, in some cases twice, with the choice of an appropriate doping level. The total change of energy discontinuities ($\Delta E_C + \Delta E_V$) with increasing temperature has been studied, and it is found that this change depends on an empirical constant and the 0°K Debye temperature of the two semiconductors. The total change of energy discontinuity can increase or decrease with the temperature. Intrinsic semiconductor-heterojunction devices are studied; band gaps of value less than 0.7 eV. are suggested if such devices are to be used at room temperature. Built-in voltage in the case of intrinsic semiconductor heterojunction devices varies slightly with the temperature. With proper choice of effective masses, the built-in voltage remains essentially constant with the temperature. The advantages of intrinsic semiconductor devices are discussed. A three-dimensional energy-distance-momentum diagram is described; it suggests a change in the wave vector of a carrier, in transit from one semiconductor to other. Thus, the probability of transition of a carrier across the heterojunction is reduced.

Chapter 1

INTRODUCTION

Historical Review

A semiconductor heterojunction is a junction between two dissimilar semiconductors where the crystal structure is continuous across the interface. Semiconductor heterojunction research is an important area of device study which has evolved from the research of the last decade. The barriers introduced into the energy-band diagram by the energy-gap difference of two semiconductors allows a new degree of freedom to the device designer. For example, in GaAs injection lasers [1], the addition of $Al_xGa_{1-x}As$ confinement barriers has resulted in a major reduction in 300°K threshold current densities. Electrooptical effects in heterojunctions that are promising include infrared to visible upconversion systems [2], the window effect [3], solar cells [4] and photo transistors [5]. High yield photo cathodes [6], cold cathodes [7] and electron multipliers [8] of the negative-electron affinity type represent yet other families of devices where heterojunctions are involved.

The heterojunction structure was first considered by Preston [9]. Gubanov [10] produced an analysis of

heterojunctions with n-n, p-p and p-n combinations. Shockley [11] proposed a circuit device incorporating a change in the magnitude of the forbidden gap in the transition region of p-n junction. Kroemer [12] proposed the use of a heterojunction as a wide-gap emitter to increase the injection efficiency of transistors. The Anderson theory [3] is discussed in the following section and a brief summary of other theories is discussed on page 7.

Anderson's Theory. The energy-band model of an ideal abrupt heterojunction without interface states was proposed by Anderson based on the previous work of Shockley. This model is important as it can adequately explain most of the transport processes, and only a slight modification [13] of the model is needed to account for non-ideal cases such as interface states. Consider the energy-band profile of the two isolated semiconductors as shown in Figure 1. The two semiconductors are assumed to have different band gaps (E_{g_1},E_{g_2}), different dielectric constants (ε_1,ε_2), different work function (ϕ_{m_1},ϕ_{m_2}), and different electron affinities (χ_{1e},χ_{2e}). The work function is defined as the energy required to remove an electron from the Fermi level (E_F) to a position just outside of

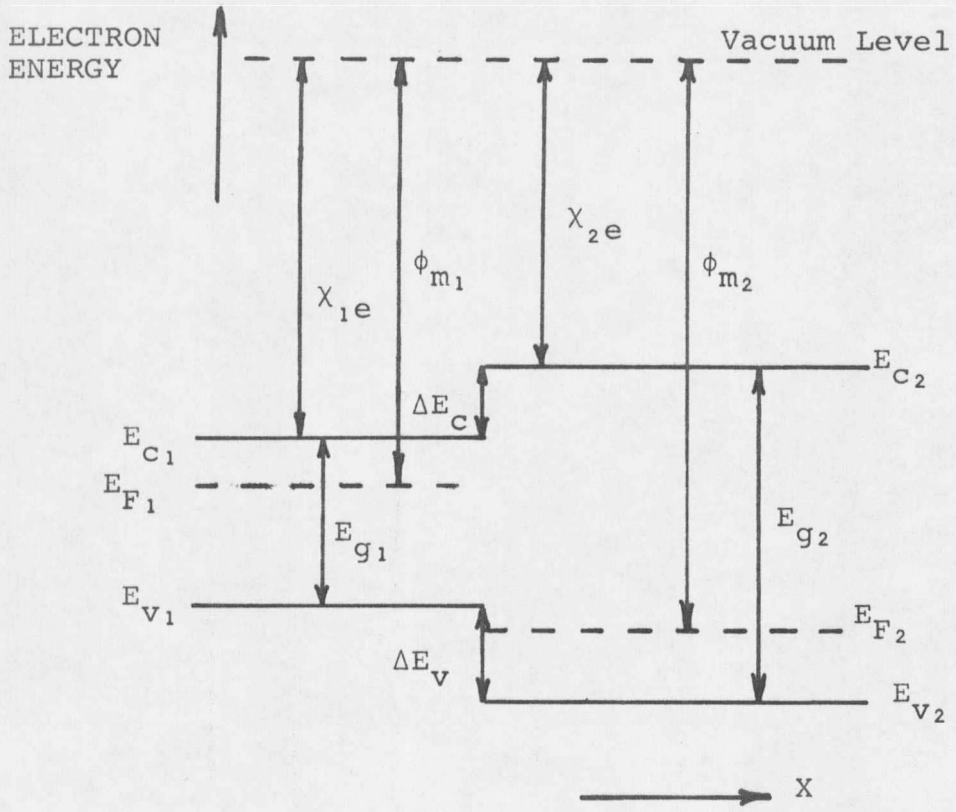


Figure 1. Energy-band diagram for two isolated semiconductors in which space charge neutrality is assumed to exist in every region.

the material. The bottom of the conduction-band is represented by E_c and the top of the valence-band is represented by E_v. The subscripts 1 and 2 refer to the narrow-gap and wide-gap semiconductors, respectively. Electron affinity is defined as the energy required to remove an electron from the bottom of conduction-band to a position just outside of the material. It is assumed that space charge neutrality exists in every region and thus band edges ($E_{c_1}, E_{c_2}, E_{v_1}, E_{v_2}$) are shown to be horizontal. The difference in energy of the conduction-band edges in the two materials is represented by ΔE_c and that in the valence-band edges is ΔE_v.

A junction formed between an n-type narrow-gap semiconductor and a p-type wide-gap semiconductor is considered. The energy-band profile of such a junction at equilibrium is shown in Figure 2.

The electrostatic potential difference between any two points can be represented by the vertical displacement of the band edges between these two points in a semiconductor. An electrostatic field can be represented by the slope of the edges on a band diagram. The total built-in voltage (V_{BHT}) is equal to the sum of the partial built-in

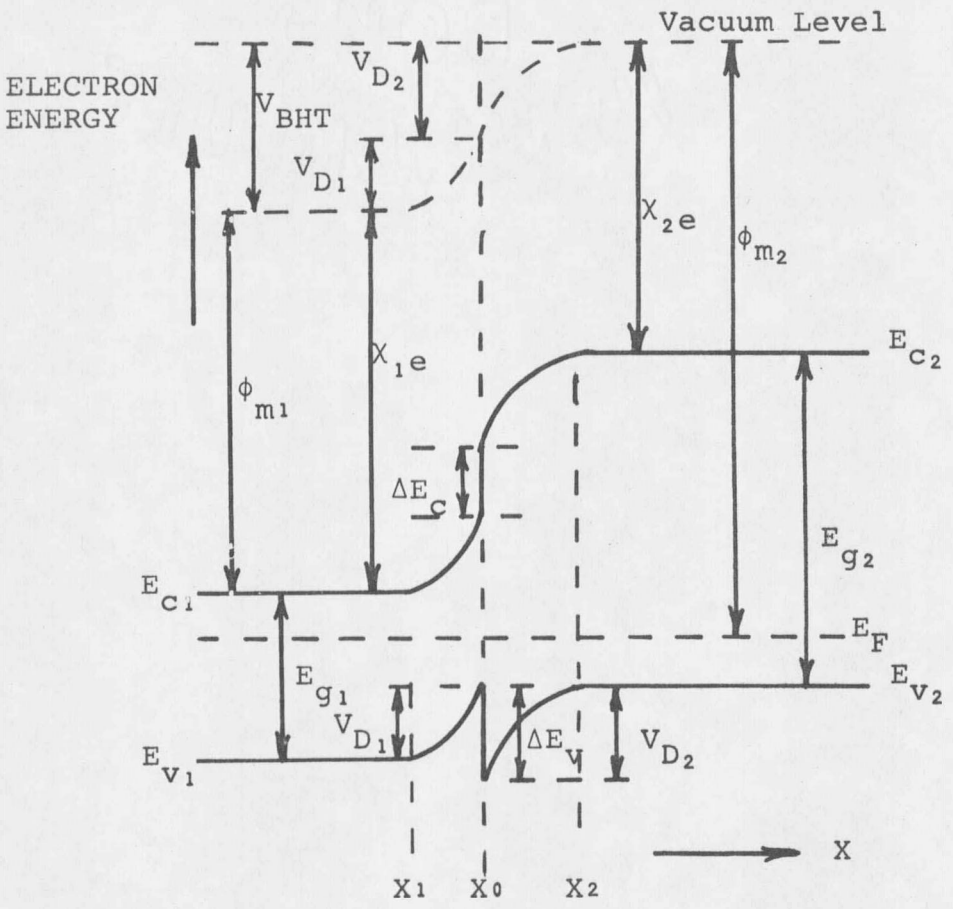


Figure 2. Energy-band diagram of n-p heterojunction in thermal equilibrium.

voltages ($V_{D_1} + V_{D_2}$), where V_{D_1} and V_{D_2} are the electrostatic potentials supported at equilibrium by semiconductors 1 and 2, respectively. Since the Fermi level must coincide on both sides in equilibrium and the vacuum level is everywhere parallel to the band edges and is continuous, the discontinuity in the conduction band edges (ΔE_c) and valence-band edges (ΔE_v) is invariant with doping in those cases where E_g and χ are not functions of doping (i.e., nondegenerate semiconductors). A degenerate semiconductor is one in which Fermi level lies in the conduction or in the valence-band. It is assumed that the electron density is zero for $X>X_1$, and the hole density is zero for $X<X_2$. Then at any point in the transition region, the space charge density is merely equal to the donor or acceptor density.

$$p_1 = q\, N_{D_1} \qquad X_1 \leq X \leq X_0$$

$$p_2 = q\, N_{A_2} \qquad X_0 \leq X \leq X_2$$

where N_{D_1} is the donor concentration in semiconductor 1 and N_{A_2} is the acceptor concentration in semiconductor 2. The solution of Poisson's equation in the one dimensional case with appropriate doping levels gives the solution for an abrupt junction as follows [3]:

Transition region width $W = (X_2 - X_0) + (X_0 - X_1)$

$$= \left[\frac{2\,\varepsilon_1 \varepsilon_2 (V_{BHT} - V)(N_{A_2} + N_{D_1})}{q(\varepsilon_1 N_{D_1} + \varepsilon_2 N_{A_2}) N_{D_1} N_{A_2}} \right]^{1/2}$$

The transition capacitance is given by a generalization of the result for homojunctions.

$$C = \left[\frac{q\, N_{D_1} N_{A_2}\, \varepsilon_1 \varepsilon_2}{2(\varepsilon_1 N_{D_1} + \varepsilon_2 N_{A_2})} \; \frac{1}{(V_{BHT} - V)} \right]^{1/2}$$

Figure 2 shows that the potential barrier for electrons is considerably greater than that for holes, and so hole current will predominate. The application of Shockley's homojunction diffusion theory, along with a diode emission model is sufficient to predict the saturation current and current-voltage characteristics of the heterojunction.

