

Chapter 2

Solid-State Physics Fundamentals of LED Thermal Behavior

Jinmin Li, Junxi Wang, Zhe Liu and András Poppe

Abstract This chapter provides the basics on the physics of light-emitting diode (LED) operation: band structures, carrier transport, different recombination mechanisms, etc. and presents the Shockley model of ideal semiconductor diodes. Device construction techniques, LED packaging styles—all affecting LED efficiency/efficacy are also discussed.

2.1 Introduction

In 1907 electrical engineer Henry J. Round touched a crystal of silicon carbide found near Niagara Falls with two wires connected to a battery. As he subsequently reported: “On applying a potential of ten volts between two points on a crystal of carborundum, the crystal gave out a yellowish light.” Though he may not have realized it at that time, Round had operated the first crude solid-state lamp, the light emitting diode.

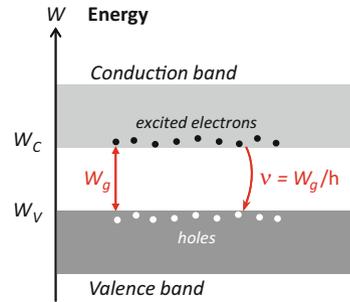
The discovery of light emission in silicon carbide was important for several reasons, but the most significant reason concerned the mechanism that caused the carborundum crystal to emit light. Most electric light sources such as incandescence of a hot electrode, luminescence of a glowing plasma, or fluorescence of a phosphor coating, are capable of producing exceedingly high luminous outputs, but each of the conventional light sources is unsuited to a broad range of potential applications because of their slow response time, inherent fragility, and short lifetime.

The light-emitting diode (LED) conveniently eliminates the limitations of conventional lamps with its solid-state reliability, speed, and compact size. Since semiconductor light sources (without any further structures like phosphor layers) emit a narrow-wavelength spectrum, various colored sources can be fabricated without the need for filters. Understanding their physical operation gives the motivation

J. M. Li (✉) · J. X. Wang · Z. Liu
Institute of Semiconductors, Chinese Academy of Science, A35, QingHua East Road,
Haidian District, 912, 100083 Beijing, P. R. China
e-mail: jmli@semi.ac.cn

A. Poppe
Budapest University of Technology and Economics, Department of Electron Devices,
Magyar tudósok körútja 2, Bldg. Q, 1117 Budapest, Hungary
e-mail: poppe@eet.bme.hu

Fig. 2.1 Light emission occurs when an excited electron falls back from the conduction band with a direct energy state transition into the valence band and recombines with a hole. The frequency of the emitted light is determined by the W_g bandgap energy



and the necessary background knowledge to design proper cooling solutions for LEDs. The purpose of this chapter is to provide this background for the reader of this book.

This chapter is built up as follows. In Sect. 2.2, the basic concepts of solid-state physics that cause certain materials to emit light are described. Section 2.3 introduces the basic elements of bandgap engineering. Section 2.4 provides details on physics of LED operation including carrier transport, recombination mechanisms, and the I-V characteristics. Section 2.5 is dedicated to the description of temperature dependence of some of the fundamental processes inside an LED. Sections 2.6 and 2.7 describe chip level and package-level construction issues of LEDs and the chapter concludes with a summary.

For a more detailed study on the semiconductor physics, the reader is advised to refer to the classical handbooks of the topic [10, 12]. Detailed description of operation of LEDs can be also found in Schubert's widely known book [13].

2.2 Basic Concept of Monochromatic LEDs: A Qualitative Overview of the Operation

A simple way to efficiently generate light in many semiconductors is to fabricate the material as a pn-junction diode. In this manner, stimulated electrons can be efficiently and directly used to generate light, and a great deal of control can be exercised over the fabrication of the light-emitting region and the efficiency of the device.