Summary of other theories. Most of the other theories have modified the ideal model of semiconductor heterojunctions as proposed by Anderson. Perlman and Feucht [13,14] used a classical kinetic-emission model to predict the current-voltage characteristic of an abrupt p-n heterojunction. This took into account the effect of changes in electron affinity, electron effective masses, dielectric constant and band gap at the junction. The p-n

heterojunction was found to have two operating modes; one similar to a homojunction, where minority-carriers build up at the depletion region edge limiting current, and another similar to a metal-semiconductor junction, where the current is limited by a potential barrier in the n-type semiconductor. On increasing the forward bias, the homojunction mode of operation changes to the metal-semiconductor mode.

Oldham and Milnes [15] studied the effect of interface states. For a lattice mismatch of the order of 2 to 4%, the dislocations were assumed to lie in a sheet and to be similar to grain boundaries. They concluded that the interface resembles a low-density,free surface with edge dislocations producing deep states in the gap. Due to the presence of interface states, the ideal model of Anderson gets distorted.

Van Ruyen, Papenhuijzen and Verhoeven [16] considered two semiconductors, each with a free surface. They considered a heterojunction with a significant number of interface states. Interface states were thought to play a decisive role since they can store sufficient charge to make the surfaces behave like a thin metal layer. Contact

between the two different surfaces leads to the formation of a dipole layer.

Prospects of Development and Technological Application

Let us consider the way in which p-n homojunctions were developed [17]. The first incomplete theory of their behaviour was given by Davydoy in 1938 and a full design theory incorporating the ideas of injection was published by Shockley in 1949. Early experimental germanium p-n junctions were described in 1950 and, as they were in the form of grown junctions, they were of high quality but suffered from a high cost of manufacture. The alloy-junction provided a cheap and usable junction but did not give high performance or good reliability. The diffusion process introduced about 1956 led to higher quality devices with better performance, but it was really the discovery in 1959 of the planar process and subsequent work on epitaxial growth which made possible the production of high quality transistors and diodes and paved the way for the integrated circuits. Excellent development and research work has been carried out on other materials but no really significant production of p-n junction devices has

been realized in materials other than silicon, with the possible exceptions of germanium alloy devices and gallium arsenide (phosphide) diodes. Silicon devices have been developed over a period of about 25 years and have now reached a state where integrated circuits of high complexity and high reliability can be manufactured in quantity.

The problems with heterojunctions are more difficult to solve both from theoretical and practical points of view. The basic theoretical ideas were published by Anderson approximately 10 years ago and since that time a considerable amount of work has been carried out. The outlook for production of heterojunctions for normal electronic devices seems uncertain. To compete with normal silicon devices which have been manufactured rather than produced by research workers in a laboratory, the heterojunction devices would need to be more reliable, or to have improved characteristics, or to be cheaper. It seems unlikely at present for heterojunctions to compete with homojunction devices for normal use. However, looking at special applications, there is a hope for technological applications. The main area for development seems to be

in the region of applying optical properties, particularly for detectors, electroluminescent lamps and lasers. Single heterojunctions lasers [18] are manufactured at present, though still in the process of development.

The device potential of the heterojunction is greatly reduced when significant numbers of interface states are present at the junction. Therefore, the development of semiconductor pairs with a very small mismatch of lattice constant is in progress, so as to minimize the number of interface states. One such developed pair is GaAs - $Al_xGa_{1-x}As$: it finds application in heterojunction lasers.

Contents of the Thesis and Assumptions

Heterojunctions are classified as graded and abrupt, depending on whether the junction is graded or abrupt. Another classification involves the predominant impurity type on either side of the junction. n-n or p-p junctions are called isotype heterojunctions and n-p or p-n junctions are called anisotype heterojunctions. A new classification is presented, considering different mutual positions of conduction and valence-band edges. Nine different classes of a semiconductor heterojunction are obtained. Each such

class has some unique properties, which are discussed, considering the effect of different work functions. Different effectives masses of carriers and different types of semiconductors (direct or indirect gap) are incorporated in the classification. These parameters provide information regarding the structure of conduction and valence-bands, and the mobility and lifetime of carriers in the two semiconductors. A large number of unique semiconductor heterojunctions are thus possible; some of them exist at present, others may develop in the future.

A general theory that describes the formation of the electric field at the junction is presented. A general expression for built-in voltage is formulated in three different forms considering nondegenerate semiconductors. The built-in voltage is split into two components. It turns out that one component is similar to that of a homojunction and can be controlled by the doping level while the other component arises as a result of energy discontinuities and different effective masses, and is fixed once the two semiconductors are chosen. Approximate distribution of the carriers near the boundary plane of an arbitrary abrupt np heterojunction in equilibrium is

plotted. A practical model of a p-n heterojunction is proposed and analyzed. In reality, due to diffusion of impurities from one semiconductor to the other, a p-n semiconductor heterojunction consists of an isotype heterojunction and a p-n homojunction.

A theory describing the temperature dependence of carrier concentration and the evaluation of Fermi level at different temperature ranges is discussed. Two different cases are considered with $\Delta E_v > \Delta E_c$, and it shows that the built-in voltage in case of a heterojunction may change sign under certain conditions of doping level and that it may change twice with another preferred doping profile in the two semiconductors. The total change of energy discontinuities ($\Delta E_c + \Delta E_v$) with increasing temperature has been studied, and it is shown that this change depends on an empirical constant and on the Debye temperature of the two semiconductors. The total change of energy discontinuities has been computed for GaP - Si, Ge - GaAs and Ge - Si heterojunctions pairs. In the case of GaP-Si and Ge-GaAs heterojunctions, the total energy discontinuity decreases with temperature, while in the case of the Ge-Si heterojunction, the total change of energy discontinuities with temperature between 0°K to 800°K.

Intrinsic semiconductor heterojunctions are possible and are studied on the basis of intrinsic conductivity. Intrinsic conductivity of a semiconductor is found to be an important consideration. For moderate conductivity, the energy-band gap of a semiconductor should be small if the device is to be used at room temperature. A band gap of a value less than 0.7 eV. is suggested. The built-in voltage changes only slightly with the temperature in the case of an intrinsic semiconductor heterojunction. It is concluded that, if the choice of effective masses is as follows:

$$\frac{m^*_{1h}}{m^*_{2h}} \cdot \frac{m^*_{2e}}{m^*_{1e}} = 1,$$

then the built-in voltage remains essentially constant with the temperature. It is suggested that a narrow-gap intrinsic semiconductor and a broad-gap extrinsic semiconductor can also be used.

Appendix A considers a three-dimensional energy-distance-momentum diagram. Since, in general, the conduction-bands and valence-bands of two semiconductors are not alike, it is concluded that the wave vector of a

carrier should change while the carrier is crossing the heterojunction.

Chapter 2

CLASSIFICATION OF SEMICONDUCTOR HETEROJUNCTIONS

Introduction

Semiconductor heterojunctions can be classified in several ways [19]. If the material changes abruptly on an atomic scale from one material to another, the junction is referred to as "abrupt". If there is an appreciable intermixing of the components, the heterojunction is referred to as "graded".

Another classification involves the predominant type of impurity on either side of the junction. In the case of similar doping on either side of a junction, the junction is called an isotype heterojunction (n-n or p-p). An anisotype heterojunction is one with different types of doping on either side of the junction. These include n-p or p-n heterojunctions.

The properties of a semiconductor heterojunction depend on the various parameters of the two semiconductors. The two fundamental parameters are electron affinities and energy gaps. A possible system of classification can be chosen using the above parameters which acquire unique positions in the energy-versus-distance diagram. On each

of such diagrams, other parameters of interest can be incorporated, such as work function, type of semiconductors (direct or indirect) and the effective masses of carriers. External influences (such as light or applied voltage), which can modify the properties of junction, are not considered in this classification.

Classification

Consider an energy-band diagram of two isolated semiconductors with energy gaps of E_{g_1} and E_{g_2} shown in Figure 1. The two electron affinities are χ_{1e} and χ_{2e}, according to the usual practice of specifying the first semiconductor as narrow-gap and the second semiconductor as broad-gap. Only nondegenerate semiconductors are under consideration. The energy discontinuities in the conduction and valence-bands are represented by ΔE_c and ΔE_v. ΔE_c is defined as positive when $\chi_{1e} > \chi_{2e}$, negative when $\chi_{1e} < \chi_{2e}$, and zero when $\chi_{1e} = \chi_{2e}$. Similarly, ΔE_v is defined as positive when $\chi_{2h} > \chi_{1h}$, negative if $\chi_{2h} < \chi_{1h}$ and zero when $\chi_{2h} = \chi_{1h}$. The above idea is introduced in order to satisfy the relation $E_{g_2} - E_{g_1} = \Delta E_c + \Delta E_v$, irrespective of the mutual position of the bands. Different values of electron affinities, band gaps and work function

give new shapes to the energy-band diagram, resulting in new properties being shown. Therefore, considering all possible different mutual positions of the bands by attributing different values of electron and hole affinity, the following different cases are obtained.

Figure 3 represents the case of $\chi_{2h} > \chi_{1h} > \chi_{1e} > \chi_{2e}$ and shows that ΔE_c and ΔE_v both are positive. The magnitude of ΔE_c and ΔE_v are not considered important for the classification proposed. This is the most common case where heterojunctions have been studied, as for example, Ge-GaAs, AlSb-GaSb, and Ge-Si.

Figure 4 represents the case where ΔE_v is zero and ΔE_c is positive; i.e., $\chi_{1h} = \chi_{2h}, \chi_{2h} > \chi_{1e} > \chi_{2e}$. The GaAs - $Al_xGa_{1-x}As$ heterojunction is an example of this class.

Figure 5 represents the case where ΔE_v is negative, ΔE_c is positive and possibly can be greater than E_{g_1}, but less than E_{g_2}; i.e., $\chi_{1h} > \chi_{2h} > \chi_{1e} > \chi_{2e}$. The CdSe-Se and InP-GaAs heterojunctions are examples of this class.