Generally speaking, light generation in a semiconductor occurs when an excited electron resumes equilibrium through a direct *conduction band to valence band* transition by recombining with a hole in the valence band (see Fig. 2.1). The frequency and the wavelength of the emitted photon is determined by the $W_g = W_c - W_v$ bandgap energy as follows:

$$\nu = W_g/h \quad (2.1a)$$

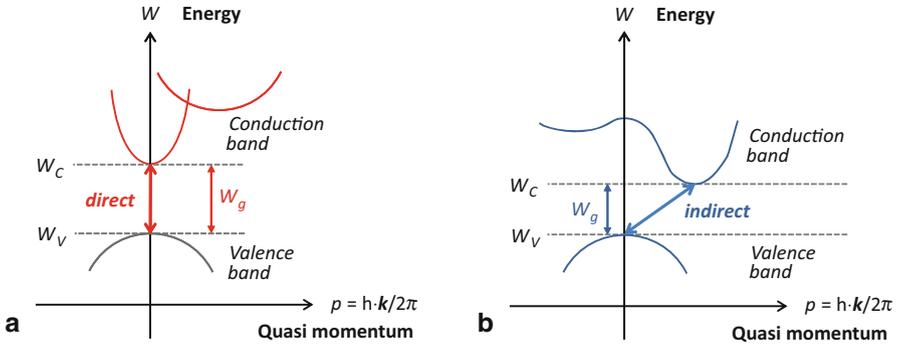


Fig. 2.2 Band structure of direct and indirect bandgap materials

$$\lambda = c/v = c \cdot h/W_g \quad (2.1b)$$

where h is Planck's constant and c denotes the speed of light.

As seen from Eqs. (2.1a) and (2.1b), the most important parameter to determine the properties of the emitted light is the energy band structure of the semiconductor material used. Semiconductor materials are classified as *direct bandgap* or *indirect bandgap* materials. Light emission takes place in pn-junctions formed in direct bandgap materials. Direct bandgap materials are characterized by an energy band structure such that the minimum energy level of the conduction band (where unlocalized electrons are free to move in the real physical space formed by the semiconductor lattice) and the maximum of the valence band (where vacancies of excited electrons called holes are left, acting as positively charged quasi-particles) are located at the same k -wavelength number (quasi momentum), as shown in Fig. 2.2a. In the Brillouin zone, electrons' and holes' energy state is characterized by k -wavelength number. In direct bandgap semiconductors, k -wavelength numbers of electrons and holes are the same. They have the same momentum. But in indirect bandgap semiconductors, k -wavelength numbers of electrons and holes are no longer the same, their momenta are different. Without the involvement of phonons, the electrons and holes cannot recombine in indirect bandgap semiconductors.

The junction diode provides a simple and efficient technique for placing large number of electrons in an excited state. In order for an electrical current to cross the potential barrier formed by the pn-junction, electrons must be raised from the valence to the conduction band. After the junction has been crossed, the electrons recombine with holes on the p side and this results in photon emission. Such a process is called *radiative recombination*. In contrast to this, in the so-called *nonradiative recombination* processes, as shown in Fig. 2.2b instead of photon emission, the energy of the recombining electron is transferred to the vibration energy of the semiconductor lattice by means of *phonons*, resulting in heat generation. Nonradiative recombination is facilitated by so-called *deep levels* or *traps* (energy levels close

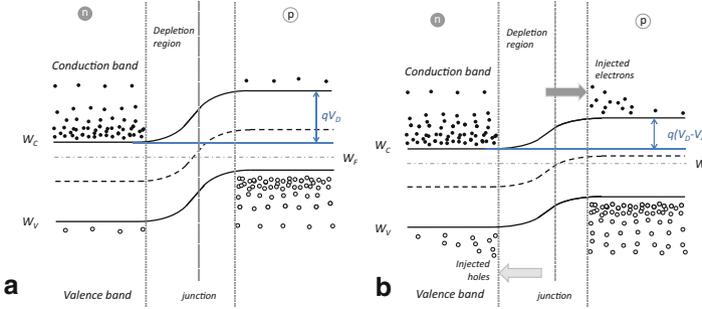


Fig. 2.3 Simplified energy-level diagrams (also known as simplified band diagrams) of a pn-junction, **a** No current flow, **b** Current flow

to the middle of the bandgap) which are called *luminescence killers*. Details of the different recombination processes will be discussed in Sects. 2.4.1–2.4.3.