Figure 6 represents two isolated semiconductors with $\chi_{1h} > \chi_{2h}$. $\chi_{2h} = \chi_{1e}, \chi_{1e} > \chi_{2e}$, i.e., $\Delta E_c = E_{g_2}$ and $-\Delta E_v = E_{g_1}$. No such pair exists at present.

Figure 7 represents the case where the mutual position of bands are such as to give $-\Delta E_v > E_{g_1}$ and

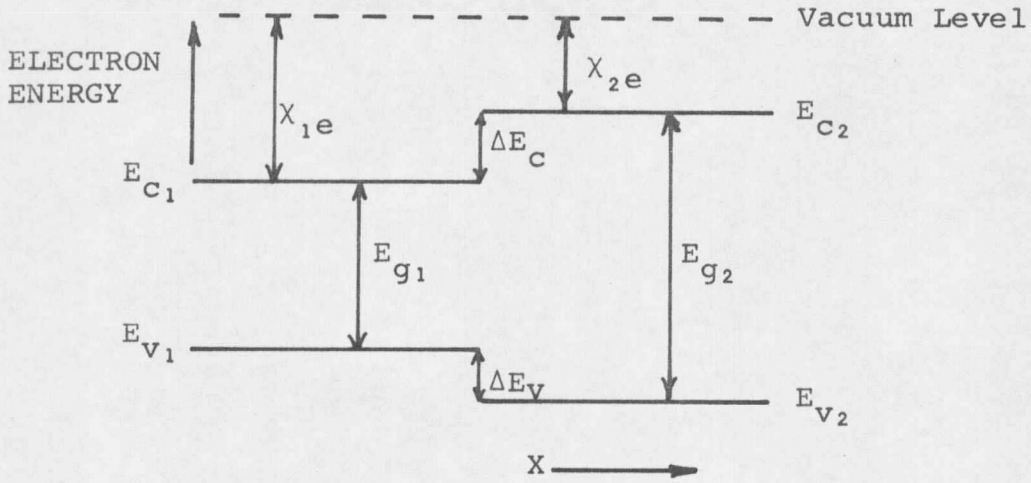


Figure 3. Energy-band diagram for two isolated semiconductors in which space charge neutrality is assumed to exist in every region with $\chi_{2h} > \chi_{1h} > \chi_{1e} > \chi_{2e}$

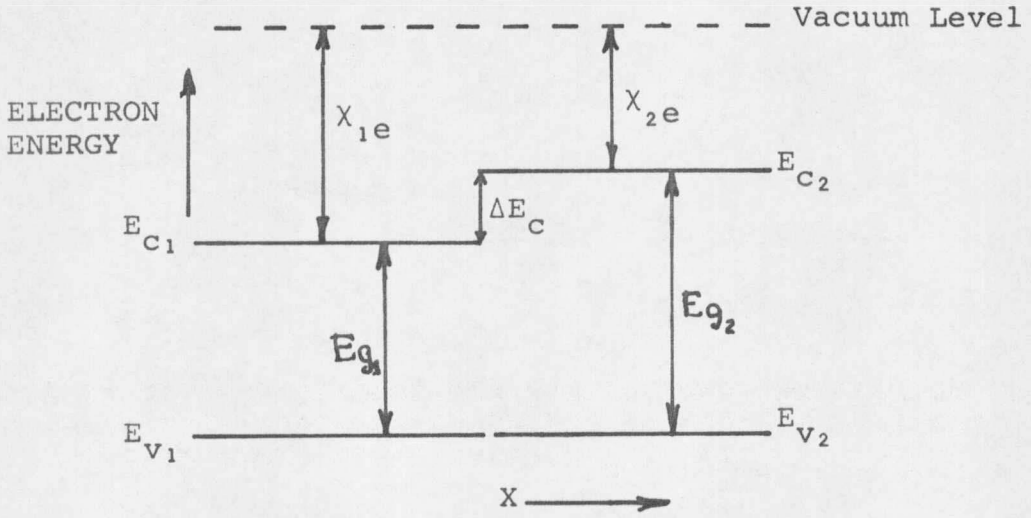


Figure 4. Energy-band diagram for two isolated semiconductors in which space charge neutrality is assumed to exist in every region with $\chi_{1h} = \chi_{2h}$; $\chi_{2h} > \chi_{1e} > \chi_{2e}$.

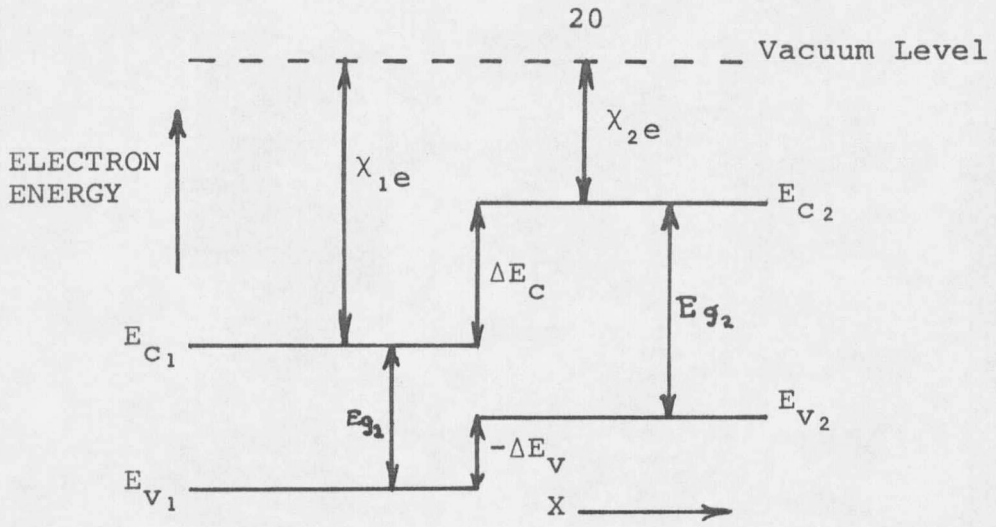


Figure 5. Energy-band diagram for two isolated semiconductors in which space charge neutrality is assumed to exist in every region with $\chi_{1h} > \chi_{2h} > \chi_{1e} > \chi_{2e}$.

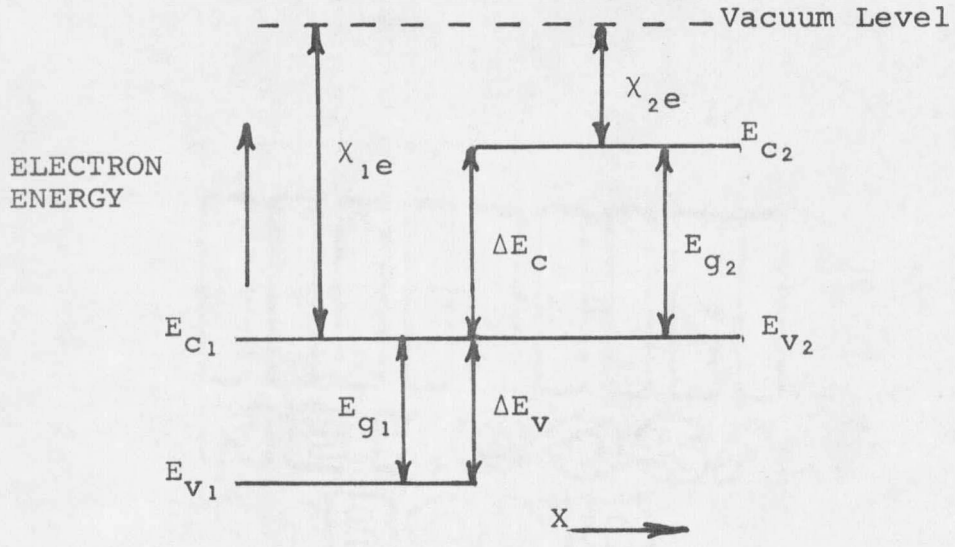


Figure 6. Energy-band diagram for two isolated semiconductors in which space charge neutrality is assumed to exist in every region with $\chi_{1h} > \chi_{1e}$; $\chi_{1e} = \chi_{2h}$; $\chi_{2h} > \chi_{2e}$.

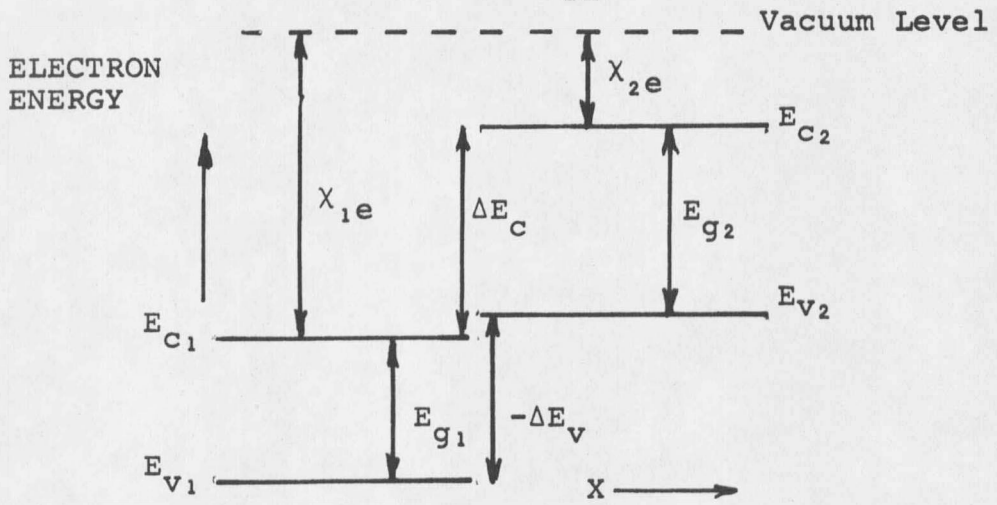


Figure 7. Energy-band diagram with two isolated semiconductors in which space charge neutrality is assumed to exist in every region with $\chi_{1e} > \chi_{2h}$.

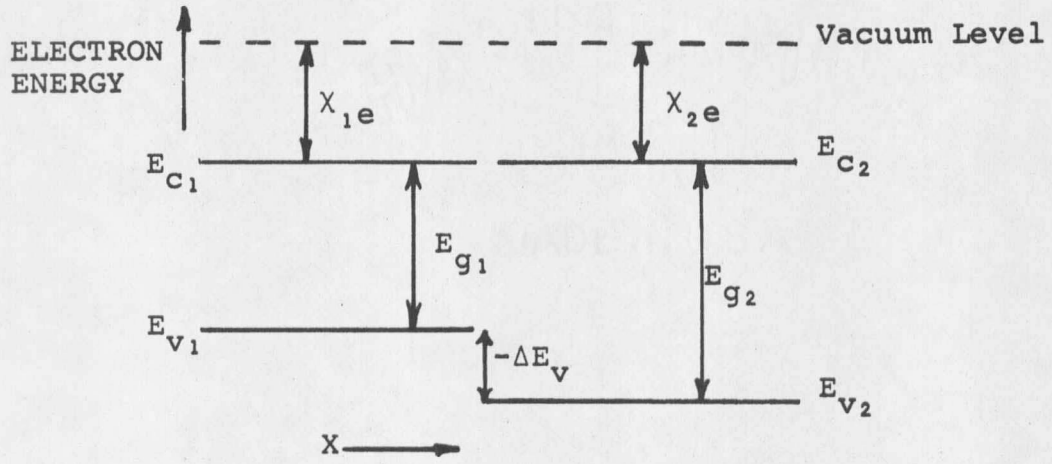


Figure 8. Energy-band diagram with two isolated semiconductors in which space charge neutrality is assumed to exist in every region with $\chi_{2h} > \chi_{1h} > \chi_{2e}$; $\chi_{2e} = \chi_{1e}$.

$\Delta E_c > E_{g_2}$. The GaAs-CsO_2 heterojunction pair is an example of this kind. Similar to Figure 4, where there was no discontinuity in the valence-band, Figure 8 represents the case where there is no discontinuity in the conduction band, $\chi_{2h} > \chi_{1h} > \chi_{2e}$; $\chi_{2e} = \chi_{1e}$ thus making $\Delta E_c = 0$ and ΔE_v is positive. The GaSb-GaAs heterojunction is an example of this class.

Figure 9 represents the case of $\chi_{2h} > \chi_{1h} > \chi_{2e} > \chi_{1e}$, resulting in a negative discontinuity in conduction-band with positive discontinuity in the valence-band as $\Delta E_v < E_{g_1}$. ZnTe-Cds heterojunction is an example of this kind.

Figure 10 represents another unique case where $\chi_{2h} > \chi_{1h}$; $\chi_{1h} = \chi_{2e}$; $\chi_{2e} > \chi_{1e}$ thus making $-\Delta E_c = E_{g_1}$, $\Delta E_v = E_{g_2}$. No such pair exists at present.

Figure 11 represents pairs of semiconductors with $\chi_{2e} > \chi_{1h}$, making $-\Delta E_c > E_{g_1}$ and $\Delta E_v > E_{g_2}$. No such pair exists at present.

Thus, nine different cases in the energy-versus-distance diagram are obtained. Let us call each one a working space. In each working space, other important constraints can be imposed, such as work function, effective masses of carriers and type of semiconductors. Hence,

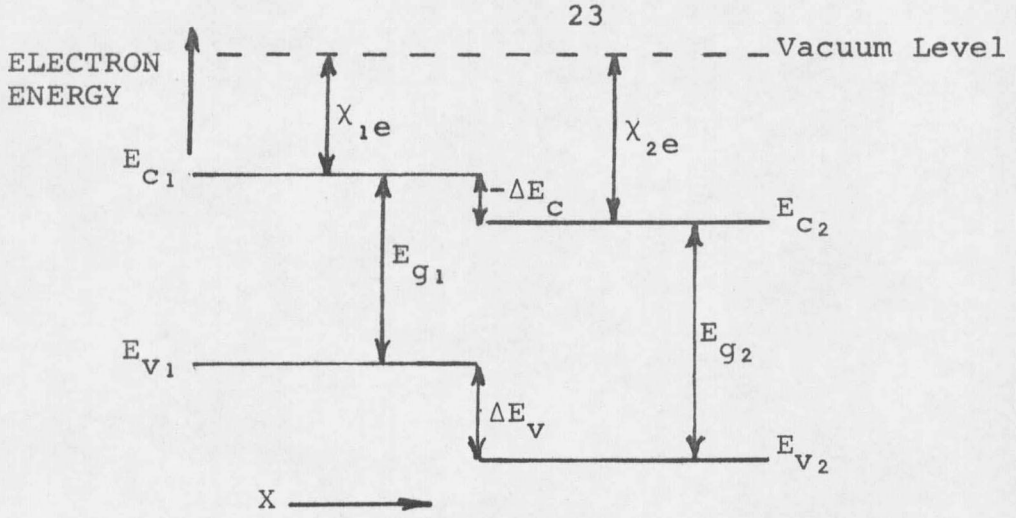


Figure 9. Energy-band diagram for two isolated semiconductors in which space charge neutrality is assumed to exist in every region with $\chi_{2h} > \chi_{1h} > \chi_{2e} > \chi_{1e}$.

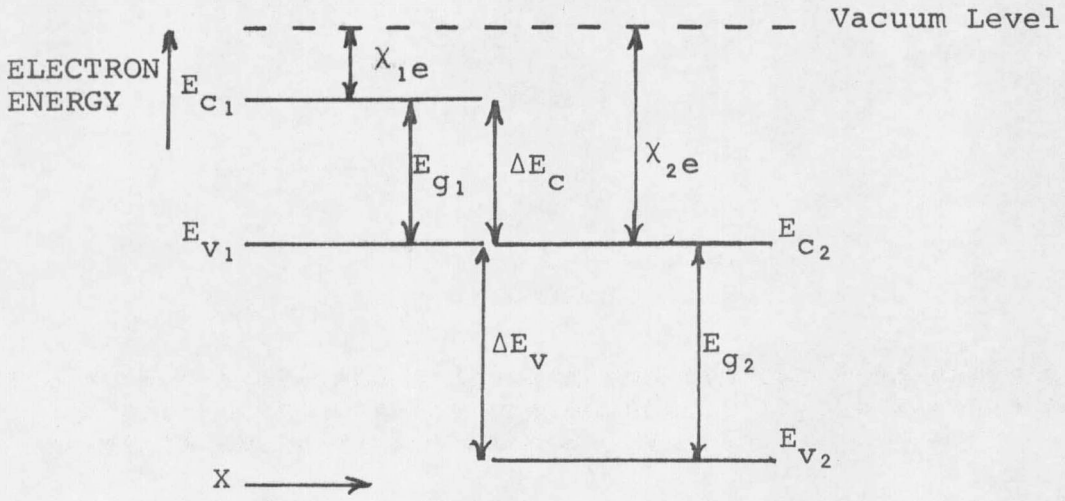


Figure 10. Energy-band diagram for two isolated semiconductors in which space charge neutrality is assumed to exist in every region with $\chi_{2h} > \chi_{1h} = \chi_{2e} > \chi_{1e}$.

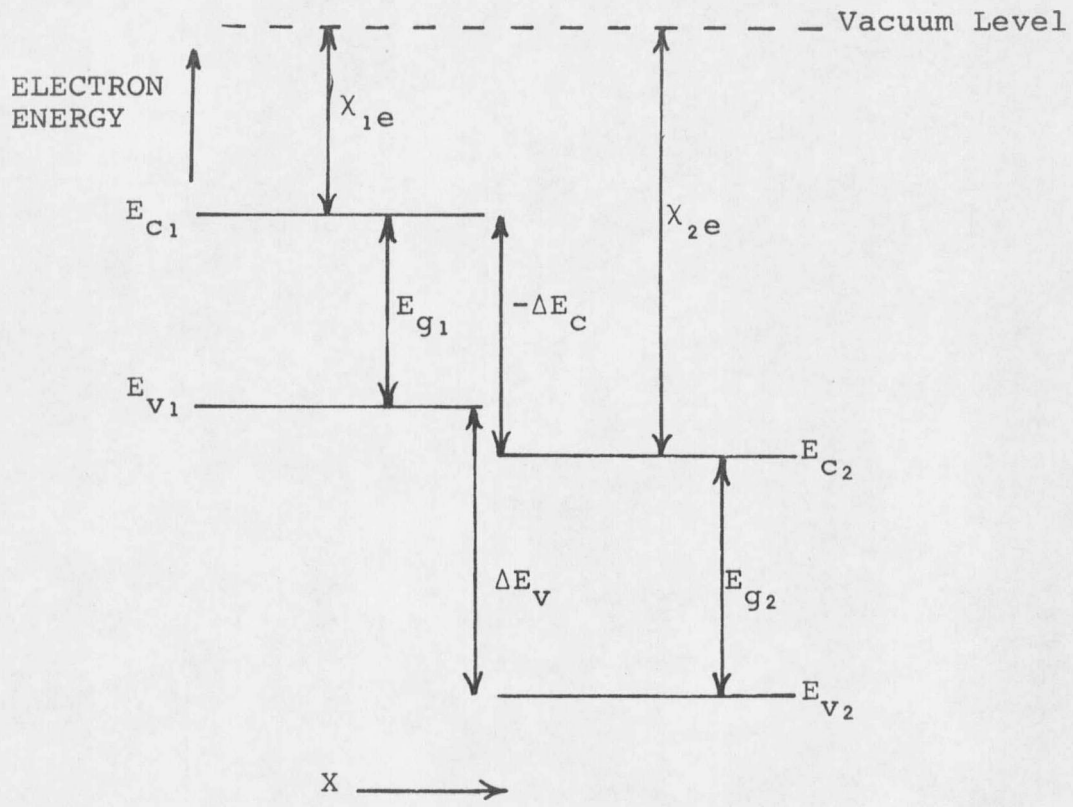


Figure 11. Energy-band diagram for two isolated semiconductors in which space charge neutrality is assumed to to exist in every region with $\chi_{2e} > \chi_{1h}$.

the following four possible constraints, each in three different forms, are possible and can be summarized as:

(1) $\phi_{m_1} > \phi_{m_2}$ or $\phi_{m_1} = \phi_{m_2}$ or $\phi_{m_1} < \phi_{m_2}$

(2) $m^*_{1e} > m^*_{2e}$ or $m^*_{1e} = m^*_{2e}$ or $m^*_{1e} < m^*_{2e}$

(3) $m^*_{1h} > m^*_{2h}$ or $m^*_{1h} = m^*_{2h}$ or $m^*_{1h} < m^*_{2h}$

(4) Direct gap semiconductor - Direct gap semiconductor, junction.

or

Direct gap semiconductor - Indirect gap semiconductor, junction.