A major goal of LED device construction is to assure the highest possible ratio of the electrical current associated with radiative recombination with respect to the total current of the LED device. The second goal is to assure that most of the photons generated at a pn-junction of an LED are coupled out to the outer space. These two aspects of LED device design will be discussed in detail in the last section of this chapter.

In this section, we start our discussion with the so-called *homojunctions*, i.e., pn-junctions consisting of a single material. Later in this chapter, *heterojunctions* and *quantum-well* structures used in today's modern power LEDs will be discussed in detail.

We further restrict our discussion now to the theory of *abrupt pn-junctions* with a homogeneous N_D donor concentration on the n-side and with a homogeneous N_A acceptor concentration on the p-side. Assuming fully ionized dopants, the concentration of the free electrons on the n-side is $n = N_D$, the concentration of the free holes on the p-side is $p = N_A$. On the n-doped side of the junction the electrons in the conduction band are the majority carriers while on the p-doped side the holes in the valence band are the majority carriers and both on the n- and p-sides of the junction there are some minority carriers also present.

In the close neighborhood of the unbiased pn-junction, the electrons of the n-side diffuse over to the p-side where they recombine with the holes. Similarly, holes diffuse to the n-side. As a result, the vicinity of the junction gets depleted from free carriers and a so-called *depletion region* is formed as indicated in Fig. 2.3. As there are no free carriers left in the depletion region, the only charge remaining there is from the ionized dopants (donors, acceptors). These dopants form a so-called *space charge region* which produces a potential that is called *diffusion potential* and which is denoted by V_D . Its value is determined by the N_A and N_D dopant concentrations and the n_i intrinsic carrier concentration of the semiconductor material:

$$V_D = \frac{kT}{q} \ln \left(\frac{N_A \cdot N_D}{n_i^2} \right) \quad (2.2)$$

where k is Boltzmann's constant, T is the absolute temperature of the semiconductor, and q is the elementary charge (charge of an electron).

Figure 2.3 shows graphs known as an *energy-level diagrams* or *band diagrams*. The valence and conduction bands on each side of the pn-junction are connected by sloping lines representing the continuous change of the energy levels as the effective doping level changes across the junction. In case of an unbiased pn-junction the diffusion potential forms an energy barrier of qV_D which blocks the flow of carriers across the junction, therefore there is no significant current flow through the junction (Fig. 2.3a). If an external V voltage is applied which results in reduction of the potential barrier electrons are injected into the p-side and holes are injected into the n-side and current flow occurs (Fig. 2.3b). This is a positive voltage between the p-side (also known as *anode*) and the n-side (referred to as *cathode*) of the pn-junction. While the injected carriers—now as minority carriers—are moving forward by diffusion, they are recombining with the majority carriers. If, for example, the recombination process is a radiative recombination (such as a direct bandgap transition), photon emission takes place.

The wavelength of the emitted radiation in such an arrangement is related to the bandgap of the semiconductor material making up the diode. *Bandgap engineering* means designing the material composition and the doping levels of an LED—which will determine the properties of the emission spectra (such as the peak wavelength) and device efficiency. Nowadays there are two widely used material systems: AlGaInP (aluminum-gallium-indium-phosphide) for producing red to yellow LEDs and InGaN/GaN (indium-gallium-nitride/gallium-nitride) for building green to blue LEDs [5–7], see Fig. 2.4.

White LEDs are typically so-called phosphor converted devices in which an LED chip with short wavelength (blue or UV) primary emission is covered by a phosphor layer which absorbs some of the short wavelength light and converts it to longer wavelength light in a wider spectral range. The fundamental physical operation of such LEDs—except the wavelength conversion process taking place in the phosphor—is the same as the operation of monochromatic LEDs, therefore it will not be discussed here separately.

2.3 Device Constructions Influencing the Energy Band Structure

The aim of this section is to provide an overview of bandgap engineering concepts resulting higher light generation efficiency.

2.3.1 Homojunction Devices

Correctly designed p and n materials are crucial to LEDs. In *homojunction* LEDs, the p and n materials are composed of the same energy gap semiconductor, as shown

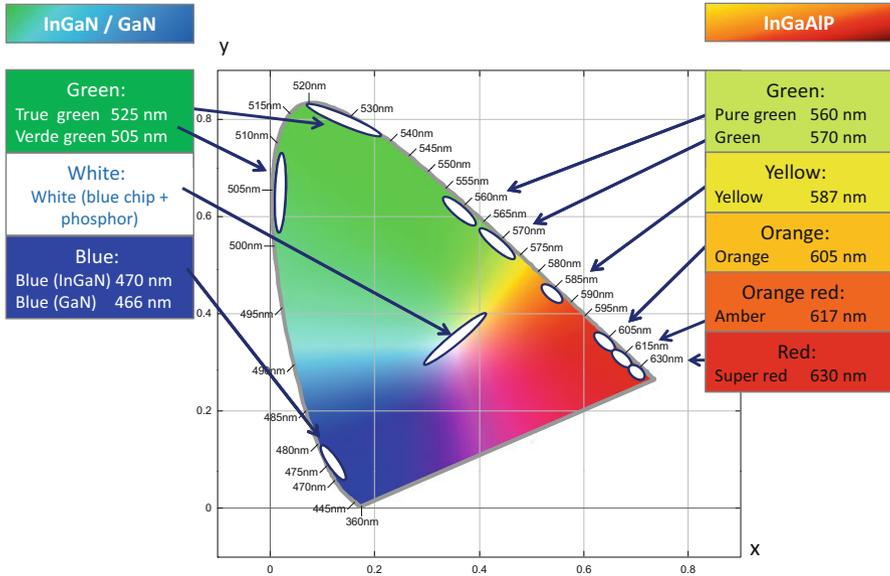


Fig. 2.4 Different LED colors realized by different material systems

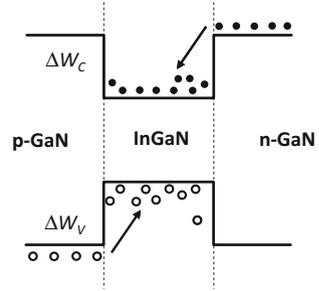
in Fig. 2.3a. The p and n materials have the same bandgap structure. Due to the different doping type, the p and n material have different *Fermi levels*. When they are joined together, uniform *Fermi levels* are established which cause energy band bending (shown as W_F in Fig. 2.3). When an electron and hole go through the *space charge region* and recombination, a photon is generated. As the diffusion distance for electrons is larger than that of the holes, the injection ratio between electrons and holes is more than one in case the dopant concentrations in the p and n material are kept the same, which may lead to a serious thermal problem during operation. Furthermore, self-absorption from the active region lowers the output efficiency in homojunction LEDs.

2.3.2 Heterojunction Devices

A *heterojunction* is defined as a semiconductor structure in which the chemical composition changes with position. Heterojunctions are able to improve the performance of semiconductor devices because they permit the device designer to modify locally the energy-band structure of the semiconductor and so control the motion of the charge carriers. In the simplest case in a heterojunction two materials of different bandgaps are used.

Bandgap engineering [1] allows the LED designer to achieve higher efficiencies in converting electric current to light than conventional homostructures. Two

Fig. 2.5 Simplified energy band diagram for a heterojunction. LEDs in which the conduction band discontinuity ΔW_C and valence band discontinuity ΔW_V provide a barrier to confine carriers into the active region



major benefits of heterostructure LEDs are the increased minority carrier injection efficiency and the ability to use wide bandgap material. The wide bandgap layers are transparent to the photon reabsorption. The radiative efficiency in heterostructure LEDs is higher than that of the homostructure LEDs because of the increased current injection efficiency at the heterojunction. Since the radiative efficiency of the p-side is usually much higher than that of the n-side, higher efficiency devices are made by eliminating the minority carrier injection into the n-type material. By creating a heterostructure at the p/n interface, a valence band discontinuity is formed which provides a hole potential barrier. Figure 2.5 illustrates the energy band diagram for heterojunction flat-band conditions where the applied bias just cancels the built-in field. The conduction band discontinuity ΔW_C and the valence band discontinuity ΔW_V , produce a barrier to electron and hole flow [2], respectively.