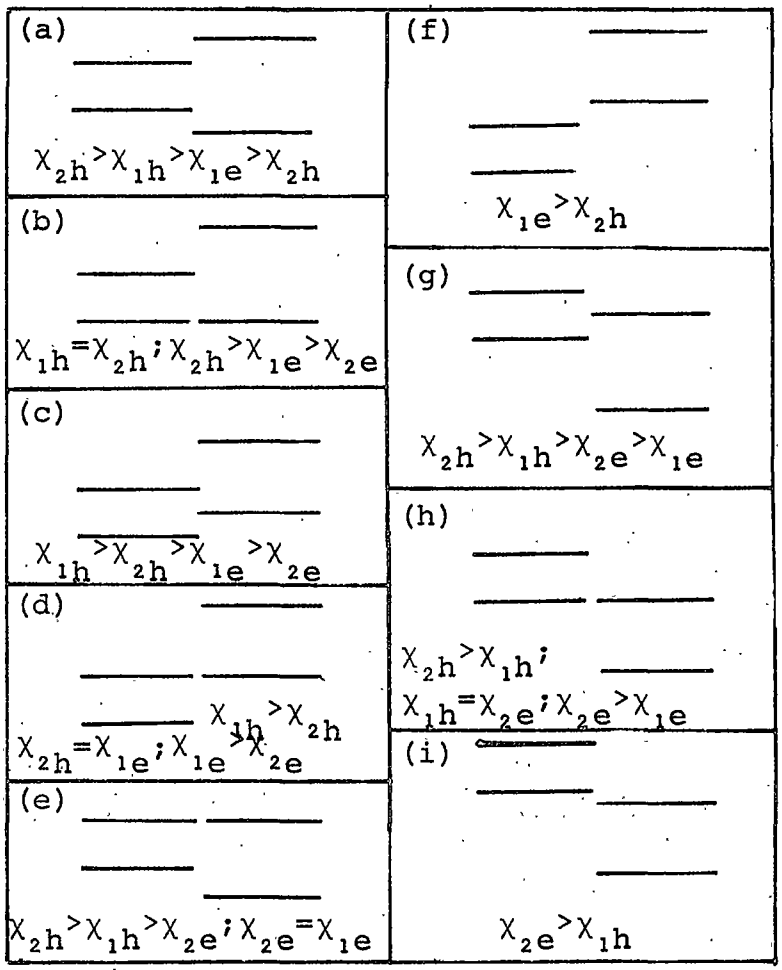
or

Indirect gap semiconductor - Indirect gap semiconductor, junction.

Hence, considering exhaustively all possible combinations of the above parameters and nine working spaces, a total number of different semiconductor heterojunctions can be $9 \times (3)^4 = 729$. Most of the junctions classified do not exist at present; however, they may exist in future with the development of new semiconductor materials with new parameters. The classification presented here is very general and unique; it includes the intrinsic semiconductor heterojunction also, and takes care of junctions yet to be discovered. On the following page a summary of the

Semiconductor Heterojunction Classification Chart

I. Working Spaces



II Fermi Level	III Electron Mass	IV Hole Mass	V Type of Semiconductor
$\phi_{m_1}>\phi_{m_2}$	$m^*_{1e}>m^*_{2e}$	$m^*_{1h}>m^*_{2h}$	Direct gap - Direct gap
$\phi_{m_1}=\phi_{m_2}$	$m^*_{1e}=m^*_{2e}$	$m^*_{1h}=m^*_{2h}$	Direct gap - Indirect gap
$\phi_{m_1}<\phi_{m_2}$	$m^*_{1e}<m^*_{2e}$	$m^*_{1h}<m_{2h}$	Indirect gap - Indirect gap

Figure 12. Semiconductor heterojunction classification chart.

semiconductor heterojunction classification chart is provided.

Importance of Classification

This section considers some of the general properties of each working space as influenced by the position of the Fermi level in the two semiconductors.

(a) Consider the most common working space described by $\chi_{2h} > \chi_{1h} > \chi_{1e} > \chi_{2e}$, as shown in Figure 3. It represents positive discontinuities in the conduction and valence-bands when isolated. On junction formation, however, the resultant barriers of electron and holes depend on the position of the Fermi level in each semiconductor. Three different cases are discussed.

If $\phi_{m_1} = \phi_{m_2}$, then on junction formation, isolated discontinuities appear as electron and hole barriers. No built-in voltage will be formed as shown in Figure 12(a).

If $\phi_{m_1} < \phi_{m_2}$, then on junction formation, the hole barrier from semiconductor 1 is independent of ΔE_c and will depend on the built-in potential supported by semiconductor 2. The electron barrier, however, will increase to $(V_{BHT} + \Delta E_v)$ as shown in Figure 13(b). In such a junction the hole current will dominate in a forward bias.

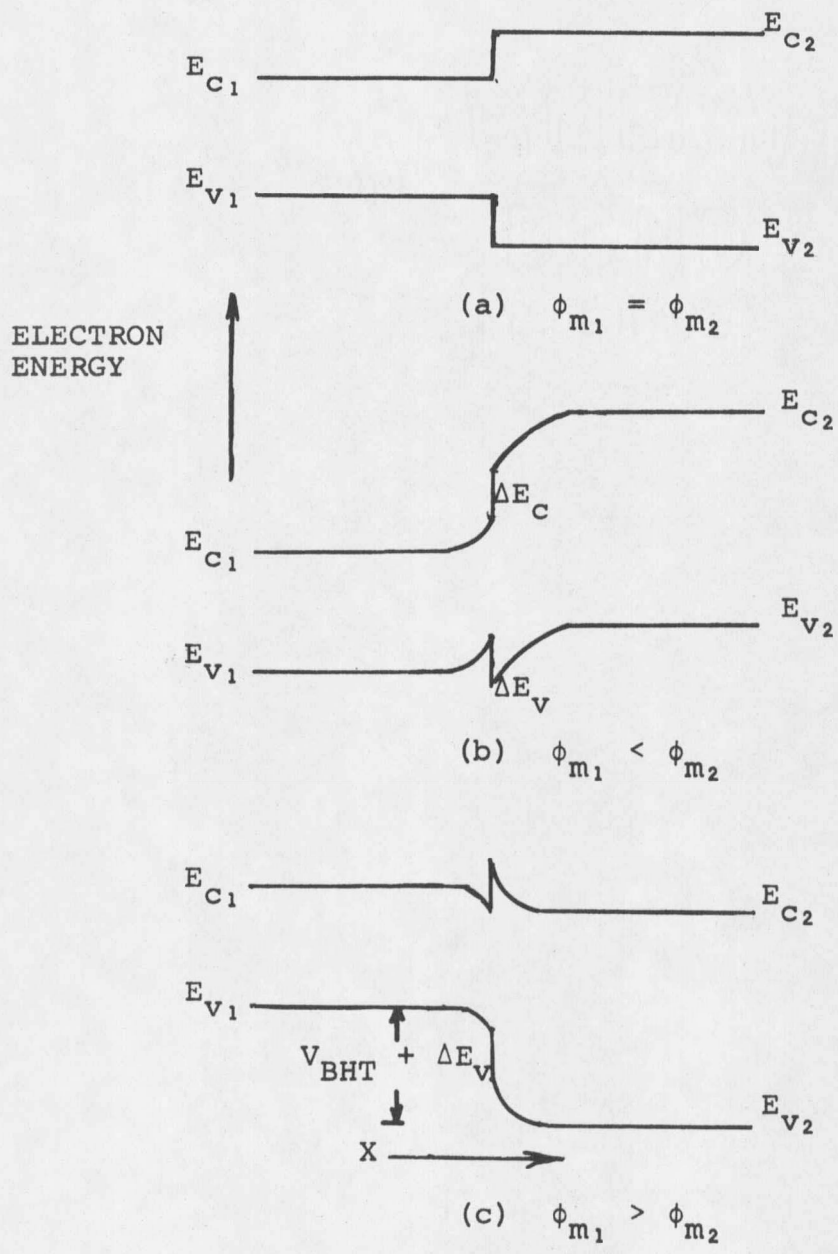


Figure 13. Energy-band diagram of two semiconductors in a thermal equilibrium that defines working space $\chi_{2h} > \chi_{1h} > \chi_{1e} > \chi_{2e}$.

This fact is also supported by Anderson [3] in connection with Ge-GaAs heterojunction. Hence, in such a structure, hole current will dominate electron current, where interface states do not play a significant role. If, however, interface states are present, the current may be governed by recombination and tunneling mechanisms via interface states [20]. Kroemer [12] has suggested that this kind of structure is useful for producing a wide-gap emitter transistor. Such a transistor will have a very high injection efficiency. A p-n-p wide-gap-emitter transistor will be obtained from such a structure. The injection efficiency of a p-n-p transistor is given by:

$$\gamma = \frac{J_p}{J_n + J_p} .$$

It can be noted from Figure 13(b) that electron current will be negligible due to the large electron barrier compared to the hole barrier. Therefore, high injection efficiency can be achieved.

If $\phi_{m_1} > \phi_{m_2}$, then on junction formation, the electron barrier from semiconductor 2 to semiconductor 1 is independent of ΔE_c and will depend on the built-in potential supported by semiconductor 2. The hole barrier, however, will

increase to $V_{BHT} + \Delta E_v$ as shown in Figure 13(c). When such a junction is forward biased, the electron current will dominate over hole current. An n-p-n wide-gap-emitter transistor can be obtained form such a structure, that will provide a high injection efficiency of electrons. Anderson [3] has obtained a photocell action from such a structure.

(b) Consider the next working space described by $\chi_{1h}=\chi_{2h}$, $\chi_{2h}>\chi_{1e}>\chi_{2e}$ (Figure 4). This configuration represents a positive discontinuity in the conduction-band and no discontinuity in the valence-band when isolated. On junction formation, however, the resultant barriers for electron and holes are obtained depending on the position of the Fermi level in each semiconductor. Three different cases are discussed.

If $\phi_{m_1}=\phi_{m_2}$, then on junction formation an isolated discontinuity appears in the conduction-band as an electron barrier. No built-in voltage will be formed as shown in Figure 14(a).

If $\phi_{m_1}<\phi_{m_2}$, then the hole barrier is V_{BHT}, with the valence-band being similar to that of the valence-band of a homojunction. The electron barrier, however, will increase to $(\Delta E_c + V_{BHT})$ as shown in Figure 14(b). This structure has been used for the confinement of electrons

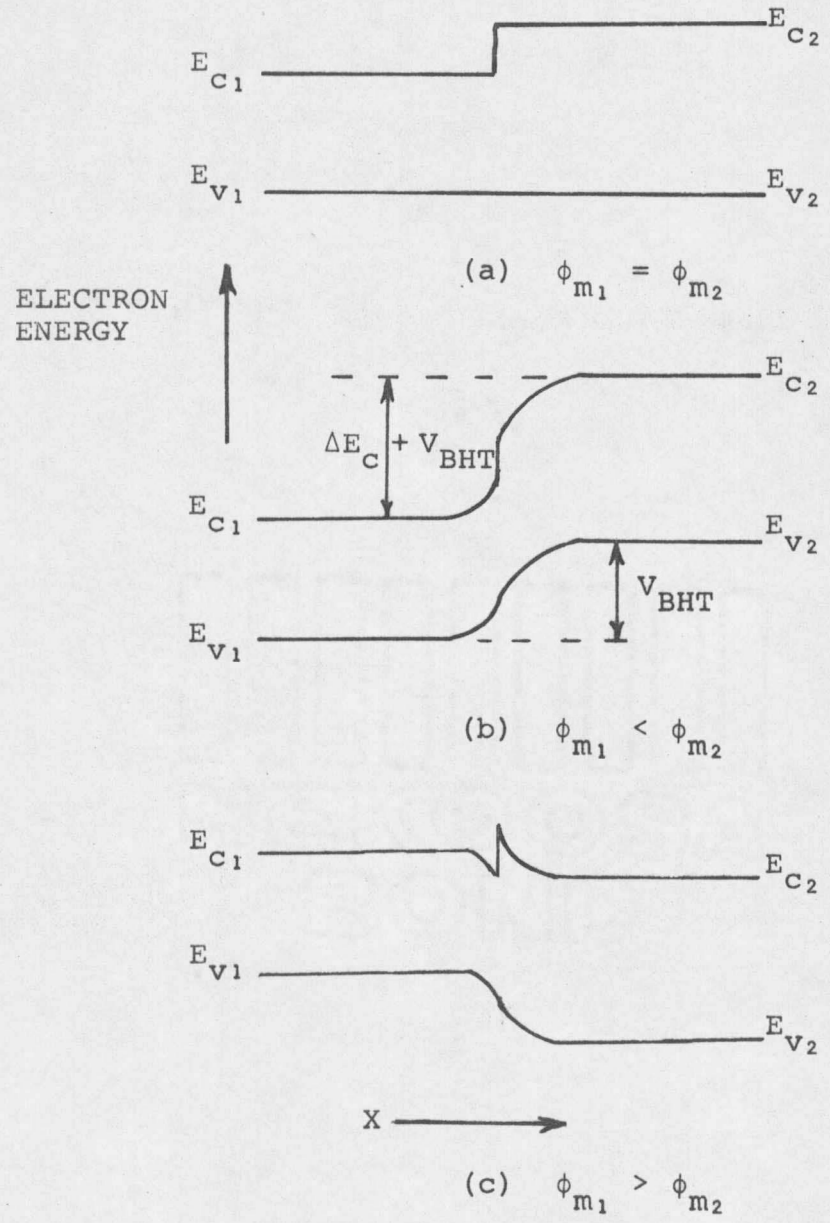


Figure 14. Energy-band diagram of two semiconductors in thermal equilibrium that defines the working space $\chi_{1h}=\chi_{2h}$; $\chi_{2h}>\chi_{1e}>\chi_{2e}$.

in a GaAs-$Al_xGa_{1-x}As$ heterostructure laser [1]. A p-n-p wide-gap-emitter transistor can be obtained from such a structure that will provide a high injection efficiency of holes.

If $\phi_{m_1} > \phi_{m_2}$, then on junction formation, the electron barrier from semiconductor 2 to semiconductor 1 is independent of ΔE_c and will depend on the built-in potential supported by semiconductor 2. The hole barrier is equal to V_{BHT} as shown in Figure 14(c).

Such a structure does not suggest any improvement in injection efficiency over a homojunction transistor, due to the absence of a positive discontinuity in valence-band.

(c) Consider the next working space described by $\chi_1 h > \chi_2 h > \chi_1 e > \chi_2 e$ as shown in Figure 5. It represents a positive discontinuity in the conduction-band and a negative discontinuity in the valence-band, when the two semiconductors are isolated. On junction formation, the resultant barrier of electrons and holes is obtained, depending on the position of the Fermi level in each semiconductor. Three different cases are shown in Figure 15 arising from $\phi_{m_1} = \phi_{m_2}$; $\phi_{m_1} < \phi_{m_2}$ and $\phi_{m_1} > \phi_{m_2}$

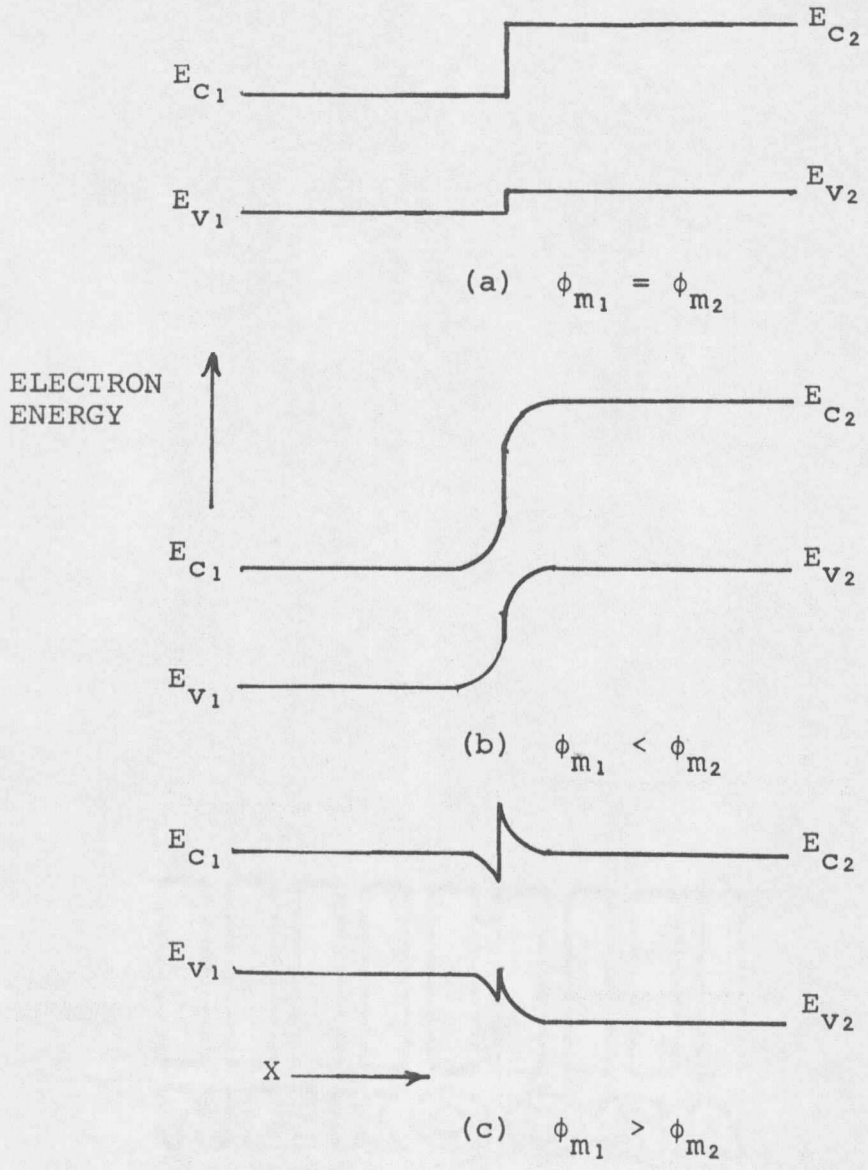


Figure 15. Energy-band diagram of two semiconductors in thermal equilbrium that defines the working space $\chi_{1h} > \chi_{2h} > \chi_{1e} > \chi_{2e}$.

It can be seen that for $\phi_{m_1} < \phi_{m_2}$, the electron barrier increases to $(\Delta E_c + V_{BHT})$ and the hole barrier also increases to $(\Delta E_v + V_{BHT})$. For $\phi_{m_1} > \phi_{m_2}$, both electron and hole barrier are independent of ΔE_c and ΔE_v.

The CdSe-Se heterojunction as studied by Moore et al. [21] and Lang et al. [22] belongs to this working space. P-Se/n-CdSe shows excellent rectification properties and a 70-volt reverse breakdown voltage.

(d) Consider the next working space described by $\chi_{1h} > \chi_{2h}$; $\chi_{2h} = \chi_{1e}$; $\chi_{1e} > \chi_{2e}$ as shown in Figure 6. It represents positive discontinuity ΔE_c equal to E_{g_2} and a negative discontinuity ΔE_v equal to E_{g_1}. On junction formation, the resultant barrier for electrons and holes is obtained depending on the position of the Fermi level.

It can be seen that $\phi_{m_1} < \phi_{m_2}$ is possible only if the narrow-gap semiconductor is a highly degenerate n-type and the broad-gap semiconductor is a highly degenerate p-type. The $\phi_{m_1} = \phi_{m_2}$ case is also possible with degenerate semiconductors only. For $\phi_{m_1} > \phi_{m_2}$, however, a heterojunction is possible with nondegenerate semiconductors as shown in Figure 16.

It can be seen that a high electric field is present in the transition region. Such a diode should,

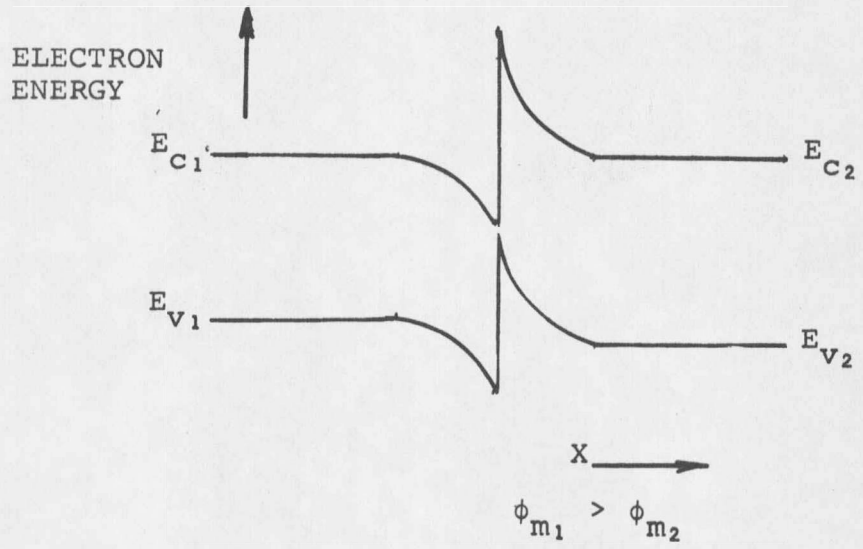


Figure 16. Energy-band diagram of two semiconductors in thermal equilibrium that defines the working space $\chi_{1h} > \chi_{2h}$; $\chi_{2h} = \chi_{1e}$; $\chi_{1e} > \chi_{2e}$.