Due to the barrier height limitation of the bandgap discontinuity, more carriers come across the pn-junction as leakage current. On the other side, nonradiative recombination through deep energy levels at the interface will reduce the radiative efficiency and causes a potential heating at the same time.

2.3.3 Quantum Well Structure

If one makes a heterostructure with sufficiently thin layers, quantum interference effects begin to appear prominently in the motion of the electrons. The simplest structure in which these may be observed is a *quantum well*, which simply consists of a thin layer of a narrower bandgap semiconductor between thicker layers of a wide bandgap material [3]. The band profile then shows a “rectangular well,” as illustrated in Fig. 2.6. The electron wave functions in such a well consist of a series of standing waves, which might be found in a resonant cavity in acoustic, optical, or microwave technologies. The energy separation between these stationary states is enhanced by the small effective mass of electrons in the conduction bands of direct-gap semiconductors. With advanced epitaxial techniques, the potential profile of the quantum well does not need to be rectangular. Because the band edge energy is usually linear in the composition, W_V and W_C will follow the functional form of the composition. The quantum states in two parabolic wells [4] are illustrated in Fig. 2.7.

Fig. 2.6 Energy band profile of a structure containing three quantum wells, showing the confined state energy levels in each well. The gaps in the lines indicating the confined state energies show the locations of nodes of the corresponding wave functions

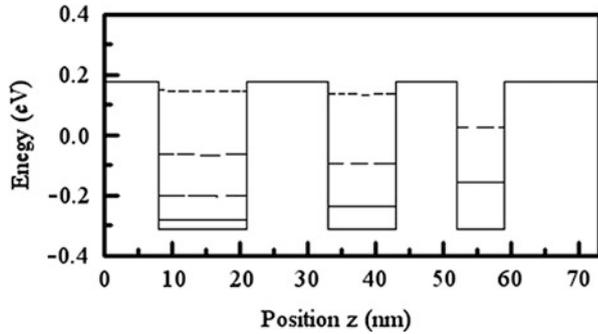
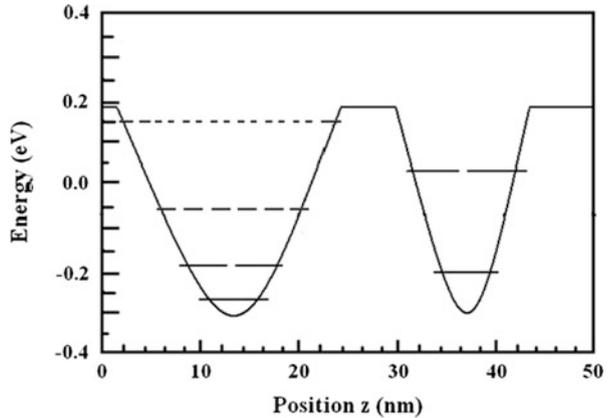


Fig. 2.7 Energy-band profile of a structure containing two parabolic quantum wells. The composition is similar to that of Fig. 2.3, and the overall widths of the wells are 20 and 8 nm



Quantum well heterostructures are key components of many optoelectronic devices, because they can increase the strength of electrooptical interactions by confining the carriers to small regions.

As for quantum well LEDs, they are a special class of heterojunction LEDs in which the thickness of the active region is less than the de Broglie wavelength of the electron in the semiconducting material. Basically, the electrons and holes are confined so tightly in space that their allowed states have discrete energy levels due to the localization. Recombination between such discrete energy levels results in very narrow emission spectra—characteristic to laser diodes.

2.4 Fundamental Physical Processes Determining the LED Behavior

This section provides a brief summary of the basic processes taking place in semiconductor pn-junctions.

2.4.1 Carrier Transport

When we talk about the carriers in the semiconductor, it means that electrons and holes are not treated separately, but are considered as macroscopic charge carrier populations or carrier concentrations. In semiconductor devices, carriers move due to different physical processes. The most important ones are: diffusion, drift caused by electric field, heat-emission, and tunneling. These processes—with a focus on recombination—are also discussed [10, 11].