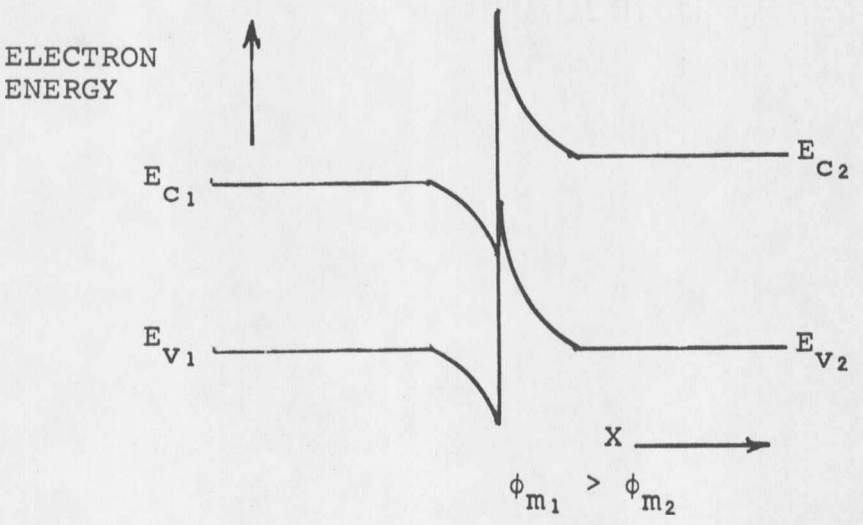


Figure 17. Energy-band diagram of two semiconductors in thermal equilibrium that defines the working space $\chi_{1h} > \chi_{1e} > \chi_{2h} > \chi_{1e}$.

therefore, have a low avalanche breakdown voltage. This electric field may also be used to produce a low work function of the narrow gap semiconductor [23] by applying a thin coating of broad-gap semiconductor on to the narrow-gap semiconductor. This produces high-yield photoemissive cathodes [23].

(e) Consider the next working space described by $\chi_{1h} > \chi_{1e} > \chi_{2h} > \chi_{1e}$, as shown in Figure 7. It represents a positive discontinuity ΔE_c greater than E_{g_2}, and a negative discontinuity ΔE_v greater than E_{g_1}.

As discussed in case (d), $\phi_{m_1} < \phi_{m_2}$ and $\phi_{m_1} = \phi_{m_2}$ may be possible in this case only with highly degenerate semiconductors. These cases will not be discussed here since the discontinuity itself will depend on the doping level [3].

For $\phi_{m_1} > \phi_{m_2}$, a high electric field can be obtained as shown in Figure 17. Such a diode should have a low avalanche breakdown voltage. The GaAs-Cs_2O heterojunction as proposed by Sonnenberg [23] is an example of this class. A thin coating of Cs_2O on GaAs can produce negative electron affinity [23] for GaAs. Such a structure may, therefore, be used to produce high-yield photoemissive cathodes.

(f) Consider the next working space described by $\chi_{2h} > \chi_{1h} > \chi_{2e}$; $\chi_{2e} = \chi_{1e}$ as shown in Figure 8. This configuration represents no discontinuity in the conduction-band but a positive discontinuity in the valence-band when isolated. On junction formation, however, the resultant barrier of electrons and holes is obtained, depending on the position of the Fermi level in each semiconductor. Three different cases are discussed.

If $\phi_{m_1} = \phi_{m_2}$, then on junction formation, an isolated discontinuity in the valence-band appears as a hole barrier. No built-in voltage will be formed as shown in Figure 18(a).

If $\phi_{m_1} < \phi_{m_2}$, then the hole barrier from semiconductor 2 to 1, depends on the built-in potential supported by semiconductor 2. The electron barrier from semiconductor 1 to 2 is V_{BHT}, as shown in Figure 18(b)

Such a structure does not suggest any improvement in injection efficiency over a homojunction transistor, because a positive discontinuity in the conduction-band is absent.

If $\phi_{m_1} < \phi_{m_2}$, then on junction formation, the electron barrier from semiconductor 2 to semiconductor 1 is

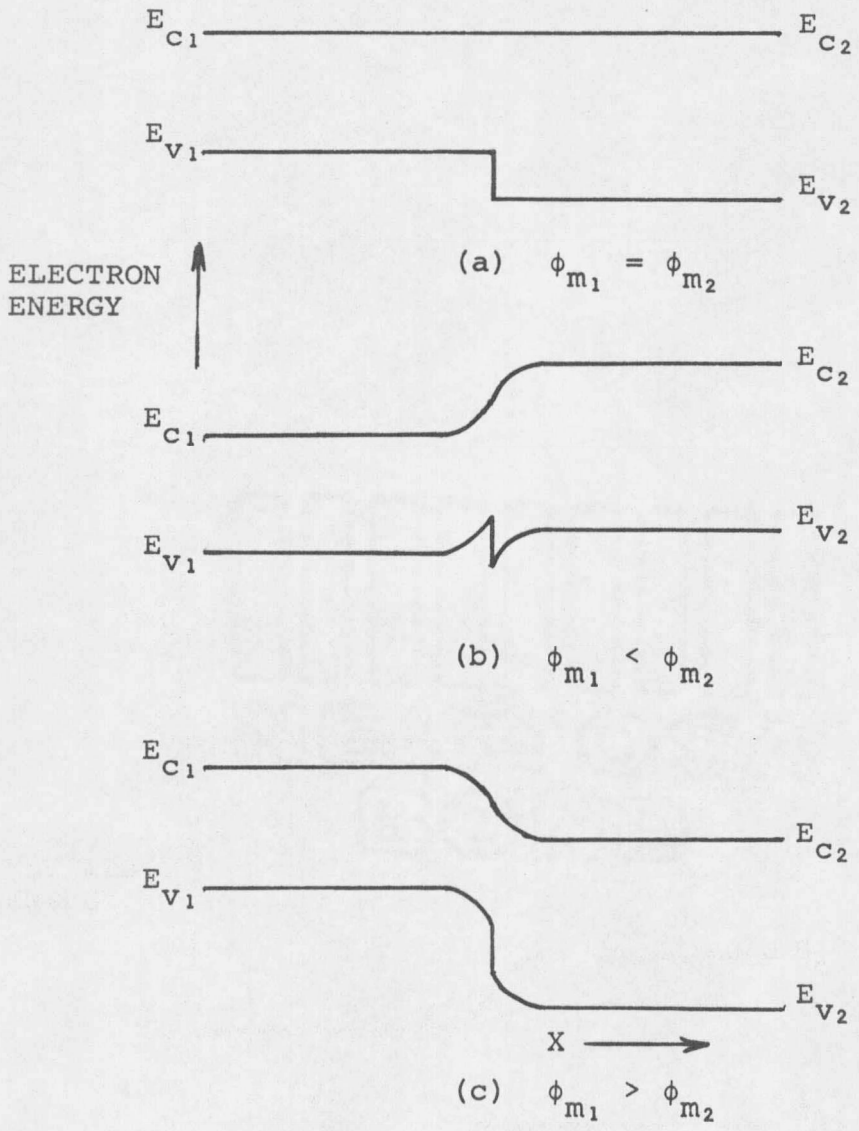


Figure 18. Energy-band diagram of two semiconductors in thermal equilibrium that defines working space $\chi_{2h} > \chi_{1h} > \chi_{2e}$; $\chi_{2e} = \chi_{1e}$.

V_{BHT}. The hole barrier will, however, become $(V_{BHT} + \Delta E_v)$, as shown in Figure 18(c). A n-p-n wide-gap emitter transistor can be obtained from such a structure that will provide a high injection efficiency of electrons.

GaSb-GaAs is an example belonging to this class. It is not of much interest, primarily because of the mismatch in lattice constants.

(g) Consider the next working space described by $\chi_{2h} > \chi_{1h} > \chi_{2e} > \chi_{1e}$, as shown in Figure 9. It represents a negative discontinuity in the conduction-band and a positive discontinuity in the valence-band, when two semiconductors are isolated. Three different cases are shown in Figure 19 arising due to $\phi_{m_1} = \phi_{m_2}$, $\phi_{m_1} < \phi_{m_2}$ and $\phi_{m_1} > \phi_{m_2}$.

It can be seen that for $\phi_{m_1} < \phi_{m_2}$ the electron and hole barriers are independent of ΔE_c and ΔE_v as shown in Figure 19(b). For $\phi_{m_1} > \phi_{m_2}$, the electron and hole barriers are $(\Delta E_c + V_{BHT})$ and $(\Delta E_v + V_{BHT})$, respectively, as shown in Figure 19(c).

A similarity of this work space with that described in Figure 5 and discussed in (c) should be noted. In this case broad-gap semiconductor is of lower electron affinity and in (c) case, the narrow-gap semiconductor is of a lower electron affinity.

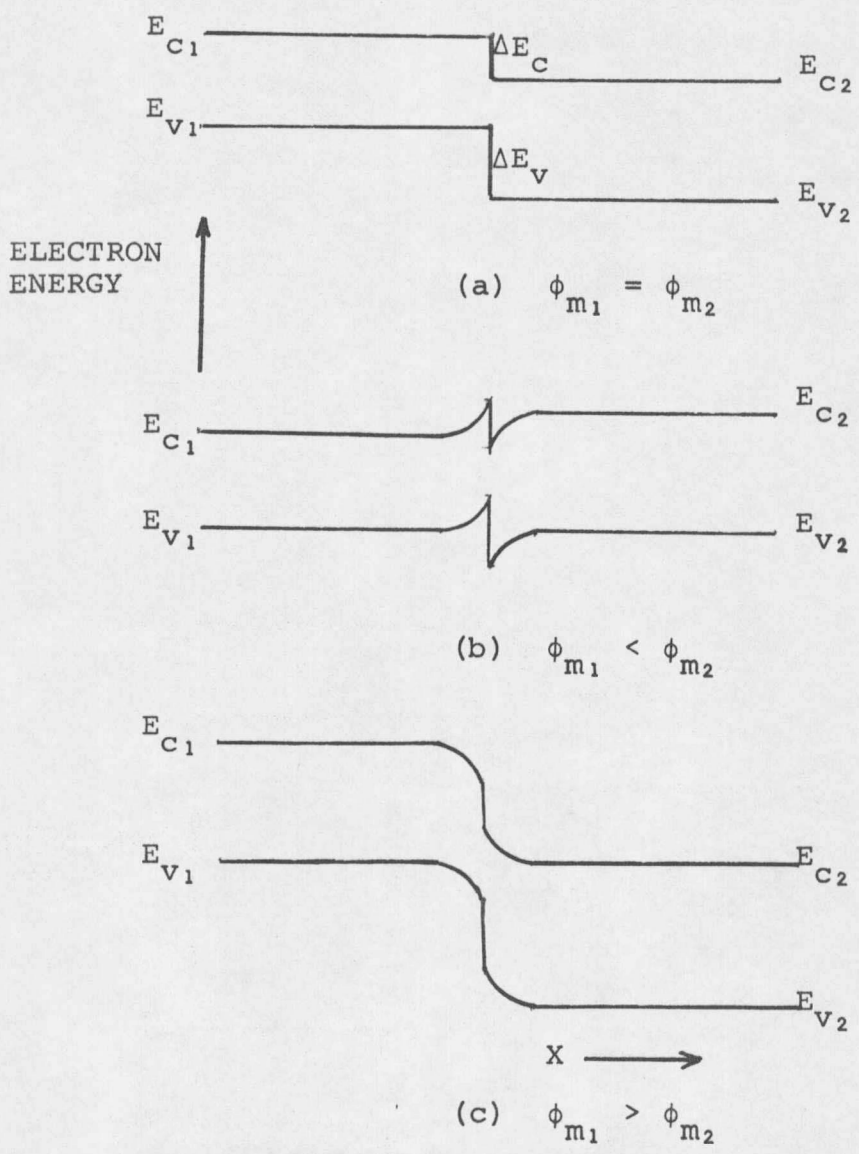


Figure 19. Energy-band diagram of two semiconductors in thermal equilibrium that defines the working space $\chi_{2h} > \chi_{1h} > \chi_{2e} > \chi_{1e}$.

A ZnTe-CdS heterojunction is an example of this class. This class has been studied by S. N. Dobrynin [24] theoretically. He has concluded that in this working space if the lower electron-affinity material is p-type and the higher electron-affinity material is n-type, the best tunneling condition are possible. Thus, a heterojunction tunnel diode is possible in this working space or in the one described in case (c), which shows improved performance over the homojunction tunnel diode.

(h) Consider the next working space described by $\chi_{2h} > \chi_{1h}$; $\chi_{1h} = \chi_{2e}$; $\chi_{2e} > \chi_{1e}$, as shown in Figure 10. It represents a negative discontinuity in the conduction-band equal to E_{g_1} and a positive discontinuity in the valence-band equal to E_{g_2}. The similarity of this working class with that discussed in (d) should be noted.

It can be seen that $\phi_{m_1} > \phi_{m_2}$ and $\phi_{m_1} < \phi_{m_2}$ is possible with degenerate semiconductors only and, hence, is not considered. For $\phi_{m_1} < \phi_{m_2}$, however, a heterojunction is possible with a nondegenerate semiconductor and as shown in Figure 20.

It can be noted that a high electric field is present in the transition region. Such a diode should, therefore, have a low avalanche breakdown voltage. This

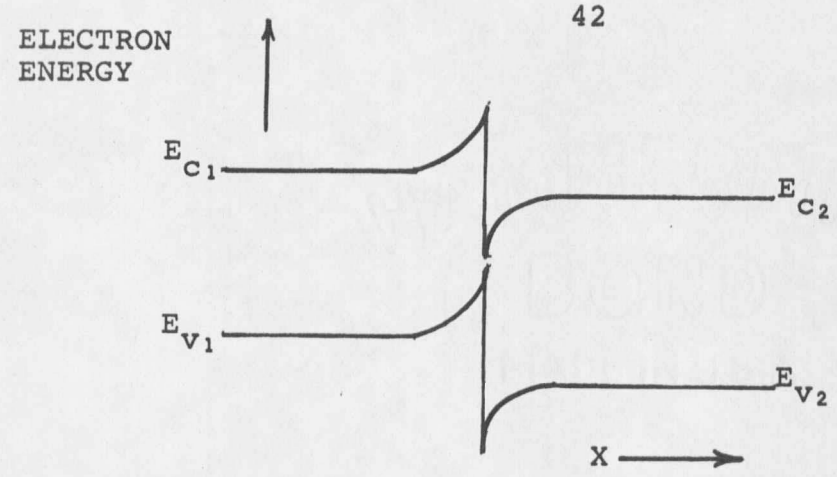


Figure 20. Energy-band diagram of two semiconductors in thermal equilibrium that defines the working space $\chi_{2h} > \chi_{1h}$; $\chi_{1h} = \chi_{2e}$; $\chi_{2e} > \chi_{1e}$.

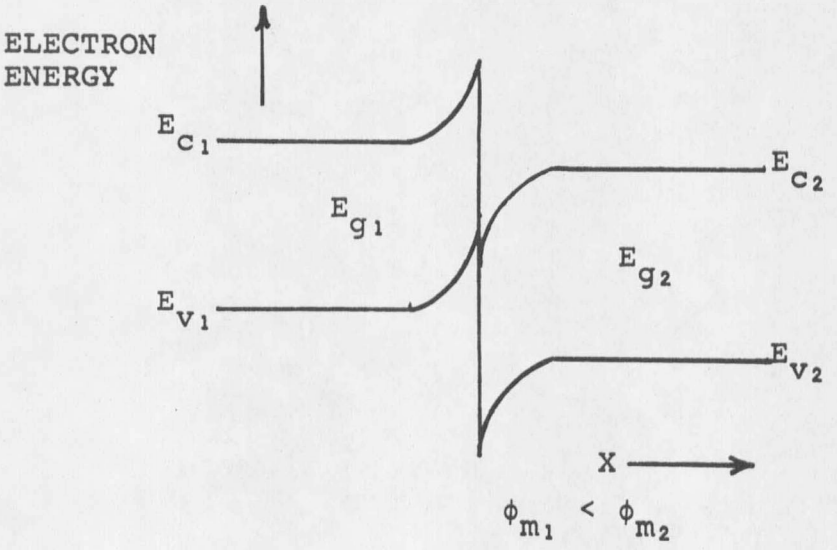


Figure 21. Energy-band diagram of two semiconductors in thermal equilibrium that defines the working space $\chi_{2h} > \chi_{1h}$.

electric field may also be used to produce low work function of the broad-gap semiconductor [23] by a thin coating of the narrow-gap semiconductor, thus producing a high-yield photoemissive cathode.

(i) Consider the next working space described by $\chi_{2e} > \chi_{1h}$. It represents a negative discontinuity of electron energy in the conduction-band, greater than E_{g_1} and a positive discontinuity of electron energy in the valence-band, greater than E_{g_2}, when isolated.

When a junction is formed, the resultant electron and hole barrier is produced which depends on the position of the Fermi levels in the two semiconductors. As discussed in (h), $\phi_{m_1} > \phi_{m_2}$ and $\phi_{m_1} = \phi_{m_2}$ are not considered here.

For $\phi_{m_1} < \phi_{m_2}$, however, a nondegenerate heterojunction is possible and as is shown in Figure 21.

It can be noted that a high electric field is present in the transition region. Such a diode should, therefore, have a low avalanche breakdown voltage. It can also be used in producing high-yield photoemissive cathodes.

Given a pair of semiconductors with known electron affinities and energy gaps, a unique working space can be assigned. Some general properties of the given heterojunction pair are thus known.

By knowing which semiconductor has the greater effective masses of electrons or holes, we get some idea about the structure of the conduction and the valence-band, respectively (see, for example, Appendix A). Effective masses also give some idea about the mobilities in the two semiconductors. A direct-gap semiconductor has a short lifetime of excited carriers compared to indirect-gap semiconductors. By knowing the type of semiconductor involved, one has more knowledge of the structure in E - K space (see Appendix A).

Chapter 3

ANALYSIS OF ELECTRIC FIELD

This chapter describes a theory of electric field formation in metal-metal junctions, homojunctions and heterojunctions. Expressions for built-in voltage are derived. Distributions of carrier concentrations are plotted, and a practical model of the n-p semiconductor heterojunction is proposed and analyzed.

General Theory

Junctions between metals, metals and semiconductors or different semiconductors with different doping levels, result in a built-in electric field at the contact. Two different metals, in general, have different conductivities at a given temperature because different concentrations of electrons are available at that temperature:

$$n_{cu} = 8 \times 10^{22}/cm^3, \; n_{Ag} = 6 \times 10^{22}/cm^3.$$

If a contact is made between two different conductors, an exchange of electrons will take place by diffusion corresponding to the concentration gradient, which causes the substance whose concentration of electrons is initially lower to become more negative than the other. Thus, a

potential is set up at the boundary. This potential does not allow additional net diffusion to take place. A semiconductor homojunction and semiconductor heterojunction are considered in the following discussion.

Semiconductor homojunction. Suppose that an abrupt p-n junction is somehow formed instantaneously by joining a uniform p-type sample to a uniform n-type sample to form a single crystal. At the instant of formation there exists a uniform concentration n_1 of mobile free electrons, and p_1 mobile free holes, on the n-side extending to the junction. On the p-side a uniform concentration of p_2 mobile holes and n_2 free electrons also extend to the junction. These concentrations are related to the net donor and acceptor densities at a temperature, while, of course, on either side the electron and hole densities satisfy the relation:

$$n_1 p_1 = n_2 p_2 = n_i^2.$$

Since the concentration n_1 of electrons on the n-side is much larger than the electron concentration n_2 on the p-side, at the instant of formation there exists an enormous

gradient in the concentration of electrons at the junction between the two regions. The same situation exists with respect to hole concentration at the junction. The large initial concentration gradients set up diffusion currents; electrons from the n-region flow down the respective concentration gradients into the region of opposite conductivity type leaving the region near the junction depleted of majority carriers. This initial diffusion flow cannot continue indefinitely because as the regions near the junction become depleted of majority carriers, the charges of the fixed donor and acceptor ions near the junction are no longer balanced by the charges of the mobile free carriers which were formerly there, and so an electric field is built up. The direction of this electric field is such as to oppose the flow of electrons out of the n-region and the flow of holes out of the p-region, and thus the magnitude of the field builds up to the point where its effect exactly counteracts the tendency for majority carriers to diffuse down the concentration "hill" into the region of the opposite conductivity type. A condition of thermal equilibrium is established in which the region near the junction is depleted of majority carriers and in which

space charge layers containing high electric fields are formed near the junction.

Alternatively, the situation can be described in terms of an energy-band model. Consider the energy-band diagram of isolated n-p semiconductors as shown in Figure 22. The position of the Fermi level in the n-type semiconductor is represented by E_{F_1} and in the p-type semiconductor is represented by E_{F_2}.

It is known that within a solid or system of solids in thermal equilibrium the Fermi level E_F has everywhere the same value. Hence, when a junction is formed, we expect an exchange of carriers. The electrons will move to the p-type semiconductor or, equivalently, the holes will move to the n-type semiconductors. The energy-band diagram of an n-p homojunction in thermal equilibrium is described in Figure 23. The electrostatic potential difference $\psi(x)$ between any two points can be represented by the vertical displacement of the band edges. The distance over which ψ changes is called the transition region. The electrostatic field can be represented by the slope of the band edges on the diagram. The built-in voltage, V_{BHM}, is represented on the diagram.