2.4.1.1 Diffusion Current

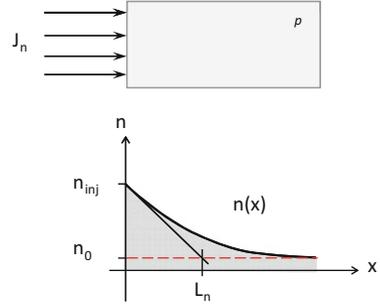
Any kind of microscopic particles, be it molecules, atoms, or electrons, could generate diffusion movement, which is merely caused by the inhomogeneity of the density of the distribution of the particles. They will diffuse from regions of high concentration to regions of low concentration. The current in this case is called *diffusion current*. Carrier diffusion could happen, whether an external field is applied or not. The basic driving force of the carrier diffusion originates from the thermal energy of the majority carriers, which leads to the constant random motion of carriers. For a piece of uniformly doped semiconductor material, carriers homogeneously distribute throughout the physical space and the random motion of carriers in all directions results in zero net macroscopic current. When the carrier distribution is not uniform in the material, which is often caused by external electrical injection in LEDs, then a carrier concentration gradient exists between the region of high carrier density and the region of low carrier density. Then the random motion shows a pronounced tendency in a particular direction resulting in a net carrier flow in that direction. Thus, this carrier flow is called *diffusion current*, the quantity of which is closely related to the gradient of the carrier concentration.

Consider a piece of homogeneously p-doped semiconductor material with a metalized surface at its left side through which we inject electrons with a current density of J_n into the material, as shown in Fig. 2.8. This situation corresponds, e.g., to the situation, when electrons from the cathode side of a forward biased pn-junction reach the p-doped, anode side of the diode and are injected into the p-type semiconductor material (as already discussed with the help of Fig. 2.3). A concentration gradient of the electrons diffusing into the p-type material develops as the diffused electrons (as minority carriers) recombine with the majority carrier holes. The concentration of the injected electrons shows an exponential decay. The average L_n penetration depth of the electrons is called *diffusion length*—see Fig. 2.8.

It is easy to understand that carriers will diffuse from the region where their concentration is highest to the region where it is lowest. The diffusion coefficients D_n and D_p are used to represent the diffusion ability for electrons and holes, respectively. Higher diffusion coefficient and longer average *carrier lifetime* result in longer diffusion length. For the electrons:

$$L_n = \sqrt{\tau_n D_n} \quad (2.3)$$

Fig. 2.8 Illustration of carrier diffusion



where τ_n denotes the average lifetime of the minority electrons in the p-type material. This lifetime depends on the net generation-recombination rate which will be discussed later. (Similar relationship between lifetime, diffusion coefficient, and diffusion length holds for the holes.)

As charge carriers diffuse from the high concentration regions to the low concentration regions along the concentration gradient, they form an electric current referred to as *diffusion current*. The diffusion current density of the electrons (for the sake of simplicity assuming only one-dimensional flow as in case in Fig. 2.8) could be expressed as:

$$J_n = qD_n \frac{dn(x)}{dx} \quad (2.4)$$

Similarly, we could obtain the diffusion current density of the holes:

$$J_p = -qD_p \frac{dp(x)}{dx} \quad (2.5)$$

In the expressions above, n and p are the densities (concentrations) of electrons and holes, respectively, q is the unit charge.

2.4.1.2 Drift Current

When an external voltage is applied on the semiconductor, the corresponding electric field would also affect the random movement of electrons: in average they will drift along the opposite direction of the electric field. The so-called *mobility* is used to represent the ability of the charge carriers to be drifted by the electric field, denoted by μ_n and μ_p for the electrons and holes, respectively (representing the average drift velocity of electrons and holes per unit electric field). The macroscopic electric current associated with this is called *drift current*. The drift current density for electrons is:

$$J_n = qn\mu_n E \quad (2.6)$$



<http://www.springer.com/978-1-4614-5090-0>

Thermal Management for LED Applications

Lasance, C.J.M.; Poppe, A. (Eds.)

2014, X, 551 p. 215 illus., 173 illus. in color., Hardcover

ISBN: 978-1-4614-5090-0